

Association for Information Systems

## AIS Electronic Library (AISeL)

---

ICEB 2016 Proceedings

International Conference on Electronic Business  
(ICEB)

---

Winter 12-4-2016

# Internet Finance: A Systematic Literature Review and Bibliometric Analysis

Xiaobei Liang

Tongji University, China, [liangxiaobei@tongji.edu.cn](mailto:liangxiaobei@tongji.edu.cn)

Jiani Wang

Tongji University, China, [wangjiani@tongji.edu.cn](mailto:wangjiani@tongji.edu.cn)

Follow this and additional works at: <https://aisel.aisnet.org/iceb2016>

---

### Recommended Citation

Liang, Xiaobei and Wang, Jiani, "Internet Finance: A Systematic Literature Review and Bibliometric Analysis" (2016). *ICEB 2016 Proceedings*. 38.

<https://aisel.aisnet.org/iceb2016/38>

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2016 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# Internet Finance: A Systematic Literature Review and Bibliometric Analysis

Xiaobei Liang, Tongji University, China, liangxiaobei@tongji.edu.cn

Yibo Yang\*, Tongji University, China, 1yangyibo@tongji.edu.cn

Jiani Wang, Tongji University, China, wangjiani@tongji.edu.cn

## ABSTRACT

Internet finance has gained growing popularity in internet plus environment. While various problems have emerged, and hindered the sustainable growth of internet finance industry. Thus, a summary of existent research and directions for future study are expected. However, few comprehensive literature reviews has been published. This paper presents a thorough bibliometric and network analysis following a systematic literature review methodology. The analysis begins by identifying 331 published studies in Web of Science. Prolific authors, institutions and nations are identified by rigorous bibliometric tools. Based on citation and co-citation analysis, influential papers from different time periods are identified. Established and emergent research clusters are identified for topological analysis by coupling analysis. Future research opportunities are pointed out.

*Keywords:* internet finance; P2P lending; peer-to-peer lending; crowdfunding; social lending.

## INTRODUCTION

Internet finance is the application internet in finance industry. It heralds an era in which all market participants can borrow and lend directly on the Internet with few information barriers. This emergent academic field has been growing for about 8 years and is well into its future of investigation. To correspond to a global trend, the notion of Internet finance has received more attention in the past few years and obtained emerging clusters of research in this field.

Research on internet finance contains two major parts: crowdfunding and peer-to-peer (P2P) lending. Crowdfunding is a new financing channel for small-sized and medium-sized enterprises and individual entrepreneurs to raise funds for innovation projects online [64]. And it has received a great deal of attention as a promising avenue to fostering entrepreneurship and innovation[8]. Online peer-to-peer (P2P) lending is a new but essential financing method for small and micro enterprises that is conducted on the Internet and excludes the involvement of collateral and financial institutions [13]. While the regulation and fraud problems are severely criticized recently.

A number of literature reviews on internet finance have been completed[10] [41] [1]. While, some of these reviews focus on specific aspect of internet finance such as crowdfunding[41] and P2P lending[1]. Cao, Chang [10] conducted a review on Chinese database. Additional analysis of this literature using strict bibliometric tools can provide further insights not previously fully grasped or evaluated.

Network analysis through bibliometric tools can prove powerful for identifying established and emerging topical areas. Besides, it can assist to identify the clusters of research and researchers showing how the various areas of thought may have emerged based on author and institutional characteristics. Identifying the more influential researchers within the clusters sets the stage for determining additional emergent study fields through capturing of more recent topics covered by these researchers. This paper presents a comprehensive evaluation of the field, focusing on forward Internet finance, starting with a pool of over 380 published studies and filtering this pool to more influential works and investigators. Using rigorous bibliometric tools, a comprehensive network analysis (e.g. citation, co-citation analyses and coupling analyses) is completed and six major clusters of research are identified. These algorithmically identified clusters set the stage for topical classification of the published models and further investigation of the evolution of these clusters over years. From these results, additional insights are also gained on the current research interests and potential directions for future research.

The remainder of the paper begins with an introduction of a structured methodology used to identify and further refine the literature that will be reviewed and evaluated in this study. Influential authors, affiliations and countries are identified by a bibliometric analysis in section 3. Section 4 presents research clusters and lead papers in internet finance through network analyses. A further dynamic investigation is conducted in section 5. Section 6 summarizes the results, presents some limitations of this study, and discusses opportunities for future research.

## RESEARCH METHODOLOGY AND INITIAL DATA STATISTICS

Literature reviews aim to evaluate and map the intellectual structure to identify research gaps and highlight the boundaries of knowledge[56]. Structured literature reviews typically start from an iterative cycle of defining appropriate search keywords[51]. Rowley and Slack proposed a structured methodology for scanning resources, designing the mind map to structure the literature review, writing the study and building the bibliography[50]. Following this structured methodology, we combine it with bibliometric analysis and network analysis to carry an iterative cycle of defining search terms, identify influential publications, determine existing topical areas of research and provide insights for future research.

---

\*Yibo Yang, Tongji University, corresponding author

### Defining the Appropriate Search Keywords

A research field usually has many relevant keywords. Traditionally, researchers collect important keywords by iteratively reading various papers. In this study, bibliometric tool is used to facilitate the iterative cycle of defining appropriate keywords and a five steps method is provided. First we use the original keywords that is "internet finance" to download the reference information in web of science. Second bibliometric tool is used to extract all the other keywords in that data set. Third, keywords are sorted by frequency. Fourth, we add all the other relevant keywords in our new search terms. Fifth, we use the new search terms to download new reference information. Sixth, we redo step two to four until that there is no more relevant keywords. The whole process give us a comprehensive search term that is "internet finance" or "P2P lending" or "peer-to-peer lending" or "crowdfunding" or "social lending".

### Database and Refinement of Data

We adopt Web of Science (WOS) core collection as the data source in this paper. WOS, covering over 10,000 high impact journals, as well as over 120,000 international conference proceedings, is the world's leading citation database. Thus, the reference and publication information in WOS is representative of the whole academic field in bibliometric analysis.

Using the search term from last step, we get 380 publications in our database. We further refine the data to articles, proceeding papers and reviews. This refinement give us a final literature set of 331 publications for further analysis.

### Initial Data Statistics

Fig. 1 indicates the trend in quantity of articles published. The results demonstrate that a geometric growth in publications is occurring, while the field is still in its early growth and expansion period. The initial statistics from table 1 show that 248 journals or sources have contributed to the publication of 331 papers. It was found that New media & Society has published 10 of these identified articles, ranked first in all journals or sources. But, no journal has dominated that list.

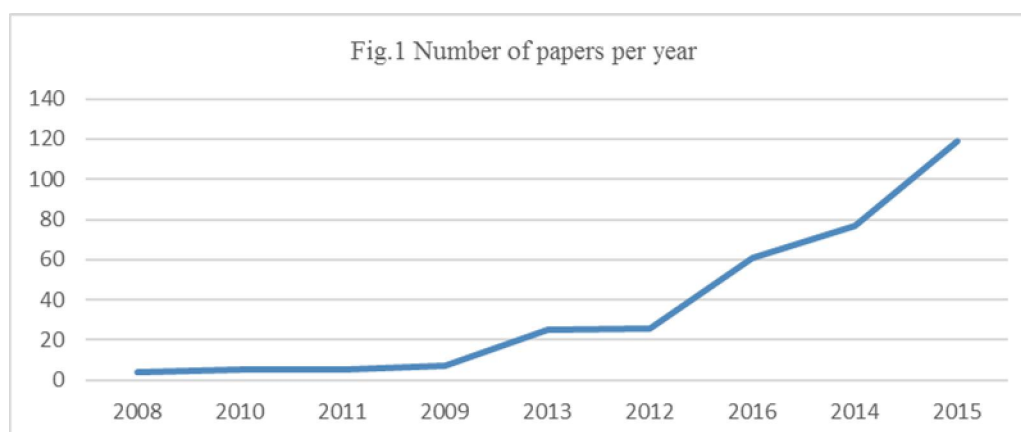


Table 1: Top 10 Journals or sources

No. of papers	Journals or sources
10	NEW MEDIA & SOCIETY
6	CALIFORNIA MANAGEMENT REVIEW
6	PROCEEDINGS OF THE 6TH (2014) INTERNATIONAL CONFERENCE ON FINANCIAL RISK AND CORPORATE FINANCE MANAGEMENT
5	MANAGEMENT SCIENCE
4	ELECTRONIC COMMERCE RESEARCH AND APPLICATIONS
4	2016 49TH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS)
4	FORBES
4	HISTORIA Y COMUNICACION SOCIAL
4	PROCEEDINGS OF THE 17TH INTERNATIONAL ACADEMIC MINDTREK CONFERENCE: MAKING SENSE OF CONVERGING MEDIA
4	ENTREPRENEURSHIP THEORY AND PRACTICE

### Data Analysis

Given the nature of a citation analysis study, we adopt an inductive approach for the purpose of data analysis[53]. The literature classification portion of this study is completed before the actual data analysis by using a deductive approach. Data analysis is composed of two parts including a "bibliometric analysis" and a "network analysis".

### BIBLIOMETRIC ANALYSIS

Bibliometric, introduced by Alan Pritchard in 1969, is an application of mathematics and statistical methods to media of communication[47]. It has been adopted in various management fields to analyze the scientific publications[63];

Ramos-Rodriguez and Ruiz-Navarro[49]; Pilkington[44]; Chai[11]; Adrian Kovacs[32]. The major assumption of bibliometric analysis on citations is that authors tend to cite those publications which contribute to their current work, thus the frequency of received citations can reflect the relevant importance of an article, a topic or a journal[26]; White and Griffith[62].

In this paper we choose BibExcel to carry the bibliometric analysis because of its compatibility with different computer applications including Excel, Ucinet, Pajek and Gephi[43]. It also has high degree of flexibility with reference data imported from various databases including Web of Science and Scopus. One drawback of Bibexcel is its relevant fixed functions. To make state of art bibliometric analysis, we use also use python to supplement Bibexcel.

### Author Influence

BibExcel can be used to analyze the frequency of occurrence of a text in different fields of the bibliographic data. The author field was extracted from the data file and frequency of appearance of all authors was recorded. Table 2 outlines the top ten contributing authors and the quantity of papers they authored or co-authored. And Table3 outlines the top paired authors. As can be seen in this Table2 and Table3, Chen,DY I is leading the list but has not coauthored often. It should be noted that Burtch, G, Wattal, S and Ghose, A have coauthored tightly with each other.

Table 2: The top 10 contributing authors

Author	No. of papers	Author	No. of papers
Chen, DY	6	Dholakia, UM	3
Lin, ZX	4	Lasrado, LA	3
Ghose, A	4	Valanciene, L	3
Zheng, HC	4	Jegeleviciute, S	3
Xu, Y	4	Rossi-Lamastra, C	3

Table 3: Top 4 paired authors

No. of cooperation	Authors	
3	Herzenstein, M	Dholakia, UM
3	Burtch, G	Wattal, S
3	Ghose, A	Wattal, S
3	Burtch, G	Ghose, A

### Affiliation Statistics

The affiliations of the authors were extracted from the TXT file and analyzed in Bib Excel. For each affiliation, the country where the organization is located was extracted for this analysis. As the origin of the Internet finance, China dominates the contributing country list followed by USA and Germany.

Table 4: Top 10 contributing countries

Region	No. of papers
China	220
USA	214
Germany	43
France	34
Canada	31
Spain	30
England	27
Italy	26
Finland	20
Israel	17

The top contributing organizations, based on geographical locations and the quantity of published papers are shown in Table 4. The geographical dispersion of these organizations demonstrates that affiliated institutions whose researches focusing on internet finance are mainly located in China and USA. The top three contributing organizations are International Space University, Southwestern University of Finance and Economics and Northwestern University as shown in Table 5.

Table 5: Top 10 affiliations

<b>Institutions</b>	<b>No. of papers</b>
International Space University	25
Southwestern University of Finance and Economics	19
Northwestern University	13
University of Turin	12
Tampere University of Technology	11
Politecnico di Milano	10
Dalian University of Technology	9
University of Toronto	9
Guizhou University of Finance and Economics	8
University of Delaware	8

Comparing this list with the top 10 contributing authors in Table 4, we observe that International Space University, Southwestern University of Finance and Economics, Northwestern University, University of Turin, Tampere University of Technology and Politecnico di Milano are represented respectively by the more prolific authors Chen, Lin, Ghose, Zheng, Xu and Dholakia. Thus, it may take the work of more than two researchers for an organization to be ranked as a top performer.

### Keywords Statistics

A similar analysis is conducted to identify the most frequently used words/phrases in the paper title and the list of keywords. The top 20 keywords used in the paper titles are summarized in Table 6. Similarly, Table 7 shows the most popular keywords from the list of keywords. A comparison between the two tables indicates that in most cases there is a consistency in the use of keywords in the title and the list of keywords. For example, the top keywords in both tables include a combination of Crowdfunding, Internet, Finance, P2P and Peer-to-Peer. Obviously, the five most popular words in Table 7 occur because they were the search keywords chosen in this study.

Table 6: Top 20 words in the title

<b>Words</b>	<b>Frequency</b>	<b>Words</b>	<b>Frequency</b>
Crowdfunding	133	model	21
Lending	61	analysis	20
Internet	42	financial	19
Finance	34	New	19
Online	34	Financing	18
Social	31	Platform	17
Research	27	China	17
P2P	27	Development	15
Peer-to-Peer	26	Based	15
Study	24	Evidence	15

Table 7: Top 20 words in the keywords

<b>Keywords</b>	<b>Frequency</b>	<b>Keywords</b>	<b>Frequency</b>
Crowdfunding	104	entrepreneurial finance	7
internet finance	33	Social networks	7
peer-to-peer lending	29	trust	6
Crowdsourcing	21	venture capital	6
P2P lending	13	Social Capital	6
entrepreneurship	12	motivation	6
Microfinance	11	Internet	5
Innovation	9	Financing	5
Social Lending	8	Kickstarter	5
social media	7	entrepreneur	5

### NETWORK ANALYSIS AND LITERATURE CLASSIFICATION

Social network analysis (SNA) is the process of investigating social structures through the use of network and graph theories.

Social networks are comprised of interactions for the exchange of resources. It involves nodes and the links or ties. Pieters and Baumgartner point out that “journal citation networks are specific social networks in which actors are journals, the valued resources are knowledge and ideas, and the interactions are citations between journals”. This method has been adopted in various management fields to analyze the scientific publications.

We use Gephi software to perform a network analysis and topical content-based classification of the existing literature of internet finance. Gephi is chosen for this study over the existing network analysis software such as Pajek[19] and VOSviewer[57] due to its easy and broad access to network data, visualization flexibility (editable and user-friendly environment), advanced filtering capabilities, specialized clustering of data, ability to work with different data formats, and several built-in network analysis toolboxes[2].

In order to use Gephi to map and visualize the citations among papers, a graph dataset needs to be generated and published papers are shown as nodes and citations are represented by the arcs/edges between the nodes. The bibliographic data is reformatted with BibExcel to represent a graph dataset, which is in “.NET” format, for Gephi analysis. In this study, the graph dataset is based on coupling analysis.

### Citation Analysis

A citation analysis is used to examine the degree of connectivity between pairs of nodes/papers in the created 331-node network. The initial citation analysis in Gephi revealed that 253 papers out of 331 have cited at least one same reference. The top 10 papers based on the number of local citations are shown in Table 8. 'Local citation' refers to the number of times a paper has been cited by other papers within the 331-paper network, and 'global citation' is the overall WOS citations for the paper. The gap between local citation and global citation values in Table 8 indicates that Internet finance has also received attention from researchers in other disciplines (i.e. they have also been cited by papers not appearing in our search results). It can also be seen that the order of papers based upon local citation does not necessarily match the global citation order. For example Pope DG is ranked 6th based on the number of local citations, but has the second highest global citation count. This result is interesting since the citation references for Pope DG seem to be more popular outside the realm of Internet finance.

Table 8: Top 10 local papers by citations

Author, Year, Journal	Local citation	Global citation
Mollick E, 2014, J BUS VENTURING	63	84
Belleflamme P, 2014, J BUS VENTURING	34	40
Zhang JJ, 2012, MANAGE SCI	25	39
Burtch G, 2013, INFORM SYST RES	25	26
Lin MF, 2013, MANAGE SCI	22	32
Pope DG, 2011, J HUM RESOUR	18	40
Herzenstein M, 2011, J INTERACT MARK	14	17
Yum H, 2012, ELECTRON COMMERCIAL	13	13
Duarte J, 2012, REV FINANC STUD	13	27
Herzenstein M, 2011, J MARKETING RES	13	9

### Pagerank Analysis

Different methods have been used in the past to measure the significance of a paper. The most common method is to count the number of citations which was presented in table 8. Ding et al. argue that in addition to popularity of a paper measured by the number of citations, prestige is another important indicator which is the number of times a paper is cited by highly cited papers. A highly-cited paper may not necessarily be a prestigious paper, although in some cases there might be a strong positive correlation between the two measures. PageRank is used as a measure for both popularity and prestige. See more details about PageRank in Fahimnia, Sarkis [24]

Table 10 shows the top papers using a PageRank measure. For this study's 331-node network, PageRank values vary between zero and 0.08808489. It can be seen that a higher number of local and global citations cannot guarantee the 'prestige' of a paper. Comparing the top ten papers based on citation and PageRank (Table 9 versus Table 10), we see that two of the highly-cited papers are not among the top ten high-PageRank papers. These include Yum H and Duarte J. In return, Chan C (and Munt A appear in the top 10 papers with highest PageRank, but they are not high-cited papers. Given that PageRank is greatly influenced by the citations from other high-cited papers, the later-published papers have obviously had a limited chance to be cited by other high-cited papers (simply because they have not been around long enough to establish citation). As the field matures and stabilizes in terms of output, PageRank will likely give a better picture of overall prestige of publications.

Table 9: Top 10 papers by Pagerank

Author, Year, Journal	Pagerank	Local citation	Global citation
Mollick E, 2014, J BUS VENTURING	0.08808489	63	84
Zhang JJ, 2012, MANAGE SCI	0.04843116	25	39
Pope DG, 2011, J HUM RESOUR	0.04571602	18	40
Herzenstein M, 2011, J INTERACT MARK	0.04506318	14	17
Burtch G, 2013, INFORM SYST RES	0.03090148	25	26
Belleflamme P, 2014, J BUS VENTURING	0.02606964	34	40
Herzenstein M, 2011, J MARKETING RES	0.02386747	13	9
Chan C, 2016, RES SOCIAL WORK PRAC	0.01775401	1	0
Munt A, 2014, J SCREENWRITING	0.01775401	1	1
Lin MF, 2013, MANAGE SCI	0.01646502	22	32

### Coupling Analysis

There are two bibliometric techniques that build on the analysis of authors' referencing behavior—co-citation analysis and bibliographic coupling. Accordingly, these techniques rely on the fundamental assumption that, to the extent that the cited references of a focal publication provide a background for this publication, they denote at least some degree of relatedness between the focal publication and the publication(s) to which it refers. Based on this assumption, citation-based indicators have been widely applied in the field of bibliometrics as a means of mapping the flow of science and the development of fields and communities[31] [60] [58]. The basic assumption underlying both bibliographic coupling and co-citation analysis is that the greater the degree of overlap in the referencing patterns of a pair of focal publications, the greater the relatedness between both publications.

We choose coupling technique to calculate the relatedness matrix for further network analysis for two reasons. First, it is more suitable to identify clusters or themes within internet finance research. While co-citation is better at identifying the knowledge base of focal research field[32]. Second, reference information is available for all publications but not all focal publications are cited especially in local citing environment on which co-citation relies. Table 10 is a sample of coupling based relatedness matrix. From the sample matrix, we can tell that article 1 and article 3 cite three same publications and the references of article 1 and 2 share no common. A python script and Bibexcel are used do make the relatedness matrix of 311 publications. Importation of the matrix to Gephi shows that 253 out of 311 publications share certain degree of references.

Table 10: Sample relatedness matrix

	Article 1	Article 2	Article 3
Article 1		0	3
Article 2	0		8
Article 3	3	8	

### Literature Classification: Data Clustering

The nodes of a network can be divided into 'clusters' or 'modules' where the connection (density of edges) is greater between the nodes of the same cluster compared to those of different clusters[17] [35] [48]. In a coupling network, a cluster can be seen as a group of publications that base on more similar knowledge base than that in other clusters or research areas. Data clustering (also termed modularity) has been used in the past as a classification tool for grouping of a set of given publications[48]. It allows for the topological analysis of a co-citation network, identifying topics, interrelations, and collaboration patterns. Data clustering has received increasing attention from scholars and research organizations turning it into a critical research field in social network analysis[4].

The default clustering tool in Gephi is based on the Louvain algorithm, an iterative optimization model that aims to determine the optimal number of partitions that maximize the modularity index. It outperform other community detection methods in terms of computation time and its accuracy is also verified. The modularity index of a partition is a scalar value between  $-1$  and  $+1$  that measures the density of links inside communities versus the links between communities. See more detail about the mathematic function of Louvain algorithm in Gephi in [4]

### Clustering result

Applying this algorithm to the 311-node coupling network in Gephi resulted in the creation of six major clusters. The modularity index in Fig. 4 is equal to 0.41 indicating strong relationship between the nodes within each cluster and yet a relatively strong relationship between the nodes of different clusters. To determine the research focus for each cluster, we need to identify the "lead papers" (top 10 papers) of each cluster. This is a common practice of other bibliometric analysis papers[24]. A PageRank measure was used for this purpose. In a co-citation network, the PageRank algorithm takes into account how many times a paper is co-cited with other papers (the 'popularity' measure) and how many times it is co-cited

with highly co-cited papers (the ‘prestige’ measure). Most of the papers in this study with high PageRank also possess a high citation count. To find out the area of research focus of each cluster, we analyzed and evaluated the contents of all lead papers of each cluster.

Although the clustering algorithm is accurate to a relative high degree, the clustering result is not perfect because of the unordinary refereeing behavior of some authors and the drawback of bibliometric analysis itself. To have a better clustering result, we conduct a deductive method based on the research focus from last step to modify clusters. Specifically, we manually check the title and abstracts of every other papers in each cluster to see whether it conforms to the research focus. If a paper is clustered by error, it will be changed to a cluster that most suits the nature of this paper.

Table 11: Major clusters and their areas of focus

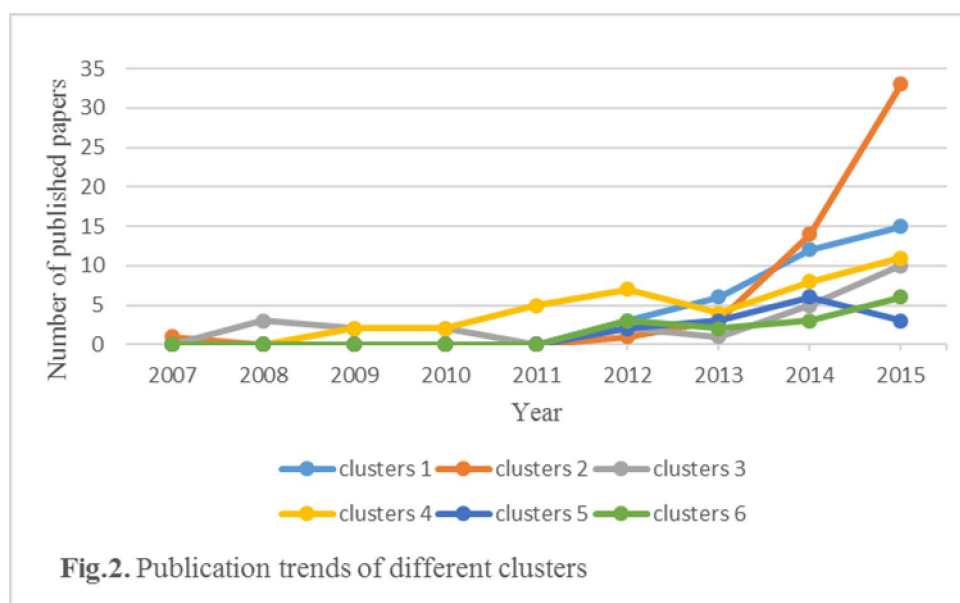
Cluster	No. of paper	Area of research focus
1	42	Application of crowdfunding in different industries and countries
2	87	Mechanisms of crowdfunders and fundraisers’ behaviors Broad-based empirical studies
3	29	Credit and risk problems in P2P lending Solve practical problems
4	45	Mechanism of lenders and borrowers’ behaviors in P2P lending
5	14	Regulation and policy issues in internet finance
6	16	Crowdfunding in media industry

The topical literature classification summarized in Table 11 shows that the research field of internet finance mainly focuses on crowdfunding (cluster 1, 2 and 6) and Peer-to-peer lending (cluster 3 and 4). Regulation and law is also a major concern of this field. However, compared to its importance, research on regulation of internet finance is still limited. Specifically, cluster one explores the possible application of crowdfunding. For example crowdfunding in scientific research [61] [29] [9], health industry [42] [20] [37] and different countries [33] [23]. Cluster two mainly focuses on mechanism of crowdfunders and fundraisers’ behaviors. Most of the research in this cluster are empirical studies. And the use of data in Proser.com is popular in this cluster. Sub research fields include geography effect [40] [39], social influence [7], behaviors in different types of crowdfunding [16] and so on. Cluster three mainly focuses on solving practical problems in P2P lending such as bidding in auction [22] [15] and credit assessing [38] [65]. Cluster four is about the mechanism of lenders and borrowers’ behaviors. Two major sub research fields are herding effect [66] [34] [12] and factors in lending decision such as demography effect [45], appearance [21] and social influence [36]. Cluster five is mainly around the Jumpstart Our Business Startups (JOBS) Act signed by Obama in 2012. Other regulations and supervision advice are also suggested for crowdfunding and P2P lending. It is noteworthy that although Chinese researcher dominate the author list, there is only one paper by Chinese researcher in this cluster. Compared with the numerous runaway events in China’s P2P industry, academic research on regulations is surprisingly lacking at least in English database. Cluster six is about the application of crowdfunding in media industry.

#### DYNAMIC ANALYSIS

To understand the evolution of internet finance research field over time, we also complete a dynamic citation analysis for the papers of all clusters. First, the growth of clusters over time for each cluster will be shown. Second, the lead papers of each clusters during three time periods will be summarized to reflect the research trend. Third, the citing pattern will be provided to help researchers to grasp the knowledge foundation of their focused research field.





As shown in figure 2, research on P2P lending (cluster 3 and 4) started earlier than that on crowd funding (cluster 1 and 2). After 2013, the growth of papers about crowd funding has surpassed that about P2P lending. Research on mechanism of participants' behavior is more popular in internet finance. Overall, all clusters have shown a growing trend except the regulation and policy research (cluster 5).

### Emergent Research Field

After a careful check of every lead papers in table 13 we find that in cluster one, research on the application of crowdfunding in scientific research [61] [9] covers both the two time periods, suggesting that researchers' interest on applying crowdfunding to raise funds lasts. New research focus include crowdfunding in corporate finance [55] and possible application of epidemic model in crowdfunding [30]. In cluster two, research focus seems to change from studying the behaviors of crowdfunders to that of entrepreneurs [40] and the underlying successful factors from the side of fundraisers [3].

In the initial stage of P2P lending, much efforts were given to research on how to build trust [14] and on the mechanism of the bidding in P2P auction [27]. Then, research on how to facilitate philanthropic donation in P2P emerged. Currently, assessing credit risk, bidding strategy and donation in P2P are three major sub research focus within cluster three. In cluster four, most attention were paid to study influence of information system [59], reputation of fundraisers [18] and type of disclosed information on crowd funders' trust [46]. During 2011 and 2013, study on the herding effect in P2P lending emerged [28] [66]. Most recently, the influence of quality of disclosed information and social network effect on lenders' trust are most studied.

In cluster five, most papers studied potential impact of the JOBS act which legalizes equity crowdfunding from 2011 to 2013. In latest time period, effect of regulations on crowdfunding in solar energy industry [6] and in gender/sexual reassignment surgeries are studied [25]. In cluster six, early papers are mainly about the use of crowdfunding on new media products. Now, much attentions are paid to leverage fans in crowdfunding new digital products [5] [52].

### Evolutionary Citing Pattern

The progress of knowledge inflows to each clusters can be inferred from those most cited papers over time. As shown in table 13, citing patterns of cluster 1 to 5 have been substantially changing. It indicates that internet finance field lacks leading authors and this research field is more propelled by business practice than by theory. Influential papers for crowdfunding share little similarity with that for P2P lending, suggesting an opportunity to leverage the findings from the other field to their own research field. Researchers who want to have a grasp of the knowledge foundation of the focused research cluster can use those top references as a convenient start.

### CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

This paper uses thorough bibliometric and network analysis to present a comprehensive and structured review on internet finance. 311 articles, proceeding papers and reviews are selected as the publication set through the use of an iteratively defined search term in Web of Science core collection. Adopting different objective measures, our study identified prolific authors, affiliations and countries, classified 6 major research clusters and illustrated the knowledge foundations as well as lead papers for each cluster over time. Researchers can use the findings to find potential coauthors, have a panorama of internet finance, track the research trends and grasp the knowledge foundation on which they can build their own research.

The following key conclusions can be drawn from this study on internet finance: (1) Two major forms of internet finance are crowdfunding and P2P lending (also called social lending). The growing of research on crowdfunding surpassed that of P2P

lending since 2013, which is partially because of the JOBS Act signed in 2012 by president Obama. (2) China and USA are two dominant contributing countries in this field. (3) Papers are sparsely distributed in various journals. No leading journals has appeared yet. (4) Internet finance research can be classified into six different categories from which we recognize that cluster one is about application of crowdfunding in different industries and countries; Cluster two focuses on mechanisms of crowd funders and fundraisers' behaviors; Cluster three pay more attention on solving practical problems in P2P lending, such as assessing credit risk and bidding strategy; Cluster four is about mechanism of lenders and borrowers' behaviors in P2P lending; Cluster five focus on policy and regulation; Cluster six cares about use crowdfunding in media industry.

Based on the analysis, we propose a key future research direction for internet finance. First, much efforts have been made to study the behaviors of lenders and crowd funders. However, a relative small number of research are about the behaviors of borrowers and fundraisers. Future research can pay more attention on the detection of potential credit default and runaway of P2P platform. Second, since herding effect is significant in P2P lending, a similar effect can be tested in crowdfunding platforms. Third, emergent interdisciplinary research fields, such as crowdfunding in corporate finance [54] and marketing [22] could be further explored. Fourth, although China is the dominant country in this field, research on regulations and policy of internet finance is few in English database. Intercountry study on regulations of internet finance can fill this gap. Fifth, the effect of appearance of entrepreneurs and the endorsement of celebrity for both fundraisers and platforms can be empirically tested.

There are also several limitations to this study. First, the data source is Web of Science. Although WOS is the biggest citation based academic database available, there are many journals and books excluded by the dataset. Second, bibliometric study relies on the assumption that researchers tend to cite more publications which contribute to their current paper. While it is not always true. Researchers also cite papers which they want to deny or criticize. Third, although coupling analysis maximize the use of reference information in network analysis, a co-citation based Pagerank analysis suffers from insufficient local citation data for latest publications.

#### ACKNOWLEDGEMENT

This work is supported by grant 71672128 of the Program of National Nature Science Foundation of China.

#### REFERENCES

- [1] Bachmann, A., et al., Online Peer-to-Peer Lending -- A Literature Review. *Journal of Internet Banking & Commerce*, 2011. 16(2).
- [2] Bastian, M., S. Heymann, and M. Jacomy, Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 2009. 8: p. 361-362.
- [3] Belleflamme, P., T. Lambert, and A. Schwienbacher, Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, 2014. 29(5): p. 585-609.
- [4] Blondel, V.D., et al., Fast unfolding of communities in large networks. *Journal of Statistical Mechanics-Theory and Experiment*, 2008.
- [5] Booth, P., Crowdfunding: A Spimatic application of digital fandom. *New Media & Society*, 2015. 17(2): p. 149-166.
- [6] Booth, S., Here Comes the Sun: How Securities Regulations Cast a Shadow on the Growth of Community Solar in the United States. *Ucla Law Review*, 2014. 61(3): p. 760-811.
- [7] Burtch, G., A. Ghose, and S. Wattal, An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets. *Information Systems Research*, 2013. 24(3): p. 499-519.
- [8] Burtch, G., A. Ghose, and S. Wattal, The Hidden Cost of Accommodating Crowdfunder Privacy Preferences: A Randomized Field Experiment. *Management Science*, 2015. 61(5): p. 949-962.
- [9] Byrnes, J.E.K., et al., To Crowdfund Research, Scientists Must Build an Audience for Their Work. *Plos One*, 2014. 9(12): p. 29.
- [10] Cao, S.Y., et al., Pope DGA Review of Current Studies on Internet Finance in China. Thirteenth Wuhan International Conference on E-Business, 2014, ed. J. Seitz. 2014, Alfred: Alfred Univ. 470-476.
- [11] Chai, K.-H. and X. Xiao, Understanding design research: A bibliometric analysis of Design Studies (1996–2010). *Design Studies*, 2012. 33(1): p. 24-43.
- [12] Chen, D.Y., et al., Herding Behavior in Online Microloan Markets: Evidence from China. Fourteenth Wuhan International Conference on E-Business, ed. J. Seitz. 2015, Alfred: Alfred Univ. 215-218.
- [13] Chen, D.Y., F.J. Lai, and Z.X. Lin, A trust model for online peer-to-peer lending: a lender's perspective. *Information Technology & Management*, 2014. 15(4): p. 239-254.
- [14] Chen, N., et al., Social Lending. 10th Acm Conference on Electronic Commerce - Ec 2009. 2009, New York: Assoc Computing Machinery. 335-344.
- [15] Chen, N., A. Ghosh, and N.S. Lambert, Auctions for social lending: A theoretical analysis. *Games and Economic Behavior*, 2014. 86: p. 367-391.
- [16] Cholakova, M. and B. Clarysse, Does the Possibility to Make Equity Investments in Crowdfunding Projects Crowd Out Reward-Based Investments? *Entrepreneurship Theory and Practice*, 2015. 39(1): p. 145-172.
- [17] Clauset, A., M.E.J. Newman, and C. Moore, Finding community structure in very large networks. *Physical Review E*, 2004. 70(6).
- [18] Collier, B., R. Hampshire, and Acm, Sending Mixed Signals: Multilevel Reputation Effects in Peer-to-Peer Lending

- Markets. 2010 Acm Conference on Computer Supported Cooperative Work. 2010, New York: Assoc Computing Machinery. 197-206.
- [19] De Nooy, W., A. Mrvar, and V. Batagelj, Exploratory social network analysis with Pajek. Vol. 27. 2011: Cambridge University Press.
- [20] Dragojlovic, N. and L.D. Lynd, Crowdfunding drug development: the state of play in oncology and rare diseases. *Drug Discovery Today*, 2014. 19(11): p. 1775-1780.
- [21] Duarte, J., S. Siegel, and L. Young, Trust and Credit: The Role of Appearance in Peer-to-peer Lending. *Review of Financial Studies*, 2012. 25(8): p. 2455-2483.
- [22] Ducarroz, C., S. Yang, and E.A. Greenleaf, Understanding the Impact of In-Process Promotional Messages: An Application to Online Auctions. *Journal of Marketing*, 2016. 80(2): p. 80-100.
- [23] Dushnitsky, G., et al., Crowdfunding in Europe: DETERMINANTS OF PLATFORM CREATION ACROSS COUNTRIES. *California Management Review*, 2016. 58(2): p. 44-71.
- [24] Fahimnia, B., J. Sarkis, and H. Davarzani, Green supply chain management: A review and bibliometric analysis. *International Journal of Production Economics*, 2015. 162: p. 101-114.
- [25] Farnel, M., Kickstarting trans\*: The crowdfunding of gender/sexual reassignment surgeries. *New Media & Society*, 2015. 17(2): p. 215-230.
- [26] Garfield, E., Citation indexes in sociological and historical research. *American documentation*, 1963. 14(4): p. 289-291.
- [27] Greiner, M.E. and H. Wang, Building Consumer-to-Consumer Trust in E-Finance Marketplaces: An Empirical Analysis. *International Journal of Electronic Commerce*, 2010. 15(2): p. 105-136.
- [28] Herzenstein, M., U.M. Dholakia, and R.L. Andrews, Strategic Herding Behavior in Peer-to-Peer Loan Auctions. *Journal of Interactive Marketing*, 2011. 25(1): p. 27-36.
- [29] Hui, J.S., E.M. Gerber, and Acm, Crowdfunding Science: Sharing Research with an Extended Audience. *Proceedings of the 2015 Acm International Conference on Computer-Supported Cooperative Work and Social Computing (Cscw'15)*, 2015: p. 31-43.
- [30] Kandhway, K. and J. Kuri, Optimal control of information epidemics modeled as Maki Thompson rumors. *Communications in Nonlinear Science and Numerical Simulation*, 2014. 19(12): p. 4135-4147.
- [31] Kessler, M.M., Bibliographic coupling between scientific papers. *American documentation*, 1963. 14(1): p. 10-25.
- [32] Kovacs, A., B. Van Looy, and B. Cassiman, Exploring the scope of open innovation: a bibliometric review of a decade of research. *Scientometrics*, 2015. 104(3): p. 951-983.
- [33] Lasrado, L.A. and A. Lugmayr, Crowdfunding in Finland - A New Alternative Disruptive Funding Instrument for Businesses. *Proceedings of the 17th International Academic Mindtrek Conference: Making Sense of Converging Media*, 2013: p. 194-201.
- [34] Lee, E. and B. Lee, Herding behavior in online P2P lending: An empirical investigation. *Electronic Commerce Research and Applications*, 2012. 11(5): p. 495-503.
- [35] Leydesdorff, L., et al., Turning the tables on citation analysis one more time: Principles for comparing sets of documents. *Journal of the American Society for Information Science and Technology*, 2011. 62(7): p. 1370-1381.
- [36] Liu, D., et al., FRIENDSHIPS IN ONLINE PEER-TO-PEER LENDING: PIPES, PRISMS, AND RELATIONAL HERDING. *Mis Quarterly*, 2015. 39(3): p. 729-+.
- [37] Loucks, D., Will crowdfunding and general solicitation spur orphan drug development for biotechs? *Formulary*, 2013. 48(10): p. 343-344.
- [38] Malekipirbazari, M. and V. Aksakalli, Risk assessment in social lending via random forests. *Expert Systems with Applications*, 2015. 42(10): p. 4621-4631.
- [39] Mendes-Da-Silva, W., et al., The impacts of fundraising periods and geographic distance on financing music production via crowdfunding in Brazil. *Journal of Cultural Economics*, 2016. 40(1): p. 75-99.
- [40] Mollick, E., The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, 2014. 29(1): p. 1-16.
- [41] Moritz, A. and J.H. Block, Crowdfunding: A Literature Review and Research Directions. 2016: Springer International Publishing. 11645-11649.
- [42] Otero, P., Crowdfunding. A new option for funding health projects. *Archivos Argentinos De Pediatría*, 2015. 113(2): p. 154-157.
- [43] Persson, O., R. Danell, and J.W. Schneider, How to use Bibexcel for various types of bibliometric analysis. *Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday*, 2009: p. 9-24.
- [44] Pilkington, A. and J. Meredith, The evolution of the intellectual structure of operations management-1980-2006: A citation/co-citation analysis. *Journal of Operations Management*, 2009. 27(3): p. 185-202.
- [45] Pope, D.G. and J.R. Sydnor, What's in a Picture? Evidence of Discrimination from Prosper.com. *Journal of Human Resources*, 2011. 46(1): p. 53-92.
- [46] Potzsch, S. and R. Bohme, The Role of Soft Information in Trust Building: Evidence from Online Social Lending, in *Trust and Trustworthy Computing, Proceedings*, A. Acquisti, S.W. Smith, and A.R. Sadeghi, Editors. 2010, Springer-Verlag Berlin: Berlin. p. 381-395.
- [47] Pritchard, R.D., Equity theory: A review and critique. *Organizational behavior and human performance*, 1969. 4(2): p. 176-211.
- [48] Radicchi, F., et al., Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 2004. 101(9): p. 2658-2663.
- [49] Ramos-Rodriguez, A.R. and J. Ruiz-Navarro, Changes in the intellectual structure of strategic management research: A

- bibliometric study of the Strategic Management Journal, 1980-2000. *Strategic Management Journal*, 2004. 25(10): p. 981-1004.
- [50] Rowley, J. and F. Slack, Conducting a literature review. *Management Research News*, 2004. 27(6): p. 31-39.
- [51] Saunders, M., P. Lewis, and A. Thornhill, *Research Methods for Business Students*. Harlow, England: Pearson Education. 2009, ISBN 978-0-273-71686-0.
- [52] Scott, S., The moral economy of crowdfunding and the transformative capacity of fan-ancing. *New Media & Society*, 2015. 17(2): p. 167-182.
- [53] Seuring, S. and M. Mueller, From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 2008. 16(15): p. 1699-1710.
- [54] Srenkel, L., M. Smorada, and P. Markovic, Crowdfunding as a Source of Capital, in *Financial Management of Firms and Financial Institutions: 10th International Scientific Conference, Pts I-Iv*, M. Culik, Editor. 2015, Vsb-Tech Univ Ostrava: Ostrava. p. 1285-1290.
- [55] Strokova, Z., Crowdfunding: the new phenomenon in corporate finance? Managing and Modelling of Financial Risks: 7th International Scientific Conference, Pts I-Iii, ed. M. Culik. 2014, Ostrava: Vsb-Tech Univ Ostrava. 723-730.
- [56] Tranfield, D., D. Denyer, and P. Smart, Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 2003. 14(3): p. 207-222.
- [57] Van Eck, N.J., et al., Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PloS one*, 2013. 8(4): p. e62395.
- [58] Vladutz, G. and J. Cook, Bibliographic coupling and subject relatedness. *Proceedings of the American Society for Information Science*, 1984. 21: p. 204-207.
- [59] Wang, H., M. Greiner, and J.E. Aronson, People-to-People Lending: The Emerging E-Commerce Transformation of a Financial Market, in *Value Creation in E-Business Management*, M.L. Nelson, M.J. Shaw, and T.J. Strader, Editors. 2009, Springer-Verlag Berlin: Berlin. p. 182-195.
- [60] Weinberg, B.H., Bibliographic coupling: A review. *Information Storage and Retrieval*, 1974. 10(5): p. 189-196.
- [61] Wheat, R.E., et al., Raising money for scientific research through crowdfunding. *Trends in Ecology & Evolution*, 2013. 28(2): p. 71-72.
- [62] White, H.D. and B.C. Griffith, Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for information Science*, 1981. 32(3): p. 163-171.
- [63] White, H.D. and K.W. McCain, Visualizing a discipline: An author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science*, 1998. 49(4): p. 327-355.
- [64] Xu, B., et al., Configurational paths to sponsor satisfaction in crowdfunding. *Journal of Business Research*, 2016. 69(2): p. 915-927.
- [65] Yao, J.R. and C. Lian, A New Ensemble Model based Support Vector Machine for Credit Assessing. *International Journal of Grid and Distributed Computing*, 2016. 9(6): p. 159-167.
- [66] Zhang, J.J. and P. Liu, Rational Herding in Microloan Markets. *Management Science*, 2012. 58(5): p. 892-912.

## APPENDIX

Table 12 Lead papers in different time periods

	2008-2010	2011-2013	2014-2016
<b>Cluster 1</b>		Wheat RE, 2013, TRENDS ECOL EVOL Perlstien EO, 2013, SPRINGERPLUS Lasrado LA, 2013, PROCEEDINGS OF THE 17TH IAMC: MAKING SENSE OF CONVERGING MEDIA	Kandhway K, 2014, COMMUNONLINE R SCI Strokova Z, 2014, MANAGING AND MODELLING OF FINANCIAL RISKS Bymes JEK, 2014, PLOS ONE
<b>Cluster 2</b>		Burtch G, 2013, INFORM SYST RES Gerber EM, 2013, ACM T COMPUT-HUM INT Schwartz AA, 2013, NOTRE DAME LAW REV	Mollick E, 2014, J BUS VENTURING Belleflamme P, 2014, J BUS VENTURING Jian L, 2014, J COMPUT-MEDIA T COMM
<b>Cluster 3</b>	Greiner ME, 2010, INT J ELECTRON COMM Chen N, 2009, 10TH ACM COEC Serrellis E, 2010, LN INST COMP SCI ISO	Balhaves M, 2012, MEDIA INT AUST Ly P, 2012, VOLUNTA S Chen XH, 2013, ADV MATH PHYS	Chen N, 2014, GAME ECON BEHAV Enekele R, 2015, APPL ECON Corazzini L, 2015, J PUBLIC ECON
<b>Cluster 4</b>	Potzsch S, 2010, LECT NOTES COMPUT SC Collier B, 2010, 2010 ACM COGSCW Wang H, 2009, LECT NOTES BUS INF	Zhang JJ, 2012, MANAGE SCI Pope DG, 2011, J HUM RESOUR Herzenstein M, 2011, J INTERACT MARK	Chen DY, 2014, INFORM TECHNOL MANAG Allison TH, 2015, ENTREPRENEUR THEORY PRACT Liu D, 2015, MIS QUART
<b>Cluster 5</b>		Sigar K, 2012, ADMIN LAW REV Stemler AR, 2013, BUS HORIZONS Bradford CS, 2012, SECUR REGUL LA W J	Murray J, 2015, JASSA Booth S, 2014, UCLA LAW REV Valanciene L, 2014, PROCD SOC BEHV
<b>Cluster 6</b>		Sorensen IE, 2012, MEDIA CULT SOC Martinez Gallardo FD, 2013, HIST COMUN SOC Baranova D, 2013, PROCEEDINGS OF THE 17TH IAMC: MAKING SENSE OF CONVERGING MEDIA	Chan C, 2016, RES SOCIAL WORK PRACT Davies R, 2015, INFORM COMMUN SOC Bennett L, 2015, NEW MEDIA SOC

Table 13 Top cited references

	2008-2010	2011-2013	2014-2016
<b>Cluster one:</b>	Ghatak M, 1999, J DEV ECON Pety R, 1986, ADV EXPT SOCIAL PSYC PETTY RE, 1981, J PERS SOC PSYCHOL PETTY RE, 1984, J PERS SOC PSYCHOL Pavlou PA, 2006, MIS QUART AKERLOF GA, 1970, Q J ECON Mishra DP, 1998, J MARKETING RES Moulton L, 2007, SOCIOL FORUM MEYER T, 2006, FINANCIAL SERVICES 2 MAYER RC, 1995, A CAD MANAGE REV	Surowiecki J, 2004, WISDOM CROWDS WHY MA Lambert T, 2010, EMPIRICAL ANAL CROWD Ordanini A, 2011, J SERV MANAGE Mollick E, 2012, DYNAMICS CROWDFUNDIN Poetz MK, 2012, J PROD INNOVAT MANAG Hazen Thomas Lee, 2012, N C L REV Hanley Dean F, 2012, INSIGHTS Cohn Stuart R, 2007, N Y U J L BUS Hoffman Sheldon Ryan, 2011, TENN L REV Loewenstein Mark J, 1999, DEL J CORP L Ghatak M, 1999, J DEV ECON Morgan J, 2009, MARKETING PROFS 0511 Murphy C, 2012, GUARDIAN Nancy G, 2006, WP0639 IMF I Morduch J, 1999, J ECON LIT Freedman S, 2008, DO SOCIAL NETWORKS S Pope DG, 2011, J HUM RESOUR Ravina E, 2008, LOVE LOANS EFFECT BE lyer R, 2009, SCREENING NEW CREDIT BANERJEE AV, 1992, Q J ECON Fisch Jill E, 1998, J SMALL EMERGING BUS Campbell Jr Ruthford B, 2006, DEL J CORP L Sjostrom Jr William K, 2001, FLA L REV Johnson Jennifer J, 2010, DEL J CORP L [Anonymous], 1982, BUS LAW	Ordanini A, 2011, J SERV MANAGE De K Buyse, 2012, FRAMEWORK EUROPEAN C Cameron P, 2013, GENOME BIOL EISENHARDT KM, 1989, ACAD MANAGE REV Mollick ER, 2013, JOURNAL OF BUSINESS Mollick E, 2014, J BUS VENTURING Belleflamme P, 2014, J BUS VENTURING Burtch G, 2013, INFORM SYST RES Ordanini A, 2011, J SERV MANAGE Zhang JJ, 2012, MANAGE SCI Berger SC, 2009, BUR BUSINESS RES Duarte J, 2012, REV FINANC STUD Lee E, 2012, ELECTRON COMM R A Herzenstein M, 2011, J INTERACT MARK Lin MF, 2013, MANAGE SCI Lin MF, 2013, MANAGE SCI Pope DG, 2011, J HUM RESOUR Yum H, 2012, ELECTRON COMM R A Duarte J, 2012, REV FINANC STUD Herzenstein M, 2011, J MARKETING RES [Anonymous], 2012, RES TECHNOL MANAGE Valanciene L, 2013, EC MANAGEMENT Sigar K, 2012, ADMIN LAW REV MASSON GAETAN, 2012, EUR PHOTOVOLTAIC IND Joint Committee of the Parliamentary Commission on Banking Standards, 2013, CHANG BANK GOOD Jenkins H, 2013, SPREADABLE MEDIA CRE Jenkins H, 2006, CONVERGENCE CULTURE Terranova Tiziana, 2000, SOCIAL TEXT Mittel J, 2013, JUST TV Pebler Luke, 2013, REVENGE OF THE FANS
<b>Cluster two:</b>			
<b>Cluster three:</b>			
<b>Cluster four:</b>			
<b>Cluster five:</b>			
<b>Cluster six:</b>			