The 16th International Society for Business Innovation and Technology Management Conference
January 25, 2015

A multi-method approach to extract social media data

Hing Kai Chan
Associate Professor in Operations Management
Nottingham University Business School China,
University of Nottingham Ningbo China.

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Co-Editor-in-Chief
Industrial Management & Data Systems
Emerald Group Publishing

Emerald
GROUP PUBLISHING

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Outline

- Introduction
- Challenges
- Research model
- Results
- Concluding remarks

Introduction

- Increasing usage of social media since the concept of Web2.0 (Kaplan and Haenlein, 2010)
- Social media sites serve as platforms for sharing and exchanging information, i.e. social media data (Akar and Topçu, 2011)
- Further boosted by the penetration of smartphone
  - Over 800 millions of Facebook users are on mobile devices (Sources: CNET)

Introduction

- Some early research studies:
  - Mangold and Faulds (2009) : changed the promotion mix in marketing communication
  - Xiang and Gretzel (2010) : Travellers are convenient to make use of social media web-sites to search for travelling information online
  - Mostafa (2013): Using text mining to analyse consumer behaviour

Introduction

- Have businesses made full use of such huge data source? (Culnan et al., 2010)
  - "Little research has been done to understand how management should respond to customer reviews in online social media" (Gu and Ye, 2013)
- Online data generated by the end consumer are often qualitative and highly unstructured
  - Multiple research methods are required in response to recent call for data-driven research in OM discipline (Simchi-Levi, 2013)

Challenges

- Difficult to quantify the return of investment in social media web-sites (Hoffman and Fodor, 2010)
- Subjectivity of the data may add additional uncertainty regarding the credibility and persuasiveness of the information being used (Cheung et al., 2009; Zhang et al., 2010)
- Data are highly unstructured

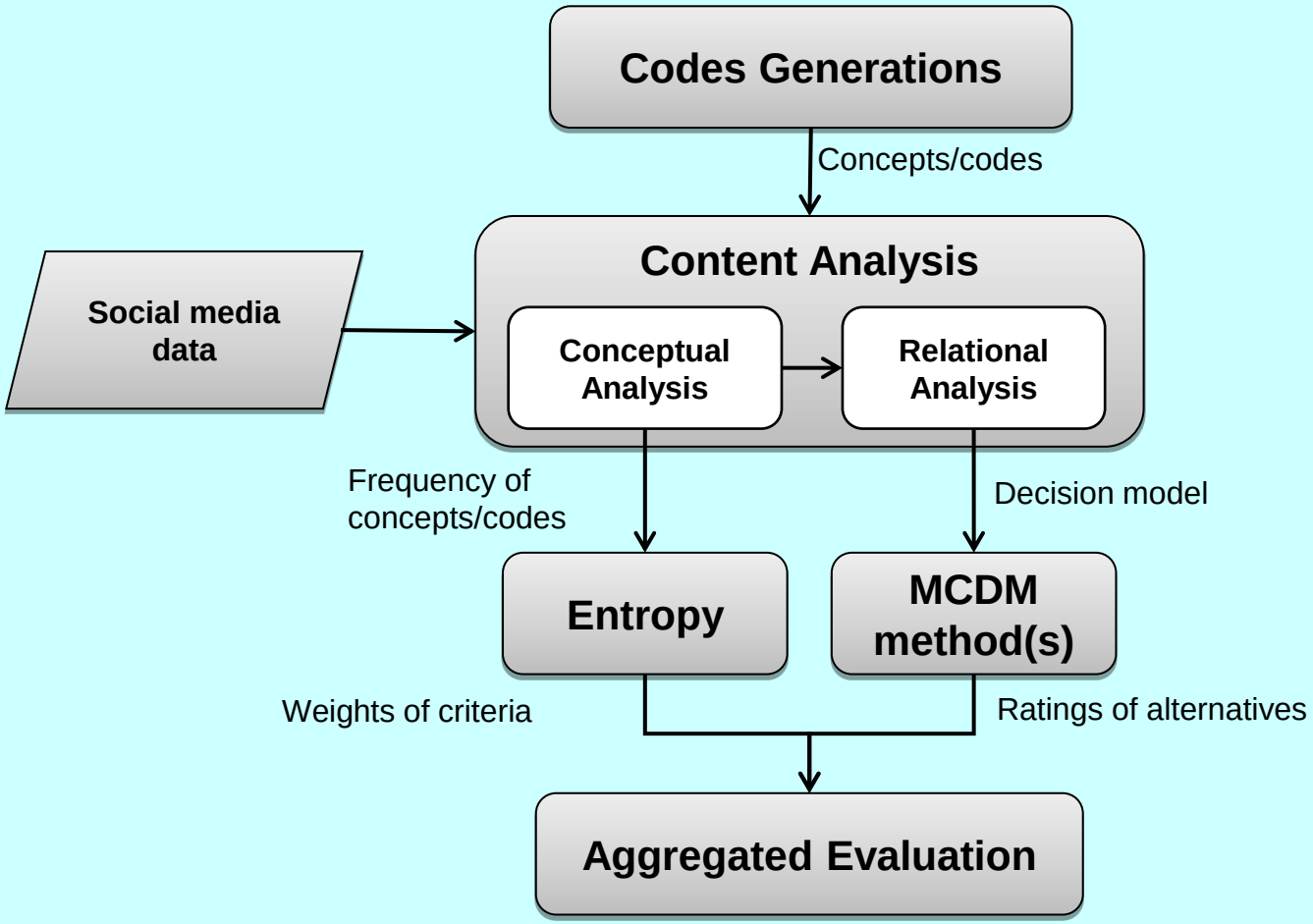
Challenges

- Social media data are not linked to operations management attributes readily
- Hierarchical models can be employed to systematic analysis of social media datasets and group online comments (Anderson and Joglekar, 2005; Tripathy and Eppinger, 2013)
  - defining such a hierarchical model for social media data is more art than science

Research Questions

- How to quantify social media data?
- How to apply the concept to new product development / innovation management?

Research model



Purpose:

- To define the codes for content analysis

- To quantify the codes and identify the relationship within the data set

- To construct the hierarchical model based on the relational analysis, and to calculate the relative weightings for decision-making

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

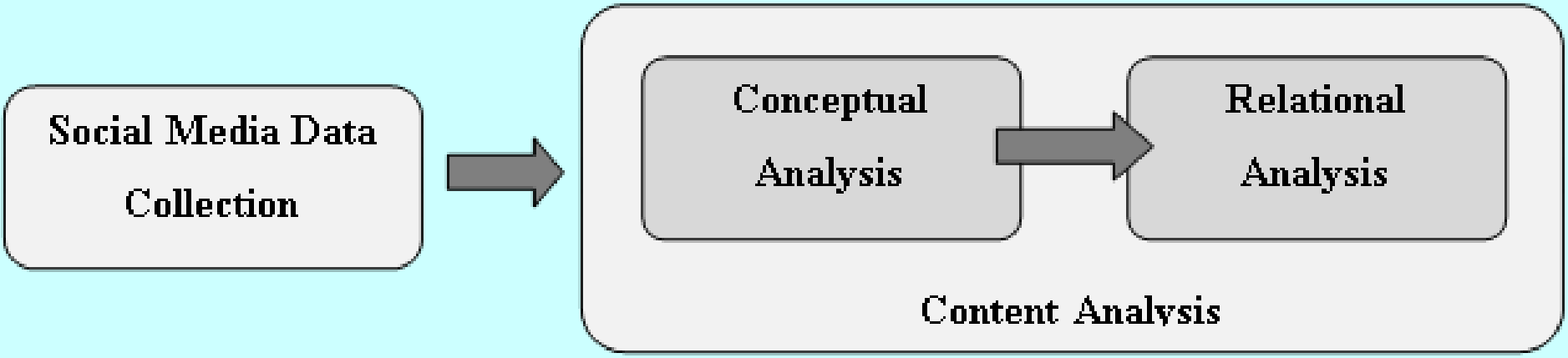
Data collection

- Data for this study was accessed from the SAMSUNG Mobile Facebook page (https://www.facebook.com/SamsungMobile)
- NCapture for NVivo 10 was used
- These data include consumers' comments from 26th September 2013 to 26th of November 2013
- Overall 86055 comments were downloaded

Codes Generation

- Five determinants of new product performance are studied:
  - Product, strategic, development process, organisational, and/or market environment factors (Montoya-Weiss and Calantone 1994; Henard and Szymanski 2001; Cho and Lee 2013)
- Table 1 & Table 2

Content Analysis: Conceptual and Relational Analysis



Content Analysis: Conceptual Analysis

- Quantifying the occurrence in the dataset of concept/codes chosen for examination
- A coding strategy was employed in order to avoid (at least minimize) the subjectivity nature of the data

Content Analysis: Conceptual Analysis

- Technological synergy - ‘update price of galaxy s3 and s4 16 GB version’- the consumer recognises congruency between the Samsung Galaxy S3 model and Samsung Galaxy S4 model.
- Legal regulation - ‘any kind of guarantee’- the consumer asks about legal restrictions related to product guarantee
- Product price - ‘tell me the update price’- consumer asks about the product price

Content Analysis: Conceptual Analysis

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

- Communication - ‘How do I use the video calling on my S4 mini please’– consumer encourages communication with the company
- Product Technological Performance - ‘use the video calling’ – consumer comment relates to product performance
- Product Innovativeness - ‘video calling’- consumer refers to new innovative feature

Content Analysis: Relational Analysis

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

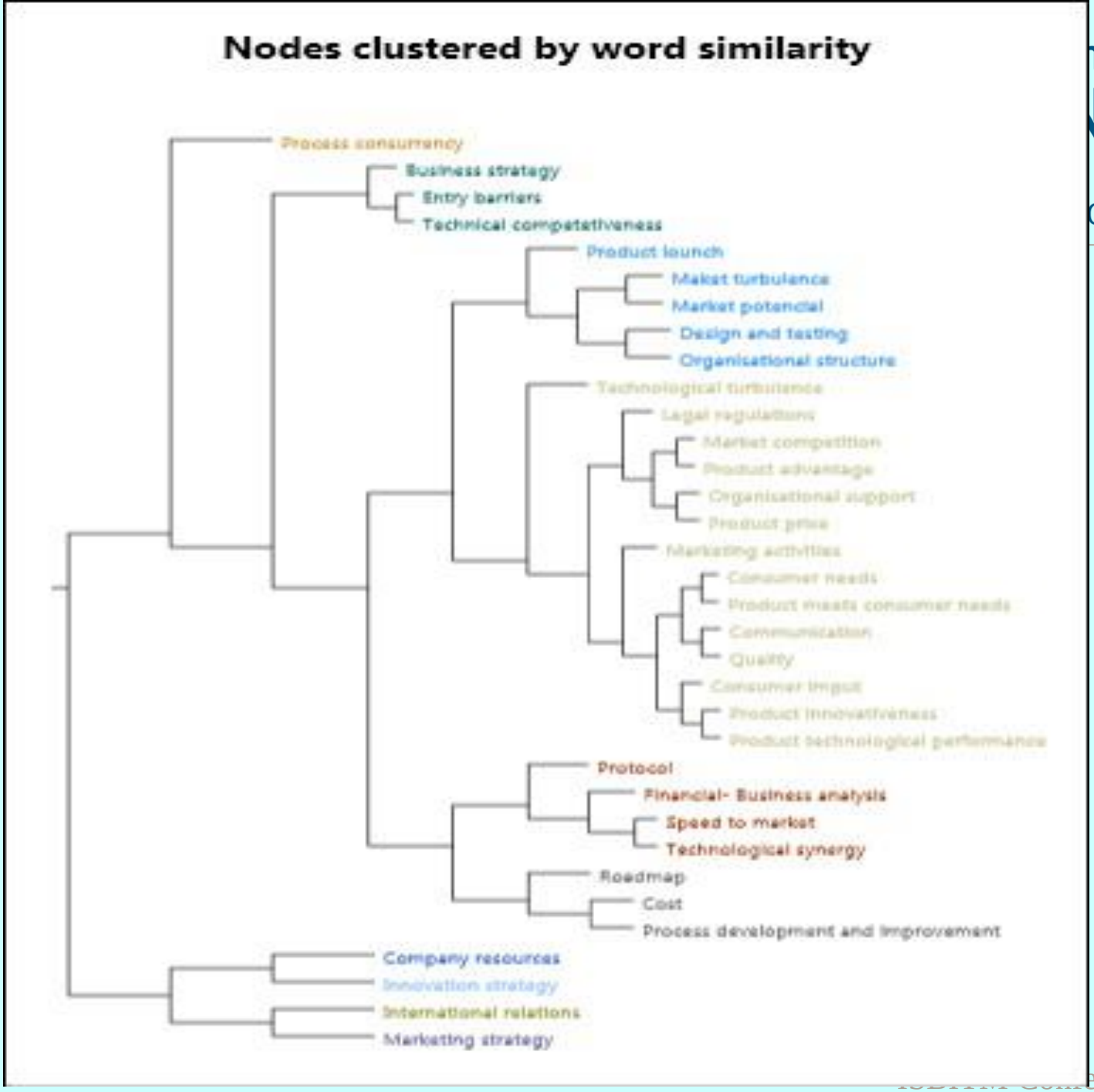
- To examine statistically accurate relationships between concepts
- Analyse a dataset without prior assumptions of the relationships of the factors of concern (Punj and Steward, 1983)
- Similar factors can be grouped (clustered) together
- Equally important is the reduction process of the number of factors of our concern (Hair et al., 2010)
- Statistical cluster analysis was employed
- Pearson coefficients are calculated (Table 3)

Research Model – Content Analysis: Relational Analysis

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

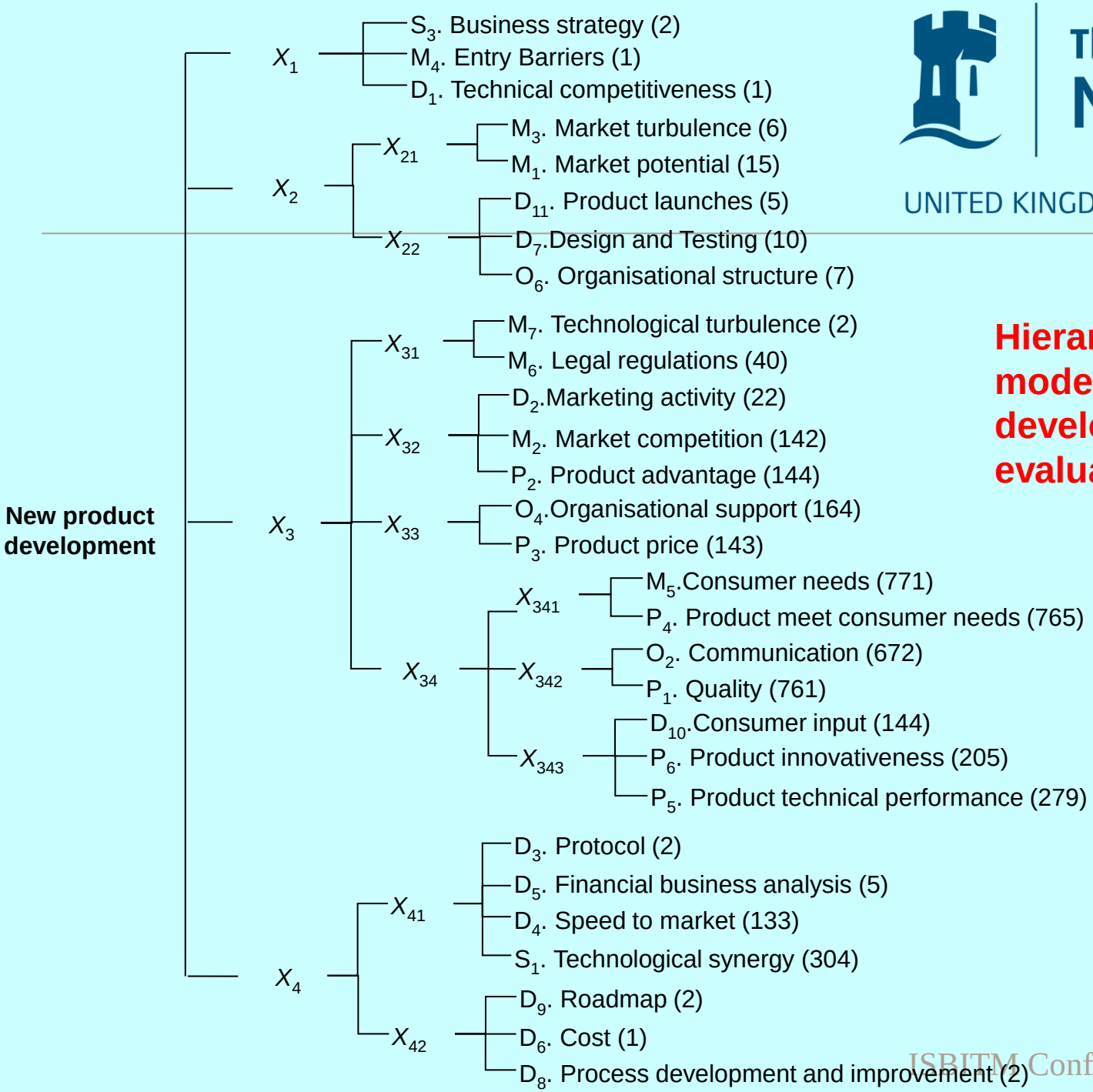
- The most correlated factors come from the 'Product' category, and are: 'Quality', 'Product advantage', 'Product price', 'Product meets consumer needs', 'Product technological performance' and 'Product innovativeness'
- Those items are correlated with two items from the 'Organisational' category: 'Communication' and 'Organisational support'.



Dendrogram of the cluster analysis

Research Model – Content Analysis: Relational Analysis

- For instance, product meet customer needs (P_4) is highly related to another measure, namely, customer needs (M_5) in the marketing category, as their Pearson Correlation Coefficient (0.998) is extremely high
- The results from the cluster analysis confirm this assertion as P_4 and M_5 are grouped as one evaluation criterion.



The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Hierarchical decision model for new product development evaluation

Weights calculation using Entropy/information theory

- A measure of how much information associated with a given state of events – quantification of information
- P(x_{ij}) is the set of all probabilities of the variables in the criterion x_i, and x_{ij} is the number of j^{th} type comments with respect to the i^{th} criterion

Weights calculation using Entropy/information theory

- The entropy of i^{th} criterion is defined as:

$$H_i = -K \sum_{j=1}^{m} P(x_{ij}) \ln P(x_{ij})$$

- $K = 1/\ln m$
- $i = 1, 2, \cdots, n$
- n is the number evaluation criteria
- m is the number of factors that form the evaluation criteria

Weights calculation using Entropy/information theory

- The weight of entropy of i^{th} criterion can then it be defined as:

$$w_i = \frac{H_i}{\sum_{i=1}^{n} H_i}$$

- In which $0 \leq w_i \leq 1$, and $\sum_{i=1}^{n} w_i = 1$
- This approach does not require an individual decision maker to rank the criteria, and the relative weight of each criterion can be obtained using rather simple calculations

Evaluate alternatives using Fuzzy AHP

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

	Entropy value (H_i)	Local weights		Entropy value (H_{ij})	Local weights		Entropy value	Local weights
X_1	0.064	0.039						
X_2	0.083	0.050	X_{21}	0.090	0.583			
			X_{22}	0.064	0.417			
X_3	1.220	0.738	X_{31}	0.140	0.060			
			X_{32}	0.434	0.186			
			X_{33}	0.632	0.270			
			X_{34}	1.134	0.485	X_{341}	1.031	0.372
						X_{342}	1.046	0.377
						X_{343}	0.698	0.252
X_4	0.286	0.173	X_{41}	0.387	0.954			
			X_{42}	0.019	0.046			

Evaluate alternatives using Fuzzy AHP

	Design 1	Design 2	Design 3	Design 4
Design 1	(1, 1, 1)	(1, 1.59, 2.08)	(0.30, 0.44, 0.79)	(0.28, 0.38, 0.63)
Design 2	(0.48, 0.63, 1)	(1, 1, 1)	(0.30, 0.44, 0.79)	(0.28, 0.40, 0.69)
Design 3	(1.26, 2.29, 3.30)	(1.26, 2.29, 3.30)	(1, 1, 1)	(0.69, 0.79, 1)
Design 4	(1.59, 2.62, 3.63)	(1.44, 2.52, 3.56)	(1, 1.26, 1.44)	(1, 1, 1)

Evaluate alternatives using Fuzzy AHP

	X_1	X_2	X_3	X_4	Aggregated index	Rankings
Weightings	0.039	0.050	0.738	0.173		
Design 1	0.170	0.233	0.178	0.170	0.179	4
Design 2	0.142	0.178	0.241	0.230	0.232	3
Design 3	0.318	0.222	0.295	0.354	0.302	1
Design 4	0.371	0.366	0.286	0.246	0.286	2

Comparative analysis of different MCDM approaches

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

	FAHP		AHP		TOPSIS	
	AI	Rankings	AI	Rankings	AI (Cc_i)	Rankings
Design 1	0.179	4	0.174	4	0.276	4
Design 2	0.232	3	0.229	3	0.363	3
Design 3	0.302	1	0.306	1	0.946	1
Design 4	0.286	2	0.291	2	0.663	2

Concluding remarks

- This research demonstrates a simply but practical approach to utilise social media data
  - The focus is put on the operations management perspective (product development)
- Demonstrates a new research arena which is data-driven research in operations management (Delage and Ye, 2010; Simchi-Levi, 2013)
- The research is multi-disciplinary in nature and defines a multi-methodological approach to blend social media research and operations management research

Practical implications

- The proposed approach allows organisations to be more economical by utilising data freely available online
- Although the case study focuses on NPD, there is a similar demand for social media data utilisation in other management areas

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Limitations

- Learning ability?
- Sentiment?
- Contribution in various methods?

Industrial Management & Data Systems

www.emeraldgrouppublishing.com/imds.htm

IMDS publishes research findings on new technologies and processes.

Indexed in

see overleaf for details of **free** sample articles

Emerald

Editors: **Dr Hing Kai Chan** and **Dr Alain Yee Loong Chong**,
Nottingham University Business School China, University of Nottingham, People's Republic of China

Double-blind peer-reviewed, with an impact factor of 1.674, *IMDS* promotes awareness of new technology for managers and facilitates better communication between functions.

Submit your research now!

We are welcoming contributions.
Submit your research now via: **http://mc.manuscriptcentral.com/imds**

Influential recent articles

Product weakness finder: an opinion-aware system through sentiment analysis
Hongwei Wang, Wei Wang

Determinants of user behaviour and recommendation in social networks : An integrative approach from the uses and gratifications perspective
Rafael Curras-Perez, Carla Ruiz-Mafe, Silvia Sanz-Blas

Set up your Emerald profile to receive table of contents alerts for this journal and more **www.emeraldgrouppublishing.com/profile/**

To contact us and find out more, visit:

www.emeraldgrouppublishing.com/imds.htm

For **Free access to topics covering Data Analytics** log on to:
www.emeraldgrouppublishing.com/tk/imds1
Valid until 5 December 2014

Emerald Group Publishing

Big Data Research for the Knowledge Economy: Past, Present, and Future

Special issue call for papers from Industrial Management & Data Systems

Knowledge Economy Big Data Research

Guested edited by Dr Xiaojun Wang and Professor Leroy White, University of Bristol, and Professor Xu Chen, University of Electronic Science and Technology of China

What is the Special Issue about?

This special issue aims to raise the awareness of big data research to the current IMDS readership. More specifically, this special issue will focus on knowledge generation and value creation from big data from various disciplines including operations, finance, marketing, innovation, tourism management, and so on. Therefore, it is hoped that the special issue can inform as well as also broaden the readership of the Journal.

Topics of interest

The scope of this call for papers is necessarily broad to provide flexibility to researchers to address current and emerging topics in big data research for the knowledge economy. We seek cutting edge developments, concepts and practices that reflect current breakthroughs in big data research. Areas in which contributions may be made include (but are not limited to):

IMDS Special Issue on Big Data Research for the Knowledge Economy: Past, Present, and Future

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Regional information

Resources:

Licensing Solutions

For Authors

For Librarians

For Engineers

Research Zone

Learning Zone

Teaching Zone

Multimedia Zone

- Sources of big data and analysis
- Methodological issues
- Social media research
- Financial services
- Tourism management
- Operations and manufacturing management
- Innovation management
- Service development and provision

Proposed Schedule

Deadline for submission: **30 March, 2015**
First round review reports: **30 May, 2015**
Revised paper submission: **30 June, 2015**
Final manuscript submissions to publisher: **August 2015**

Submissions Guidance

- Manuscripts should be submitted no later than ***30 March 2015*** using the IMDS author guidelines
- Please submit your article using the IMDS Emerald manuscript central portal and select ***"Special Issue: Big Data Research for the Knowledge Economy: Past, Present, and Future"*** when it prompts to indicate the "Article Type" in the submission.
- Please contact Dr Xiaojun Wang if you have any questions prior to submitting your article

More information about the topic of the Special Issue

While there has been a growing academic as well as practitioner interest in 'big data', i.e. large, fast-moving, complex and open datasets, there are also countless organisations using and producing big data that are not just traditional IT companies. There is increasing enthusiasm for exploiting big data, and making better use of quantitative and qualitative data from a range of "open" and administrative sources. Previous studies, however, have focused heavily on data mining algorithms, and associated applications on the available data, and so on. Research regarding big data leading to insights for knowledge economy is under-studied. Many disciplines such as information management, operations management, finance, innovation management, etc., have contributed to vast knowledge generation, and yet the links between big data and this proliferation of knowledge is not well understood. The main purpose of this special issue is to reflect the recent developments made in this respect and it is our aim to link the past to the future by both looking back to what has been done and by looking forward to what needs to be done in big data research for the knowledge economy.

Thank you!
Questions?!