Big Data Research 2 (2015) 28–32

Contents lists available at ScienceDirect

Big Data Research

www.elsevier.com/locate/bdr

Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix ☆

Shaokun Fan [a,*], Raymond Y.K. Lau [b], J. Leon Zhao [b]

[a] *College of Business, West Texas A&M University, Canyon, USA*
[b] *Department of Information Systems, City University of Hong Kong, Hong Kong*

A R T I C L E   I N F O

Article history:
Received 8 November 2014
Received in revised form 12 February 2015
Accepted 12 February 2015
Available online 18 February 2015

Keywords:
Big data analytics
Business intelligence
Marketing intelligence
Marketing mix
Survey versus log data

A B S T R A C T

Big data analytics have been embraced as a disruptive technology that will reshape business intelligence, which is a domain that relies on data analytics to gain business insights for better decision-making. Rooted in the recent literature, we investigate the landscape of big data analytics through the lens of a marketing mix framework in this paper. We identify the data sources, methods, and applications related to five important marketing perspectives, namely people, product, place, price, and promotion, that lay the foundation for marketing intelligence. We then discuss several challenging research issues and future directions of research in big data analytics and marketing related business intelligence in general.

1. Introduction

Recent technological revolutions such as social media enable us to generate data much faster than ever before [28]. The notion of big data and its application in business intelligence have attracted enormous attention in recent years because of its great potential in generating business impacts [12]. "Big Data" is defined as "the amount of data just beyond technology's capability to store, manage and process efficiently" [21]. Big data can be characterized along three important dimensions, namely volume, velocity, and variety [38].

In marketing intelligence, which emphasizes the marketing-related aspects of business intelligence, data relevant to a company's markets is collected and processed into insights that support decision-making [19]. Marketing intelligence has traditionally relied on market surveys to understand consumer behavior and improve product design. For example, companies use consumer satisfaction surveys to study customer attitudes. With big data analytic technologies, key factors for strategic marketing decisions, such as customer opinions toward a product, service, or company, can be automatically monitored by mining social media data [35].

However, while accessibility to big data creates unprecedented opportunities for marketing intelligence, it also brings challenges to practitioners and researchers. Big data analytics is mainly concerned with three types of challenges: storage, management, and processing [21]. For typical marketing intelligence tasks such as customer opinion mining, companies nowadays have many different ways (social media data, transactional data, survey data, sensor network data, etc.) to collect data from a variety of information sources. Based on the characteristics of collected data, different methods can be applied to discover marketing intelligence. Analysis models developed based on a single data source may only provide limited insights, leading to potentially biased business decisions. On the other hand, integrating heterogeneous information from different sources provides a holistic view of the domain and generates more accurate marketing intelligence. Unfortunately, integrating big data from multiple sources to generate marketing intelligence is not a trivial task. This prompts exploration of new methods, applications, and frameworks for effective big data management in the context of marketing intelligence.

We investigate different perspectives of marketing intelligence and propose a framework to manage big data in this context. We first identify popular data sources for marketing intelligence perspectives. Then, we summarize the methods that are suitable for different data sources and marketing perspectives. Finally, we give examples of applications in different perspectives. The proposed framework provides guidelines for companies to select appropriate data sources and methods for managing vital marketing intelligence to meet their strategic goals.

☆ This article belongs to Comp, bus & health sci.

* Corresponding author.
E-mail addresses: sfan@wtamu.edu (S. Fan), raylau@cityu.edu.hk (R.Y.K. Lau), jlzhao@cityu.edu.hk (J.L. Zhao).

http://dx.doi.org/10.1016/j.bdr.2015.02.006
2214-5796/

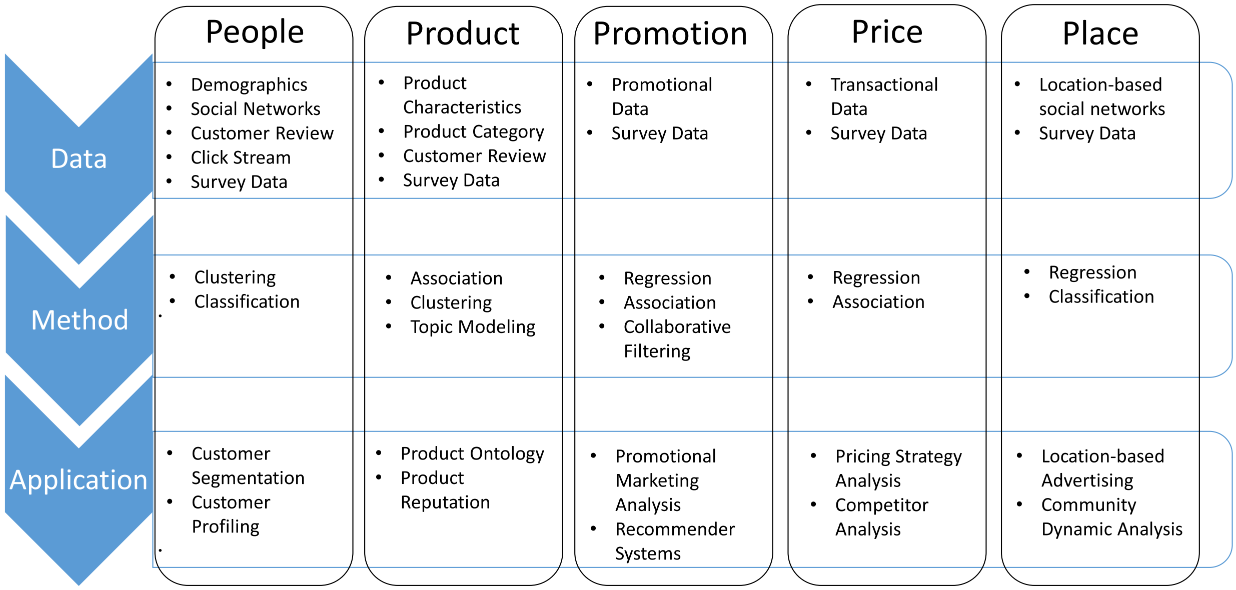

Fig. 1. A marketing mix framework for big data management.

2. A big data management framework

The marketing mix framework is a well-known framework that identifies the principal components of marketing decisions, and it has dominated marketing thought, research, and practice [6]. Borden [5] has been recognized as the first to use the term "marketing mix" and he proposed a set of 12 elements. McCarthy [29] regrouped Borden's 12 elements to four elements or 4Ps, namely product, price, promotion, and place. The 4P model has been considered to be most relevant for consumer marketing. However, it has been criticized as being a production-oriented definition of marketing, and researchers proposed a fifth P (people) [18]. We adopt the 5P model of the marketing mix framework in this paper because these perspectives play critical roles in developing successful marketing strategies in the information age.

In this paper, we propose a marketing mix framework to manage big data for marketing intelligence. This model classifies the research in marketing intelligence into five perspectives according to the marketing mix framework. Further, we identify common data, methods, and applications in each perspective and highlight the dominating big data characteristic with respect to each perspective. This framework provides guidelines for marketing decision-making based on big data analytics. Fig. 1 is an overview of the proposed big data management framework for marketing intelligence. First, data from various sources are retrieved and utilized to generate vital marketing intelligence. Second, a variety of analytics methods are applied to convert raw big data to actionable marketing knowledge (intelligence). Finally, both data and methods are combined to support marketing applications with respect to each perspective of the marketing mix model.

2.1. Data

Researchers use various methods to collect data, such as surveys, interviews, focus groups, observations, and archives [2]. Note that data collection methods are different from research methods. For example, experiments are a widely used research method in marketing, but researchers rely on surveys, observations, or interviews to collect experimental data [27]. Surveys and logs are the two most common methods to acquire data for business intelligence [22]. A survey is defined as "collecting information in an organized and methodical manner about characteristics of interest from some or all units of a population using well-defined concepts, methods and procedures, and compiles such information into a useful summary form" [10]. Firms use surveys to collect data for various purposes, such as understanding customers' preferences and behaviors. For example, Apple has sent surveys to customers who recently purchased an iPhone to gain feedback about their purchase and their experience with the product [16]. Log data is generated by information systems that capture transactional records and user behavior [20]. For example, Walmart has started to explore analyzing social media data to gain customer opinions about the company or a particular product [7]. Log data and survey data can be different in terms of size, quality, frequency, objectives, contents, and processing techniques [37]. The two data collection methods complement each other in various business contexts. Surveys can be useful when we want to collect data on phenomena that cannot be directly observed. Log data are preferred when real-time conclusions about users' actual behavior are required. The two methods can be combined when we want to study the relationship between user intention and user behavior. There are advantages and disadvantages to both methods, and we believe big data management should take both methods into consideration.

2.2. Methods

Marketing intelligence refers to developing insights from data for marketing decision-making. Data mining techniques can help to accomplish such a goal by extracting or detecting patterns or forecasting customer behavior from large databases. According to the data mining literature, common data mining methods include association mining, classification, clustering, and regression [31]. We need to select appropriate data mining methods based on the data characteristics and business problems [25].

2.3. Applications

2.3.1. Customer segmentation and customer profiling

For effective marketing, it is essential to identify a specific group of customers who share similar preferences and respond to a specific marketing signal. Customer segmentation applications can help identify different communities (segments) of customers who may share similar interests. Kim et al. [23] proposed clustering customer groups with respect to lifecycle characteristics. Usually, various clustering and classification techniques are applied to customer segmentation and user profiling. However, customer segmentation is becoming increasingly challenging under a big data environment. For instance, to differentiate among customer groups for telecommunication applications, it is necessary to analyze their call data apart from their demographics [1]. The volume of call data is huge (e.g., the communication time between each pair of customers on each day), and a variety of data should be taken into account (e.g., both qualitative demographic data and quantitative call records). In fact, for the most fine-grained targeted marketing (e.g., one-to-one marketing), we are not talking about identifying groups of similar customers, but the "profiling" of each individual customer such that the most suitable products/services are marketed to the most appropriate individual given a steam of customer service consumption data generated in real-time [1].

2.3.2. Product ontology and product reputation management

To alleviate the shortcoming of retrieving limited product reputation via survey data, Morinaga et al. [30] developed an automatic framework to monitor the reputation of a variety of products by mining Web contents. Clustering and association mining techniques are among the most common methods employed to support reputation management applications. More recently, Di et al. [14] proposed a reputation management method which not only mines text-based reputation data from the Web but also considers the graphical images of products posted to the Web. Nevertheless, by the time of this writing, twenty billion images have been uploaded to Instagram.[1] Given such an extraordinary size of images archived online, it is extremely challenging to analyze the sheer volume of images for product reputation management, not to mention the variety of formats of source data (e.g., text versus images). To carry out an automatic analysis of the textual comments posted to the Web for product reputation management, it is essential to develop a rich computer-based representation of product information for subsequent product reputation analysis. Recently, an automated product ontology mining method that is underpinned by latent topic modeling has been explored to build product ontologies based on textual descriptions of products extracted from online social media [24]. The automatically constructed product ontologies can be used as the basis to support product reputation management applications and other marketing intelligence applications. However, given the computational complexities involved in automated product ontology extraction from online social media, new computational methods must be developed to cope with the volume, velocity, and variety issues of big social media data.

2.3.3. Promotional marketing analysis and recommender systems

In the increasingly competitive business environment, billions of dollars are spent on promotions each year [34]. Thus, promotional marketing analysis has attracted a lot of attention from practitioners and researchers. Effective promotional strategies are one of the key success factors for companies to increase their sales and revenue [4]. Promotional data usually includes information about promotion types (price cut or coupons), promotion time, and purchase records during the promotional period. Early work related to promotional marketing analysis mostly focused on analyzing how different types of customers respond to different promotional strategies or how different categories of products affect the effectiveness of promotional strategies [33]. Most existing work uses regression methods to study promotions in different contexts [4].

In the big data environment, more log data becomes/is available for promotion analysis. A recent work studied WOM derived from both customer reviews and promotions [26]. The authors found a substitute relationship between the WOM volume and coupon offerings, but a complementary relationship between WOM volume and keyword advertising. Promotional marketing analysis can also include factors from other perspectives, such as price and place. For example, enabled by mobile technologies and location-based services, companies can use customers' location information to improve their promotion strategy and select targeted customers.

To improve product awareness and promote products to potential customers, recommender systems have been widely used in the e-commerce context [15]. User rating-based collaborative filtering methods or content-based association mining methods are commonly applied to develop recommender systems. However, existing methods may not scale up to big data. For instance, given N user ratings, the general computational complexity of a collaborative filtering method is N^2 [9]. Therefore, it is quite challenging to scale up existing recommender systems to cope with big data (e.g., $N =$ tens of millions) and generate appropriate recommendations to potential customers in real-time as expected in e-commerce settings. This is the reason why "velocity" is one of the most challenging issues for the "promotion" perspective in the context of marketing intelligence.

2.3.4. Pricing strategy and competitor analysis

There has been much research on what pricing strategies managers should follow under various situations. Traditionally, empirical research on pricing strategies uses survey data and regression methods. For example, researchers used a national mail survey to study the determinants of pricing strategies [32]. They found different pricing strategies are preferred under different marketing situations. The growth of e-commerce has made price information available on websites and researchers started using log data to study pricing strategy in e-commerce websites. For example, a recent study uses a method to estimate demand levels from sales rank and derive demand elasticity, variable costs, and the optimality of pricing choices directly from publicly available e-commerce data [17]. Based on the data derived from various log data sources, they can study the optimality of price discrimination. While regression methods are widely used for price prediction applications, association mining methods are applied to competitor analysis applications. An automated competitor analysis application does not simply identify the potential competitors of a company; it also effectively discovers the potentially competitive products and the product contexts [3]. This type of application has proven useful to facilitate the "price" aspect of the marketing mix model. However, the sheer volume of product pricing information on the Web has also posed new challenges to scale up existing applications with big data.

2.3.5. Location-based advertising and community dynamic analysis

Place is also an important dimension in marketing analysis. Research on place-based marketing focuses on the impact of places on marketing strategies. For example, researchers used a survey to collect customer data and study different levels of place-based marketing in the form of region of origin strategies used by wineries in their branding efforts [8].

With the widespread use of mobile technology, location-based services (LBS) can provide users personalized information in a specific location at a specific time. Location-based advertising has been proposed as an efficient marketing strategy [13,27]. Location is one of the most important solutions to meet consumers' need and it is a valuable source for personalized marketing information. In location-based advertising, customers can get timely advertisements or product recommendations based on their current position or predicted future position. Location-based advertising provides a new tool for companies to attract more customers and

[1] http://instagram.com/press/#.

enhance brand value. One challenge for location-based advertising is how to accurately predict customers' locations. Both spatial and temporal data should be taken into consideration (temporal moving pattern mining for location-based service). We need to process a large volume of spatial and temporal data within a short time period before customers move to new locations. Thus, the "velocity" issue of big data is also one of the most challenging aspects for location-based advertising.

Researchers explored the log data in location-based social networks to uncover user profiles; these automatically discovered user profiles have the potential to be subsequently applied to location-based targeted marketing [36]. Regression and classification methods are often utilized for location-based marketing applications. In another study, Castro et al. [11] leverage the GPS traces of individuals to uncover the location-based dynamics of different communities. Through analyzing the dynamics of local communities, it is possible to predict their changing product/service preferences. As a result, effective marketing strategies can be developed with respect to both the place and time dynamics of a group of customers. Nevertheless, this type of application needs to deal with both the "variety" and "velocity" issues of big data. For instance, both the relational data among users in location-based social networks and GPS signals need to be analyzed to uncover the location-based dynamics of a local community. In addition, since individuals may constantly move around different places, location-based marketing applications must be able to respond quickly in order to maintain the location sensitivity with respect to the constantly moving customers.

3. Future research directions

We propose to use a marketing mix framework for guiding research in big data management for marketing intelligence. We identify the data sources, methods, and applications in different marketing perspectives. We further discuss the challenging issues related to big data management in the context of various marketing perspectives. Based on the framework, we highlight future research directions in big data management.

1. Study how to select appropriate data sources for particular goals. The amount of available data is increasing. Current techniques do not allow us to process all data available in a timely manner. Thus, data selection is a critical decision for managing marketing intelligence. How to select data that can provide the most value to business decision-making requires future research on the alignment between data and marketing intelligence goals.
2. Analyze how to select appropriate data analysis methods. There are many types of methods that can be used to process data. Given a particular data set, many methods may be applicable. Regression and classification are usually used for prediction, while clustering and association rule mining are used for description. Further, big data brings issues such as imbalanced data distribution and large number of variables, which cannot be efficiently handled by existing data mining methods. We need to improve existing methods to increase the efficiency and accuracy.
3. Inquire how to integrate different data sources to study complicated marketing problems. Most existing studies use data from one single data source. However, some complicated business problems require combining data from different sources. For example, in order to study the impact of social media behavior on purchase behavior, we may need to combine social media data and transaction records.
4. Investigate how to deal with the heterogeneity among data sources. For example, both customer reviews and social media data can be used to study customer opinions toward a company or a product. However, data collection and analysis methods may be different due to different structure, quality, granularity, and objective. Further, survey data and log data can also be used to study the same marketing problem. How to conduct surveys in social media and confirm the survey result with log data in social media will become an important topic in e-commerce research and applications.
5. Examine how to balance investments in marketing intelligence techniques. Big data-enabled marketing intelligence will become a competitive source for consumer behavior and product planning; therefore all companies must invest in big data infrastructure including data scientists and big data platforms.
6. Explicate how to refine the framework as the big data technology evolve continuously. Data of a variety of formats and qualities will continue to grow and be digitized. Even though peta-scale data (e.g., petabytes of customer records) may be considered big data now, the same volume of data may not be considered big in a few years. It is important to continuously refine the framework, methods, and techniques that we discuss in this paper in order to meet the challenges for more advanced business intelligence in the next generation of big data management.

Acknowledgements

Dr. Lau's work is partly supported by a grant from the Science Technology and Innovation Committee of Shenzhen Municipality – Basic Research Program (Project: JCYJ20140419115614350). Dr. Zhao's work has been partially supported with the NSFC project #71110107027 and the SAP Grant – CityU Project No. 9220059.

References

[1] M.A. Abbasoğlu, B. Gedik, H. Ferhatosmanoğlu, Aggregate profile clustering for telco analytics, Proc. VLDB Endow. 6 (12) (2013) 1234–1237.
[2] W.G. Axinn, L.D. Pearce, Mixed Method data Collection Strategies, Cambridge University Press, 2006.
[3] S. Bao, R. Li, Y. Yu, Y. Cao, Competitor mining with the web, IEEE Trans. Knowl. Data Eng. 20 (10) (2008) 1297–1310.
[4] D.R. Bell, J. Chiang, V. Padmanabhan, The decomposition of promotional response: an empirical generalization, Mark. Sci. 18 (4) (1999) 504–526.
[5] N.H. Borden, The concept of the marketing mix, J. Advert. Res. 4 (2) (1964) 2–7.
[6] F. Brassington, S. Pettitt, Principles of Marketing, FT Prentice Hall, 2005.
[7] E. Brown, Mining the social data stream for deeper customer insight, http://www.zdnet.com/, 2012.
[8] J. Bruwer, R. Johnson, Place-based marketing and regional branding strategy perspectives in the California wine industry, J. Consum. Mark. 27 (1) (2010) 5–16.
[9] Y. Cai, R.Y. Lau, S.S. Liao, C. Li, H.-F. Leung, L.C. Ma, Object typicality for effective Web of Things recommendations, Decis. Support Syst. 63 (2014) 52–63.
[10] Survey Methods and Practices, Statistics Canada, 2010.
[11] P.S. Castro, D. Zhang, C. Chen, S. Li, G. Pan, From taxi GPS traces to social and community dynamics: a survey, ACM Comput. Surv. 46 (2) (2013) 17.
[12] H. Chen, R.H. Chiang, V.C. Storey, Business intelligence and analytics: from big data to big impact, MIS Q. 36 (4) (2012) 1165–1188.
[13] S. Dhar, U. Varshney, Challenges and business models for mobile location-based services and advertising, Commun. ACM 54 (5) (2011) 121–128.
[14] W. Di, N. Sundaresan, R. Piramuthu, A. Bhardwaj, Is a picture really worth a thousand words?: on the role of images in e-commerce, in: Proceedings of the 7th ACM International Conference on Web Search and Data Mining, ACM, 2014, pp. 633–642.
[15] M.B. Dias, D. Locher, M. Li, W. El-Deredy, P.J.G. Lisboa, The value of personalised recommender systems to e-business: a case study, in: Proceedings of the 2008 ACM Conference on Recommender Systems, ACM, Lausanne, Switzerland, 2008, pp. 291–294.
[16] D. Etherington, Apple sends out iPhone survey, seeks feedback on Android, touch ID and more, http://techcrunch.com/, 2014.
[17] A. Ghose, A. Sundararajan, Evaluating pricing strategy using e-commerce data: evidence and estimation challenges, Stat. Sci. (2006) 131–142.
[18] C.L. Goi, A review of marketing mix: 4Ps or more?, Int. J. Mark. Stud. 1 (1) (2009) 2.
[19] H. Hedin, I. Hirvensalo, M. Vaarnas, The Handbook of Market Intelligence: Understand, Compete and Grow in Global Markets, John Wiley & Sons, 2014.

[20] A. Jacobs, The pathologies of big data, Commun. ACM 52 (8) (2009) 36–44.
[21] S. Kaisler, F. Armour, J.A. Espinosa, W. Money, Big data: issues and challenges moving forward, in: 6th Hawaii International Conference on System Sciences, (HICSS), 2013, IEEE, 2013, pp. 995–1004.
[22] M. Kaptein, P. Parvinen, E. Poyry, Theory vs. data-driven learning in future e-commerce, in: 46th Hawaii International Conference on System Sciences, (HICSS), 2013, IEEE, 2013, pp. 2763–2772.
[23] S.-Y. Kim, T.-S. Jung, E.-H. Suh, H.-S. Hwang, Customer segmentation and strategy development based on customer lifetime value: a case study, Expert Syst. Appl. 31 (1) (2006) 101–107.
[24] R.Y. Lau, C. Li, S.S. Liao, Social analytics: learning fuzzy product ontologies for aspect-oriented sentiment analysis, Decis. Support Syst. 69 (2014) 80–94.
[25] S.-H. Liao, P.-H. Chu, P.-Y. Hsiao, Data mining techniques and applications–a decade review from 2000 to 2011, Expert Syst. Appl. 39 (12) (2012) 11303–11311.
[26] X. Lu, S. Ba, L. Huang, Y. Feng, Promotional marketing or word-of-mouth? Evidence from online restaurant reviews, Inf. Syst. Res. 24 (3) (2013) 596–612.
[27] X. Luo, M. Andrews, Z. Fang, C.W. Phang, Mobile targeting, Manag. Sci. 60 (7) (2013) 1738–1756.
[28] A. McAfee, E. Brynjolfsson, T.H. Davenport, D. Patil, D. Barton, Big data, the management revolution, Harv. Bus. Rev. 90 (10) (2012) 61–67.
[29] E.J. McCarthy, Basic Marketing, Richard D. Irwin, Homewood, IL, 1964.
[30] S. Morinaga, K. Yamanishi, K. Tateishi, T. Fukushima, Mining product reputations on the web, in: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2002, pp. 341–349.
[31] E.W. Ngai, L. Xiu, D.C. Chau, Application of data mining techniques in customer relationship management: a literature review and classification, Expert Syst. Appl. 36 (2) (2009) 2592–2602.
[32] P.M. Noble, T.S. Gruca, Industrial pricing: theory and managerial practice, Mark. Sci. 18 (3) (1999) 435–454.
[33] K. Pauwels, D.M. Hanssens, S. Siddarth, The long-term effects of price promotions on category incidence, brand choice, and purchase quantity, J. Mark. Res. 39 (4) (2002) 421–439.
[34] S. Srinivasan, K. Pauwels, D.M. Hanssens, M.G. Dekimpe, Do promotions benefit manufacturers, retailers, or both?, Manag. Sci. 50 (5) (2004) 617–629.
[35] W. Tan, M.B. Blake, I. Saleh, S. Dustdar, Social-network-sourced big data analytics, IEEE Internet Comput. 17 (5) (2013) 62–69.
[36] M.A. Vasconcelos, S. Ricci, J. Almeida, F. Benevenuto, V. Almeida, Tips, dones and todos: uncovering user profiles in foursquare, in: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, ACM, 2012, pp. 653–662.
[37] J.L. Zhao, S. Fan, D. Hu, Business challenges and research directions of management analytics in the big data era, J. Manag. Anal. 1 (3) (2014) 169–174.
[38] P. Zikopoulos, C. Eaton, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw–Hill/Osborne Media, 2011.