

# **Analysing customer behaviour in mobile app usage**

## **Abstract**

**Purpose** – Big data produced by mobile apps contains valuable knowledge about customers and markets and has been viewed as productive resources. This study proposes a multiple methods approach to elicit intelligence and value from big data by analysing customer behaviour in mobile app usage.

**Design/methodology/approach** – The big data analytical approach is developed using three data mining techniques: RFM (Recency, Frequency, Monetary) analysis, link analysis, and association rule learning. We then conduct a case study to apply the approach to analyse the transaction data extracted from a mobile app.

**Findings** – The approach can identify high-value and mass customers, and understand their patterns and preferences in using the functions of the mobile app. Such knowledge enables the developer to capture the behaviour of large pools of customers and to improve products and services by mixing and matching functions and offering personalised promotions and marketing information.

**Originality/value** – The approach used in this study balances complexity with usability, thus facilitating corporate use of big data in making product improvement and customisation decisions. The approach allows developers to gain insights into customer behaviour and function usage preferences by analysing big data. The identified associations between functions can also help developers improve existing, and design new, products and services to satisfy customers' unfulfilled requirements.

**Keywords:** big data, mobile app, customer behaviour

**Article Classification:** Research paper

## **1. Introduction**

Big data refers to “datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse” (Manyika et al., 2011, p.1). Such data are transforming the way that companies, governments, and health organisations operate (Briody 2011; Chen et al., 2012). Big data is characterised by volume, velocity, variety, and veracity and can be generated through multiple information technologies and systems, including sensor networks, enterprise resource

planning systems, social media, Internet clicks, and mobile apps (Goes, 2014; George et al., 2014; Bhat and Quadri 2015; Chan et al., 2016). Big data often contains rich information about customer opinions and behaviour, making it an important factor of production for companies (Chen et al., 2012). Data-driven decision making increases a company's competitive advantage by enabling it to better understand markets and customers, improve problem-solving, create customised and innovative products and services, and develop new organisational capabilities (Briody 2011; Davenport et al., 2012). Researchers have found that companies applying data-driven decision making are 5% more productive and 6% more profitable than their competitors (McAfee and Brynjolfsson, 2012). Big data is, thus, not simply a technological tool but can also play a transformative role in companies' fundamental operations (Li et al., 2015; Zhong et al., 2016). However, big data is difficult to manage and analyse, making value extraction an elusive goal for many companies (Ross et al., 2013; Goes, 2014). Managers also find that it is difficult to use big data to enhance design, production, and service process efficiency (Li et al., 2015; Chen et al., 2015). Further, a lack of studies exists on how to capture and harvest the potential value embedded in big data (George et al., 2014; Tan et al., 2015).

Due to the prevalence of smartphones and mobile devices, billions of people are contributing to the amount of existing big data (Manyika et al., 2011; Chan et al., 2016; Ding and Chai, 2015). Mobile devices provide users both entertainment and productivity through mobile apps (Hew et al., 2015; Shi et al., 2012). The number of connected mobile devices will reach 10 billion by 2020, and smartphones and apps are transforming various facets of society (Chen et al., 2012; Ding and Chai, 2015). Data-driven decision making is becoming a critical tool allowing mobile app developers to not only understand customer behaviour but also to design customised and innovative products and services in response to consumer requirements and changes in demand (Shi et al., 2012; Zhong et al., 2016). Studies have suggested that learning from customers, customer involvement, and customer orientation are critical for generating and disseminating market intelligence about products and services (Nambisan, 2002; Menor and Roth, 2008; Lai et al., 2012; Wang et al., 2015; Zhang et al., 2015a). Many mobile app developers now face the challenge of acquiring knowledge about customer behaviour and preferences for products and services (Hew et al., 2015). This process involves efficiently and effectively understanding customers' evolving needs and expectations (Menor et al., 2002; Zhang et al., 2015b).

The software industry is characterised by short product life cycles and high clock speeds (Davenport, 2013). Traditional customer surveys lack timeliness and specificity, making them insufficient for informing timely and targeted product development and customisation decisions. Big data enables mobile app developers to obtain knowledge without asking customers explicitly and to make data-driven decisions, for example, to segment customers into groups, understand customers' tacit and unmet needs, and discover opportunities to customise or develop new products and services (Chen et al., 2015; Davenport et al., 2012; Shi et al., 2012).

The objective of this study is to develop an analytical approach to analyse big data extracted from a mobile app for product improvement and customisation. The proposed multiple methods approach enables developers to utilise big data by understanding high-value (i.e., fans) and mass (i.e., followers) customers' behaviour. This approach also identifies core functions and the associations among functions, thereby facilitating evidence-based product improvement and customisation decision-making (Davenport, 2013; Manyika et al., 2011).

This study is organised as follows: Section 2 proposes a multiple methods approach for analysing big data for product improvement and customisation. Section 3 outlines the approach used to analyse big data extracted from a mobile app. Section 4 discusses theoretical contributions and managerial implications. Section 5 summarises this study and highlights its limitations and future research directions.

## **2. Research method**

The big data analytical approach is developed using three data mining techniques: RFM (Recency, Frequency, Monetary) analysis, link analysis, and association rule learning. These methods are well-known for use in conducting in-depth analysis and extracting knowledge from large volumes of data (Berry and Linoff, 2004; Hughes, 1996). Figure 1 presents the proposed procedure, and the subsequent sections discuss its details.

-----  
Insert Figure 1 here  
-----

### **2.1. RFM analysis**

RFM analysis is a data mining technique quantifying customer value by examining how recently (recency), often (frequency), and much (monetary) a

customer purchases (Hughes, 1996). This approach is effective for grouping customers to analyse their behaviour and predict their attitudes toward product customisation and improvement.

In RFM analysis, customers are sorted by the length of the time between their last purchase and a given date in ascending order (recency), by the number of transactions (frequency) in descending order, and by the amount of money spent in a given time period (monetary) in descending order. The definition and computation of the recency, frequency, and monetary are determined by the research problem (Güçdemir and Selim, 2015). In this study, numerical values of 1 to 5 are assigned to customers based on their recency, frequency and monetary value rankings: the top 20% are assigned a 5; the second 20% are assigned a 4, and so on, with the bottom 20% being assigned a 1. The RFM score can then be defined as

$$RFM\ score = recency \times weight_R + frequency \times weight_F + monetary \times weight_M$$

where weights are determined by expert opinions. A high RFM score indicates that a customer has high value.

## **2.2. Link analysis**

Link analysis is a data mining technique based on graph theory (Berry and Linoff, 2004). A graph has two parts: nodes (i.e., the elements in the graph that have relationships) and edges (i.e., the pairs of nodes connected by a relationship). Link analysis has been widely used by social scientists to find and evaluate patterns in relationships between organisations, people, and transactions. This method is useful for visualising and exploiting these relationships and for creating derived characteristics for further mining (Berry and Linoff, 2004).

Centrality refers to the extent to which a node lies in the centre of a network and reflects its importance in the network (Kilduff and Tsai, 2004). In this study, node centrality is evaluated by the degree and betweenness of a node's position in a network (Freeman 1979; Kilduff and Tsai, 2004). A node's degree indicates how connected it is within a graph (Fox et al., 2013). Degree can be defined as the total number of incoming and outgoing directional links (Kilduff and Tsai, 2004). The former are the direct linkages from other nodes to the focal node, while the latter are direct linkages from the focal node to other nodes.

$$Degree = in\ degree + out\ degree$$

Betweenness refers to the extent to which a node mediates the geodesic paths; these are the shortest routes linking one node to the other or connecting pairs of nodes. Betweenness can be measured as the number of shortest geodesic paths between other nodes for which the focal node lies in the middle (Freeman, 1979). For node  $v$  in the graph, betweenness centrality is defined as

$$C_b(v) = \sum_{i < j}^n \frac{g_{ij}(v)}{g_{ij}}; i \neq j \neq v$$

where  $g_{ij}$  is the number of distinct geodesics from  $i$  to  $j$  and  $g_{ij}(v)$  is the number of geodesics from  $i$  to  $j$  that pass through  $v$ . High betweenness centrality indicates that a particular node plays an important role in connecting nodes and in controlling information flows in the network (Fox et al., 2013).

### **2.3. Association rule learning**

Association rules represent patterns in groups of products and services (Agrawal et al., 1993; Hussein et al., 2015). According to these rules, the presence of certain products in a shopping cart implies a high probability that certain other products are present. Association rule learning has been widely used to identify interesting correlations between sets of items in transaction databases (Agrawal and Srikant, 1994; Yan et al., 2009). It can provide high-quality and actionable information for the innovative or customised bundling of products and services. The importance of an association rule can be evaluated by support, confidence, and lift (Porkodi et al., 2009; Qodmanan et al., 2011). In a large transaction database, managers are usually only interested in the items that are frequently purchased together (Agrawal et al., 1993). Hence, minimal support, confidence, and lift are specified by managers to drop trivial association rules because strong association rules can bring more meaningful intelligence about markets and customer behaviour (Qodmanan et al., 2011).

Let  $D = \{T_1, T_2, \dots, T_n\}$  be a set of transactions and  $I = \{i_1, i_2, \dots, i_m\}$  be a set of items. Each transaction is a set of items, and  $T_i \in I$ . An association rule can be defined as an implied form of  $X \rightarrow Y$ ,  $X, Y \in I$  and  $X \cap Y = \emptyset$  (Agrawal et al., 1993). The antecedent is  $X$ , and the consequence is  $Y$ . The support of  $X$  or  $Y$  is the percentage of transactions in  $D$  containing  $X$  or  $Y$ . The support of rule  $X \rightarrow Y$  is the proportion of transactions in  $D$  that contain both  $X$  and  $Y$ . Higher support indicates that  $X$  and  $Y$  bundles are more popular in the transaction set (Porkodi et al., 2009). The confidence

of rule  $X \rightarrow Y$  is the proportion of transactions in  $D$  that contain  $X$  that also contain  $Y$ . Lift measures the strength of rule  $X \rightarrow Y$  over the random co-occurrence of  $X$  and  $Y$ , given their individual supports. This measurement provides information about the change in the probability of  $Y$  given the presence of  $X$ . Lift values greater than 1 indicates that transactions containing  $X$  tend to contain  $Y$  more often than transactions not containing  $X$ . It shows that the occurrence of  $X$  has a positive effect on the occurrence of  $Y$  and that  $X$  is positively correlated with  $Y$  (Yan et al., 2009; Hussein et al., 2015). Let  $P(X)$  and  $P(Y)$  be the probability of the appearance of item set  $X$  and  $Y$  in  $D$  respectively.  $P(X \cup Y)$  be the probability of the appearance of item sets  $X$  and  $Y$  in  $D$ . Support, confidence and lift can be defined as

$$\text{Support}(X) = P(X)$$

$$\text{Support}(Y) = P(Y)$$

$$\text{Support}(X \rightarrow Y) = \text{Support}(X \cup Y) = P(X \cup Y)$$

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{Support}(X \cup Y)}{\text{Support}(X)}$$

$$\text{Lift}(X \rightarrow Y) = \frac{\text{Confidence}(X \rightarrow Y)}{\text{Support}(Y)} = \frac{\text{Support}(X \cup Y)}{\text{Support}(X) \times \text{Support}(Y)}$$

Apriori is the best-known algorithm for identifying association rules within transaction databases (Agrawal and Srikant, 1994). The Apriori algorithm contains two steps. First, frequent item sets with supports greater than a given minimum are identified. Apriori uses a “bottom up” approach wherein frequent subsets are extended individually, and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Second, association rules are generated from the frequent item sets, which highlight general database trends (Porkodi et al., 2009).

### 3. Analysis and results

A case study is conducted to apply the proposed approach to analyse the transaction data extracted from a mobile app (APP) to elicit knowledge about customer behaviour. The developer of APP is a leading Chinese provider of intelligent speech and language technologies and related software and hardware; APP is one of their key products. Launched in December 2012, APP is available on both Android and IOS operating systems and currently has more than 1 million customers.

APP is a voice assistant similar to Apple's Siri. It has more than 30 functions, enabling customers to use voice commands for tasks such as making phone calls, messaging, buying lottery tickets, subscribing to ring tones, playing music, downloading applications, reading novels and news, and searching for various types of information including weather conditions, mobile phone usage, restaurant recommendations, and airline schedules.

APP automatically sends data to the developer's server when customers are online and make a request or use a function. If a customer is offline, APP records the customer's operations and sends the data to the server upon connection to the Internet. On average, 1 to 2 G of data is sent to the server every hour, and the developer can collect more than 30 G of data about customer behaviour every day. The database includes transaction data on which customers are using which functions and when. Based on APP developer's suggestions, we extract one month of data (i.e., April 2014), including 41 million records, to conduct the following analyses.

### ***3.1. Identifying fans and followers***

Customers use APP in different ways: fans spend extensive time on APP and use many functions frequently, while followers only use some functions occasionally. Understanding the preferences and behaviour of fans and followers can help the developer design customised products and services bundles for these different groups of customers. We use RFM analysis to identify fans and followers. In the analysis, recency is operationalised as the length of time between the last use and 1 May 2014. Frequency is determined by how often a customer uses the APP from 1 April 2014 to 30 April 2014. The monetary value of each function is measured by two dimensions: profit opportunities and impacts on customer loyalty. An expert group, including managers from the APP developer's product development, marketing, and data analysis departments, is asked to evaluate each function of APP using a 5-point Likert-type scale (0 = "strongly unimportant"; 4 = "strongly important") on the two dimensions. The average score of the two dimensions is used to measure each function's ( $V_F$ ) monetary value. We calculate how many times each customer uses each function during the month ( $N_F$ ). The monetary value for a customer is then operationalised as  $\sum_{i=1}^p V_{Fi} \times N_{Fi}$ , where  $p$  is the number of functions used by the customer. The recency, frequency, and monetary values are then standardized using a 1-5 score. In particular, customers are ranked in ascending order in terms of recency,

and in descending order in terms of frequency and monetary. In each ranking, the top 20% are assigned a 5; the second 20% are assigned a 4, and so on, with the bottom 20% being assigned a 1. The expert group is also asked to evaluate the relative importance (i.e., weight) of recency, frequency, and monetary value. The results are  $\text{weight}_R = 0.3$ ,  $\text{weight}_F = 0.6$ , and  $\text{weight}_M = 0.1$ . SPSS Modeler is then used to calculate the RFM score and the customers are ranked in descending order by the score. Based on the Pareto principle (i.e., the 80–20 rule), the top 20% customers (214 thousand) are selected as fans and the middle 60% (642 thousand) are selected as followers.

### 3.2. Identifying core functions

Link analysis is applied to identify APP’s core functions based on each function’s centrality in the network. Gephi interactive network analysis software is used to explore the underlying relationship structures between functions and to visualise fans’ and followers’ behaviour (Bastian et al., 2009).

APP functions are treated as nodes in a network. If a fan/follower uses two functions together, an edge will exist between the two functions. We use Python to transform function usage data into source-target form, which is then visualised using Gephi (Bastian et al., 2009). Fan and follower function usage patterns are presented in Figures 2 and 3, respectively. We use degree and betweenness centrality to evaluate function importance; the results are reported in Table 1.

-----  
 Insert Figure 2 here  
 -----

-----  
 Insert Figure 3 here  
 -----

-----  
 Insert Table 1 here  
 -----

We find that fans focus on the Weather, Dialog, Music, News, Telephone, and Contacts functions. These customers have the highest degree and betweenness centrality among all functions. The degree for each of the six functions is 72, indicating that fans are centrally located in the network and directly connected with all APP functions (Fox et al., 2013). The in and out degree values indicate that fans may use any function before and after they use the six functions. The highest betweenness value shows that the six functions are positioned as important intermediaries in the network and play critical roles in linking all other functions (Fox



et al., 2013). Hence, we conclude that Weather, Dialog, Music, News, Telephone, and Contacts are the core APP functions for fans. For followers, the Weather and Dialog functions have a degree of 72 and the highest betweenness value, making them core functions for followers. The results indicate that both fans and followers frequently use APP to check weather information (Weather) and to chat with the voice assistant (Dialog). Fans also use APP to operate mobile phones for tasks such as making calls (Telephone), searching contact details (Contacts), listening to music (Music), and following the news (News).

### **3.3. *Generating association rules***

We use the Apriori algorithm and SPSS Modeler to generate association rules for fans and followers. The core functions are used as antecedents in rule generation. An expert group, including managers from the APP developer's product development, marketing, and data analysis departments, is asked to decide the minimum support, confidence, and lift based on their experiences (Agrawal et al., 1993). They decide that the minimum support, confidence, and lift should be 10%, 20%, and 1 respectively. Hence we only report those rules with support, confidence, and lift greater than the minimal values to ensure their practical significance and predictive power (Table 2) (Porkodi et al., 2009).

-----  
Insert Table 2 here  
-----

Association rules provide insights into customer behaviour and function usage preferences. These insights can help developers design and customise products and services. The results show that downloading application software (Apps), browsing community question answering services (Cqa), location (Map) or general information (Other) searches, opening a picture (Picture), setting an alarm (Schedule), and playing a video (Video) are each associated with the usage of core functions. For example, the rule {Telephone, Music, Dialog → Cqa } shows that 12.202% of fans' function usage contain Telephone, Music, Dialog and Cqa (support = 12.202%). Additionally, 86.742% of fans' function usage containing Telephone, Music, and Dialog also contain Cqa (confidence = 86.742%). This rule indicates that if a fan has used the Telephone, Music, and Dialog functions, then he/she is very likely to use the Cqa function. Hence, if a fan has used the Telephone, Music, and Dialog functions, the developer should customise services by sending the fan promotional messages or offers related to the

Cqa function. Developers could also use such knowledge to optimise interface and function designs. For example, a developer could position the Cqa function close to the Telephone, Music, and Dialog functions. We also find that when a fan has used the {Contact, Dialog, Weather}, {Telephone, Music, Weather}, {News, Music, Dialog}, {News, Dialog, Weather}, and {News, Music, Weather} function bundles, he/she is very likely to use the Cqa function. Developers should also send marketing information related to the Cqa function when these bundle functions are detected. Associations are found between App, Map, Other, Picture, Schedule, and Video functions and different function bundles. For example, if a fan has used the {Dialog, Weather} function bundle, the developer can expect that he/she may also use the Map (24.950% possibility), Picture (22.685% possibility), or Schedule (22.229% possibility) functions. Thus, the developer may consider sending the fan marketing information and offers related to these three functions. These associations could also be considered when designing new services. Only two association rules are generated for followers. In particular, 28.169% of followers play music when using the Dialog function, and 56.957% browse Cqa when using the Dialog and Weather functions together. A large portion of customers may use these functions together. Therefore, developers should place the Cqa function close to the Dialog and Weather functions and the Music function close to the Dialog function.

The findings reveal that fans and followers have different APP usage preferences. Followers have various reasons for using only certain functions occasionally. In contrast, fans utilise APP as their main mobile app. Fans can, thus, provide much more intelligence than followers about the associations between functions.

## **4. Discussion**

### ***4.1. Theoretical contributions***

This study contributes to literature by providing a new perspective on how to learn from customers and benefit from customer knowledge by analysing big data. Knowledge obtained from customers can be used to anticipate and analyse market trends and to establish a product improvement strategy in both design and service processes (Wang et al., 2015). The proposed approach offers an efficient and effective way to dynamically analyse and interpret customer-generated data. Although the RFM analysis, link analysis, and association rule learning have been widely applied in academia and industry, they have been mainly applied as separate methods (Agrawal

et al., 1993; Berry and Linoff, 2004; Hughes, 1996). This study develops an approach to integrate the three methods and demonstrates its application through a case study, providing insights into how to combine different data mining techniques to fully reap their benefits in harvesting big data (Tan et al., 2015). By analysing the transaction data of APP, this study reveals that APP users have different capabilities in providing market knowledge and fans can provide more knowledge on product improvement and customisation. The product development literature has suggested companies learn and elicit knowledge from customers and directly involve them in the design, production, and delivery of products and services (Menor et al., 2002; Nambisan, 2002; Menor and Roth, 2008). However, a company can only involve a limited number of customers. Not every customer who possesses valuable product and market knowledge can contribute to product improvement. Customer participation also increases project uncertainties and risks, and requires the company to invest in building and maintaining customer relationships (Nambisan, 2002). This study shows that a company can identify high-value customers (i.e., fans) through big data analysis and focus on building close relationships and obtaining knowledge from them. The proposed approach thus provides a cost-efficient method of generating customer behaviour and preference intelligence for companies, helping them optimize resources investment decisions to foster fans' willingness to contribute new product ideas (Zhang et al., 2015a). By analysing big data, companies can obtain valuable knowledge from customers and understand their needs without the costs of establishing and maintaining relationships with every customer. By using customers as a source of product improvement, this approach enables companies to acquire customer knowledge to address major challenges and continuously improve products and services.

The findings of APP data analysis also reveal that knowledge obtained from fans provides a new means to customise products and services to satisfy personalised preferences. The product customisation literature argues that acquiring customer knowledge and understanding customer expectations, preferences, and usage contexts are key challenges in serving markets of one (Salvador et al., 2009; Lai et al., 2012; Zhang et al., 2015b). Information technologies and systems, such as web-based tools, user toolkits, and choice menus, can be used to interact with customers, identify what they want, and determine how much they are willing to pay (Salvador et al., 2009; Zhang et al., 2015a). However, customers may have difficulties of determining and

articulating their preferences or be unable to select from the plethora of options, and thus become frustrated or dissatisfied (Zhang et al., 2015b). The analysis shows that the proposed approach allows APP developer to obtain information about the associations between APP functions from users based on observations of actual transactions without explicitly asking them. It thus allows the APP developer to process and visualise high-value customers' demand data and to consider their voices when making customisation decisions. Therefore, the proposed approach can help companies discover associations between customer requirements and preferences and predict their behaviour and preferences. As a result, companies can tailor messages and customise product and service bundles. This customisation can be performed transparently by utilising knowledge about customers' behaviour and preferences deduced from big data.

#### ***4.2. Managerial implications***

This study provides managerial guidelines on how to use big data to make product improvement and customisation decisions. In particular, we find that APP developer can obtain valuable intelligence by categorising customers and analysing transaction data obtained from customers. We thus suggest that companies invest in information technologies and systems and develop mechanisms and procedures to collect and record large volumes of customer transaction data. For example, companies could use online communities and social media sites to interact with customers and obtain feedback. The data could be stored in databases and analysed in real-time. Training programs should be developed to enhance employees' knowledge of information systems and data mining techniques. In addition, we find that the associations identified based on fans and core functions allow the APP developer to improve and customise products and services. We suggest companies implement data-driven decision-making methods to guide product improvement and customisation decisions. In particular, managers should define and identify high-value and mass customers using big data. Big data analytical techniques, such as RFM analysis, link analysis and association rule learning, should be used in combination to analyse and interpret customer behaviour and preferences. These techniques should also be used to identify core products and services that are critical for high-value and mass customers. Companies should analyse the linkages between their products and services based on the differences in high-value and mass customer behaviour. Such knowledge improves companies' understandings of how their products and services create value

for customers. Designing new products and services based on high-value customer preferences can help companies develop long-term relationships with and transform some mass customers into high-value customers. Marketing strategies, such as promotion and advertising decisions, should also incorporate knowledge of customer behaviour. Companies should compare and contrast the behaviour of high-value and mass customers using big data analytical techniques. Product and service bundles should be customised to fulfil the requirements of different customer groups.

## **5. Conclusions**

This study proposes a big data analytical approach and conducts a case study by using the approach to analyse the data extracted from a mobile app. The results show that big data analysis enables the mobile app developer to learn customer preferences and patterns of function usage. The proposed approach enables the identification of fans, followers, and core functions, and the discovery of associations among functions. We also find that fans and followers use APP functions in different ways and fans can provide more valuable market intelligence. By identifying core functions and fans, the approach used in this study ensures that the generated association rules are not trivial. The information generated from this approach enables developers to design function bundles that improve the satisfaction of existing customers and attract new customers.

Although this study makes significant contributions to the literature, some limitations exist that open up avenues for future research. First, this study analyses structured data. Future studies could investigate how to combine unstructured (e.g., social media) with structured data to understand customer behaviour. Such an analysis could be used to shape product improvement and customisation decision-making (Chan et al., 2016). Second, this study applies the proposed analytical approach in a mobile app developer. Future studies could extend the findings of this study by implementing the proposed approach to companies in other industries with different business and market environments (Tan et al., 2015; Zhong et al., 2016). Third, the analysis is based on one month's data. A longitudinal study could be conducted to compare users' behaviour or other metrics before and after the launch of a new version of APP or marketing campaigns. Future studies could also examine how to use time-series data to track individual customers and analyse how their behaviour dynamics influence product development and customisation decisions (Goes, 2014; George et al., 2014). Fourth, big data implementation is not only a technical issue but

also a managerial challenge (McAfee and Brynjolfsson, 2012; Ross et al., 2013). Future studies could investigate the organisational and cultural challenges associated with big data implementation (Manyika et al., 2011).

## References

- Agrawal, R., Imielinski, T. and Swami, A. (1993), "Mining association rules between sets of items in large databases", *ACM SIGMOD Record*, Vol.22, No.2, pp.207-16.
- Agrawal, R., and Srikant, R. (1994), "Fast algorithms for mining association rules", in the proceeding of the 20<sup>th</sup> *International Conference on Very Large Data Bases* in Santiago, Chile, pp.487-499.
- Bastian, M., Heymann, S. and Jacomy. M. (2009), "Gephi: An open source software for exploring and manipulating networks", in the proceedings of *the Third International ICWSM Conference 2009*, AAAI Press, California, pp.361-362.
- Berry, M. and Linoff, G.S. (2004), "*Data mining techniques for marketing, sales, and customer relationship management*", Wiley Publishing, Indianapolis, Indiana.
- Bhat, W.A. and Quadri, S.M.K. (2015), "Big Data promises value: Is hardware technology taken onboard?", *Industrial Management & Data Systems*, Vol.115, No.9, pp.1577 - 1595.
- Briody, D. (2011), "Big data: Harnessing a game-changing asset", available at :[http://www.sas.com/resources/asset/SAS\\_BigData\\_final.pdf](http://www.sas.com/resources/asset/SAS_BigData_final.pdf) (accessed 23 November 2014).
- Chan, H.K., X. Wang, X., Lacka, E. and Zhang, M. (2016), "A mixed-method approach to extracting the value of social media data", *Production and Operations Management*, Vol.25, No.3, pp.568-583.
- Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly* Vol.36, No.4, pp.1165-88.
- Chen, K., Li.X. and Wang, H. (2015), "On the model design of integrated intelligent big data analytics systems", *Industrial Management & Data Systems*, Vol.11, No.9, pp.1666 - 1682.
- Davenport, T.H., Barth, P. and Bean, R. (2012), "How 'Big Data' is different", *MIT Sloan Management Review*, Vol.54, No.1, pp.43-6.

- Davenport, T.H. (2013), "Analytics 3.0", *Harvard Business Review*, Vol.91, No.12, pp.64-72.
- Ding, Y. and Chai, K.H. (2015), "Emotions and continued usage of mobile applications", *Industrial Management & Data Systems*, Vol.115, No.5, pp.833 - 852.
- Fox, G.L., Smith, J.S., Cronin, J.J. and Bruscol, M. (2013), "Weaving webs of innovation", *International Journal of Operations & Production Management*, Vol.33, No.1, pp.5-24.
- Freeman, L.C. (1979), "Centrality in social networks conceptual clarification", *Social Networks*, Vol.1, No.2, pp.15-39.
- George, G., Haas, M. and Pentland, A. (2014), "Big data and management", *Academy of Management Journal*, Vol.57, No.2, pp.321-6.
- Goes, P.B. (2014), "Big data and IS research", *MIS Quarterly*, Vol.38, No.3, pp.3-8.
- Güçdemir, H. and Selim, H. (2015), "Integrating multi-criteria decision making and clustering for business customer segmentation", *Industrial Management & Data Systems*, Vol.115, No.6, pp.1022 - 1040.
- Hussein, N., Alashqur, A. and Sowan, B. (2015), "Using the interestingness measure lift to generate association rules", *Journal of Advanced Computer Science & Technology*, Vol.4, No.1, pp.156-162.
- Hew, J., Lee, V.H., Ooi, K.B. and Wei, J. (2015), "What catalyses mobile apps usage intention: An empirical analysis", *Industrial Management & Data Systems*, Vol.115, No.7, pp.1269 -1291.
- Hughes, A. M. (1996), *The complete database marketer: Second generation strategies and techniques for tapping the power of your customer database*, Chicago, Irwin.
- Kilduff, M. and Tsai, W. (2004), *Social networks and organizations*, Thousand Oaks, CA, Sage.
- Lai, F., Zhang, M., Lee, D.M.S. and Zhao, X. (2012), "The impact of supply chain integration on mass customisation capability: An extended resource-based view", *IEEE Transactions on Engineering Management*, Vol.59, No.3, pp.443-456.
- Li, J., Tao, F., Cheng, Y. and Zhao, L. (2015), "Big data in product lifecycle management", *International Journal of Advanced Manufacturing Technology*, Vol.81, No.1, pp.667-684.

- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011), "Big data: The next frontier for innovation, competition, and productivity", McKinsey Global Institute, Available at [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation) (accessed 23 November 2015).
- McAfee, A. and Brynjolfsson, E. (2012), "Big data: The management revolution", *Harvard Business Review*, Vol.90, No.10, pp.60-8.
- Menor, L.J., Tatikonda, M. and Sampson, S.E. (2002), "New service development: Areas for exploitation and exploration", *Journal of Operations Management*, Vol.20, No.2, pp.135-157.
- Menor, L.J. and Roth, A. (2008), "New service development competence and performance: An empirical investigation in retail banking", *Production and Operations Management*, Vol.17, No.3, pp.267-284.
- Nambisan, S. (2002), "Designing virtual customer environment for new product development: Toward a theory", *Academy of Management Review*, Vol.27, No.3, pp.392-413.
- Porkodi, R., Bhuvaneshwari, V., Rajesh, R. and Amudha, T. (2009), "An improved association rule mining technique for XML data using Xquery and Apriori algorithm", In the proceeding of the *IEEE International Advance Computing Conference*, Patiala, India, pp. 1510 - 1514.
- Qodmanan, H.R., Nasiri, M. and Minaei, B. (2011), "Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence", *Expert Systems with Applications*, Vol.38, No.1, pp.288-298.
- Ross, J.W., Beath, C.M. and Quaadgras, A. (2013), "You may not need big data after all", *Harvard Business Review*, Vol.92, No.9, pp.58-68.
- Salvador, F., De Holan, P.M. and Piller, F. (2009), "Cracking the code of mass customisation ", *MIT Sloan Management Review*, Vol.50, No.3, pp.71-78.
- Shin, C., Hong, J. H., and Dey, A. K. (2012), "Understanding and prediction of mobile application usage for smart phones", in the Proceeding of the *2012 ACM Conference on Ubiquitous Computing*, New York, pp. 173-182.
- Tan, K., Zhan, Y., Guojun, J., Ye, F., and Chingter, C. (2015), "Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure



- based on deduction graph", *International Journal of Production Economics*, Vol.165, pp. 223-233.
- Wang, Q., Voss, C., Zhao, X. and Wang, Z. (2015), "Modes of service innovation: A typology", *Industrial Management & Data Systems*, Vol.115, No.7, pp.1358 - 1382.
- Yan, X., Zhang, C. and Zhang, S. (2009), "Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support", *Expert Systems with Applications*, Vol.36, No.2, pp. 3066–3076.
- Zhang, M., Lettice, F., and Zhao, X. (2015a), "The impact of social capital on mass customisation and product innovation capabilities", *International Journal of Production Research*, Vol.53, No.17, pp.5251-5264.
- Zhang, M., Qi, Y., Zhao, X. and Duray, R. (2015b), "Mass customisation system: Complementarities and performance consequences", *International Journal of Logistics Research and Applications*, Vol.18, No.6, pp.459-475.
- Zhang, S., Wang, Z., Zhao, X., and Zhang, M., (2016), Effects of institutional support on innovation and performance: Roles of dysfunctional competition, *Industrial Management & Data Systems*, DOI: 10.1108/IMDS-10-2015-0408
- Zhong, R. Y., Lan, S., Xu, C., Dai, Q., and Huang, G. Q. (2016), "Visualization of RFID-enabled shopfloor logistics big data in cloud manufacturing", *The International Journal of Advanced Manufacturing Technology*, Vol.84, No.1, pp.5-16.