Dealing with the paradoxes of customer opinion for effective decision support in churn management

Abstract
The process of analysing opinion expressed about a brand, organization or product in churn management has been taking on a new dimension in recent times. A subjective opinion, which hitherto could be rendered as positive or negative, can now be positive and negative. Accordingly, the act of channelling an effective acquisition or retention strategy for customers becomes more complex with increasing overhead costs. Moreover, the increase in false positive and negative churn classification via sentiment analysis affects organizational knowledge, which oftentimes causes other societal challenges. As churn prediction analysis transits dramatically from local transactional data analysis, to social network content analysis for real time churn decision support, the paradoxical nature of opinion may render sentiment analysis as a tool for opinion mining less effective for decision support. This is because the sentiment analysis approach handles customers’ opinion as an independent entity; while oftentimes, user opinions are used only once for defining the behavioural class of a customer. These processes neglect the relative exclusive association, and the influence of community members over one another in churn management. This paper presents a context-based clustered conversation model for churn complexity management via index-driven opinion within a community of social network users. The essential aim is to cluster opinion about knowledge organization that is scalable, incremental and contextual for an appropriate churn management decision support tool.

1. Introduction
In today’s competitive market, knowledge and the capacity to create and utilize knowledge are considered the most important source of a firm’s competitive advantage (Hosseini et al. 2012). However, many companies are unable to function as knowledge-based organizations because they suffer from learning disabilities, and cannot create, identify, and value their knowledge assets (Lai 2007). In a world of dynamic change, with the influence of the Internet on everything from marketing and customers to knowledge discovery, the pace at which organizations are constantly seeking ways to adapt themselves to new conditions in order to survive and flourish in a competitive marketplace cannot be underestimated (Chi 2011). Through several micro blogging platforms like Facebook, Twitter (Gruzd et al. 2011), the volume of unstructured data keeps increasing everyday with users expressing their opinion on different issues at different times and for different purposes (Gupta 2014). Twitter, with its a vast number of tweets, and millions being added every day, serves as a worthy platform for sentiment analysis due to its large user base drawn from a variety of social and cultural regions worldwide, but aside from this fact, it can also be used for prediction (Liu et al. 2011). Such clusters of unstructured data consist of a gathering of qualities and variables related in a particular sense and varying in another. Most organizations typically use this data once and then lose it, despite the savings they could realize by
reusing it. With this in mind, some have concluded that it is necessary to view knowledge organization through the lens of big data attributes as well as through the functional goals of the user/customer (Michael and Carolyn 2013). In addition, a key to maximizing the values of decision support in churn management through knowledge organization is to be able to detect potential churners before they leave the service and target them exclusively for such campaigns. No doubt, churn prediction analysis (Dasgupta et al. 2008) towards effective customer acquisition or retention strategy has become an integral part of business strategy. In customer relationship management, a current business approach, which used to be managed intrinsically (Yan et al. 2004) through customer profiles and/or inherent features of the service, is also now being extrinsically managed to portray the service in terms of the value it accrues through its social roles (Karnstedt et al. 2011). This is necessary because humans are influenced by the actions of others in the society. Thence, the idea of social network analysis (SNA) is to pay close attention to the actions of an immediate social circle through behavioural analysis (Richter et al. 2010). Consequently, one of the biggest challenge will be a matter of turning big data into knowledge discovery, an insight that will enable organizations to make cogent decisions. So far, the methods employed via sentiment analysis have been less effective due to the independent analysis of each opinion without a communal influence. While conflicting, noisy, and missing data (Holzinger et al. 2013), as found in the increasing volume of unstructured data, is also a factor, the paradoxical nature of opinions are mostly not considered by the existing sentiment analysis approach. The paradoxical attributes of an opinion show that feelings about a brand sometimes seems self-contradictory or absurd but in reality may also express a possible truth, which can be complex. Thus, mere classification of sentiments inherent in this opinion can be vague, such that optimal decisions derived from respective churn class may not be accurate because the complexities in processing customers’ available information can also produce more complexities in decision support. This, process leads to an increase in false positive (FP) and false negative (FN) for various customer churn classification analyses conducted via social media other than feature-based approaches. Solving this problem requires an approach to organizing and indexing contents (opinions) for a sustainable period, to handle the complexities in decision making for respective churn classes. In addition, for the complexities to be adequately managed, effective decision support paradigms for potential churners most be put in place. Here, a scalable, incremental, contextual clustered conversation (SICC) algorithm is proposed to cluster organized opinions about a product, or brand of a company, over a period. Accordingly, the organised indexed tweet is clustered based on context, response(s) and influence on the members of the community. The essence of the context-based conversation cluster of opinions is to enable the almost real-time determination and classification of customers’ churn class via the customers’ affinity network analysis and expressed opinion. The churn class of a customer informs the
decision to be taken by the organization at a given time. Therefore, in section 2, a literature review on related works is presented, while in section 3, the research motivation is discussed and the proposed context-based conversation cluster is presented in section 4. In section 5, sample experiment and evaluation on users’ experience and response expressing opinions about product or brand of an organization is presented using a sample of 1,350 active tweeter account handlers. Finally, the research is concluded in section 6 with a brief overview of future research intentions.

2. Literature review

Since the emergence of social media on the Internet as an important evolving knowledge asset (Apoorv et al. 2011), the field of customer relationship management (CRM) and marketing through these platforms has created many opportunities by attracting customers with engaging content and creating communities to leverage their products. This rapid development of digital systems and associated information technologies has provided a reliable digital CRM system, which has begun to enhance opportunities to understand customers (Xu 2011) through customer churn management. Historically, CRM and knowledge management initiatives have been directed towards the delivery of continuous improvements for customers (Gebert et al. 2003) to build profitable, long-term relationships. However, to achieve the knowledge-enabled customer relationship management goal, an effective churn prediction must be in place. Customer churn, which is often referred to as customer attrition or customer turnover, is the loss of existing customers to another company or service provider (Kerdprasop et al. 2013). This process has been extensively studied in various domains such as banking, employee management (Saradhi and Palshikar 2011), online gaming, airlines, and telecommunication services (Hossain and Suchy 2013) among others. Based on this, various data mining technologies, such as Clustering (Popović and Bašić 2009), Decision Tree (Rahul et al. 2011), Neural Networks (Kamalraj and Malathi 2013), Regression (Amin et al. 2014), Support Vector Machine (Kim et al. 2012) and an ensemble of hybrid methods (Kumar and Ravi 2008) have been used for churn predictions using intrinsic customer features such as demography, age, pricing, service failure rate, etc. For a successful data mining process to take place, the inherent noise in the dataset must be well taken care of. Feature selection (Verónica et al. 2011) is a process of removing the irrelevant and redundant features from a dataset in order to improve the performance of its learning algorithm. The approach (Ibitoye and Onifade 2016) is required to filter and select the most optimal data subset that is required for churn prediction management. No doubt, the result of a churn prediction model is largely dependent on the success of the future selection model. While the intrinsic factors used in churn prediction do not account for the role of social ties between individuals in affecting the propensity to churn, the extrinsic factors (Verbeke et al. 2011) portray churn prediction service in terms of the value it accrues through its social
roles e.g. community opinion, effect of word of mouth. The idea that similar people tend to interact more frequently, otherwise known as homophily, offers the possibility to predict a person’s behaviour based on the observed behaviour around them (Zhang et al. 2012). One great platform where this may be achievable is through social media via social network analysis (Shuang-Hong et al. 2014). This is because social media websites provide a public forum that gives individual consumers their own voice, as well as access to product information that facilitates their purchase decisions (Kozinets et al. 2010). Examples of social media that are popular among all level of consumers include blogs, YouTube, MySpace, Facebook (Sin et al. 2012). However, based on its scalability, and accessibility to all, Twitter, a fast growing online social network, allows users to upload short text messages, also known as tweets, with up to 140 characters. It is often used for customer behavioural analysis (Liu et al. 2011) conducted for effective customer relationship management. Consumer behaviour is the study of individuals, groups or organizations about their process of selecting, securing, using and disposing products, services, experiences or ideas, to satisfy needs and the impact of these processes on the consumer and society (Nan-chan 2009). Since customer behaviour modelling has gained increasing attention in the operation management community (Shen 2007), especially via extrinsic factors, a significant field of study that analyses people’s attitudes towards entities, individuals, organizations, products, services, events, and topics, and their attributes (Liu 2011), is referred to as opinion mining or sentiment analysis. It is a well-suited process for various types of business intelligence (Shahheidari et al. 2013) while its timely manner, at a low cost, do have a great influence on customer perceptions and behaviours. The process of building a system to collect and examine opinions can be applied at four different levels; (Tan et al. 2011) the document level, sentence level, aspect level and the level of user or social relationships between different users. These levels can be adequately processed through natural-language processing, (NLP) lexicon or machine learning solutions, by labelling text manually or through a hybrid model, which requires the effort of human experts in labelling and training the data sets. While the machine learning-based approaches for opinion mining are supervised learning tasks, which utilise textual feature representation coupled with classification algorithms to infer the opinions expressed in the text (Adam et al. 2015), the lexicon-based technique is an unsupervised approach, which relies on the assumption that the collective polarity of a sentence is the sum of the polarities of the individual words or phrases that constitute it. (Chetan 2014) The objectives of these methods are to define positive, negative, or neutral feeling through text (Medhat et al. 2014) as a tool for customer retention or acquisition incentives. However, when such personalized marketing is implemented in CRM strategies as a sole entity without the influence of the community, the value of the knowledge of the social media structure cannot be fully realized. Thus, majority of incentives have a huge potential to be wrongly targeted at customers. Therefore, the need for an organized
knowledge flow from the company to the customer is a prerequisite. No doubt, when knowledge management is being upheld by information technology, organizational performance through enhanced efficiency will greatly improve. However, for an organization to recognize the requirements and expectations of customers through customer knowledge management, organizations must cluster carefully and process the customers’ implicit knowledge of their brand or product as represented through expressed opinion on social media, to find a better understanding about customer needs and expectations. This organized knowledge for effective decision support in churn prediction is presented in section 4 as a contextual clustered conversation model.

3. Research motivation

As events unfold daily on microblogs, they display real-world occurrences over space and time. Although exchanging information is not a new concept on the Internet, much of the information exchanged on social networking sites has always been conversational in nature. However, its particular importance to churn management has not been adequately explored towards effective knowledge organization. Here, we define conversation retrieval as a subset of event detection paradigms that intercept information retrieval and social network analysis. Over the years, several existing conversation-clustering algorithms have focused on the mere use of packages of words through a single pass incremental model that does not cater for evolving tweets with respect to initial conversations. Moreover, the methods are not scalable enough to identify multiple contexts through multiple conversation clustering detection analysis. Thus, extracting meaningful information is a major challenge of tweet mining. At the same time, attempts by several sentiment analysis approaches have been more complex and less effective due to the paradoxical nature of opinion. In solving these problems, a context-driven conversation clustered around indexed pre-processed tweets is presented in this research. The process of achieving this goal is further illustrated in section 4 below.

4. Clustering Tweet conversations based on social context

In order to engender organised knowledge; using unstructured user opinions, which are generated in real time via social network analysis, in this research, a semi-supervised algorithm for event monitoring is used to build a structured cluster of 1 to n conversations. The proposed organized knowledge for opinions as illustrated in figure 1 below is contextual (based on type), incremental (no time limit) and scalable (functions better when volume increases) with and without any prior knowledge.
From figure 1, \( t_{st_i} \) is the tweet streaming at time \( t_i \), while \( TC_{ci} \) is the tweet conversation class identifier. At first, tweets are pre-processed in order to clean tweets from noisy tokens and characters, normalize the language used, and generalize the vocabulary used for the expression. Thereafter, the tweet words are indexed to keep track of the tweet for onward processing. Since tweets are represented as word vector spaces, during the process of clustering conversation, a word can belong to one or more conversation/document vector spaces, the goal of indexing is to enhance the contextual semantics of tweets through direct ordering analysis. Then, a scalable, increment and contextual conversation (SICC) algorithm begins by finding the root of a conversation for the text clustering processes. Usually, the first tweet on a subject matter defines the root of a conversation among a given set of social network users. Here, the inherent special features, like the hashtag (#), @reply, mention (@username), are used to determine the context of tweets on arrival while other text through a word vector space decomposition is used to enhance the contextual clusters. Therefore, we define a set of context labels based around company name, company products, product prices, product brand, and product qualities, among others, in order to generate multiple contexts of conversation. The algorithm listing to achieving this goal is explained briefly in table 1 below:
Algorithm Listing 1: Scalable and Incremental Conversation Clustering algorithm for Tweets

Input: Let \( p_{rt_i} \) be the indexed Pre-processed Tweet
Output: Context Based Conversation Clusters.

1. Let \( C_i \) (where \( C_i \) is from 1 to n) be a set of clustered conversation
2. Let \( p_{rt_0} \) be the first tweet (root) published in the clustered conversation \( C_i \)
3. while (type (\( p_{rt_i} \)) != root) do
   4. For every \( p_{rt_i} \):
      5. Extract # from \( p_{rt_i} \) by matching its field in relative to \( C_i \) status. (if any)
      6. Extract @reply from \( p_{rt_i} \) by matching its field in relative to \( C_i \) status. (if any)
      7. Extract mention from \( p_{rt_i} \) by matching its field in relative to \( C_i \) status. (if any)
      8. Extract other features in \( p_{rt_i} \) then
      9. Find the conservation sequence of \( p_{rt_i} \) using Hidden Markov Model
     10. Add \( p_{rt_i} \) text to cluster \( C_i \) while \( S_{T_i} \) is the set of all words in \( C_i \)
    11. End for
    12. End while

The above algorithm in listing 1 goes beyond contextual classification of tweet with hashtag to defining the context of the clustered conversation by finding the social relationship between users through the @mention and @reply (RT) functions. Since long conversations are scarce on Twitter, the algorithm process also monitors the sequence of the dialogue by using the Hidden Markov Model (HMM). This is necessary because every user on a social network like Twitter is entitled to their opinion. Thereby, making information drift from ongoing conversation to starting a new discussion is the norm on the platform. Hence, to have a robust set of contextual conversation cluster, for every incoming tweet, the system checks and extracts it hashtag (if any), @reply (if any) and a mention (if any) to determine if a relevant conversation cluster(s) to the tweet is in existence. If a clustered context is already defined for the parameters obtained from the incoming tweet, the HMM function is used to determine the tweet sequence level by measuring its structural dependencies between existing utterances in various clusters. Thereafter, the tweet is added to the most appropriate conversation cluster using the probability function in equation (1) below.

\[
p(C_i | T_i) \propto p(C_i | c_{T_{i-1}}) \cdot p(C_i | T_{i+1}) \cdot p(T_i | C_i) \tag{1}
\]

While \( C_i \) is the cluster of a tweet \( T_i \), \( T_{i-1} \) denotes the previous tweets while \( T_{i+1} \) is the next tweet in a given cluster \( C_i \). \( \propto \) is the hyper-parameter that determines how likely a new cluster is created and \( p(T_i | C_i) \) is the probability that \( T_i \) is generated from \( C_i \). Thus, from equation 1,

\[
p(T_i | C_i) = \sum_{x \epsilon W} \frac{\text{count}(t_i, x)}{\text{count}(t_i, x)}
\]

Where \( x \) are the selected features used for context definition, \( W \) is the entire word vector present in a tweet. \( \text{count}(t_i, x) \) is a function that returns the number of occurrences of a feature \( x \) for a tweet or a cluster. Then, from equation 2,
The essence of using this approach is to derive concise clusters of tweets that are contextually related and trackable as the volume of tweet increases over time. Oftentimes, a transition probability $p(C_n|C_i)$ is required in engendering a new cluster. This is defined using the equation (4) below.

$$p(C_n|C_i) = \frac{\text{transitions} (C_i, C_n) + \gamma}{\sum_{x=1}^{n} \text{transitions} (C_i, C_x) + n \cdot \gamma + \alpha}$$

(4)

Where $n$ is the number of occupied clusters, $\gamma$ is the flooring value to avoid zero probability, the transition $(C_i, C_n)$ returns the number of transitions from $C_i$ to $C_n$. Thus, a new conversation cluster $(C_i = \text{new})$ is generated if any of the extracted parameters do not match any of the available context conversation cluster using the probability function in equation 5 below

$$p(C_{\text{new}}|C_{T_{i-1}}) \cdot p(C_{T_{i+1}}|C_{\text{new}}) \cdot p(T_{i}|C_{\text{new}})$$

(5)

Where $p(C_{\text{new}}|C_{t_{i-1}})$ and $p(C_{t_{i+1}}|C_{\text{new}})$ are derived by using equation 6 below

$$p(C_{\text{new}}|C_{T_{i-1}}) = \frac{\alpha}{\sum_{x=1}^{n} \text{transition} (C_{T_{i-1}}, C_x) + \alpha}$$

(6)

Based on these processes, an organized knowledge for contextual clusters of conversation is built from customers’ opinion on company brands and products. Based on this, appropriate decision support tools can be applied to target respective customers based on their churn class from expressed opinion mining. Therefore, to evaluate the need for the proposed organized opinion through social media (especially twitter) in churn prediction and decision support for customer relationship management, in section 5, an experiment on how well customers experiences in the use of social media have met their expectations is presented.

### 5. Experiments and evaluations

Based on a high percentage of active social media users, the following experiments were conducted with the goal of evaluating and establishing the need for organized knowledge in decision support for customer relationship management. Initially, we chose to find out much social media presence our respondents have. The results as presented in figure 2 reflect that over 40% have four social media accounts, 36.8% have just 3, 10.5% have 2, with 7.9% having just 1 account.
A further probe was to determine how often they use any of their social media accounts. A high percentage of our respondents use social media daily with few doing so only on a rare basis. Figure 3 further illustrates the values obtained.

Then, as part of a key factor to organizing opinions, we found out how frequently our respondents express their views about products or brands on social media. Quite a high percentage, as presented in figure 4 below, are passive on opinion expression on brands, i.e. they watch for a while before passing a comment.
Beyond mere expression of opinion, we also found out how often our respondents receive feedback from a company about their expressed opinions. This is important to determine the need for decision support in opinion clustering. The results are indicated in figure 5 below.

Then, in figure 6, we chose to evaluate how our respondents feel if a company contacted them based on their expressed opinion. Then, in figure 7, we present their rate of satisfaction with the decision supplied by the company in response to their expressed opinion.
In figure 8, to establish the need of organized knowledge for effective decision support, we decided to discover how our respondents were influenced (if at all) by members of their social network in the choice of a brand or product. 46.8% of them are strongly influenced while the rest are presented below.
Through the developed organized contextual clustered conversations, in figure 9 we present the decision from the customers based on whether customers should decide the cost of a new data bundle before it is finally pronounced. 44.1% strongly agreed, 20% agreed, 3% are indifferent while 24.3% and 8.6% disagreed and strongly disagreed, respectively, as presented below.

Table 1 shows the features of each brand that the proposed algorithm was implemented on; in a bid to determine the overall decision rating of customers’ opinion on the product.
Table 1: Product features and description

<table>
<thead>
<tr>
<th>S/N</th>
<th>FEATURES</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bundle ID</td>
<td>The ID of the data bundle</td>
</tr>
<tr>
<td>2</td>
<td>Name</td>
<td>The name of the data bundle whether accepted or not</td>
</tr>
<tr>
<td>3</td>
<td>Size</td>
<td>The value of the data in megabyte if considered or not</td>
</tr>
<tr>
<td>4</td>
<td>Usage</td>
<td>Is it frequently used?</td>
</tr>
<tr>
<td>5</td>
<td>Price</td>
<td>Is the price moderate or not?</td>
</tr>
<tr>
<td>6</td>
<td>Preferred</td>
<td>Is it preferred?</td>
</tr>
<tr>
<td>7</td>
<td>Plan</td>
<td>Is it 24/7, weekend only, midnight or daytime plan?</td>
</tr>
<tr>
<td>8</td>
<td>Purchase</td>
<td>What is the frequency of purchased the data bundle</td>
</tr>
</tbody>
</table>

6. Conclusion

Since social network have massively enhanced globalization, communal beliefs, norms and values cannot be overlooked in customer relationship management. While customers are free to express their opinions about brands and products, companies are also investing hugely in how to make cost effective decisions. Several existing approaches to this process are not organized and do not consider community members’ opinions. However, they independently mine customer opinion to determine their sentiment class, which can be either positive or negative. In this research, an organised contextual clustered conversation of opinion is presented. The essential aim is to provide opinion clustering based on type, which can function better when volume increases and most importantly does not have a time limit) with and without any prior knowledge. The result of the experiments conducted further strengthens the need for organized knowledge to aid decision support in today’s customer relationship management.

References


