Measuring the Effectiveness of Multi-Channel Marketing Campaigns Using Online Chatter

Haggai Roitman, Gilad Barkai, David Konopnicki, Aya Soffer
IBM Research - Haifa
Haifa University Campus, 31905, Israel
{haggai,gilad,davidko,ayas}@il.ibm.com

ABSTRACT

Measuring the effectiveness of marketing campaigns across different channels is one of the most challenging tasks for today's brand marketers. Such measurement usually relies on a combination of key performance indicators (KPIs), used for assessing various aspects of marketing outcomes. Recently, with the availability of social-media sources, new options for collecting KPIs have become available and numerous social-media monitoring tools and services have emerged. Yet, given the vast media spectrum, which goes beyond social-media channels, existing solutions fail to generalize well and the curation of marketing performance KPIs for most marketing channels still relies on labor intensive means such as surveys and questionnaires. Trying to address the challenges, we propose to demonstrate a novel solution we have developed in IBM: *Multi-channel Marketing Monitoring Platform* (M3P for short). M3P is better tailored for the marketing performance domain, where online chatter is being harnessed for effective collection of meaningful marketing KPIs across all possible channels. We describe M3P's main challenges and review some of its novel KPIs. We then describe the M3P solution, focusing on its KPI extraction process. Finally, we describe the planned demonstration using a real-world marketing use-case.

Categories and Subject Descriptors: H.3.3 [Information Search and Retrieval]: Information Filtering

General Terms: Design, Algorithms

Keywords: Crowd-sensing, Social-Media, Marketing Research

1. MOTIVATION

Measuring the effectiveness of marketing campaigns across various earned, owned and paid media channels is one of the most challenging tasks for today's brand marketers. With a limited campaign budget and numerous channel possibilities for campaign execution, the assessment of marketing effectiveness per channel becomes an essential requirement for deriving "optimal" marketing strategies and more f exible cross-channel budget allocations.

Measurement of marketing campaigns' effectiveness usually relies on a combination of key performance indicators (KPIs), used for assessing various aspects of marketing outcomes (e.g., response, leads generation, brand image, etc). KPI measurements can be then aligned against various marketing objectives set by the brand's executive level, e.g., return on marketing investment (ROMI), cost per lead, lift in sales or brand's image, etc.

Recently, with the availability of social-media sources, such as social networks (e.g., Facebook), blogs and microblogs (e.g., Twitter) and online media sites (e.g., YouTube), new options for collecting KPIs have become available [4, 7] and numerous social-media monitoring tools and services have emerged (e.g., Sysomos[1], Radian6[2], Topsy[3]).

Yet, given the vast media spectrum, which goes beyond social-media channels (e.g., TV, radio, etc), existing social-media monitoring solutions fail to generalize well, mainly due to two reasons. First, many of existing solutions are very generic, not well designed for the multi-channel marketing performance monitoring domain. Second, only simple KPIs such as campaign impressions, trends and sentiments are currently supported. The usefulness of KPIs extracted by existing solutions has been recently questioned[4], challenging the research community to discover more meaningful social media KPIs for measuring the effectiveness of marketing campaigns.

Nowadays, the curation of marketing performance KPIs for most marketing channels, and specif cally for "traditional" media channels (e.g., TV, radio, newspapers, billboards, etc), still relies on labor intensive means such as surveys and questionnaires, which usefulness strongly depends on data availability and the gathering of enough representative samples, and retrospective analysis of actual sales.

2. M3P

Trying to address the challenges, we propose to demonstrate a novel solution we have developed in IBM: *Multi-channel Marketing Monitoring Platform* (M3P for short). M3P is better tailored for the marketing performance domain, where online chatter is being harnessed for effective collection of marketing KPIs across all possible channels.

M3P is motivated by the basic fact that many users nowadays more and more tend to share their life experiences online, among others their encounters with ongoing marketing campaigns on various channels and opinions towards them. Figure 1 provides few examples of user posts published on Twitter containing impressions towards various campaigns executed on three different "traditional" channels (TV, radio, billboards).

[1]http://www.sysomos.com/
[2]http://www.salesforcemarketingcloud.com/
[3]http://topsy.com/
[4]http://thetysonreport.com/coca-cola-facebook-likes-dont-affect-sales/

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577027.

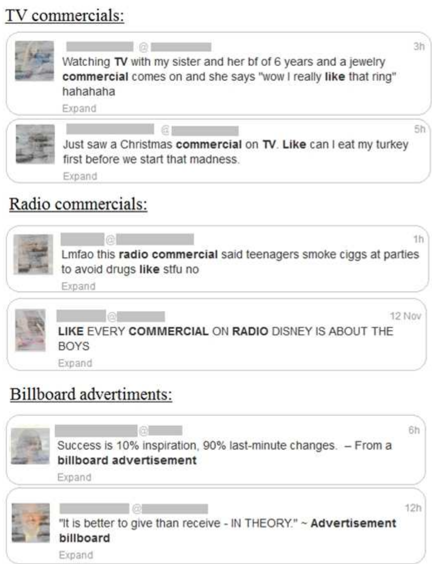

Figure 1: Examples of user impressions on various campaigns posted on Twitter. User details are obfuscated for privacy reasons.

The power of "crowd-sensing" from online chatter has been long recognized [5] and was previously studied in many other domains. Among others, online chatter was utilized to predict stock values [3], movie ratings [8], product adoptions [2] and sensing events of various scales [11, 10].

Additionally to existing social-media KPIs, M3P further extracts a novel set of KPIs from online chatter that go beyond the capabilities of existing social-media monitoring solutions, aiming at providing more statistical meaningful evidence on campaigns' performance on various channels. Furthermore, M3P extracted KPIs can be further leveraged to devise various predictive models such as the prediction of sales and future demands or the derivation of more effective media mix marketing policies.

2.1 Main Challenges

We now shortly list M3P's main challenges in harnessing online chatter for extraction of meaningful marketing performance KPIs.

• **Configurability**: marketers wishing to collect marketing response KPIs from online chatter should be provided with an easy way to configure the data collection, analysis and KPI reporting process. Such configuration should include among others definitions of the various campaign characteristics and channels to be monitored.

• **Discovery**: online chatter contains lots of noise, embedded within the natural language and openness characteristics of this data, including usage of private jargon, stenographic language (e.g., Twitter) and even non-lingual terms (e.g., emoticons).

• **Contextualization**: gathered user generated content should be classified based on its relevance to the brand and its products in general and relevance to specific monitored campaigns. In addition, channels that are being attributed to in user impressions should be detected and used to segment KPI extraction on a channel level.

• **Representativeness**: reports of various KPIs should be based on enough representative samples. Furthermore, campaign coverage within the brand's audience should be estimated to evaluate the significance (confidence) of KPI values.

• **Efficiency**: large scale amounts of online chatter data may be gathered and require efficient frameworks for data processing and timely KPI reporting.

3. TOWARDS MEANINGFUL KPIS

Recently, the usage of existing social-media monitoring solutions for the assessment of marketing campaigns' performance has been seriously criticized by many marketing professionals. We identify two main pitfalls in existing solutions. First, the relatively simple KPIs (e.g., number of impressions, sentiments, sharings, etc) that are currently extracted by existing solutions fail to provide satisfactory evidence for a campaign's performance. The main cause of such failure is due campaign independent factors that are currently ignored by most existing solutions, such as the brand's general popularity, which may govern extracted KPIs "readings" and add unobserved "noise". Second, existing KPIs do not provide sufficient correspondence with marketing business goals such as sales lift, ROMI, etc. For KPIs to be effective, they must provide clear evidence on the campaign's ability to generate new leads or improve the brand's image during its execution.

Trying to address such issues, M3P aims at providing marketers with a novel set of KPIs, built on solid statistical foundations, which are not available in existing social-media monitoring solutions.

We now shortly list three such example KPIs that are currently available in M3P's advanced KPIs suite. M3P provides an effective mechanism for designing such "advanced" KPIs using KPI expressions (see Section 4.3).

• **Campaign Focus**: This KPI captures the relative focus a given campaign receives from the brand's targeted audience. Borrowing ideas from the association rules mining literature [6], such KPI is measured as the confidence of the association rule: $campaign \Rightarrow brand$, defined as the relative number of brand's impressions that also mention the campaign. The higher the confidence is, the more we attribute the campaign for influencing the brand's targeted audience beyond its general popularity.

• **Leads Attribution**: This KPI aims at providing hints on the most essential question that bothers today's marketers: "*can the campaign be truly attributed for generating leads?*". Modeling leads as rare events, this KPI can be measured by the log of odds-ratio [5] between the probability of a lead's occurrence or its absence, conditioned on the existence of an associated campaign impression or without it. The higher and the positive the KPI value is, the stronger is the campaign's effect on leads generation, while a zero value implies no effect and a negative value further implies an opposite effect.

• **Leads Lift**: This KPI captures the relative lift on leads generation gained by executing the campaign; It is measured by the lift [6] of the association rule $campaign \Rightarrow lead$, given by the ratio of impressions that include both campaign and lead mentions to those that include either one independently. The higher the KPI value is above 1, the higher is the lift.

4. OVERALL SOLUTION

Figure 2 depicts the high level architecture of M3P. Online chatter generated by thousands of people in various online sources is constantly being gathered. Relevant impresions towards a monitored campaign are intercepted and referenced channels to which the impressions are attributed to are determined. M3P aggregates over such "readings" and efficiently collects various marketing per-

[5]http://en.wikipedia.org/wiki/Odds_ratio

[6]http://en.wikipedia.org/wiki/Lift_(data_mining)

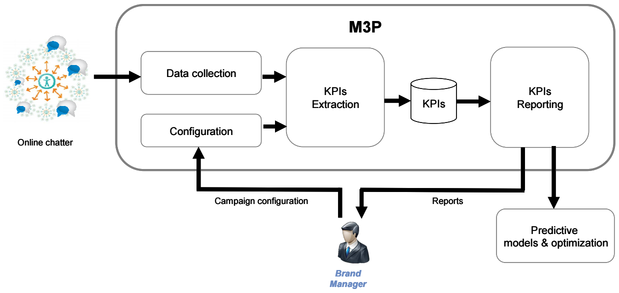

Figure 2: M3P high-level architecture

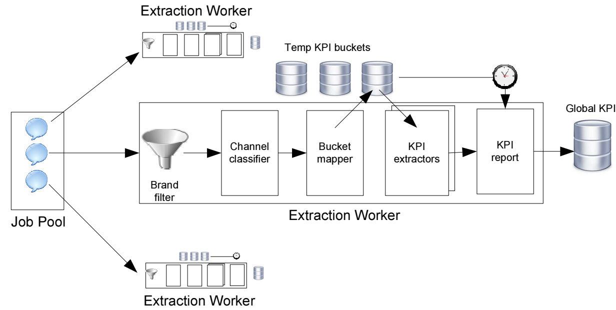

Figure 3: KPI extraction process

formance KPIs, helping marketers to get quick feedback on ongoing marketing campaigns almost as soon as they go live[7].

We now describe the various parts of M3P's end-to-end solution in more details.

4.1 Configuration

M3P provides a convenient configuration UI, used by campaign managers and marketers for defining their brand, its products and monitored campaigns' characteristics. A campaign's characteristics may include descriptors such as its slogans used to push the campaign message, mentioned entities such as names (e.g., product or presenter names), and its time period. The user is further allowed to choose channels to be monitored for comparison and the time granularity for KPI reporting (e.g., every two hours).

4.2 Data Collection

Once a configuration is defined, M3P starts to collect relevant online chatter data. M3P currently utilizes the Twitter's stream as its primary source for data collection. Brand and product descriptions provided in the user's configuration are used for accelerating the collection of relevant data. It is desired that collected data will contain enough samples of both tweets that are relevant to the brand and its products in general and specifically relevant to the monitored campaign. Such collection is crucial for obtaining statistical significant estimates of the campaign's performance. In addition, to reduce possible noise that may be acquired during data collection from Twitter's stream (e.g., topic-drift), M3P employs several topic detection and tracking (TDT) [1] and stream filtering techniques [9].

4.3 KPI Extraction

M3P's KPI extraction has two main processing challenges. First, a popular brand may generate large volumes of streaming user generated data, imposing both efficiency and memory constraints on KPI extraction. Second, some of the streaming samples can be collected out-of-order, making it more challenging to maintain consistent KPI values over time. To address the first challenge, we enforce several constraints on KPIs:

1. KPI computation is a single-pass in-memory process.
2. Efficient per-record KPI computation: each KPI must be processed in at most sub-second time. While some of M3P's KPI require simple additive operations, others may include several pre-processing steps such as parsing, annotation, filtering and classification (e.g., sentiments extraction).
3. KPIs are associative and commutative operations. Such constraint allows to distribute the computation and boost efficiency.
4. KPIs are computationally independent. Exceptional are KPI expressions defining more complex KPIs, such as those described in Section 3. For efficiency, KPI expression computation is done only on synchronous reporting points.

To address the second challenge, M3P employs a bucketing strategy where KPI computation is distributed into buckets. Each bucket represents a different data collection time period, and the number of buckets can be configured to control the tradeoff between efficiency and memory. The time granularity of buckets depends on the user's pre-configured reporting interval and the amount of available memory. Buckets have a limited time-life and survive as long as data records with relevant timestamps are streamed and some threshold is met (e.g., no relevant record was seen in the last k records that were streamed).

Figure 3 further illustrates how KPI extraction is distributed in M3P. Data records are inserted into a FIFO-queue. M3P maintains a pool of KPI extraction workers, each is assigned with a single record for processing. As a first step, the record's relevance to the brand and its products is determined using a threshold-based inverted filter built from the user's descriptions. Next, a pretrained classifier is used to determine the referred channel (or none at all). Next, the record's timestamp is hashed and the relevant bucket is determined. Next, KPIs are updated within the bucket. Finally, a synchronous KPI reporting operation is called to determined whether the current bucket needs to be "flushed" out and global KPI values (including KPI expressions) should be updated.

Our experiments with M3P by far have demonstrated satisfactory performance. As an example, using a single 8-cores machine with 16GB memory and a pool with 16 workers, we managed to process more than 10,000 tweets per second (TPS), evaluating more than 20 different KPIs in parallel. Such performance is quite reasonable, given that Twitter currently streams about 6,000 TPS.[8]

4.4 Reporting

Figure 4 depicts M3P's reporting dashboard. The dashboard includes several levels of KPI summarization that can help to better understand how well the monitored campaign is performing. The dashboard includes summary of main KPIs and leads generated on each monitored channel, and reports on the evolution of the various

[7]Note that, the significance of KPIs values over time depends on the collection of enough statistics. Hence, M3P further measures confidence internals for each KPI used to determine the significance of its reported KPI values. Details are omitted due to space limitations.

[8]https://blog.twitter.com/2013/new-tweets-per-second-record-and-how

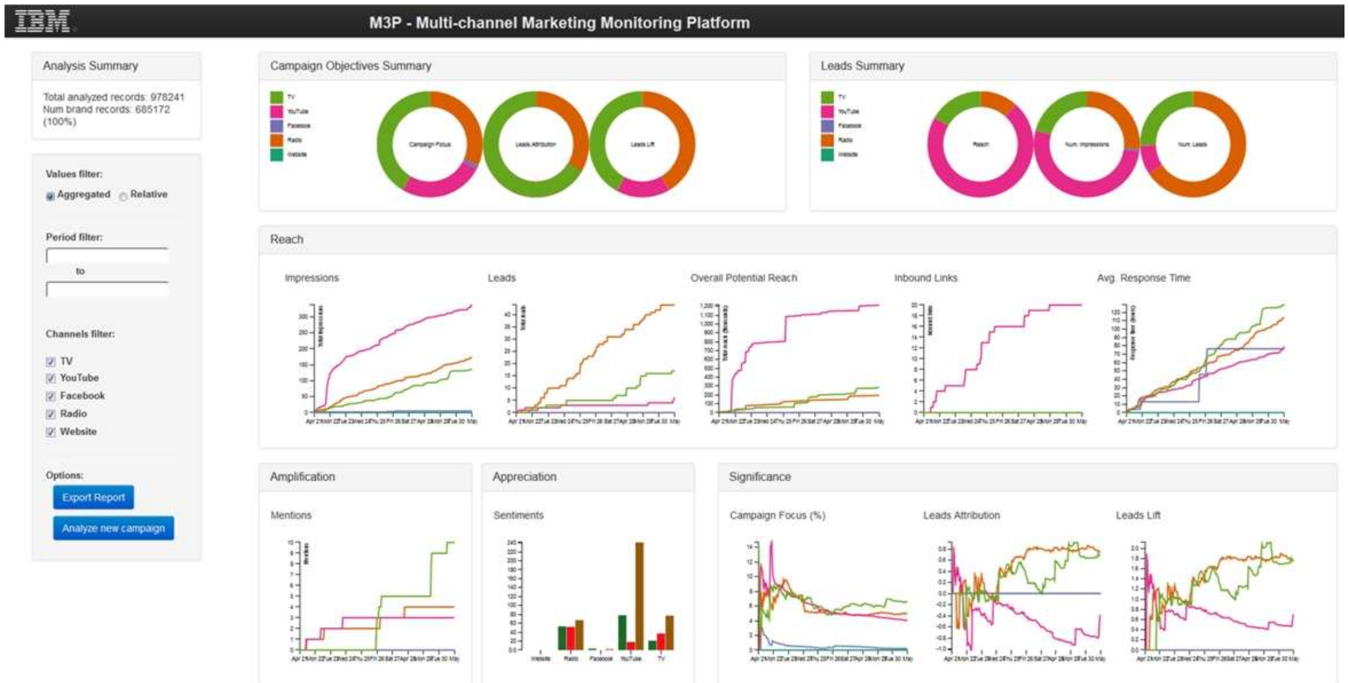

Figure 4: M3P reporting dashboard

KPIs over time. The dashboard also allows the user to explore KPI trends over time and further focus the comparison on a subset of channels or time periods of interest.

5. DEMONSTRATION

The demonstration will use online chatter data that was collected from Twitter in the last week of November 2013 during the "Black Friday" and "Cyber Monday" post-Thanksgiving shopping events. The data contains impressions on various campaigns that were executed by the top-30 US retailers during that week, among others campaigns on clothing, electronics, household, etc.

Using this data and the "Black Friday" and "Cyber Monday" campaigns' effectiveness monitoring as a use-case, we will demonstrate the various M3P's multi-channel marketing performance monitoring capabilities.

We will start by introducing the domain and its main challenges, then brief y review existing solutions and identify the existing gaps. Using M3P's reporting dashboard, we shall introduce both existing social-media KPIs and the new KPIs that are unique to M3P. We will explain the importance of such KPIs in capturing the effectiveness of campaigns executed on various earned, paid and owned media channels. As an illustration, we shall further demonstrate that simple KPIs such as impressions and sentiments may not always provide a clear picture on the success or failure of campaigns, and further illustrate how M3P's novel KPIs are better f tted to this domain.

The demonstration audience will be encouraged to try M3P's reporting dashboard and explore how various "Black Friday" and "Cyber Monday" campaigns were performing over time, relatively to each other and across the different channels.

Finally, we shall describe the various technical considerations in the design and implementation of the M3P monitoring tool, focusing the discussion on its eff cient stream-based KPI extraction capabilities.

6. REFERENCES

[1] James Allan, Jaime Carbonell, George Doddington, Jonathan Yamron, Yiming Yang, Umass Amherst, and James Allan Umass. Topic detection and tracking pilot study, 1998.

[2] Rushi Bhatt, Vineet Chaoji, and Rajesh Parekh. Predicting product adoption in large-scale social networks. In *Proceedings of CIKM'10*.

[3] Johan Bollen, Huina Mao, and Xiao-Jun Zeng. Twitter mood predicts the stock market. *CoRR*, abs/1010.3003, 2010.

[4] Irena Pletikosa Cvijikj and Florian Michahelles. Understanding social media marketing: A case study on topics, categories and sentiment on a facebook brand page. In *Proceedings of MindTrek'11*.

[5] Daniel Gruhl, R. Guha, Ravi Kumar, Jasmine Novak, and Andrew Tomkins. The predictive power of online chatter. In *Proceedings of KDD'06*.

[6] Jiawei Han, Micheline Kamber, and Jian Pei. *Data mining: concepts and techniques*. Morgan kaufmann, 2006.

[7] W. Glynn Mangold and David J. Faulds. Social media: The new hybrid element of the promotion mix. *Journal of Business Horizons*, 2012.

[8] Andrei Oghina, Mathias Breuss, Manos Tsagkias, and Maarten de Rijke. Predicting imdb movie ratings using social media. In *Proceedings of ECIR'12*.

[9] Haggai Roitman, David Carmel, and Elad Yom-Tov. Maintaining dynamic channel prof les on the web. *Proc. VLDB Endow.*, 1(1):151–162, August 2008.

[10] Haggai Roitman, Jonathan Mamou, Sameep Mehta, Aharon Satt, and L.V. Subramaniam. Harnessing the crowds for smart city sensing. In *Proceedings of CrowdSens'12*.

[11] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proceedings of WWW'10*.