

Measure Social Metrics with Sodatech: A Monitoring and Analysis Platform of Big Data

Yingmin Li, Christine Balagué

Institut Mines-Télécom, Ecole Management, 9 Rue Charles Fourier, 91000 Évry, France
yingmin.li@telecom-em.eu, christine.balague@telecom-em.eu

Abstract

Social Network Services (SNS) nowadays require marketing research study using Big Data tools. In this paper we present Sodatech platform: a monitoring and analysis tool of big data to help researchers and SNS community managers to measure effectiveness of social media marketing activities.

Sodatech platform provides three functions: collect massive data from social network; analyze marketing metrics with big data tools; display intuitive data in a dynamic and interactive interface.

Introduction

Social media brings together Internet users by all kinds of relations/interests and offers open API, which makes SNS turns to be a marketing research experimenting ground. Facebook and Twitter provide analytics for brand management, and there are tons of online social media analysis services, e.g., Google Analytics, Simply Measured, Koozai, Social Baker, etc.

We create our own analysis platform Sodatech to support various research needs. The first is to cumulate and export massive structured dataset for the research in different areas, for example, multi-culture SNS user behaviors comparison (vs Chinese SNS). The second is to combine commercial transactional data with SNS data for Client Relation Management (CRM) and recommendation algorithm development. The third is to create, verify, and integrate new research models and metrics to support community management.

Our R&D currently focus on explicit marketing activities on main Social Networks Services (SNS), like Facebook and Twitter. The aim is to define and measure the common effectiveness of community manager's publications. With Sodatech we can analyzes 3 metrics: engagement, influence, and earned media.

Engagement describes the intensity of a user's participations and connection with a brand's social media contents, either user or brand initiate (Vivek, Beatty, and Morgan 2012). Influence measures how the brand's social media contents affects the user behaviors on SNS (Trusov, Bodapati, and Bucklin 2010). Earned media measures the SNS users' reactions on publications of community managers.

System Architecture

Sodatech platform integrates several Big Data open source applications for monitoring and analyses large amount of brands on SNS (see Figure 1). It provides 3 main functions:

- Continuously collect in continue massive brands' contents from social networks through REST/STREAMING API and store them in document or in graph NoSQL database cluster.
- Analyze a series of marketing metrics with big data tools taking into account SNS media and transactional data offered by enterprise clients.
- Text mining for polarity/opinion detection and topic classification.
- Display the intuitive data in a dynamic and interactive interface framework.

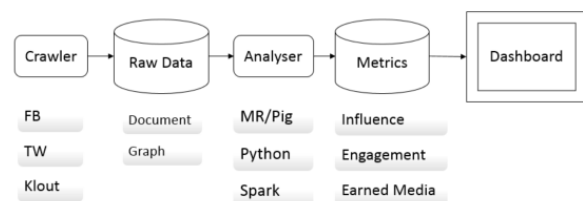


Figure 1: Sodatech platform architecture.

Social Metrics

Different SNS allow similar social actions (paired in Table 1), which allow us to easily adapt metrics equations for each SNS. Variable $B_{i,u}(t)$ defines the user (u)'s engagement contribution (number sum) of each action on date (t). Variable $R_{i,u,p}(t)$ defines the content (p)'s influence contribution of each u 's action on date (t). Variable P_m is the number of content created by community manager; and M is the number of reactive media contents.

Facebook	Twitter	Engagement(t)	Influence(t)	EM
Post	Tweet	$B1,u(t)$	$R1,u,p(t)$	P_m
Comment	Reply	$B2,u(t)$	$R2,u,p(t)$	M
Like	Retweet	$B4,u(t)$	$R3,u,p(t)$	0
Share		$B3,u(t)$	$R4,u,p(t)$	0

Table 1: Social action pairs.

More specifically, the engagement E of user (u) is the integration on date (t) of the weighted (α) sum of scores $B_{i,u}(t)$ attenuated by date. The influence I of content p on each user u is the integration on t of the weighted (β) sum of scores $R_{i,u,p}(t)$ attenuated by date. The earned media EM measures the average number of media contents aroused by social media contents

Visualization interface

Sodatech aims to offer intuitive insights on effectiveness of community management. One can collect the own social brands and all their competitor brands to compare the engagement, influence, and earned media scores. Figure 2 and 3 detail the statistics and bar chart on timeline of a brand's actions of community management and reactions of users. In figure 3, we can clearly see the target of community manager is to arouse bigger blue part by creating relatively smaller orange part.

	Web Users	Managers
Actions	9 412	1 010
Posts	208	510
Comments	1 346	491
Likes of posts	6 817	3
Likes of comments	1 041	6
Reactions	958	3 497
Likes of posts	509	2 974
Comment on posts	410	1 414
Shares of posts	0	391
Likes of comments	449	132
Metrics		
Earned Media		2.64

Figure 2: Brand's actions and reactions statistics.

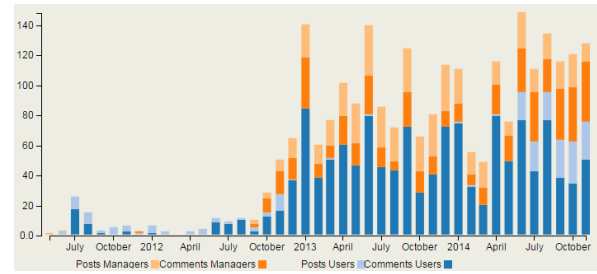


Figure 3: Brand's actions and reactions on timeline.

We also compare the part (%) of publications of absolute top 20 engagement users, relative top 20% engagement users, absolute top 20 influent users, and relative top 20% influent users, to test if it fits Pareto principle on SNS: the 20% top engaged users publish about 80% contents.

Future work

As the future work, Sodatech will cumulate dataset, develop crawlers for more SNS, and more metrics (in Figure 4).



Figure 4: Current (inner) and future (outer) applications and functions of Sodatech platform.

References

- Vivek S. D., Beatty S. E. & Morgan, R. M. 2012. *Customer engagement: Exploring customer relationships beyond purchase*, Journal of Marketing Theory and Practice, 20, 2, 127-145.
- Trusov M., Bodapati A. and Bucklin R. E. 2010, *Determining Influential Users in Internet Social Networks*, Journal of Marketing Research, Vol. 47 (4), 643-658.