

Real-Time Tracking and Mining of Users' Actions over Social Media *

Ejub Kajan¹, Noura Faci², Zakaria Maamar³, Mohamed Sellami⁴, Emir Ugljanin⁵,
Hamamache Kheddouci², Dragan H. Stojanović⁵, and Djamel Benslimane²

¹ State University of Novi Pazar
Vuka Karadžica bb, 36300 Novi Pazar, Serbia
dr.ejubkajan@gmail.com

² Univ Lyon, Université Claude Bernard Lyon 1, LIRIS
69622, Villeurbanne Cedex, France
firstname.lastname@univ-lyon1.fr

³ Zayed University
Po Box 19282, Dubai, U.A.E
zakaria.maamar@zu.ac.ae

⁴ Télécom SudParis, SAMOVAR, Institut Polytechnique de Paris
91011, Evry Cedex, France
mohamed.sellami@telecom-sudparis.eu

⁵ University of Niš
Aleksandra Medvedeva 14, 18106 Niš, Serbia
emirugljanin@gmail.com, dragan.stojanovic@elfak.ni.ac.rs

Abstract. With the advent of Web 2.0 technologies and social media, companies are actively looking for ways to know and understand what users think and say about their products and services. Indeed, it has become the practice that users go online using social media like Facebook to raise concerns, make comments, and share recommendations. All these actions can be tracked in real-time and then mined using advanced techniques like data analytics and sentiment analysis. This paper discusses such tracking and mining through a system called *Social Miner* that allows companies to make decisions about what, when, and how to respond to users' actions over social media. Questions that *Social Miner* allows to answer include what actions were frequently executed and why certain actions were executed more than others.

Keywords: Data analytics, Facebook, Sentiment analysis, Social media.

1. Introduction

Business Processes (BP) are a cornerstone to the success of any company that wishes to sustain its growth and remain competitive. According to [25], *"a process is nothing more than the coding of a lesson learned in the past, transformed into a standard by a group of experts and established as a mandatory flow for those who must effectively carry out the work"*. More precisely, a BP consists of tasks (t) connected to each other according to

* This is an extended version of a 2-page WETICE2016 demo paper [36].

a process model defining, at design-time, who does what, when, and where. At run-time, tasks are assigned to persons (p) and/or machines (m) so for execution.

Since the advent of social media, the traditional view of how companies operate has completely changed. Contrary to top-down commands and bottom-up feedback that limit innovation and creativity in companies, interactions that social media allows to happen in companies are crossing all levels and occurring in all directions. This organizational shift reduces cost, improves efficiency, facilitates innovation, among other benefits [7],[35]. Social media is also impacting the design of BPs. Earlier, we looked into this design to shed light on social interactions between tasks ($t2t$), between persons ($p2p$), and between machines ($m2m$) in BPs [14],[21]. These interactions reveal for instance, which task is “easy” to replace with other tasks, which person is mostly solicited for partnership with other persons, and which machine works well with other machines.

As a follow-up to our work on BP social-design we also looked into bridging the gap between the business world (hosting BPs) and social world (hosting Facebook as an illustrative application of social media) [20]. While the business world continuously attracts the R&D community’s attention [13],[23], the social world’s surface is barely “scratched” and hence, several opportunities are untapped. To reverse this trend we raise many questions that need responses such as who are the social world’s stakeholders, what actions can the stakeholders perform, and how to track the interactions in the social world. While we detail in [10] the actions that stakeholders perform in the context of social media, we focus in this paper on the interactions that arise in the social world in response to both events triggered and actions taken in the business world and then, to what extent these interactions would impact BPs. For instance, increasing the delivery fees of goods in a process could raise concerns over social media that decision makers in the business world would like to be aware of, should corrective actions need to be taken to address these concerns. For this purpose, we associate the business world with control flow and the social world with social flow, define the constituents of each flow, and manage the cross-flow interactions. We present the design and development of Social Miner, a **real-time** tracking and mining tool on top of Facebook.

The rest of this paper is organized as follows. Section 2 gives a short overview of social media mining in favor of business. Section 3 presents a motivating scenario to stress out the gap in examining the interactions between the business and social worlds. Section 4 briefly discusses data mining (with focus on sentiment analysis) and business process modeling. Section 5 details the real-time tracking and mining of users’ actions over social media where Facebook is used for illustration purposes. Section 6 implements this tracking and mining through a **real** marketing campaign on Facebook. Finally, Section 7 discusses the importance of connecting the business and social worlds together. Finally, Section 8 concludes the paper and identifies some future work.

2. Related work

Mining social networks in favor of business application is not new. Bonchi et al. present an overview of key problems in this domain and the techniques in social network analysis in an infant stage and emphasize, among others, potential business benefits and technical challenges [5]. These challenges include data preparation, network dynamics, propaga-

tion, evaluation, etc, whilst value-added benefits may be summarized in a short sentence: *"Social networking may allow increased revenue"*.

In essence, 2 types of social network mining strategies are in use. The first strategy mimics social networks by tracking activities of employees and their relationships inside an organization during business process execution in order to make them more flexible and productive. Examples include, but not limited to, MiSoN (Mining Social Networks) and SUPER (Social-based bUusiness Process managEmEnt fRamework). In MisoN, Aalst et al. use events logs for social networks mining [1]. These event logs are made by employees of an enterprise when they use some of enterprise information systems (e.g., ERP and CRM) and transform into sociograms by MisoN, which are later use for workflow analysis. SUPER is based on social relations between employees who are in charge to execute particular BPs. These relations are delegation, substitution and peering and results of their use are reported in [14,21]

The second strategy goes beyond an enterprise and use public social networks (e.g., Facebook and Tweeter) to mine customers' opinions about companies' products and services to facilitate CRM via customer needs anticipation and reputation monitoring, identify their churn and reasons for it, find experts, etc. [5]. These findings are obtained using sentiment analysis, opinion mining, and at large scale social influence mining. For the later, Tang et al. present how this mining may help expert finding [33]. A general approach that allows to identify the node that influences others is presented in [4].

3. Motivating scenario

GreenUtility is a utility company that is going to launch an awareness campaign about renewable energies on Facebook. To achieve the campaign's targets like increasing the number of green advocates, the marketing team must look after this campaign's business aspects (e.g., develop the campaign's content and layout, secure the necessary approvals, and assess the results of the campaign) and social aspects (e.g., announce the campaign on Facebook page, engage in discussions with this Facebook page's subscribers, post, and refresh the marketing content if necessary). Both business and social aspects become intertwined when the campaign is in a full swing. For instance, posting content on Facebook needs the marketing director's approval. And, collecting subscribers' comments from Facebook permits to decide on extending the campaign.

After securing the necessary approvals to launch the campaign, GreenUtility's Facebook page is updated with details like tips for saving energy and actions for contributing to a green world. Afterwards, the subscribers to this Facebook page **could** (in fact, the subscribers are not obliged) react to the campaign by posting responses, initiating new communication threads, and expressing feelings over some ecological incidents (e.g., Gulf of Mexico oil spill).

Putting all the social actions (e.g., post comments and invite others) that subscribers perform on Facebook page together should permit to develop social flows that would give GreenUtility better insights into what is being discussed over its Facebook page instead of relying on some quantitative performance indicators (e.g., number of visitors), only. Mining the social flows would help GreenUtility answer many questions like when is it appropriate to post a campaign so that a good response rate is achieved, who are the main

supporters/opponents of/to a campaign so that their feedback/concerns are shared/dealt with, and who should respond to supporters/opponents so that fruitful discussions occur.

4. Some preliminaries

This section is an overview of data mining and business process modeling.

Data mining. The phenomenal growth of online social media (e.g., discussion fora and blogs) is backed by the abundance and richness of user content such as comments, reviews, and feeds that could correspond to opinions on events like visited places, tried restaurants, and consulted books. Opinions can also be associated with sentiments reflecting users' attitudes and feelings towards events like anger and happiness. For a better understanding of opinions, many Sentiment Analysis (SA) techniques and tools are reported in the literature (e.g., Batrinca and Treleaven [3] and Tang et al. [32]). Generally speaking, Pang and Lee suggest 3 stages that SA goes through: *opinion retrieval*, *sentiment classification*, and *opinion summarization* [26]. The first determines which textual sources (e.g., documents, posts, blogs, and news) should be considered when looking for opinionated material with respect to a certain granularity level (e.g., entire documents, phrases, and separate words). The second identifies the overall sentiment that each textual source conveys. Finally, the third provides an integrated view of sentiments expressed by multiple textual sources. Fig. 1 depicts a comprehensive view of SA's outcomes according to the type of source to analyze (single *versus* multiple).

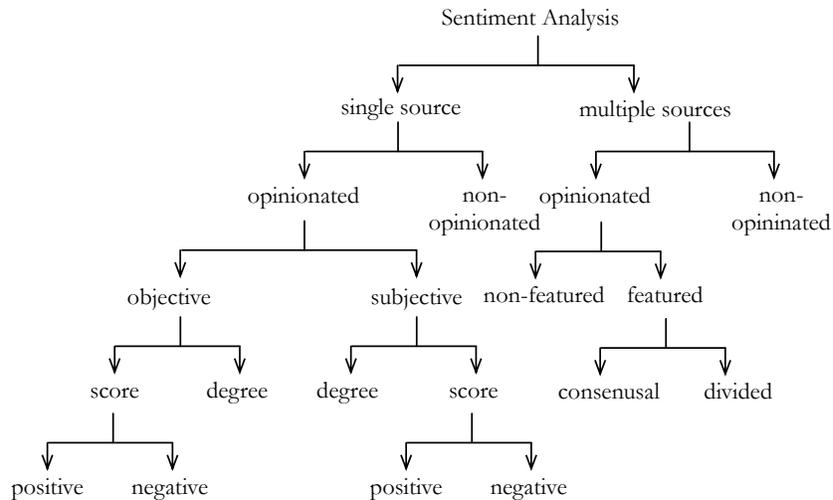


Fig. 1. Sentiment analysis decomposition

The aforementioned 3 stages can require different techniques and tools that are at the crossroad of Natural Language Processing (NLP), Information Retrieval (IR),

and Machine Learning (ML) [26]. Since textual sources could be opinion-free, the *opinion retrieval* stage relies on IR techniques to rank content in terms of *relevancy* (i.e., whether the content is topic-related) and *opinionatedness* (i.e., whether the content contains opinions) (e.g., Luo et al. [18] and Soboroff et al. [30]). The *sentiment classification* stage adopts NLP-based ML techniques to agree on content's *subjectivity/objectivity* and *polarity* (e.g., Liu [16] and Ting et al. [34]). Polarity that can be either qualitative or quantitative, permits to tag opinions with positive/negative sentiment scores (e.g., good/bad and like/dislike) or degrees (e.g., good/excellent and bad/worst). Last but not least, the *opinion summarization* stage uses IR techniques related to a content's *featurability* and NLP-based ML techniques related to *consensus* and *division* (e.g., Archak et al. [2,17]). The former identifies important features that would characterize some events. And, the latter decides whether multiple sources contain similar and/or contradictory opinions related to features so that opinions are differentiated.

Despite the benefits of existing SA techniques, many shortcomings can be reported. Indeed, they consider textual sources as one block and, thus, overlook nested exchanges in blocks. Plus, they do not capture users' attitudes towards received content (should the user respond, delay to respond, or ignore). In this work, we move one step-forward by analyzing a content's structure and growth along with users' sentiments so that we understand what happened, might happen as well as when opinion changes happened.

Business process modeling in brief. BP modeling is about documenting and displaying BPs graphically to help stakeholders analyze process models and find possible ways of improvement. A modeling language consists of 3 parts: (i) syntax that provides constructs and rules to combine constructs, (ii) semantics that gives meaning to constructs, and (iii) notation that includes graphical symbols to visualize constructs. In [27], Pourshahid et al. state that all 3 parts together should allow to model various aspects of a BP such as tasks, events, resources, roles, constituents, functions, organization, and hierarchy. Over the years, Business Process Model and Notation (BPMN) has rapidly become a standard for BP modeling as per the Object Management Group (OMG) [24]. BPMN provides graphical elements to develop multiple flows like control flow to show the partial order (i.e., conditional and concurrent) between tasks in a BP and communication flow to show the exchange of messages between a BP's stakeholders. In addition to BPMN flows, other types of flows are reported in the literature. Sadiq et al. use data flow for process specification following a data/artifact-centric perspective and process verification according to 3 properties: *correctness*, *soundness*, and *variability* [29]. Reichert and Dadam use control flow and data flow to specify BPs following a process-centric and data-driven perspective, and verify a BP's correctness properties like *reachability* and *termination* [28]. Finally, Maamar et al. develop and synchronize control, communication, and navigation flows to monitor BP execution [22].

Completely different from flows for BP modeling that are structured and known in advance, social flows are built on-the-fly and capture social actions over social media. We expect tapping into social flows to understand why users execute certain social actions, what business tasks triggered certain social actions, which social actions are

triggered because of some social actions, what is the execution chronology of social actions, and what content like feelings and opinions does social actions convey.

5. Tracking and mining approach

This section details the approach for real-time tracking and mining users' actions over social media. It starts with some definitions and examples and then, discusses how putting social actions together lead into the development of social flows.

5.1. Foundations

To formalize our approach, we first, refer readers to [19] where a definition for social action (e.g., send, co-author, and tweet) is proposed. It is an operation that a Web 2.0 application allows users to execute whether online or offline. Also, as per the same reference, a social action falls into one of the following categories (Table 1): *communication* (e.g., chat and poke), *sharing* (e.g., publish and upload), and *enrichment* (e.g., comment and tag). In this paper A is the set of all social actions available for execution over all Web 2.0 applications ($A = \{\text{poke, chat, send, co-author, tweet, post, comment, reply, tag, upload, } \dots\}$) and $A_{app2.0}$ ($A_{app2.0} \subseteq A$) is the set of all social actions available for execution over a particular Web 2.0 application ($app_{2.0}$), Facebook in our case ($A_{Facebook} = \{\text{poke, chat, send, post, upload, comment, tag, } \dots\}$).

Table 1. Representative categories of social actions ([19])

Category	Description	Examples of social actions
Communication	Includes actions that establish back-and-forth interactions between users, which should engage them in joint operations	Chat with a user or group of users, Poke someone, Send direct messages to a user's inbox
Sharing	Includes actions that establish one-way interactions and allow to create and edit shared content and to facilitate this content's consumption	Co-author a text/media on a Wiki, Publish a post on a Blog Web site, Upload a photo/video on a public repository, or any other data (e.g., sensor reading)
Enrichment	Includes actions that provide additional [meta] data on shared content by providing opinions and/or ranking	Comment a post, Tag users' photos, videos, activities, etc.

Fig. 2 depicts how the control and social flows are anchored to the business and social worlds, respectively, with focus on the interactions from the business to social worlds that trigger forming social flows. Interactions from the social to business worlds are the result of mining social flows. For clarity purposes, these interactions are not represented in Fig. 2 but are handled through metrics defined in Section 6.2. It happens that the successful execution of some business tasks in the control flow makes users execute some social actions over Web 2.0 applications (e.g., after approving the marketing content, GreenUtility posts the content on Facebook). The outcomes of executing these social actions could motivate (same and/or other) users to execute additional social actions and so

on. This execution chain of social actions expands until some termination conditions are met (e.g., deadline has passed and response-rate has become low). By putting all the social actions together with respect to when they occurred, social flows are obtained. Fig. 2 shows 1 control flow (F_1^c) and 3 social flows (F_1^s , F_2^s , and F_3^s). Later it will be shown that social flow could branch into sub social-flows (*aka* nested flows).

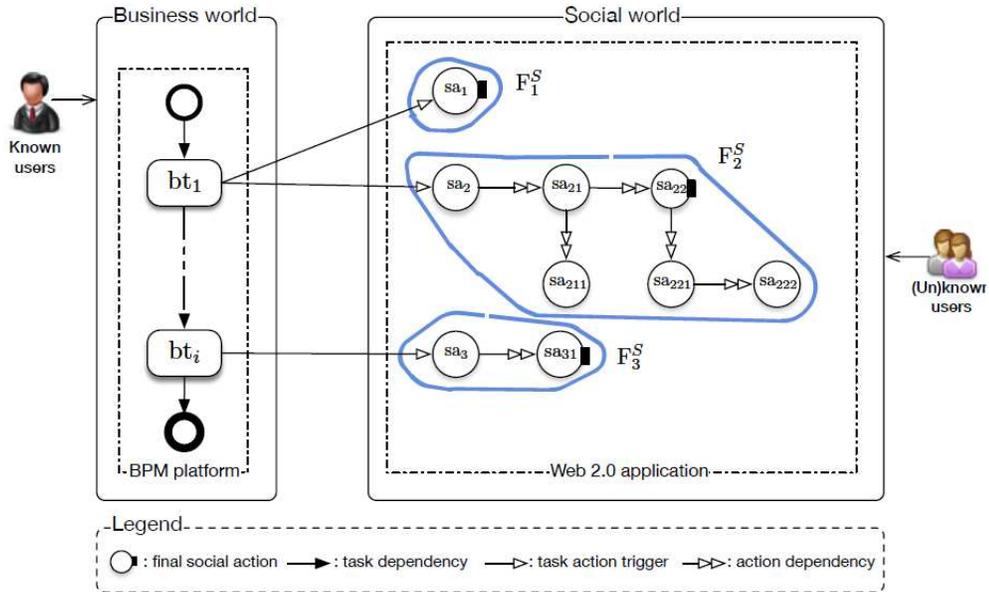


Fig. 2. Business and social worlds from a flow perspective

5.2. Definitions and examples

In the following all examples are drawn from the motivating scenario and Fig. 2. The definitions given in this section are integrated into the automatic building of social flows as demonstrated in Section 6.

Definition 1. Control Flow (F^c). It represents the process model of a BP and consists of business tasks (bt) and dependencies between business tasks. Formally, F^c is a 4-tuple $\langle T^c, D^c, IT^c, FT^c \rangle$ where: T^c contains all business tasks in a BP; $D^c \subseteq T^c \times T^c$ is the set of all dependencies between business tasks; $IT^c \subseteq T^c$ is the set of all initial business tasks; and, $FT^c \subseteq T^c$ is a set of all final business tasks.

Example: $F^c = \langle \{bt_1, bt_2, \dots, bt_i\}, \{(bt_1, bt_2), \dots, (bt_{i-1}, bt_i)\}, \{bt_1\}, \{bt_i\} \rangle$ where bt_1 = secure-necessary-approvals and bt_2 = develop-campaign-design.

Definition 2. Business-to-Social Link ($L(b2s)$). It captures the statement that “upon the successful execution of a business task, a user **may** execute social actions in response to this execution”. We aim at conditionally mapping each business task in a control flow onto a set of initial social actions that will form the roots of the future social flows. Formally, $L(b2s) : T^c \times C \rightarrow 2^{A^{s_0}}$ is a function where: T^c contains all business tasks in a BP, C is a set of conditions; and, A^{s_0} is a set of all initial social actions in the social flows.

Example:

- $L(b2s) : (bt_1, [cond_{bt_1}]) \rightarrow \{sa_1, sa_2\}$ where $bt_1 =$ secure-necessary-approvals, $cond_{bt_1} =$ is-campaign-approved?, $sa_1 =$ post-benefits-of-the-campaign-on-Facebook, $sa_2 =$ post-benefits-of-the-campaign-on-Twitter; sa_1 and sa_2 are executed if the campaign is approved.
- $L(b2s) : (bt_2, \phi) \rightarrow \phi$ where $bt_2 =$ develop-campaign-design; no social action is associated with bt_2 .
- $L(b2s) : (bt_3, [cond_{bt_3}]) \rightarrow \{sa_3\}$ where $bt_3 =$ analyze-customer-application, $cond_{bt_3} =$ is customer application rejected?, and $sa_3 =$ post-a-note-on-customer-wall; sa_3 is executed if the customer application is rejected.

Definition 3. Social Flow (F^s). It is a set of social actions put together on-the-fly. One of these social actions is initial (i.e., linked to a business task as per Definition 2) and the rest are either intermediaries or finals. First, the connection between social actions is dependent on (i) the authorized relations that Web 2.0 applications allow to have between their social actions (Section 5.3) and (ii) the nested levels of exchange that a Web 2.0 application allows to happen⁶. Second, the selection of the next social actions to execute is based on contextual elements that do not fall into this paper’s scope. Formally, F^s is a 4-tuple $\langle A_{app2.0}^s, STR_{app2.0}^s, sa_0^s, FA^s \rangle$ where $A_{app2.0}^s \subseteq A_{app2.0}$ contains those social actions in a Web 2.0 application that end-users have **voluntarily** decided to execute; $STR_{app2.0}^s : A_{app2.0}^s \times L_{app2.0} \rightarrow A_{app2.0}^s$ is a function that corresponds to a time-stamped authorized relation connecting a social action, that occurred at a certain level of exchange ($l \in L_{app2.0}$), to another social action; $sa_0^s \in A_{app2.0}^{s_0}$ is the initial social action; and, $FA^s \subseteq A_{app2.0}^s$ is a set of final social actions.

Example:

- $F_1^s = \langle A_{app2.0_1}^s, STR_{app2.0_1}^s, sa_{0_1}^s, FA_1^s \rangle$ where $A_{app2.0_1}^s = \{sa_1\}$, $STR_{app2.0_1}^s$ not applicable, $sa_{0_1}^s = sa_1$, and $FA_1^s = \{sa_1\}$.
- $F_2^s = \langle A_{app2.0_2}^s, STR_{A_{app2.0_2}^s}^s, sa_{0_2}^s, FA_2^s \rangle$ where $A_{app2.0_2}^s = \{sa_2, sa_{21}, sa_{221}, \dots\}$, $STR_{A_{app2.0_2}^s}^s = [(sa_2, sa_{21}), (sa_{21}, sa_{22}), [(sa_{21}, sa_{211})], [(sa_{22}, sa_{221}), (sa_{221}, sa_{222})]]$, $sa_{0_2}^s = sa_2$, and $FA_2^s = \{sa_{21}, \dots\}$. In this flow, 2 levels of exchange represented by [] and [[]], respectively, exist. Note that F_2^s contains 1 primary sub-flow referring to social actions between [] and 2 secondary sub-flows referring to social actions between [[]].

⁶ E.g., 1 nested-level of exchange would support 2 types of social flows: primary and secondary. In Facebook *comment* and *reply* can trigger *post*. Thus, *comment* and *reply* are part of the primary social-flow and *post* is part of the secondary social-flow.

- $F_3^s = \langle A_{app2.0_3}^s, STR_{app2.0_3}^s, sa_{0_3}^s, FA_3^s \rangle$ where $A_{app2.0_3}^s = \{sa_3, sa_{31}\}$,
 $STR_{app2.0_3}^s = \{(sa_3, sa_{31})\}$, $sa_{0_3}^s = sa_3$, and $FA_3^s = \{sa_{31}\}$
- $A_{app2.0}^{s_0} = \{sa_1, sa_2, sa_3\}$ for $F_{1,2,3}^s$.

In addition to control flow, business-to-social link, and social flow definitions, extra concepts and definitions are deemed necessary to allow the mining of social flows. Among these concepts we cite *scores* of social actions (*nodes* for short) that are calculated while the different social flows are under-development. As per Fig 1, each content to analyze is related to a single user, and is both opinionated and subjective. So, a *score* is either *local* that is about user's feedback, *global* that aggregates local scores using direct neighbors' scores, or cumulative that aggregates global scores using direct neighbors' scores, as well.

Definition 4. Local Score Function (LS). It quantifies a user's feedback on a social action using for instance, sentiment analysis techniques such as CoreNLP ([31]). CoreNLP's scores are -2 for very negative, -1 for negative, 0 for neutral, 1 for positive, and 2 for very positive. With respect to social actions that do not have content such as "like" and "wow", their sentiment values are assigned using "common sense", for example. Formally, $LS : A_{app2.0}^s \rightarrow \mathbb{Z}$ is defined as per Equation 1:

$$LS(sa) = \begin{cases} \text{sentimentAnalysis}(sa(\text{feedback})), & \text{sa has content} \\ \text{selfAssignment}(sa()), & \text{sa has no content} \end{cases} \quad (1)$$

where *selfAssignment* assigns a sentiment value to *sa* based on "common sense".

Definition 5. Global Score Function (GS). It represents the cumulative feedback of a social action's free-of-content and secondary neighbors at time *t*. The number of secondary neighbors depends on the Web 2.0 application's nested levels (1 for illustration purposes). Formally, $GS : A_{app2.0}^s \times T \rightarrow \mathbb{Z}$ is defined as per Equation 2:

$$GS(sa, t) = \begin{cases} \text{sign}(sa_{parent}) \times LS(sa) \\ \text{sign}(sa_{parent}) \times LS(sa) + \sum_{i=1}^k GS(sa_i, t), & \text{sa is secondary} \\ LS(sa) + \sum_{i=1}^k GS(sa_i, t) + \sum_{j=1}^m GS(sa_j, t), & \text{sa is primary and } \neq sa_0^s \text{ (Definition 3)} \end{cases} \quad (2)$$

where sa_{parent} is a social action's parent; *sign* is a function that returns +1 if $LS(sa_{parent})$ is positive or neutral, otherwise -1; sa_i is a free-of-content neighbor of *sa*; and, sa_j refers to all secondary neighbors of *sa*. Note that $GS(sa_0^s, t) = \sum_{j=1}^m GS(sa_j, t)$

where sa_j refers to all primary neighbors of sa_0^s .

From an implementation perspective, Algorithm 1 illustrates how Equations 1 and 2 for local and global score calculations, were programmed. At the end of each execution round, each social action gets its respective local and global score, that is used later for analysis and presentation. After a certain time of launching a campaign, collecting necessary details about social actions begins allowing to proceed as follows:

- Lines 4-8 list all reactions on the initial post. The self assignment value for each reaction is multiplied by the sign of the post to get the global score of the reaction, and added to the global score of the post.
- Lines 9-28 list all the comments on the initial post including reactions, replies, and reactions on replies, as well. All of these affect the global score of posts and comments via the nested loops. A comment's sentiment value is checked using CoreNLP tool, multiplied by the post's sign to get its global score and then added to the post's global score.
- Lines 12-26 proceed with the same analysis targeting this time reactions on comments, replies to comments, and reactions to replies, respectively.

Algorithm 1 Local/Global score calculation

```

1: for all posts as p do
2:   p.LS = 0
3:   p.GS = 0
4:   for all reactionsOnPost as rp do
5:     rp.LS = selfAssignment(rp)
6:     rp.GS = sign(p) * rp.LS
7:     p.GS = p.GS + rp.GS
8:   end for
9:   for all commentsOnPost as cp do
10:    cp.LS = sentimentAnalysis(cp.feedback)
11:    cp.GS = sign(p) * cp.LS
12:    for all reactionsOnComment as rc do
13:      rc.LS = selfAssignment(rc)
14:      rc.GS = sign(cp) * rc.LS
15:      cp.GS = cp.GS + rc.GS
16:    end for
17:    for all replyOnComment as rpc do
18:      rpc.LS = sentimentAnalysis(rpc.feedback)
19:      rpc.GS = sign(cp) * rpc.LS
20:      for all reactionsOnReply as rr do
21:        rr.LS = selfAssignment(rr)
22:        rr.GS = sign(rpc) * rr.LS
23:        rpc.GS = rpc.GS + rr.GS
24:      end for
25:      cp.GS = cp.GS + rpc.GS
26:    end for
27:    p.GS = p.GS + cp.GS
28:   end for
29: end for

```

Definition 6. Cumulative Score function (CS). It represents the cumulative score of a social action's direct neighbors at time t . The number of direct neighbors depends on the

Web 2.0 application's nested levels (1 for illustration purposes). Formally, CS is defined as per Equation 3:

$$CS(sa) = \begin{cases} GS(sa) + CS(sa_i, t), & sa \text{ is secondary} \\ GS(sa) + CS(sa_i, t) + CS(sa_j, t), & sa \text{ is primary} \end{cases} \quad (3)$$

where sa_i is both a secondary node and a certain direct neighbor of sa ; and, sa_j is both a primary node and a certain direct neighbor of sa . Note that Equation 3 provides different cumulative scores whether the social action is primary or secondary, and, therefore, considers the social flow's structure. Regarding the initial social actions, their CS is computed recursively using all nodes' cumulative scores. When a change happens in a social flow (e.g., new like, new comment, and new share), CS automatically changes.

When the development of a social flow is in progress, some scores automatically change (e.g., if like is connected to comment at time $t+1$, like's score at $t+1$ will be different from time t). This change would impact other nodes in the social flow through score propagation. We rely on asynchronous self-stabilization principle to propagate impacted global scores after each update (i.e., a newly-added social action to the social flow) [9]. This principle consists of re-computing the scores of initial social actions' neighbors. Each node checks its direct neighbors and detects any change of scores among their neighbors. If there is a change, the node computes its score again and again. Thanks to this domino effect, all nodes update their scores until reaching all initial social actions.

5.3. Authorized relations between social actions

The ongoing expansion of social flows is dependent on the authorized relations that a Web 2.0 application supports in order to connect social actions together ($STR_{app2.0}^s$ in Definition 3). Each Web 2.0 application allows a limited number of (next) social actions from which users can select for execution. Although these relations are not explicitly shown in Web 2.0 applications, we expose them for 2 reasons: enumerate the next possible social actions and recommend some next possible social actions to users with respect to what has been executed earlier. Enumerating the next possible social actions is relevant when building social flows; it permits to track exchanges online and to connect social actions on the fly.

Table 2 suggests examples of next possible social actions in some representative Web 2.0 applications. In this table, $[0/1.. * (resp.1)]sa$ means zero/one to many (resp. only one) social action(s) will be executed, and $(||)$ and (\oplus) are *or* and *xor* logical operators, respectively. To define some authorized relations in Web 2.0 applications, we analyzed Decker and Lesser's coordination relations between tasks namely *facilitates*, *cancels*, *inhibits*, *constrains*, *enables*, and *causes* [8]. Due to the inappropriateness of the first 3 relations for our work, we discuss the remaining ones:

1. $enables(sa_i, \{sa_j\})$: upon the successful execution of a social action sa_i , the Web 2.0 application activates other social actions $\{sa_j\}$ from which users can execute some (i.e., zero to many) and many times. Examples are $enables(share, \{like\})$ and $enables(post, \{share, like, comment\})$ in Facebook and $enables(tweet, \{reply, retweet, post - to - Facebook\})$ in Twitter.

2. $constrains(sa_i, \{sa_j\})$: upon the successful execution of a social action sa_i , the Web 2.0 application activates other social actions $\{sa_j\}$ from which users can execute one social action sa_j , only. Examples are $constrains(request, \{confirm, delete\})$ in Facebook and $constrains(follow, \{accept, deny\})$ in Instagram.
3. $causes(sa_i, \{sa_j\})$: upon the successful execution of a social action sa_i , another social action sa_j is automatically executed. Example is $causes(add, \{follow\})$ in Facebook.

Table 2. Illustration of some authorized relations between social actions

Web 2.0 application	Social action	Next possible social actions
Facebook	post	[0..*]like [0..*]comment [0..*]share [0..1]delete
	share	[0..*]like [0..1]delete
	like	[0..1]unlike
	follow	[0..1]unfollow
	comment	[0..*]reply [0..1]edit [0..1]delete
	friend request	[0..1]confirm \oplus [0..1]delete
Twitter	tweet	[0..*]reply [0..*]re-tweet [0..1]like [0..1]delete
	reply	[0..*]reply [0..*]like [0..*]re-tweet [0..1]delete
	quote tweet	[0..*]reply [0..*]re-tweet [0..1]delete
	like	[0..1]unlike
Instagram	post	[0..*]send to [0..*]comment [0..*]like [0..*]share-to
	comment	[0..*]reply [0..*]like [0..1]delete
	follow	[0..1]accept \oplus [0..1]deny
	send to	[0..1]like \oplus [0..*]comment

Let us consider GreenUtility and Facebook's social actions defined in Table 2. When no social action is executed, GreenUtility administrator executes post so that texts, images, or videos are displayed on the company's Facebook page. This post enables the administrator and other (un)known Facebook members to like that post (like), comment that post (comment), and/or share it (share). This happens because of $enables(post, \{like, comment, share\})$ authorized relation. In the same way, executing one of the recently enabled social actions will allow executing other social actions in a chain reaction. For instance, executing comment after post enables to like that comment (like) and/or to reply to that comment (reply) in compliance with $enables(comment, \{like, reply\})$ authorized relation.

5.4. Illustration

Let us apply the different definitions to GreenUtility. On the one hand, the campaign's business aspect refers to a control flow that includes many business tasks (e.g., prepare-campaign-material and approve-campaign-material) and dependencies between these

tasks (e.g., approve-campaign-material requires finalize-material before). On the other hand, the campaign's social aspect refers to multiple social flows initiated depending on the outcomes of executing certain business tasks. Let us assume that approve-campaign-material is successfully executed. Next is to share this material with the community on Facebook. The administrator logs into GreenUtility's Facebook page and executes post as a social action. Now that the campaign's material is online, subscribers of GreenUtility's Facebook page can share the material with others, comment the material, or like the material as per the authorized relations associated with post (Table 2). If a person makes a comment, then comment as a social action is executed and will be connected to the first social action that is post. At this stage, the under-development social flow consists of two social actions: post then comment.

In Fig. 3a, we show GreenUtility's post on Facebook at time t . This post has resulted into executing additional actions by people like Alison, Bob, and the administrator of GreenUtility's Facebook page. In Fig. 3a, we map this execution onto an under-development social flow that will grow over time. Since Facebook supports 1 nested level of exchange, the social flow is represented as 1 primary (level 0) *caterpillar* (i.e., a tree such that its internal vertices constitute a path and the other vertices are the "hairs" of the tree and 2 secondary (level 1) *caterpillars* [13]. The primary *caterpillar* has GreenUtility's post as a root with 2 like and 2 subsequent comment while the first secondary *caterpillar* has Bob's reply as a hairless root. The nodes are labeled with 3 values (sentiment score, global score, and cumulative score). All under-development social flows are acyclic and temporal.

6. System development

This section consists of 2 parts. The first part describes the architecture supporting real-time monitoring and mining of users' social actions over Facebook. The second part describes the experiments that were carried out along with the results of these experiments.

6.1. Architecture

We developed a tool, named Social Miner (SM), for tracking and mining users' actions over social media with Facebook as a targeted Web 2.0 application. SM's architecture is given in Fig. 4 (a demo video is available at <https://youtu.be/crBsEk2pSzo>) and consists of 4 modules: *dashboard*, *social-action manager*, *social-action tracker*, and *social-flow analyzer*.

The *dashboard* is the interface provided to employees and BP engineers to manage campaigns on GreenUtility's Facebook page like launching a new campaign with the assistance of the *social-action manager* (1.1) and to perform the necessary analysis (1.2). The *social action tracker* "keeps-an-eye" on any change over this Facebook page while the *social-flow analyzer* obtains insights into the activities over the Facebook page such as, which subscribers are (un)supportive of a campaign and which campaign is most attractive according to our mining analysis. To this end, the *social action tracker* uses Webhooks⁷ to subscribe to changes in the Facebook page. These changes are stored in the *social actions*

⁷ developers.facebook.com/docs/graph-api/webhooks..

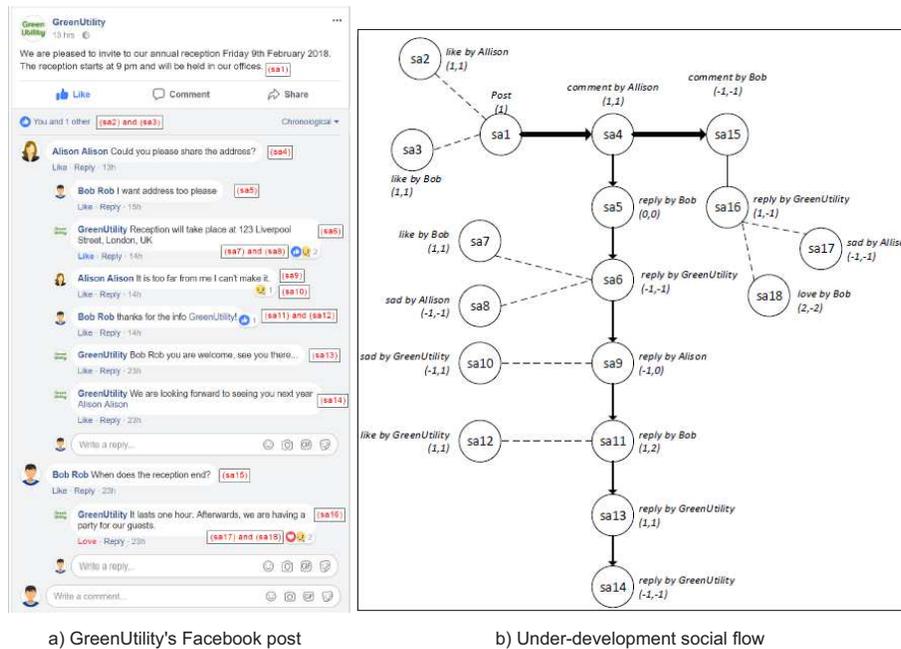


Fig. 3. GreenUtility's Facebook page and its associated social flows

repository and then made available to the *social-flow analyzer* for building the necessary *social flows* so they are mined at a later stage.

1. The *social-action manager* uses Facebook SDK library for PHP so that requests (e.g., publish campaign and reply to some comment message) are submitted to Facebook Graph API⁸ and published on a Facebook page.
2. The *social-actions repository* is a MySQL database that stores details like time about the social actions executed over GreenUtility's Facebook page.
3. The *subscribers repository* is another MySQL database that stores details about the subscribers (e.g., user id, user name, and weight) who take part in the discussions over GreenUtility's Facebook page.
4. The *social action tracker* includes 2 modules: *CoreNLP* and *Score calculator*. On top of time-stamped details about the executed social actions (e.g., user id, page id, post id, and parent id) collected via Facebook Webhooks' notifications, CoreNLP as a sentiment-analysis tool annotates these actions with sentiment scores. The result of CoreNLP analysis is formatted as JavaScript Object Notation (JSON) and then, translated into a relational format (through an in-house script) prior to storing it in the *social-actions repository*. The relational format has eased the storage of different details in multiple tables and running queries over these tables when building and analyzing the social flows. Any notification from Webhooks (e.g., newly added/updated social actions) triggers the *score calculator* that computes new scores (Equa-

⁸ developers.facebook.com/docs/graph-api.

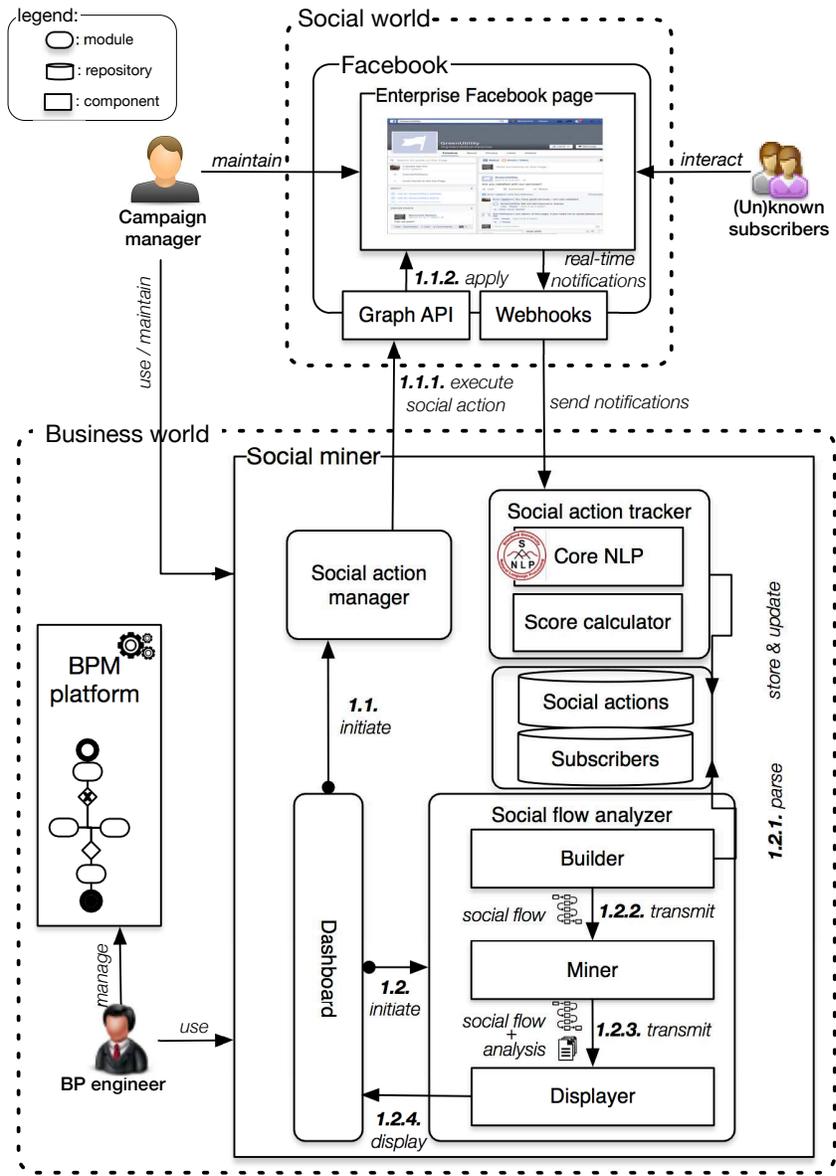


Fig. 4. Architecture of Social Miner

tions 4 and 5) and/or revises some existing scores as per the score propagation algorithm (Definition 6). For local scores, the *score calculator* considers subscribers' weights (e.g., reputation) using the *subscribers repository* (Definition 4).

- The *social flow analyzer* includes 3 sub-modules: *builder*, *miner*, and *displayer*. The *builder* parses the content of the *social actions repository* (1.2.1) to generate the necessary social flows enriched with scores and transmit the enriched social flows to the *miner* for further analysis (1.2.2). The *miner* performs 3 types of metrics and one analysis discussed in Section 6.2. The *displayer* visualizes real-time social flows along with the obtained analysis transmitted by the *miner* (1.2.3) on a browser showing how GreenUtility promotes its services to customers and seeks their feedback through Facebook (Fig. 2). The *displayer* uses Cytoscape graph-theory library for analysis and visualization [11]. It also highlights with assistance of the *miner* the relevant social actions that form the social flows. Different shapes are used to differentiate the social actions: star for post, rectangle for comment, hexagon for reply and triangle for reaction. The *displayer* also uses colors to emphasize whether a social action is positive, negative, or neutral so that a manager can easily identify the points of interest. GreenUtility uses the different flows to identify what social actions that (un)known subscribers have executed over its Facebook page. This could lead into reviewing BPs if their feedback were deemed relevant. Finally, the *displayer* is developed in HTML 5 and JavaScript while the *builder* and *miner* are developed as PHP programs and deployed on an Apache Web server.

Besides installing Facebook Graph API, companies interested in using Social Miner do not require any additional installation or configuration to track their campaigns on Facebook.

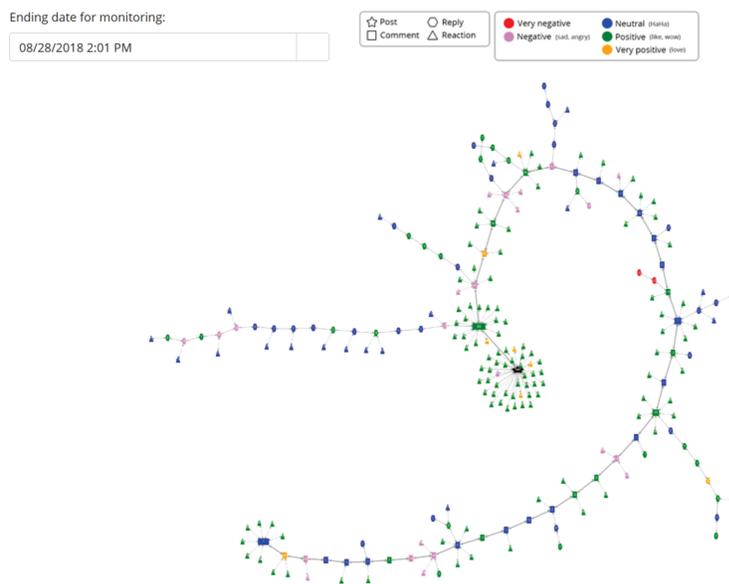


Fig. 5. Example of social flow

6.2. Experiments

To evaluate the benefits of using SM to decision makers, we carried out many experiments associated with a real campaign known as “*we are announcing Universe 11 plus, the greatest phone ever*” that was active from March 11, 2018 to March 16, 2018 on Facebook. The metrics that result out of these experiments are discussed below and assessed over one-day long time intervals. These metrics are implemented in the *miner* sub-module, part of the *social flow analyzer* (Fig. 4).

1. *Campaign attractiveness* metric (M_1 , Fig. 6) defines how appealing a campaign was to respondents by tracking their positive, neutral, and negative responses over different time intervals. Formally, M_1 is defined by Equation 4. By considering attractiveness, managers could extend campaigns, for example.

$$M_1(t_i) = \frac{new(t_i)}{returning(t_i) + new(t_i)} \quad (4)$$

Where t_i is a certain time interval [from, to] that could be days, weeks, months, etc., $new(t_i)$ is the number of new respondents who executed some social actions during t_i , and $returning(t_i)$ is the number of returning respondents who first, executed some social actions during t_i and second, were included in the previous time interval (t_{i-1}) that was used for defining the attractiveness metric. At t_0 , all respondents are treated as new. We rely on the time-stamped authorization relation $STR_{Facebook}^s$ (Definition 3) to compute $new(t_i)$ and $returning(t_i)$.

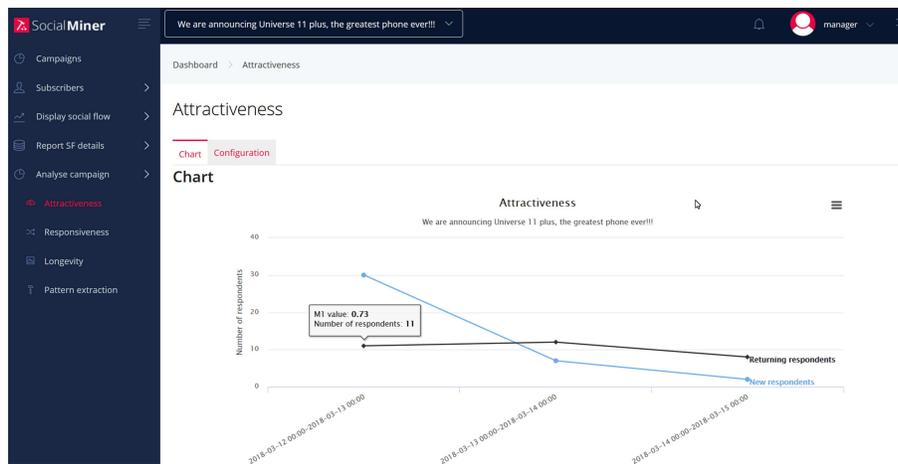


Fig. 6. Chart associated with campaign attractiveness

Since M_1 enables a campaign’s manager to discuss attractiveness from a global perspective, the focus is on the number of (new and returning) respondents. It would be useful for the manager to study attractiveness from a local perspective by identifying

the social actions that led for instance, to a major increase/decrease in the number of new/returning respondents both compared to previous interval times. To this end, we define *local-attractiveness* metric (M'_1) with focus on returning respondents' responsiveness levels (Equation 5):

$$M'_1(sa, t_i, t_{i+1}) = \frac{\text{returning}(sa, t_i, t_{i+1})}{\text{returning}(sa, t_{i-1}, t_i)} \quad (5)$$

Where sa is a certain social action that is subject to analysis, t_{i-1}, t_i, t_{i+1} are 3 homogeneous time intervals such that t_i happened earlier than t_{i+1} , $\text{returning}(sa, t_i, t_{i+1})$ is the number of returning respondents who were "new" at t_i and executed any action that came after sa during t_{i+1} , and finally, $\text{returning}(sa, t_{i-1}, t_i)$ is the number of returning respondents who were "new" at t_{i-1} and executed any action that came after sa during t_i . Similar to $\text{new}(t_i)$ and $\text{returning}(t_i)$, $\text{returning}(sa, t_i, t_{i+1})$ and $\text{returning}(sa, t_{i-1}, t_i)$ are computed based on $STR_{Facebook}^s$. Since a social action may appear many times in the time interval t_{i+1} , the manager points the desired social action that he wishes to analyze using first occurrence, for example.

2. *Campaign responsiveness* metric (M_2 , Fig. 7) indicates how a campaign is perceived by the community of respondents based on their feedback, whether positive (supportive), negative (opponent), or neutral. We rely on the local score functions (Definition 4) to formally define M_2 with focus on positive feedback in Equation 6:

$$M_2(t_i) = \frac{|sa|_{\text{positive}}}{|sa|_{\text{positive}} + |sa|_{\text{neutral}} + |sa|_{\text{negative}}} \quad (6)$$

Where t_i is a certain time interval [from, to], $|sa|_{\text{positive}}$ is the number of social actions executed during t_i such that $\text{sign}(GS(sa, t))$ is positive ($t \in t_i$), $|sa|_{\text{negative}}$ is the number of social actions executed during t_i such that $\text{sign}(GS(sa, t))$ is negative, and $|sa|_{\text{neutral}}$ is the number of social actions executed during t_i such that $LS(sa)$ is zero.

3. *Campaign longevity* metric (M_3 , Fig. 8) indicates how long a campaign remained "alive/active", i.e., respondents have continuously (without "big" gaps) provided feedback on the campaign so that a certain activity level over Facebook for example, is maintained. The longevity metric refers to the longitudinal dispersion of the provided feedback over a certain time interval t_k ([from, to]) that shall fall into a certain accepted activity level set by the campaign's manager (e.g., minimum number of actions in a day). We define this activity level with respect to a standard deviation (σ) upon which the decision of putting the campaign either on hold (suspend) or offline (stop). We rely on the time-stamped authorization relation $STR_{Facebook}^s$ to formally define this standard deviation as per Equation 7:

$$\sigma = M_3(t_k) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (7)$$

where t_k is a time interval [from, to] sliced into n equal time intervals (t_i) (e.g., in days and in weeks), x_i is the number of social actions executed during the slice

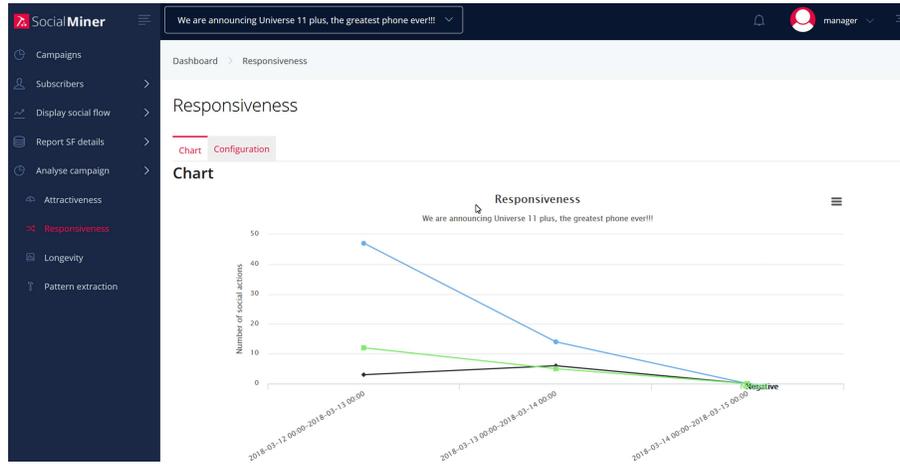


Fig. 7. Chart associated with campaign responsiveness

$t_i \subset t_k$ where $i \in [1, n]$, and \bar{x} is the average number of all social actions executed during the time interval $t_k (\sum_{i=1}^n x_i/n)$. The number of social actions over a certain time interval remains “acceptable” if this number falls into $[\bar{x} - \sigma, \bar{x} + \sigma]$. “Acceptable” could be defined over time and by benchmarking different time intervals together. We, thus, analyze the longevity metric according to \bar{x} and σ so that this metric’s time space corresponds to the set of time intervals where the number of social actions is declared “acceptable”.

In addition to the different metrics that SM produces, a *reversal trend analysis* of a campaign is implemented to help a manager identify the reasons behind a change in a campaign’s perception, for example. This perception could be based on a series of positive/negative and then negative/positive feedback. SM relies on the social flows’ secondary caterpillars to look for potential patterns such as 2 consecutive positive feedback followed by 3 consecutive negative feedback, and so on. To achieve the reversal analysis, we adopted gSpan algorithm for mining labeled graphs [37]. This algorithm uses a set of graphs \mathbf{D} and the minimum frequency (i.e., number of subgraphs before claiming that these subgraphs are repetitive) as inputs. In our case, \mathbf{D} could be a set or portions of secondary caterpillars. Because many social actions can be executed over time, we “cleaned” the secondary caterpillars from irrelevant time intervals (e.g., those where the campaign is inactive) for quality purposes. In Fig. 9, the minimum support threshold is 3 as an example, and all the social actions until August 28, 2018 are considered when extracting the patterns.

7. Discussions

From a business point of view, according to Zion Market Research, “*Global enterprise 2.0 technologies market expected to reach around USD 14,955 million by 2024, at a*

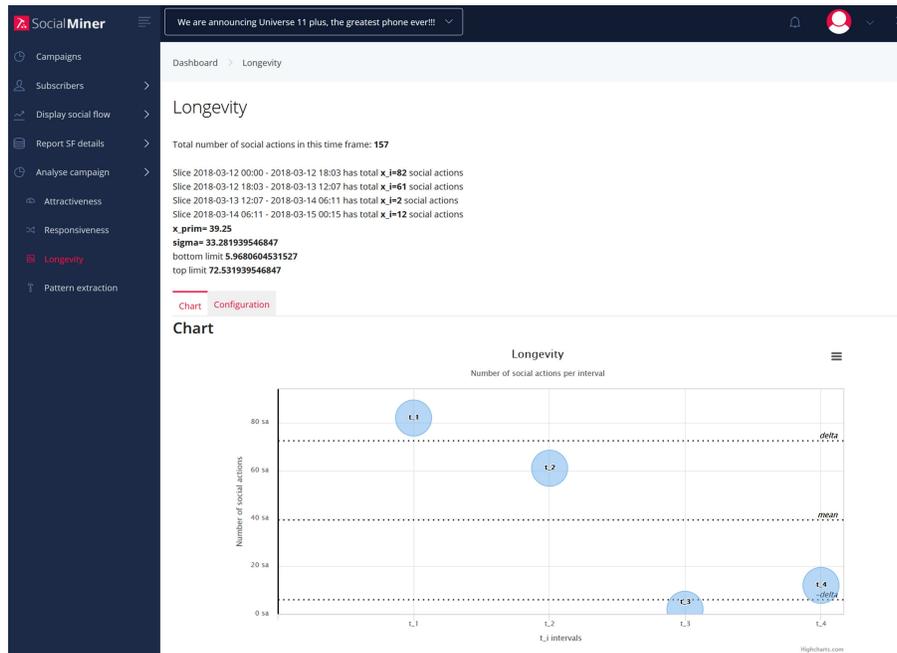


Fig. 8. Chart associated with campaign longevity

CAGR approximately 46.87 for the forecast period from 2018 to 2024”⁹. Despite this heavy investment and the social fever that has caught every single activity of people’s daily lives, many are still reluctant to embracing social media whether for personal use or for business use. Different concerns are continuously raised, including whether social media is bringing any value-added to companies. Gartner clearly states that “... many large companies are embracing internal social networks, but for the most part, they are not getting much from them” [15]. And, social media is also seen as the source of new forms of security threats, privacy breaches, and distraction to employees [10]. Contrarily to these “skeptical” views, a London-based think tank, Demos, encourages companies to allow their employees to embrace social network applications in order to establish and foster contacts with stakeholders such as colleagues, customers, and suppliers [12]. Striking the right balance between social media’s pros and cons requires strict guidelines that social media users should comply with. Burégio et al. present such guidelines from 3 perspectives known as technology, organization, and management [6]. The technology perspective identifies the appropriate type of social media that should sustain a company growth and fall into its mission. The organization perspective puts in place the necessary procedures that should ensure an efficient use of social media to avoid misuses, for example. Finally, the management perspective identifies the metrics (or key performance

⁹ www.zionmarketresearch.com/report/enterprise-technologies-market.., last visited November, 14, 2019

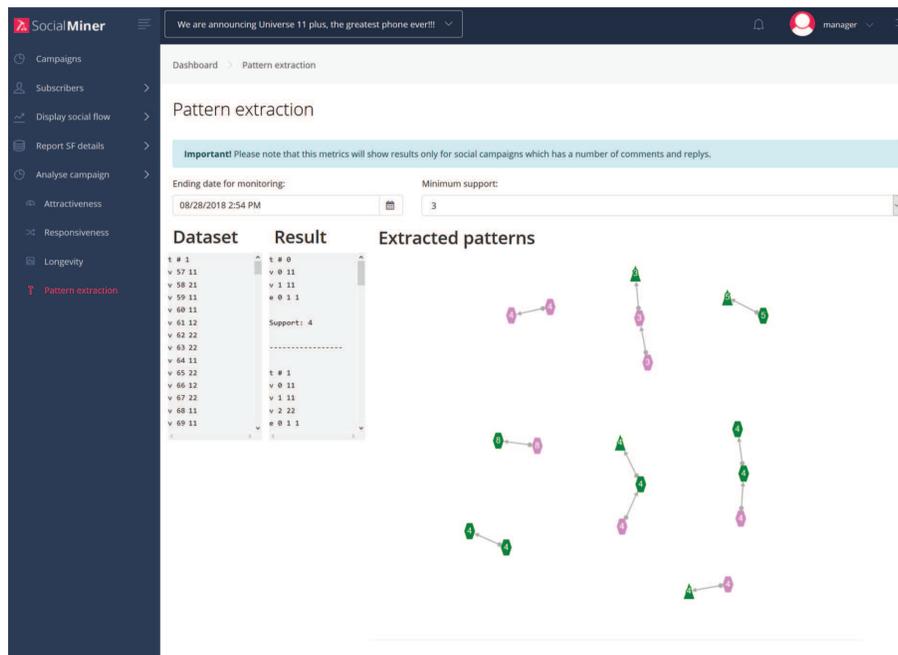


Fig. 9. Pattern extraction during a campaign

indicators) that should permit to evaluate the efficient use of social media based on some tangible benefits.

8. Conclusion

We presented an approach for developing flows in the context of companies that wish to tap into social media's opportunities. The flows are specialized into control and social. The former consists of tasks that form business processes. The latter consists of social actions that are executed over social media in response to specific events. Social flows are enriched with scores based on sentiment analysis so that companies would secure a better understanding of what-happened and what-might-happen in their ecosystems. For validation purposes, we developed a tool, Social Miner, on Facebook allowing to track and mine users' posts, comments, responses, etc. The system permits to answer questions like what actions were frequently executed, why certain actions were executed more than others, and when such actions were executed. In term of future work we would like to examine the impact of contextual factors on the next social actions to execute and the deployment of Social Miner on another social media such as Twitter.

References

1. van der Aalst, W., Song, M.: Mining Social Networks: Uncovering Interaction Patterns in Business Processes. In: Proceedings of the Second International Conference on Business Process

- Management (BPM'2004). Potsdam, Germany (2004)
2. Archak, N., Ghose, A., Ipeirotis, P.: Show me the Money!: Deriving the Pricing Power of Product Features by Mining Consumer Reviews. In: Proceedings of the 13th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD'2007). San Jose, CA, USA (2007)
 3. Batrinca, B., Treleven, P.: Social Media Analytics: A Survey of Techniques, Tools, and Platforms. *AI Soc.* 30(1) (2015)
 4. Bonchi, F.: Influence Propagation in Social Networks: A Data Mining Perspective. *IEEE Intelligent Informatics Bulletin* 12(1) (2011)
 5. Bonchi, F., Castillo, C., Gionis, A., Jaimes, A.: Social Network Analysis and Mining for Business Applications. *ACM TIST* 2(3) (2011)
 6. Burégio, V.A., Maamar, Z., Meira, S.L.: An Architecture and Guiding Framework for the Social Enterprise. *IEEE Internet Computing* 19(1) (2015)
 7. Cross, R., Gray, P., Cunningham, S., Showers, M., Thomas, R.: The Collaborative Organization: How to Make Employee Networks Really Work. *IEEE Engineering Management Review* 39(1) (2011)
 8. Decker, K., Lesser, V.: Generalizing the Partial Global Planning Algorithm. *Int. J. Cooperative Inf. Syst.* 1(2) (1992)
 9. Dolev, S.: Self-stabilization. MIT press (2000)
 10. Faci, N., Maamar, Z., Burégio, V., Ugljanin, E., Benslimane, D.: Web 2.0 Applications in the Workplace: How to Ensure their Proper Use? *Computers in Industry* 88 (2017)
 11. Franz, M., Lopes, C., Huck, G., Dong, Y., Sümer, S., Bader, G.: Cytoscape.js: A Graph Theory Library for Visualisation and Analysis. *Bioinformatics* 32(2) (2016)
 12. Howe, J.: Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business. *Crown Business* (2009)
 13. Jalali, A.: Aspect Mining in Business Process Management. In: Proceedings of the 13th International Conference on Perspectives in Business Informatics Research (BIR'2014). Lund, Sweden (2014)
 14. Kajan, E., Faci, N., Maamar, Z., Loo, A., Pljaskovic, A., Sheng, Q.Z.: The Network-based Business Process. *IEEE Internet Computing* 18(2) (2014)
 15. Kanaracus, C.: Gartner: Social Business Software Efforts Largely Unsuccessful, For Now (2013), <https://tinyurl.com/y5lc9eg4>
 16. Liu, B.: Sentiment Analysis and Subjectivity. Taylor and Francis Group, Boca (2010)
 17. Lu, Y., Zhai, C.: Opinion Integration Through Semi-supervised Topic Modeling. In: Proceedings of the 17th International Conference on World Wide Web (WWW'2008). Beijing, China (2008)
 18. Luo, Z., Osborne, M., Wang, T.: Opinion Retrieval in Twitter. In: Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM'2012). Dublin, Ireland (2012)
 19. Maamar, Z., Burégio, V.A., Faci, N., Benslimane, D., Sheng, Q.Z.: "Controlling" Web 2.0 Applications in the Workplace. In: Proceedings of the 19th IEEE International Conference on Enterprise Distributed Object Computing (EDOC'2015). Adelaide, Australia (2015)
 20. Maamar, Z., Burégio, V., Sellami, M., Rosa, N.S., Peng, Z., Subin, Z., Prakash, N., Benslimane, D., Silva, R.: Bridging the Gap Between the Business and Social Worlds: A Data Artifact-Driven Approach. *T. Large-Scale Data- and Knowledge-Centered Systems* 35 (2017)
 21. Maamar, Z., Faci, N., Kajan, E., Sakr, S., Boukhebouze, M., Barnawi, A.: How to Make Business Processes "Socialize"? *EAI Endorsed Trans. Indust. Netw. & Intellig. Syst.* 2(5) (2015)
 22. Maamar, Z., Faci, N., Sellami, M., Boukadi, K., Yahya, F., Barnawi, A., Sakr, S.: On Business Process Monitoring using Cross-Flow Coordination. *Service Oriented Computing and Applications* 11(2) (2017)

23. Martinho, D., Rito Silva, A.: Non-intrusive Capture of Business Processes Using Social Software. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) Business Process Management Workshops help in conjunction with BPM'2012. Tallinn, Estonia (2012)
24. OMG: Business Process Model and Notation (BPMN) version 2.0. Tech. rep., Object Management Group (2011)
25. OpenKnowledge: Social Business Process Reengineering. Tech. rep., OpenKnowledge (2012), <http://socialbusinessmanifesto.com/social-business-process-reengineering>
26. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2(1-2) (2008)
27. Pourshahid, A., Amyot, D., Peyton, L., Ghanavati, S., Chen, P., Weiss, M., Forster, A.J.: Business Process Management with the User Requirements Notation. *Electronic Commerce Research* 9(4) (2009)
28. Reichert, M., Dadam, P.: ADEPT_{flex}-Supporting Dynamic Changes of Workflows without Losing Control. *J. Intell. Inf. Syst.* 10(2) (1998)
29. Sadiq, S.W., Orlowska, M.E., Sadiq, W., Foulger, C.: Data Flow and Validation in Workflow Modelling. In: Proceedings of the Fifteenth Australasian Database Conference (ADC'2004). Dunedin, New Zealand (2004)
30. Soboroff, I., Ounis, I., Macdonald, C.: Overview of the TREC-2008 Blog Track (Jan 2010)
31. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP'2013). Seattle, Washington, USA (2013)
32. Tang, J., Chang, Y., Liu, H.: Mining social media with social theories: A survey. *SIGKDD Explor. Newsl.* 15(2) (2014)
33. Tang, J., Sun, J., Wang, C., Yang, Z.: Social influence analysis in large-scale networks. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'2009). Paris, France (2009)
34. Ting, S., Ip, W., Tsang, A.: Is Naive Bayes a Good Classifier for Document Classification? *International Journal of Software Engineering and Its Applications* 5(3) (2011)
35. Turban, E., Bolloju, N., Liang, T.: Enterprise Social Networking: Opportunities, Adoption, and Risk Mitigation. *J. Org. Computing and E. Commerce* 21(3) (2011)
36. Ugljanin, E., Faci, N., Sellami, M., Maamar, Z.: Tracking Users' Actions over Social Media: Application to Facebook. In: Proceedings of the 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'2016). Paris, France (2016)
37. Yan, X., Han, J.: gSpan: Graph-Based Substructure Pattern Mining. In: Proceedings of the 2002 IEEE International Conference on Data Mining!(ICDM'2002). Maebashi City, Japan (2002)

Ejub Kajan is an Associate Professor of Computer Science at State University of Novi Pazar, Serbia. His research interests include social computing, interoperability, service computing, business process management and IoT. He has a PhD in computer science from University of Nis, Serbia. He has published with almost all major publishers like ACM, Elsevier, IEEE, Springer, Wiley, etc. He is a Senior Member of ACM.

Noura Faci is an Associate Professor at the University Lyon 1, France since October 2008. Her current research interests are service computing, social computing, business process management, and IoT. She has published several papers in high-quality journals

and conferences and actively contributes to the IEEE TSC, IEEE IC, and Computer journal's review process, as well.

Zakaria Maamar is a Professor in the College of Information Technology at Zayed University, Dubai, United Arab Emirates. His research interests include Web services, social networks, and context-aware computing. He has a PhD in computer science from Laval University, Quebec City, Canada.

Mohamed Sellami (http://www-public.imtbs-tsp.eu/sellam_m/) is an Associate Professor of Computer Science at Telecom SudParis member of Institut Polytechnique de Paris and the SAMOVAR Laboratory (Evry, France). He received his PhD in computer science from Telecom SudParis in 2011. His main research interests are related to service computing and cloud computing. His publication list includes international journals and conferences and he serves on the program and organizing committees of numerous international conferences and workshops.

Emir Ugljanin is a PhD student at the University of Nis. He has MA degree from Technical faculty at the University of Novi Sad and engineer's degree from the State University of Novi Pazar. His current research interests are Internet of Things, social web and business process management.

Hamamache Kheddouci is full Professor in Computer Science at Lyon 1 University since 2004. He received his PhD degree in Computer Science from Paris XI University in 1999. In 2003, he obtained his research supervision habilitation in Computer Science from the Burgundy University, Dijon. His research interest includes combinatorial and algorithmic aspects of graphs and their applications, in particular, in big data and social networks. For more details, see: <http://perso.univ-lyon1.fr/hamamache.kheddouci/>

Dragan Stojanović is a Professor in Computer Science, at the Faculty of Electronic Engineering, University of Niš, Serbia. His research interests include Big Data processing and analytics, spatio-temporal and mobility data management, as well as mobile and ubiquitous systems and IoT in Smart Cities. He has published widely in those and related topics.

Djamal Benslimane (<http://www710.univ-lyon1.fr/dbenslim/>) is currently a Professor at Lyon 1 University in France. He is the head of the computer science department at the IUT Lyon 1. His research interests lie in the areas of Services computing and database interoperability. His recent works were published in IEEE Transaction on Knowledge and Data Engineering, IEEE Internet Computing, IEEE Transactions on Services Computing, IEEE Transactions on Systems, Man, and Cybernetics, Communications of the ACM, ACM Transactions on Internet Computing, ACM Transactions on Software Engineering and Methodology, and WWW Journal.

Received: August 20, 2019; Accepted: February 20, 2020.