Available online at www.sciencedirect.com

ScienceDirect

Procedia Economics and Finance 27 (2015) 636 – 645

Procedia Economics and Finance

www.elsevier.com/locate/procedia

22nd International Economic Conference – IECS 2015 "Economic Prospects in the Context of Growing Global and Regional Interdependencies", IECS 2015

Markov Model for Tweets Geographic Distribution Characterization

Eduard Alexandru Stoica[a,]*, Antoniu Gabriel Pitic[b], Marian Pompiliu Cristescu[a]

[a]*Lucian Blaga University of Sibiu, Faculty of Economic Sciences, 17 Dumbrăvii Street, Sibiu 550324, Romania*
[b]*Lucian Blaga University of Sibiu, Faculty of Engineering, 4 Emil Cioran Street, Sibiu 550025, Romania*

Abstract

In this paper we will continue our researches regarding e-Business and e-Government modeling on Social Media presented in (Stoica, Pitic, & Mihaescu, 2013). Among message and user parameters we add a new parameter used to describe the geographical dispersion of Twitter messages. This new parameter will characterize the way one set of messages will spread in Social Graph from the physical word point of view. The first model, presented as "A Novel Model for E-Business and E-Government Processes on Social", will be extended with the geographical parameter P_G. We will define and we will describe the Markov Model used to organize the messages gathered from social media. The main idea of building the Markov Model is to assign a geographical location to each user who send a message and every re-broadcast will define a transition.

Peer-review under responsibility of Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu"

Keywords: Markov model, social graph, social media, Twitter;

1. Introduction

Social Media offers a huge amount of data which was not available before technological revolution. The data is used by data mining scholars in new ways, like the user interactions on social networks (Macskassy, 2012), (Wilson, Boe, Sala, Puttaswamy, & Zhao, 2012), (Viswanath, Mislove, Cha, & Gummadi, 2009) or (Wilson, Sala, Puttaswamy, & and Zhao, 2012), the community structure (Leskovec, Lang, Dasgupta, & Mahoney, 2008), (Weng, Menczer, & Ahn, 2013) , the marketing on social networks (Arthur, Motwani, Sharma, & Xu, 2009), (Hartline, Mirrokni, & Sundararajan, 2008) or privacy related researches (Hader & Brown, 2010).

* Corresponding author. Tel.: +4 0740.060.967
E-mail address: eduard.stoica@ulbsibiu.ro (E.A. Stoica), antoniu.pitic@ulbsibiu.ro (A.G. Pitic), marian.cristescu@ulbsibiu.ro (M.P. Cristescu)

Peer-review under responsibility of Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu"
doi:10.1016/S2212-5671(15)01066-7

Data gathering from social networks can be achieved using free or paid tools as Tweetarchivist or Sproutsocial. In our studies we used TW-Crawler, a software presented in (Stoica, Pitic, & Tara, 2012).

2. Used entities

The models that are proposed can be used, theoretical, in any social network. However, the practical way of obtaining entities will be network dependent. Some of the definitions was presented in (Stoica, Pitic, & Mihaescu, 2013).

- **Definition 1.** We note with U the user of a social network.
- **Definition 2.** We note with M the message that is transmitted by a user to some U.
- **Definition 3.** We note with @, @ , the multitude of users mentioned in a message. We have @={U_1, …, U_n}.
- **Definition 4.** We note with #, # $\in M$, the terms of interest from a message. We have # = {T1, ..., Tn}, where, Ti = 1, n, is a term of interest.
- **Definition 5.** We note with B the message body, B = M {@, #}.
- **Definition 6.** We note with #ext the set of terms that are similar or derived from an original term
- **Definition 7.** We note with SG (Social Graph) the directed graph with nodes formed of users, which is derived from a term (#), a user (U) or geographical informations (G).
- **Definition 8.** We note with C a generic classification of users or messages.
- **Definition 9**. We note with G information related to geographic location of a user (U).

3. General architecture

We extend the architecture presented in (Stoica, Pitic, & Mihaescu, 2013). In the same paper we present the modules used to obtain the P_M and the P_U parameters of the model.

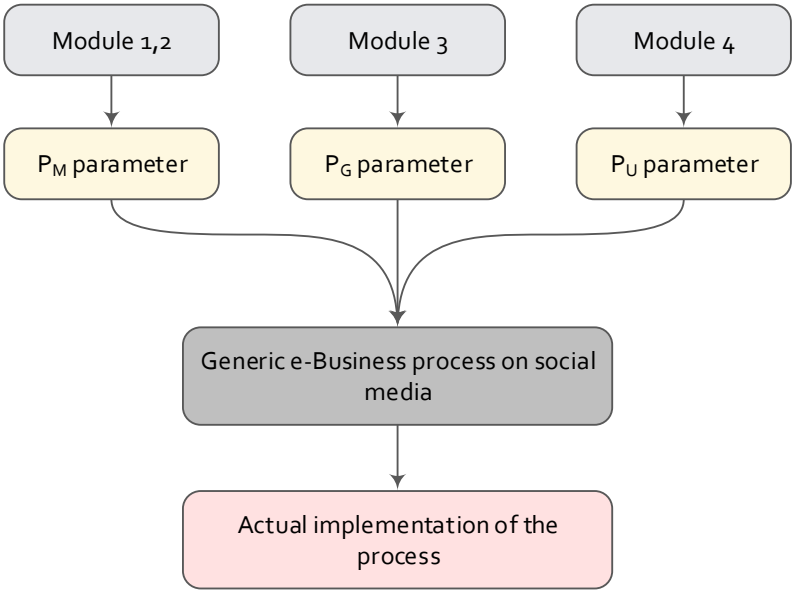

Figure 1. The general architecture