Information Diffusion and Provenance of Interactions in Twitter: Is it only about Retweets?

Io Taxidou[*] Peter M. Fischer[*] Tom De Nies[‡]
Erik Mannens[‡] Rik Van de Walle[‡]

[*]{taxidou,peter.fischer}@cs.uni-freiburg.de
University of Freiburg, Germany
[‡]{tom.denies,erik.mannens,rik.vandewalle}@ugent.be
Ghent University - iMinds - Data Science Lab, Belgium

ABSTRACT

This paper sheds light on the different interaction types among social media users that benefit information diffusion and provenance analysis. In particular, we identify explicit and implicit interactions in Twitter, including informal conventions applied by users. In our empirical evaluation considering only retweets, the most common means of information propagation in Twitter, we can infer 50% of message provenance. However, if we consider other types of interactions, we can explain another 13%. Accordingly, we enrich the PROV-SAID model for information diffusion, which extends the W3C PROV standard for provenance.

1. INTRODUCTION

Social media facilitate information diffusion through the social connections they support. The analysis of information diffusion mechanisms include the understanding of the *provenance* of such information, i.e., how it reached its current state. This provenance typically includes the information sources and the intermediate steps that were taken to produce it. However, this information is currently not fully exposed by social media providers.

Information diffusion and provenance mechanisms are mainly studied using one obvious propagation means – e.g., retweets in Twitter, reposts in Facebook, revines in Vine. However, implicit diffusion mechanisms adopted by social media users that offer a more complete picture are being ignored. Complementary, the combination of different means of propagation has not been investigated thoroughly, resulting again in incomplete provenance.

In this paper, we empirically identify implicit mechanisms and conventions of information diffusion that will allow researchers to incorporate them into their analysis. We focus on Twitter, one of the most well studied networks. The knowledge of how users interact implicitly and explicitly, and the combination of these, facilitate the deeper understanding of information diffusion and provenance. According to our insights, we expand PROV-SAID [2], our model for information diffusion extending the W3C PROV Data Model.

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2889393.

2. TYPES OF INTERACTIONS

We discern two main categories of interactions: the well-known *explicit* interactions that are produced by using social media features (in our case Twitter), whereas *implicit* interactions are generated by users' conventions. As a result, implicit interactions are harder to identify and more challenging to study. For explicit interactions, Twitter provides the 1) *Retweet*, 2) *Reply*, and 3) *Quote* features. While Retweets and Replies are well-studied mechanisms, the newly created feature of Quote, resembles the Retweet but with the possibility to comment while propagating messages. For retweets, the provenance information that is enclosed includes the root of the retweet cascade, while the intermediate steps are not being exposed; for replies and quotes, the previous step is exposed, the root is not. The combination of such features further complicates things: for example when a user quotes a retweet, the retweet root gets lost.

For implicit interactions, we identified the following categories by observing sets of similar messages in Twitter datasets: 1) *user influence*: a) with *explicit credit*, b) *without credit*, 2) *external influence*, 3) *self-influence*: a) *delete and rewrite*, b) *promotion*

1) As a first case, we observed that a user is influenced by another user by propagating similar messages. Here, two cases can occur. a) Sometimes, users prefer to give *explicit credit* to the initial contributor by mentioning the username within the message text (with "@" or "via"). This behaviour was adopted by users before the retweet feature was released. b) In case there is no explicit credit within the message text and there is still high similarity between two messages, we check if the users are neighbours in the social graph. Since users are exposed to the messages by their connections, there is a high chance that they are influenced by them.

2) If no social graph connection exists between two users that emitting highly similar messages, we observed that there is either an *external* event that drives the similarity (e.g., a football match) or influence from the public trends.

3) As a last case, we observed that many highly similar messages share the same author. For Twitter, we discerned two categories: a) highly similar messages from the same author with the earlier being deleted, as the result of users who *delete and rewrite* their own messages, since there is no edit feature provided by Twitter; b) *promotion* of existing information by the same user, e.g., for advertisement. Figure 1 depicts the workflow we use to discern these cases.

Empirical Evaluation.

To illustrate the identified interactions and their combination, we empirically evaluated a controlled dataset that was crawled during the ISWC 2015 conference. The dataset contains 3909 messages,

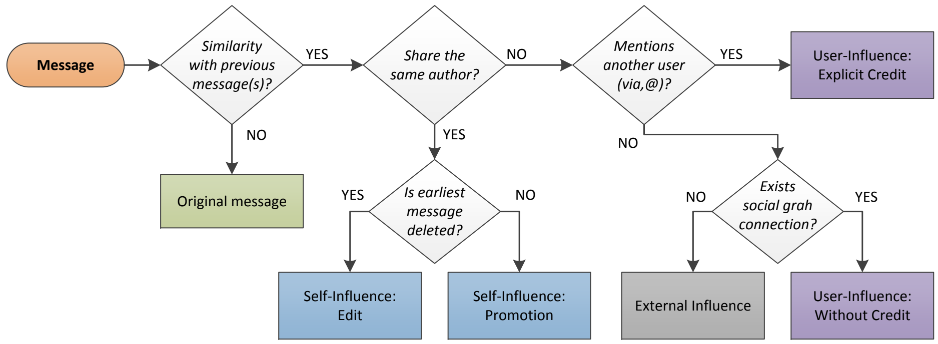

Figure 1: Flow Diagram for implicit interactions

consisting of 2068 retweets, 198 quotes, and 93 replies. As for the combination of these interactions, half of the quoted messages also received retweets, more than a third of the messages that received replies received also retweets and 17 messages are part of a reply chain. This means that we cannot study different interaction types in isolation. In order to identify implicit interactions and reconstruct the missing provenance we used our previously proposed approach [1]; its core assumption is that "if two messages are highly similar, there is a high probability that they share a part of their provenance". We observed that messages that exhibit more than 0.4 similarity in our dataset possibly share some provenance.

By using a similarity-based clustering algorithm with this threshold, we were able to reconstruct the provenance for 192 messages, that was not possible to identify with the explicit Twitter means. 87% of these messages share provenance by mentioning a username in the text (with 'RT', '@', or 'via'), while 90% of the messages refer to the events of the conference, thus demonstrating external influence. 16 pairs of similar messages share the same author, out of which 11 were deleted and re-written (edited) and the remaining 5 were information promotion. Following the workflow in Figure 1, we successfully reconstructed the provenance for these messages that were diffused by implicit interactions, which was not possible using only our previous work.

3. EXTENDING PROV-SAID

After identifying the different cases of interactions, we use them to extend our PROV-SAID model [2] to assert the provenance of information diffusion on social media by extending the PROV-DM. By using such a model we can express and incorporate different types of interactions and influence in the same model. In addition, it allows the combination of data from different (social media) sources that do not share the same concepts and notations.

In this model, users represent *agents* who are eligible for certain *activities*, and as a result *influence relationships* are established. PROV-SAID already supports *user influence*, by providing two sub-types for the generic concept *prov:Influence*:
1) One agent is influenced by another, by establishing a unidirectional (follow) relationship. This type of activity is annotated as *prov-said:FollowActivity* and the influence relationship as *prov-said:FollowRelationship*. 2) An agent is influenced by another user, by having copied or revised messages of the latter. This type of influence is denoted using the activity type *prov-said:InteractionInfluenceActivity*, and the influence type *prov-said:InteractionInfluenceRelationship*. We extend the PROV-SAID model[1] to capture *self-influence* and *external influence*. More specifically, we define two new subtypes of *prov:Activity*: *prov-said:SelfInfluenceActivity*, and *prov-said:ExternalnfluenceActivity*. Two corresponding subtypes of *prov:Influence* are also added: *prov-said:SelfInfluenceRelationship* for self-influence and *prov-said:ExternalnfluenceRelationship* for external influence. For the self-influence, we express the constraint that two agents associated with a *prov-said:SelfInfluenceActivity* should be the same. For external influence, the agent who generates the influence may not be specified, since we can detect some influence but its source cannot be identified (e.g., in the case of external events). We provide an example of usage for the case of external influence.

Example: External Influence.

```
// User @Bob was influenced by an external event
// and the influencing agent is not specified
prov:wasInfluencedBy(twitter:Bob, -,
   [prov:type=
     'prov-said:ExternalInfluenceRelationship'])

// A prov-said:ExternalInfluenceActivity was
// started (and ended) at the moment user @Bob
// emitted the message having external influence
activity(bob-influenced-externally,
       2015-01-09T13:05:00, 2015-01-09T13:05:00,
       [prov:type=
         'prov-said:ExternalInfluenceActivity'])
```

By extending the PROV-SAID model, we allow more types of influence and means of propagation to be incorporated, thus enhancing the model's expressiveness and reusabality. This does not only provide new options for our provenance reconstruction approach, such as generic modeling but also allows the interaction of datasets originating from different social media.

4. CONCLUSION AND FUTURE WORK

We empirically identified explicit and implicit interactions in Twitter. We have shown that by relying solely on Twitter features, we are ignoring valuable provenance. As a consequence, we need to investigate into combinations of interactions and user conventions. Also, automating such processes is necessary to gain a deeper understanding into provenance and information diffusion. As social media are constantly evolving, so must our PROV-SAID model. Therefore, we continue to investigate and model new propagation means across multiple social media platforms.

5. REFERENCES

[1] T. De Nies et al. Towards multi-level provenance reconstruction of information diffusion on social media. In *CIKM*, 2015.
[2] I. Taxidou et al. Modeling information diffusion in social media as provenance with W3C PROV. In *WWW Comp*, 2015.

[1] http://semweb.mmlab.be/ns/prov-said/