You ain't from around here, are you? Message authoring guidance in social media

Pablo N. Mendes, Chris Kau, Alfredo Alba, Steve Welch, Daniel Gruhl, Neal Lewis, Clemens Drews

IBM Research Almaden, USA

Abstract. We have developed an application that helps users to tailor a message to a particular audience. It uses a message resonance model that rates the wording of a message based on an analysis of millions of pieces of data with regard to their historical success in the context of a target audience. We have shown that higher resonance messages are twice as likely to be retweeted. Based on a semantic concept expansion component, our application is able to find similar words that resonate better in the same context and propose word replacements to improve the overall likelihood of success.

For this demonstration, we are using Twitter messages and an audience of Cloud Computing experts and enthusiasts. Links to the Web application and supporting material are available from `http://swc14.pablomendes.com`. The username is 'reviewer2' and the password is 'gardacheluna' (without quotes).

1 Pitch

Brand language [17] is a recognized critical part of marketing, and word choice is an early lesson to anyone looking at writing [16]. The words we choose reveal aspects of personality [4], and often the more complex word choice favored in some work environments is not desired social media.

While the positive impact of high quality messaging is very difficult to quantify (albeit widely acknowledged [6]), the negative impact of low quality messaging is quite clear. From athletes losing hundreds of thousands of dollars in scholarships [3], to celebrities losing their jobs over a social media misstep [5], to a retailer unsuccessfully trying to be funny [9], there is a critical need for systems that can help a company not just avoid pitfalls but present a better message. This is especially true with large companies, as 2/3 of CMOs fear their companies are unprepared for social media [15].

CMOs are divided on who should "own" social media (PR or Marketing) [15], but there is a growing recognition that employees messaging about their employer in social media is a reality [10] and if done correctly, employees can be the best advocates for company brands [13]. While it is impractical to have a company's department overseeing each and every social media post created by their workforce, it would be clearly beneficial if some level of coaching could be provided.

We have created an application that, analogously to a grammar checker, points out portions of a message that you may wish to reconsider *before* you post. The **Message Authoring Guidance** (MAG) application is a cognitive system that semantically analyzes language use from millions of pieces of content to best position your messaging to resonate with an intended audience. Through the use of MAG before publishing a message, business professionals and content curators can best position their messaging to resonate with a specific audience through validation against millions of pieces of content.

MAG relies on two main components: Message Resonance and Concept Expansion. The **Message Resonance** (MR) component analyzes words based on their historic performance in social media and rates words in a message according to a score that is built on three aspects of message diffusion - volume, prevalence, and sustain. When the message being composed has one or more words with low resonance score, the user can request replacement words to try to identify other wording that resonates better. The **Concept Expansion** (CE) contextual learning service takes example terms of a desired concept (a "seed" list) and identifies other terms like it to expand your terminology set. This expansion can happen automatically or semi-automatically, based on an ontology and a set of semantic relatedness models, and improves over time with user interaction.

2 Technical Description

Consider the following user story (illustrated in Figure 1). A company employee is writing a social media post (e.g. tweet): "Watch: The addition of @SoftLayer into the IBM portfolio of has added a robust capability around cloud computing." They would like to maximize the reach and impact of this message, so it is important to choose the words that will "resonate" with an audience – i.e. be well understood and that will pick their audience's attention. The app will highlight words from the user's message that have been found to perform well/poorly with a given audience. In this example, words in orange ('robust' and 'addition') show poor results. Scores are colored lowest to highest in the following order orange, blue, green, and purple (the app includes a legend). The user can select a word or phrase (e.g. 'robust') and request suggestions of alternative words/phrases that would likely improve that message's success. The app will suggest related or nearly equivalent words (e.g. 'sturdy', 'vigorous', 'strong', etc.), along with the calculated historical performance for each word. Words that are better "fits" for the message and have higher resonance scores will be shown closer to the top in the list of suggestions. The app is based on cognitive systems that "learn" with user interaction and improve their suggestions over time. As the user selects 'strong', the system stores user feedback. This feedback will be used next time the system semantically expands a concept. Our application has the following components:

App: A Message Authoring Guidance app lets users type a message in a text box, shows dialogs with candidate term replacements, allows modifications to the message and notifies the server of user choices.

MR: A Message Resonance service calculates the resonance of words given an audience.

KG: A Knowledge Graph containing ontologies, thesauri and lexicons that are used to obtain equivalent and related terms to suggest as replacements to the user. The KG also stores feedback collected from the user in order to allow continuous evolution of the suggested terms.

CE: A Concept Expansion service finds semantically related terms, to expand the suggestions offered by the KG.

CXR: A Contextual Replacements service ranks synonyms, related terms and conceptually related phrases based on their semantic "fitness" in context – i.e. for a given user, audience and intended message.

The following subsections provide more details on each of the components.

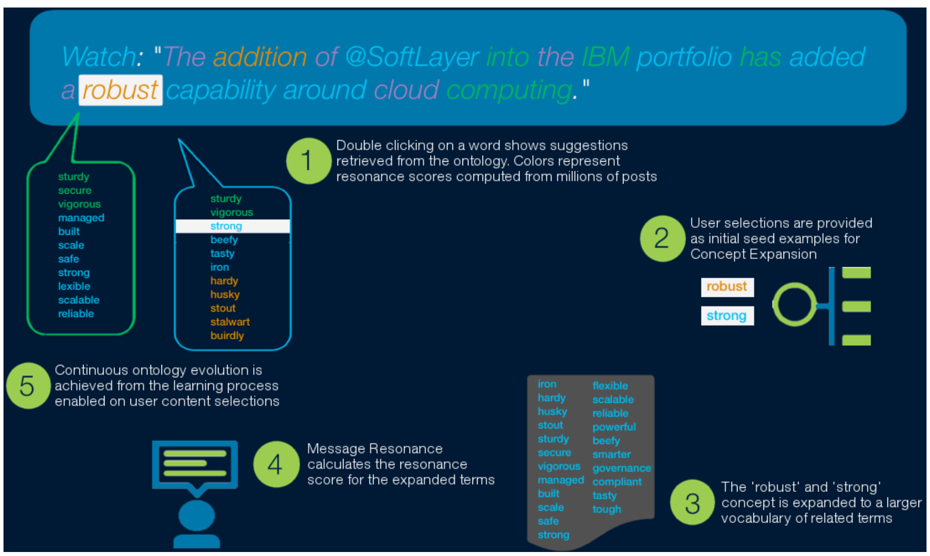

Fig. 1. An illustration of the message authoring guidance. Users type a message, words are analyzed for resonance (different word colors), get suggestions from ontology for replacement (1), user interaction is captured (2), vocabulary is expanded based on semantic context (3), expanded suggestions are evaluated (4) and suggestions are continuously improved (5).

2.1 Message Authoring Guidance User Interaction

Implemented as a simple Twitter posting client (see Figure 1), the application allows a user to type a proposed tweet (given a target audience), and get a real-time estimate of how well it will resonate. The design of the user interface (UI) was guided by 3 main themes established by the iOS Human Interface Guidelines

[7] (a) **deference:** the user interface helps the user to understand and interact with the content, but never competes with it; (b) **clarity:** text is legible at every size, icons are precise and lucid, adornments are subtle and appropriate, and a sharpened focus on functionality motivates the design; (c) **depth:** visual layers and realistic motion impart vitality and heighten users delight and understanding.

More specifically we wanted the user interface to: (i) provide immediate feedback on the effectiveness of a tweet while composing it; (ii) provide scores for individual words; (iii) present the user with a list of synonyms with higher scores to replace a word with a low score. To achieve the first objective the UI makes use of color to show the score for each word as well as for the overall tweet to provide the user with an immediate understanding of the effectiveness of each word in the tweet. Additionally the overall score for the tweet is shown in numerical representation, as is the number of characters remaining from the 140 character limit for a tweet. Tapping a word twice highlights the word and provides the detailed score for the selected word. If synonyms are available they are presented in a context menu layered above or beyond the selected word. Tapping one of the synonyms replaces the selected word with the selected synonym. We chose to use a context menu due to it being a familiar interaction pattern for text manipulation on the iOS platform. Figure 1 demonstrates message composition using the above outlined techniques.

2.2 Message Resonance

Message Resonance is a service that analyzes the popularity of a given word measured against a collection of social media content for a topic or audience of interest that is defined by potentially millions of tweets and blogs. The Message Resonance service teaches the potential of content optimized to cause a large and sustained discussion by returning scores that indicate the content resonance. By using this ranking system, it is possible to recommend more resonant content to use when crafting messages targeted at specific audiences and interatively refining the content before publication.

We have evaluated the impact of wording on the propagation of a message throughout its intended audience on Twitter, and shown that a retweet predictor that includes 'wording' features improves by .07 F1 points over a classifier that includes only social and Twitter-metadata features to achieve 0.82 F1 [12]. Moreover, in an evaluation with 1 million random draws of pairs of Twitter messages, we found that tweets with higher resonance score are twice as likely to be more successful.

2.3 Concept Expansion

Concept Expansion is a learn-by-example service that uses representative concept samples that you provide and finds related concepts by their use in similar contexts. The Concept Expansion service employs an interactive human-machine partnership to guide the concept expansion from the suggested "seed list" using an efficient speed pattern matching algorithm - enabling rapid development of rich entity dictionaries. The algorithm has been shown to perform very well in medical data [2]. The current implementation has generalized, extended, and improved the performance of the earlier algorithm.

The CE service is well suited for expanding a wide variety of entity categories and is ideal in cases where the unstructured source text is not well formed language (e.g., social media data, professional progress notes, email, helpdesk reports, and other less formal communications). It can create concepts out of user-provided seeds, instances of entity types defined in an ontology such as DBpedia [8], a thesaurus such as WordNet [14], or sentiment lexicons such as [18]. We have shown that the expansions provided by this service increase the coverage and accuracy of algorithms that rely on these sources for tasks such as Concept Based Sentiment Analysis [1].

2.4 Knowledge Model

All of the data used in the system is kept in a knowledge base that we call the Message Authoring Guidance Knowledge Graph (MAG-KG). The MAG-KG stores Linked Open Data, ontologies, thesauri, and semantic lexicons that provide semantic context and candidate replacements for the app. Terms and concepts are stored along with their alternative surface forms and topic signatures [11], as well as their distributional contexts [2].

This enriched collection of ontologies and other semantic assets enables us to answer questions of the type: what is the concept most related to "Cloud Computing", or what are the most similar words to 'strong', or what are the words or concepts that best fit "has added a `()` capability around cloud computing." by substituting the placeholder "`()`". The MAG-KG also stores and enables us to leverage feedback obtained from users to suggest one's preferred replacements, or even to tune the algorithms to suggest replacements that are globally more successful.

3 Conclusion

We have presented a message authoring guidance application for social media. The application may best be thought of as a "grammar checker" that is often right, but still consults a human to make the decisions on best phrasing. Suggested replacements are derived from an ontology and a number of semantic relatedness models. Words are then evaluated for their resonance within a community, and presented to the user for confirmation. User feedback is stored and used to continuously evolve the suggestions offered by the system. Links to the live demo and supporting material are available from `http://swc14.pablomendes.com/`. The username for access is 'reviewer2' and the password is 'gardacheluna' (without quotes, pun intended).

Appendix: Evaluation Requirements

Our application meets at least the following requirements, minimal criteria and desirable features of the Semantic Web Challenge:

- **Commercial value** and **Novelty**. Marketing campaigns can easily run to several millions of dollars, especially if coordinated across several media. Poorly worded campaigns run the risk of coming across negatively, especially to niche audiences. An appropriately worded message means that the money invested in marketing will get better use, and help prevent inadvertently damaging remarks. To our knowledge, we are the first to analyze message effectiveness and propose semantically similar words to help users improve the analyzed message.
- **User base**. As corporate leaders are becoming more active on behalf of their companies, message resonance systems are a tool to help "coach" those new to the medium to more quickly find their voice and become effective advocates for their brands and companies. Thus the user base of a message resonance tool is potentially the entire workforce of a modern company.
- **Usefulness**. Marketing campaigns are not just for products; some of the best campaigns are ones which help non-profits reach out to a new set of potential donors, making them aware of an issue or need in their community and connecting with them to make a difference.
- **Semantic Web Technologies**. The entity sets, entity types and relationships between entities allow us to derive semantic models that are used to suggest related terms to users. Moreover, from semantically annotated text corpora, we are able to derive models that analyze how well each concept "fits" in a given message.
- **Heterogeneity**. Our app relies on a wide variety of datasets, including Twitter data, rule-based definitions of audiences, user-captured data, ontologies, thesauri and semantic lexicons as well as other test corpora. All of this information is stored in our Knowledge Graph and used by our algorithms to enhance message authoring guidance.
- **Scalability**. The application uses a corpus with millions of posts that is quickly growing from a stream. In order to efficiently process the information, we implemented our data processing algorithms on Apache Spark [19] for distributed processing.
- **Evaluations**. We have shown that a retweet predictor that includes 'wording' features improves by .07 F1 points over a classifier that includes only social and Twitter-metadata[12]. Moreover, we found from 1 million random draws, that tweets with higher resonance score are twice as likely to be more successful.
- **Contextual**. Our application performs contextual processing of information according to three contexts: Audience Context, Temporal Context and Semantic Context. First, the audience of a message defines the context for resonance. Since the MR component finds phrases that resonate in the context of a particular community, the same phrase will have different resonance

scores if the community context is switched. Second, since the corpus is continuously updated, the same phrase may have different resonance scores in two different time windows. Third, the CE component finds semantically similar concepts based on the contexts in which they are mentioned within a corpus. Moreover, it is not enough that the concepts are similar, a replacement must also make sense in the context of the message being written. Therefore, the Message Authoring Guidance app takes suggestions that resonate well and rate them with regard to how well they fit the context of a particular message.
- **Dynamicity** and **Real World Data**. The app uses a corpus that is continuously updated from a stream of social media posts (e.g. Twitter). As new posts become available, there is a renewed opportunity for expanding the ontology and adapting its vocabulary to what is more commonly used by an audience. Moreover, as users interact with the system and select their preferred replacements, a feedback call is made to the server to keep 'teaching' the system about which replacements are preferred over others.
- **Usability** The design of the user interface (UI) was guided by 3 main themes established by the iOS Human Interface Guidelines: deference, clarity and depth [7]. It was designed to be as simple as possible, functioning analogously to a grammar checker, but learning with every user interaction.
- **Flexibility and Accessibility** The application is available through a Web interface tested for desktops and laptops, responds to touch actions for tablets and also includes an iOS client optimized for the iPhone (not demonstrated here). The MR and CE components are available as Web services, and can be easily reused in other applications. Both services have minimum reliance on Natural Language Processing technology (only tokenization), which makes them easily adaptable to other languages.

References

1. Coden, A., Gruhl, D., Lewis, N., Mendes, P.N., Nagarajan, M., Ramakrishnan, C., Welch, S.: Semantic Lexicon Expansion for Concept-based Aspect-aware Sentiment Analysis. In: ESWC 2014 Concept-Level Sentiment Analysis Challenge (2014)
2. Coden, A., Gruhl, D., Lewis, N., Tanenblatt, M.A., Terdiman, J.: Spot the drug! an unsupervised pattern matching method to extract drug names from very large clinical corpora. In: HISB. pp. 33–39 (2012)
3. DiVeronica, J.: One bad tweet can be costly to a student athlete (2014), `http://www.democratandchronicle.com/story/sports/high-school/2014/09/11/social-media-student-athletes-twitter/15473399/`, [Online; accessed 16-September-2014]
4. Dönges, J.: What Your Choice of Words Says about Your Personality (2009), `http://www.scientificamerican.com/article/you-are-what-you-say/WhatYourChoiceofWordsSaysaboutYourPersonality`, [Online; accessed 16-September-2014]
5. Emmanuel Gorby, P.: Aflac Fires The Guy Who Does The Voice Of Its Duck For Making Tasteless Japan Jokes On Twitter (2011), `http://www.`

businessinsider.com/gilbert-goffried-fired-afl-2011-3, [Online; accessed 16-September-2014]
6. Hutchins, B.: 4 Social Media Lessons via Russel Crowe & Noah (2014), https://www.linkedin.com/today/post/article/20140404164710-974181-4-social-media-lessons-via-russell-crowe-noah, [Online; accessed 16-September-2014]
7. Inc., A.: iOS Human Interface Guidelines (2013), https://developer.apple.com/library/ios/documentation/userexperience/conceptual/mobilehig/, [Online; accessed 3rd-October-2013]
8. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P., Hellmann, S., Morsey, M., van Kleef, P., Auer, S.: DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. Semantic Web Journal (2014)
9. Li, A.: American Apparel Angers Twittersphere With 'Hurricane Sandy Sale' (2012), http://mashable.com/2012/10/30/american-apparel-sandy/, [Online; accessed 16-September-2014]
10. Meister, J.: Social Media Training Is Now Mandatory (2012), http://www.forbes.com/sites/jeannemeister/2012/10/31/social-media-training-is-now-mandatory/, [Online; accessed 16-September-2014]
11. Mendes, P., Jakob, M., Bizer, C.: DBpedia for NLP: A Multilingual Cross-domain Knowledge Base. In: Chair), N.C.C., Choukri, K., Declerck, T., Doan, M.U., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (eds.) LREC'12. European Language Resources Association (ELRA), Istanbul, Turkey (may 2012)
12. Mendes, P.N., Gruhl, D., Drews, C., Kau, C., Lewis, N., Nagarajan, M., Alba, A., Welch, S.: Sonora: A Prescriptive Model for Message Authoring on Twitter. In: WISE 2014, Part II. pp. 17–31 (2014), http://dx.doi.org/10.1007/978-3-319-11746-1_2
13. Mike Ambassador Bruny: 4 ways to make your employees your brand's best champions (2014), http://www.fastcompany.com/3035546/hit-the-ground-running/4-ways-to-make-your-employees-your-brands-best-champions, [Online; accessed 16-September-2014]
14. Miller, G.A.: Wordnet: A lexical database for english. Commun. ACM 38(11), 39–41 (Nov 1995), http://doi.acm.org/10.1145/219717.219748
15. Sass, E.: CMOs: Social Media Key, But We Are Clueless (2014), http://www.mediapost.com/publications/article/221666/cmos-social-media-key-but-we-are-clueless.html, [Online; accessed 16-September-2014]
16. at UNC Chapel Hill, T.W.C.: Word Choice (2014), http://writingcenter.unc.edu/handouts/word-choice/, [Online; accessed 16-September-2014]
17. Wikipedia: Brand Language (2014), http://en.wikipedia.org/wiki/Brand_language, [Online; accessed 16-September-2014]
18. Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., Patwardhan, S.: Opinionfinder: A system for subjectivity analysis. In: Proceedings of HLT/EMNLP 2005 Interactive Demonstrations. pp. 34–35. Association for Computational Linguistics, Vancouver, British Columbia, Canada (October 2005), http://www.aclweb.org/anthology/H/H05/H05-2018
19. Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S., Stoica, I.: Spark: Cluster computing with working sets. In: HotCloud'10. pp. 10–10. USENIX Association, Berkeley, CA, USA (2010), http://dl.acm.org/citation.cfm?id=1863103.1863113