Analysis of Bloggers' Influence Style within Blog

Luke Kien-Weng Tan*

Wee Kim Wee School of Communication & Information
Nanyang Technological University, Singapore
E-mail: w080078@ntu.edu.sg

Jin-Cheon Na

Wee Kim Wee School of Communication & Information
Nanyang Technological University, Singapore
E-mail: tjcna@ntu.edu.sg

ABSTRACT

Blogs are readily available sources of opinions and sentiments which allows bloggers to exert a certain level of influence over the blog readers. Previous studies had attempted to analyze blog features to detect influence within the blogosphere, but had not studied in details influence at the blogger-level. Other studies studied bloggers' personalities with regards to their propensity to blog, but did not relate the personalities of bloggers to influence. Bloggers may differ in their way or manner of exerting influence. For example, bloggers could be active participants or just passive shares, or whether they express ideas in a rational or subjective manner, or they are received positively or negatively by the readers. In this paper, we further analyze the engagement style (frequency, scope, originality, and consistency of the blog postings), persuasion style (appeals to reasons or emotions), and persona (degree of compliance) of individual bloggers. Methods used include similarity analysis to detect the sharing-creating aspect of engagement style, subjectivity analysis to measure persuasion style, and sentiment analysis to identify persona style. While previous studies analyzed influence at blog site level, our model is shown to provide a fine-grained influence analysis that could further differentiate the bloggers' influence style in a blog site.

Keywords: Influence style, influence detection, blogger analysis

1. INTRODUCTION

A blog or weblog is a specialized web site that allows an individual or a group of individuals to express their thoughts, voice their opinions, and share their experiences and ideas. The entries are called blog posts, while the individuals who authored the blog posts are referred to as bloggers. A blog site could contain a sin-

gle blogger, and is referred as a personal blog, while a blog site containing a group of bloggers is known as a multi-author blog as shown in Figure 1. Multi-author blogs have certain characteristics amongst their bloggers such as common shared goals and guidelines, and the editor would guide and organize the activities of the blog. The reasons for having multiple authors on a blog are the wider variety of opinions and ideas that could cater to more readers. Nevertheless, the greater variety of opinion expressions could lead to differences in the bloggers' influence styles. Influence style refers to the manner in which the blogger exert influence through the blog postings. For example, a blogger could be subjective in persuasion influence style as seen in the expression *"The much anticipated iPad mini by Apple is finally here."* as compared to the non-*subjective expression "Apple has launched the iPad mini."* indicating an objective style with both expressions communicating the same message.

Our Staff

Arnold Kim
Edtorial Director
Email • Twitter

Eric Slivka
Edtor in Chief
Email • Twitter

Jordan Golson
Editor
Email • Twitte

Marianne Schultz
Editor
Email • Twitte

Looking for job? We are hiring.

Fig. 1 Bloggers of Multi-author Blog Site MacRumors (http://www. macrumors.com) Retrieved on 29 Oct 2012

Previous studies had attempted to analyze the personality of bloggers (Guadagno, Okdie, & Eno, 2008; Yarkoni, 2010). However, the studies measured the personality of bloggers manually in relation to their propensity to blog, and had not considered the influence styles of bloggers. Other studies had proposed using blog features (Adar & Adamic, 2005; Agarwal & Liu, 2008), similarity comparison (Matsumura, Yamamoto, & Tomozawa, 2008; Song, Chi, Hino, & Tseng, 2007), and community detection (Agarwal, Liu, Tang, & Yu, 2008; Ghosh & Lerman, 2008) to detect blog influence. However, the studies did not consider the

subject matter of the blog post, and thus, would not measure influence accurately. This is because influence is subjective in nature and is often dependent on the contextual details. Recent studies further considered the context of blog posts by analyzing the sentiments between linked blog posts (Leskovec, Huttenlocher, & Kleinberg, 2010; Cai, Bao, Yang, Tang, Ma, Zhang, & Su, 2011; Li, Bhowmick, & Sun, 2011). These studies focused on a single notion that influence exists between blogs with hypertext links, but ignored the details in the other aspects of influence. Influence is a complex concept that consists of various components such as the different influence styles, and could not be described using a single detached feature. Influential bloggers are not restricted to a monolithic description of influence and may differ in influence style. By identifying each blogger's influence style we provide more in-depth profiles of the bloggers. For example, bloggers' influence could be better described through identifying the bloggers as active participants or just passive sharers, whether they focus on specific topics or products, as seen in Figure 2, or they are broad in their scope, whether they express ideas in a rational or subjective manner, or whether their posts are received positively or negatively by the readers. Identification of blogs' influence style could be applied in several applications. For example, we could determine whether influential
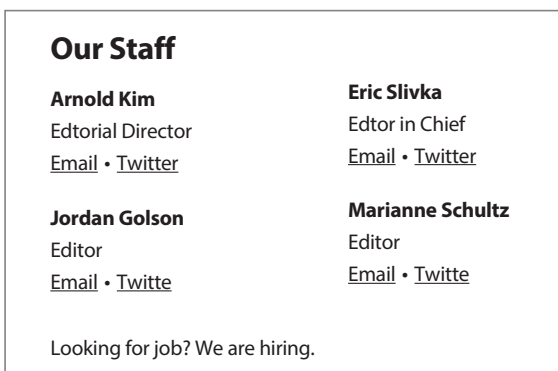
MacRumors

MacRumors attracts a broad audience of both consumers and professionals interested in the latest technologies and products. We also boast an active community focused on purchasing decisions and technical aspects of the iPhone, iPod, iPad, and Mac platforms.

Fig. 2 Description of Blog Site MacRumors (http://www.macrumors.com) Retrieved on 29 Oct 2012

product-marketing blogs are persuasive or objective, or whether influential financial blogs tend to post actively or are usually dormant, but would post few highly relevant postings. The identified influence style would give a better description and understanding of the influence exerted by the blogs, and could be used as

features to improve influence detection performance. On the other hand, previous studies which assumed a monolithic definition of influence existence between linked blog posts would not be able to differentiate the varying influence styles of the blogs and individual bloggers.

In this paper, we aim to profile the influence styles of blogs, specifically at blogger level, by automatic detection and analysis of the influence styles that exist between linked blog posts. The influence style model used was adopted from (Tan, Na, Theng, and Chang, 2012b), where we define the influence of blogs and bloggers from an objective and social perspective with an expanded description of influence consisting of three possible styles, mainly engagement style, persuasion style, and persona of the bloggers. Engagement style indicates the frequency, scope, originality, and consistency of the blog postings. Persuasion style refers to appeals to reasons or emotions displayed in the blog posts, and persona is the degree of compliance exerted by the bloggers. Additionally, as influence is time dependent (Adar & Adamic, 2005; Agarwal & Liu, 2008) and topic constrained (Agarwal, Liu, Tang, & Yu, 2008), our model evaluated blogger influence within the analyzed time period, and specific to the blog post topics. In contrast to (Tan, Na, Theng, and Chang, 2012b), this paper conduct an in-depth analysis of the individual bloggers influence style to differentiate the manner in which each blogger exert their influence.

The next section describes related work followed by the research design where details of the methods used are given. Next, we present the experiment procedures and results, followed by the discussion and conclusion.

2. LITERATURE REVIEW

Previous studies had attempted to classify blogs through content analysis. Herring, Scheidt, Bonus, and Wright (2004) analyzed blog content to identify and quantify structural and functional properties of blogs according to their genre. The study analyzed the demographic characteristics of the blog authors from the blogs, and they conducted structural analysis of blogs on features such as number of links, images,

presence of search features, and advertisements, as well as frequency of blog updates to categorize blogs based on their respective genres. Other studies (Guadagno, Okdie, & Eno, 2008; Yarkoni, 2010) measured the personality of bloggers in relation to predicting blogging. Costa and McCrae (1992) evaluated personality based on clinical studies through five key traits: neuroticism, extra-version, agreeableness, openness to experience, and conscientiousness. Neuroticism represents an individual's tendency to experience distress. Extraversion is the dimension underlying a broad group of traits, including sociability, activity, and the tendency to experience positive emotions. Agreeableness is a dimension of interpersonal behavior where high agreeableness individuals are trusting, sympathetic, and cooperative. Openness to Experience refers to traits that describes individuals as imaginative and sensitive to art and beauty and have a rich emotional life. Conscientiousness is a dimension that contrasts scrupulous, well-organized, and diligent people with lax and disorganized individuals. Guadagno, Okdie, and Eno (2008) measured the personality of bloggers based on the five key traits by (Costa & McCrae, 1992) to predict blogging. Their studies indicated that people who are high in openness to new experience and high in neuroticism are likely to be bloggers. The results indicated that personality factors impacted the likelihood of being a blogger and have implications for understanding who blogs. Similarly, Yarkoni (2010) analyzed blogger personality on word use in their blog postings to study the association between personality and language of bloggers. However, these studies had manually measured bloggers' characteristic and personality with regards to their propensity to blog, and did not relate the characteristic and personality of bloggers to their influence styles. Moreover, the five key personality traits by (Cost & McCrae, 1992) are not directly applicable for measuring influence style. In our study, we study bloggers' traits in relation to their influence style exerted on other bloggers through automatic analysis of the blog post content.

Influence is often related to conformity, which is the act of matching attitudes, beliefs, and behaviors to group norms (Hogg & Vaughan, 2005). Kelman (1958) identified conformity as form of compliance, which is the expression of agreement towards the people or

groups who are influential to the individual, while possibly keeping one's own original beliefs. In our study, we measure compliance through analyzing similar sentiments on target topics in the textual content between the linked blog posts. Cialdini (2001) defined several human characteristics in relation to influence. *Authority* refers to subject experts in our context. *Reciprocity* is the behavior whereby people tend to return a favor. People are usually *committed* and *consistent,* and do not like to be self-contradictory. Once they commit to an idea or behavior, they are averse to changing their minds without good reason. In addition, *scarcity* of resources will generate demand. Characteristics used in our study to evaluate the engagement style include authority, reciprocity, commitment and consistency, and scarcity. Cialdini and Goldstein (2004) observed that how people react to beliefs held by others is often contingent on one's perceptions of the level of consensus for those beliefs. That is, bloggers would tend to share content of blog posts which they agree with. The objective consensus approach (Mackie, 1987) states that individuals are more likely to systematically process a majority-endorsed message because people assume that the majority view reflects reality and people believe they share similar attitudes to the members of the majority. This means that bloggers would also tend to share blog post content of influential blogs. We consider the behavior of sharing blog posts' content as a *Sharing-Creating* type in the engagement style evaluation. Content similarity was used in (Gruhl, Guha, Liben-Nowell, & Tomkins, 2004; Song, Chi, Hino, & Tseng, 2007; Agarwal & Liu, 2009) to detect influence in the blogosphere. The approaches used in these studies involved measuring document similarity between linked blog posts using information retrieval techniques, such as cosine similarity. These approaches mainly checked whether the linked blog posts were discussing the same topics as part of the influence detection method. However, content similarity is not a good measurement of direct influence, as the in-link blog author could copy the content to be used as reference, and express contrasting opinions on the common content. In our study, similarity comparison is used to determine the engagement style of bloggers based on the level of content creation or sharing, and

not to detect influence directly.

Brehm (1966) observed that emotion and disposition may affect likelihood of conformity or anti-conformity. The arousal and affective states have an effect of discrete emotions on targets' cognitions as well as on the eventual outcome of the influence attempt (Cialdini & Goldstein, 2004). Persuasion is a process of influence through appeals to reason or appeals to emotion. *Subjectivity,* that is, opinionated or emotional phrases expressed in the blog content could provide clear evidence of the blogger's persuasion style. Our study considers persuasion as a component of influence and evaluates it through analyzing the subjectivity expressed in the blog posts content.

Previous studies that used blog features to detect influence were mainly graph-based. Herring, Kouper, Paolillo, Scheidt, Tyworth, Welsch, Wright, and Yu (2005) investigated the extent to which blogs are interconnected using quantitative social network analysis, visualization of link patterns, and qualitative analysis of references and comments of the blogs. Their results showed that the blogosphere is partially interconnected and sporadically conversational. However, the interconnectivity of the blogosphere does not indicate the influence of the bloggers as there could be disagreement between linked blog posts. Adar and Adamic (2005) related the number of blog links to influence. The study applied link inference techniques to find implicit graph links that could further identify influence. In the study by Agarwal and Liu (2008), an influential blogger was defined based on the number of in-links, comments for the posts, and the number of out-links. A high authority value based on a large number of in-links to the blog postings indicates high number of followers. However, having many in-links may not necessary indicate influence. For example, a linking blog post opposing the ideas and opinion of a high authority blog post could express disapproval and not be influenced, while connected to the linked blog post. Hence, number of in-links alone could not accurately describe the influence exerted by a blogger. Community identification was also used to detect influence in the blogosphere. Ghosh and Lerman (2008) generalized the notion of network connectivity to the number of paths that exist between two nodes in relation to blogosphere influence. Agarwal, Liu, Tang,

and Yu (2008) proposed to detect blog communities through aggregating similar individual blogs to identify influential bloggers. However, community identity alone would not detect influence accurately as blogs within a community could differ in opinions on certain topics, while blogs outside of the community could also share similar opinions.

The above influence studies utilized positive interactions (e.g., agreement and trust) between individuals, and ignored the negative relationships in the blogosphere. Milgram (1963) observed instances of dissent can greatly wane the strength of an influence, indicating that detecting both agreement (positive sentiments) and disagreement (negative sentiments) is equally important in influence analysis. Recent studies had included the consideration for both positive and negative links in the analysis of influence. Leskovec et al. (2010) adapted a framework of trust and distrust in an attempt to infer the attitude of one user toward another using the observed positive and negative relations. Tan, Na, and Theng (2011) observed that analyzing the positive, neutral, and negative sentiments of linked blog posts on common topics could improve influence propagation detection. Cai et al. (2011) defined social influence through positive and negative social relationships. The study was inspired by the idea of persona in sociology studies where influence is the approvals gained from other reliable people. Granovetter (1973) and Krackhardt (1992) observed that the effects of the social influence by users of varying personae are different. The three general types of persona identified were positive, negative, and contro-

versial. Li et al. (2011) similarly considered the positive and negative edges of the nodes, and computed the influence index and conformity index in their attempt to detect influence in the social network. However, these studies had focused on a single notion of the influence of bloggers in detecting influence, that is, whether influence exists between linked blog posts, and had not studied in details the influence styles of bloggers. Each individual blogger could exhibit his own influence style, and the overall blog site influence style may not be representative of all bloggers in the blog site. Our study further describes influence through the engagement style, persuasion style and persona of the individual bloggers.

3. RESEARCH DESIGN

In the influence style model adapted from Tan, Na, Theng, and Chang (2012b), influence is further described in terms of *Engagement style, Persuasion style*, and *Persona style* in profiling the blog sites and individual bloggers. An overview of the influence style model is shown in Table 1.

The influence style model adapted the four general influence types from the Klout Influence Matrix, which include *Listening-Participating, Broad-Focused, Sharing-Creating,* and *Consistent-Casual* to describe the engagement style of the blog sites and bloggers. Engagement style refers to the frequency, scope, originality, and consistency in which the influencer presents the blog post content. A blogger who is a subject

Table 1. Influence Style Model

Influence Style	Type	Description
Engagement	Listening-Participating	Describes the blogs' or bloggers' participation level in posting specific target topics.
	Broad-Focused	Describes the scope of the topics discussed in the blog posts.
	Sharing-Creating	Describes the level of blog posts' content originality.
	Consistent-Casual	Describes the consistency and duration of blogs' or bloggers' ability to garner in-links.
Persuasion	Subjective-Objective	Describes the types of Persuasion style
Persona	Positive-Negative-Controversial	Describes the types of Persona

expert would tend to garner more in-links as well as post frequently within the specific target topics scope indicating a high level of authority (Cialdini, 2001). The bloggers' authority could be derived from the depth and extent of dominance in the target topics, measured through the *Listening-Participating, Broad-Focused,* and *Consistent-Casual* type respectively in our study. The *Listening-Participating* type describes the bloggers' participation level in posting specific target topics. The listening type usually do not post much, but rather follow the blog posts with a low level of posting activities, while a participating type would actively post articles on target topics and share information readily. The *Broad-Focused* type describes the scope of the topics discussed in the blog posts. A broad type blogger would post a wide range of topics, and a focused type would concentrate on specific topics which they are domain experts. *Sharing-Creating* type indicates the originality of the blog posts' content. A sharing blogger tends to restate the content of the linked blog posts. On the other hand, a creating blogger posts original content. For the *Consistent-Casual* style, the *consistent* type refers to bloggers that could continually attract in-links over a duration of time, whereas a casual blogger is irregular or intermittent in their number of in-links. The extent to which one's commitments are made actively is a powerful determinant of the likelihood of compliance (Cialdini & Trost, 1998). Our model measures commitment through the consistency of the target blogger's ability to garner in-links.

Persuasion is a process of influence through appeals to reason or appeals to emotion (Brehm, 1966). Subjectivity, that is, opinionated or emotional phrases expressed in the blog content could provide clear evidence of the blogger's persuasion style (Cialdini & Goldstein, 2004). Our study evaluates persuasion through analyzing the subjectivity expressed in the blog posts content.

Kelman (1958) defined *Compliance,* a form of conformity as the expression of agreement towards the people or groups who are influential to the individual, while possibly keeping one's own original beliefs. The degree of compliance towards the blog site is measured based on the persona of the bloggers in our model. Persona style shows the degree of compliance exerted

by the influencer, and is assessed through analyzing the sentiments on common topics between the in-link blog post and target blog post. *Positive persona* describes bloggers with high positive influence, where their links from others often indicate approval and agreement. Milgram (1963) observed that dissent within a social network can greatly wane the strength of an influence. Our model measures dissent through evaluating the negative persona of bloggers. *Negative persona* represents bloggers with high negative influence, and their links from others usually express disagreement or distrust. The *Controversial persona* represents bloggers that are both likely to be challenged or supported by many (Cai et al., 2011), which is shown in the high number of agreement and disagreement in-link blog posts. Blog influence is specific to the topics or aspects discussed in the blog posts (Agarwal, Liu, Tang, & Yu, 2008; Somasundaran & Wiebe, 2010). Our model considers the bloggers' influence in relation to specific target topics by identifying the topics, aspects, and features discussed in the blog posts. Bloggers' influence could also change with time (Adar & Adamic, 2005; Agarwal & Liu, 2008). In our model, the influence styles of the bloggers are defined within the analyzed time period.

3.1. Influence Style Analysis Framework

Figure 3 shows the influence style analysis framework. In-links are incoming links to the target blog post, while out-links are links to other blog posts from the target blog post. We extracted the in-link blogs URL from the Technorati.com blog directory, and downloaded the in-link blog posts from the respective in-link blog sites. The *"search for more reactions"* feature found in Technorati.com was used to extract the in-link to the target blog site. This is followed by downloading of the target blog posts using the target blog URLs extracted from the in-link blog posts. Subsequently, the out-link blogs URLs are extracted from the target blog posts, and used to download the blog posts from the out-link blog sites. The *Listening-Participating, Broad-Focus, Consistent-Casual* type of the Engagement style were analyzed from the target blog posts, while the *Sharing-Creating* type was analyzed between the target blog posts and the out-link blog posts. The *Persuasion* style is determined based
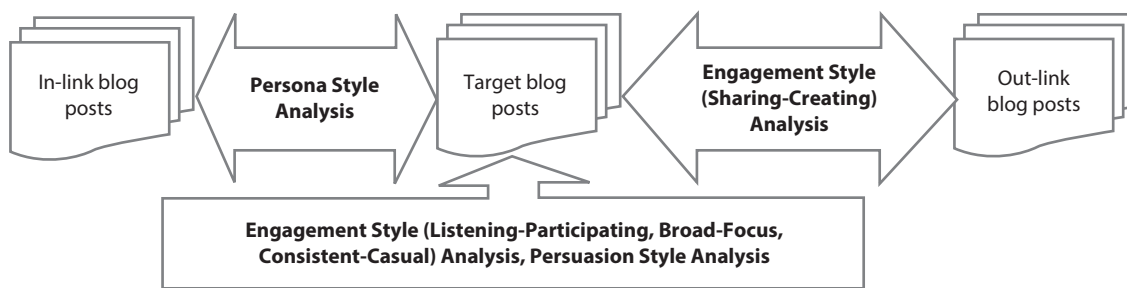
Fig. 3 Influence Style Analysis Framework

on the subjectivity study on the target blog posts, and the *Persona* style analysis is performed between the in-link blog posts and the target blog posts.

We measure the *Listening-Participating* type by the number of blog post titles that contain the target topic and related feature terms. In this study, we limit the scope of the target topics to three main topics: iPad, iPhone, and Mac. The count value is then normalized by the total number of posts for each respective blog site to give the Listening-Participating score. A high percentage of target topic or related feature postings would indicate that the blogger is of Participating type, and conversely, a low percentage shows a Listening type. The Broad-Focused type is measured by the number of unique target topics and their related features found in the blog post titles. A wide range of topics and their related features discussed would indicate a Broad type, while a limited number of topics and their related features discussed refer to a Focused type. The *Sharing-Creating* type is evaluated through analyzing the similarity between the target blog posts content and its out-link blog posts content to give the Sharing-Creating score based on the Jaccard coeffi-

cient defined as the size of the intersection divided by the size of the union of the two linked posts content (A and B) shown as : $J(A_{linking\ blog\ post}, B_{linked\ blog\ post} = |A \cap B|/|A \cup B|$. A high similarity value would mean that the blogger shares most of the out-link blog post content. On the other hand, the target blog post contains more original content if the similarity value is low. We further separate the analysis for self-links found in the blog posts to differentiate the sharing-creating style between self-link and non-self-link posts. Self-links refer to blog posts having links within the same blog site. The overall Sharing-Creating type is evaluated by taking the average of the self-linked and non-self-linked Sharing-Creating scores. The *Consistent-Casual* type is determined by the number of target blog site postings in a near-term, mid-term, and long-term duration linked by the in-link posts within the analyzed time period.

We detect *Persuasion* style by analyzing the subjectivity expressed by the bloggers in their blog posts through counting the number of subjective terms in the target blog posts. This is done by matching the subjectivity terms from Wilson, Wiebe, and Hoffmann

*iOS redesigned App Store has leaked a **major forthcoming** update to Twitter's official iPhone application in its update history section. Twitter in recent months has **come under fire** due to potential new changes to its mobile application API and the **lack of** power-user features in its official mobile device apps. It appears that this controversy may change in the coming days as leaked release notes show some **impressive** new changes……*

(http://9to5mac.com/2012/07/08/major-twitter-for-iphone-update-incoming-more-interactive-enhanced-notifications/). Retrieved on 2 Dec 2012.

Fig. 4 Example of Subjective Phrases

The Electronic Product Environmental Assessment Tool (EPEAT) is a program that ranks computers and displays based on environmental attributes in accordance with IEEE 1680.1-2009. A clear and consistent set of environmental performance criteria for the design of personal computer products including notebook computers desktop computers and computer displays is provided thereby providing an opportunity to secure market recognition for efforts to reduce the environmental impact of electronic products. This standard is intended to provide a tool for government institutional and corporate purchasers. Product manufacturers may also use this tool to earn recognition in the consumer market re- cognizing that certain criteria may not be applicable to all types of purchasers……

(http://9to5mac.com/2012/07/07/apples-new-designs-force-it-to-abandon-epeat-green-standard-potentially-cutting-off-big-customers/). Retrieved on 2 Dec 2012.

Fig. 5 Example of Objective Phrases

(2005) with the target blog posts terms. The matched number of subjectivity terms is then normalized with the length of the blog post to give the Persuasion Style score. Wiebe, Wilson, Bruce, Bell, and Martin (2004) observed that the percentage of opinionated words over total number of word instances for subjective content ranged from 7% to 12%. The 7% benchmark is used as the threshold for subjectivity in our model. Examples of subjective and objectives phrases found in the blog posts are given in Figure 4 and Figure 5. It can be seen from Figure 4, subjective terms such as *'major forthcoming', 'come under fire', 'lack of'* and *'impressive'* expressed subjectivity in the post, resulting in the blog post being subjective in persuasion style. On the other hand, the absence of subjective terms in Figure 5's post

made the blog post's style objective and reporting in nature.

The bloggers' Persona is identified by detecting the sentiments expressed between the in-link and target blog posts. In our model, the sentiment analysis rules based on typed dependency rules were adapted from Tan, Na, Theng, and Chang (2012a). The typed dependency polarity pattern rules used include the adjectival modifier (AMOD), adverbial modifier (ADVMOD), direct object modifier (DOBJ), and nominal subject (NSUBJ) rules. The typed dependency polarity pattern rule is denoted as "typed-dependency(governor-term:[polarity], dependent-term:[**polarity**])→[**polarity**]". For example, the adjectival modifier pattern rule "AMOD([**neutral**], [+])→[+]" will be used for the
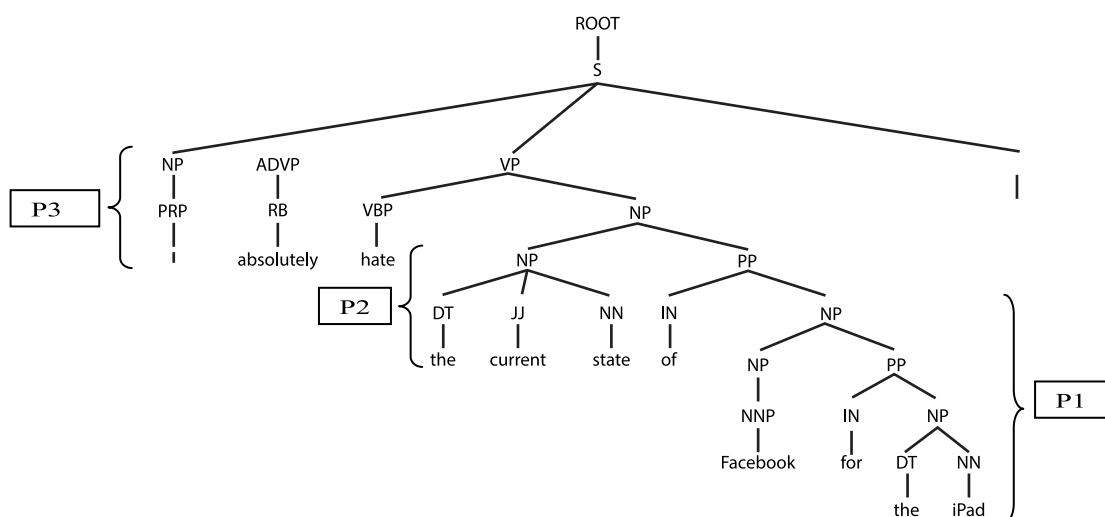

Fig. 6 Phrase structure tree for the clause *"I absolutely hate the current state of Facebook for the iPad."*

Typed dependencies in P1	Typed dependencies in P2	Typed dependencies in P3
prep_of(state, Facebook)	det(state, the)	**nsubj(hate, I)**
det(iPad, the)	**amod(state, current)**	**advmod(hate, absolutely)**
prep_for(Facebook, iPad)		**dobj(hate, state)**

Fig. 7 Eyped dependencies of the clause *"I absolutely hate the current state of Facebook for the iPad."*

phrase "great phone": AMOD(phone:[**neutral**], great :[+]) to give a positive polarity value. The typed dependency polarity pattern rules are evaluated using a bottom up approach within a clause. For example, starting from the bottom phrase in the phrase structure tree, as shown in Figure 6 for the clause *"I absolutely hate the current state of Facebook for the iPad."*, the initial noun phrases (NP) (indicated by P1) would give a neutral output as there are no evaluated typed dependency polarity pattern rules. The typed dependencies generated for the clause are shown in Figure 7.

For the subsequent phrase (P2), the adjectival modifier pattern rule "AMOD([**neutral**], [**neutral**])→[**neutral**]" for the phrase "current state": AMOD (state :[**neutral**], current:[**neutral**]) would give a neutral output. In the recursive evaluation, the preceding lower level sentiment polarity output is used as the governor term polarity of the subsequent typed dependency polarity pattern rule. Hence, the governor term "state" in the subsequent direct object pattern rule would inherit the neutral polarity of the preceding rule "AMOD(state:[neutral], current: [neutral])→[neutral]" evaluation output. With this, the direct object pattern rule "DOBJ([-], [**neutral**])→[-]" for the phrase "hate state": DOBJ(hate:[-], state:[**neutral**]) gives a negative polarity sentiment output within the verb phrase indicated by P3. The subsequent adverb modifier pattern rule "ADVMOD([-], [**intensify**])→[**intensified** -]" for

the phrase "absolutely hate": ADVMOD(hate:[-], absolutely:[**intensify**]) will also yield a negative polarity. The intensifier term "absolutely" further increase the negative value of the polarity, which is recursively input into the governor term in the nominal subject pattern rule ""NSUBJ([-], [**neutral**])→[-]" for the phrase "I hate": NSUBJ(hate:[-], I:[**neutral**]) to give the eventual negative sentiment polarity output for the clause in the final recursive step. The overall blog post's sentiment on the target topic is derived by aggregating the individual clause sentiments for the target topics.

The target topic lexicon was created by listing the aspects and features of the target topics from Apple product related websites (e.g., www.apple.com), with examples of the lexicon records shown in Table 2. In this study, we limit the scope of the target topics to three main topics: iPad, iPhone, and Mac. We identify the target topic by matching the target terms and feature terms in the lexicon list with terms found in the noun phrases of the blog post sentences parsed using the Stanford parser. Feature terms are words describing the product's functionalities and properties, while aspect terms describe the general characteristics of the product. For example, in the clause *"Despite the many advantages of the iPhone's multi touch-screen, a lack of tactile feedback remains its biggest disadvantage"*, the identified feature *"touch-screen"* would map to the display aspect of the iPhone target topic. The predicted

Table 2. Examples of Target Topics and Their Overall Aspects and Features

Target Topics	Overall Aspects	Features
iPad, iPhone, Mac	Connectivity	Thunderbolt, Network, Wifi, etc
	Display	Touch-Screen, Pixel, Retina, etc
	Application	Safari, iTunes, Siri, iCloud, MobileMe, etc
	Operating system	iOS, Android, Windows, etc

triplet output (*target topic, aspect, sentiment polarity*) for the clause would be (*"iPhone", "Display", "Negative"*). *The overall blog post's sentiment on the target topic i*s derived by aggregating the individual clause's sentiments for the target topics. Compliance could be inferred from the similar sentiments expressed on common target topics and related features between the linked blog posts, even if the sentiments are negative. This can be seen from the linked blog posts that exhibited negative sentiments on the common feature term *"MobileMe"* as shown in Figure 8. On the other hand, high dissimilarities in sentiments show disagreement towards the target blog posts, which in turn indicate the linked bloggers' negative persona. This is seen in Figure 9, where the disagreement is shown in the negative sentiments and positive sentiments expressed by the in-link blog post and the target blog post respective-

ly on the feature term *"iOS 6"*.

In order to differentiate the influence styles of each individual bloggers quantitatively, we carefully chose the threshold values of each influence style type based on the corresponding scales shown in Table 3. The scales were derived from the distribution of scores for each influence type, having the majority score distribution classified into the normal scale. The Listening-Participating score, Sharing-Creating score, and Persuasion Style score are used to determine the scoring scale for the respective influence style types. The Persona Style score is based on the percentage of similar sentiments postings, while the Broad-Focused type and Consistent-Casual type are evaluated through relative comparison amongst the bloggers as the features are specific to the blog sites' context.
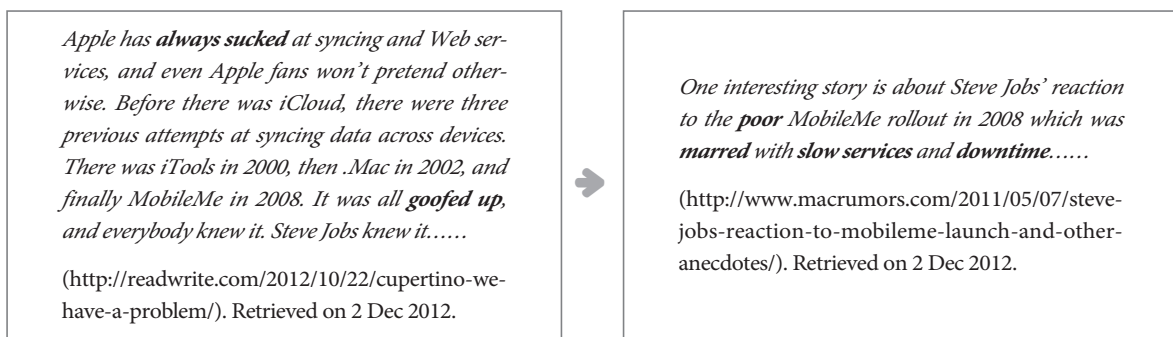
> *Apple has **always sucked** at syncing and Web services, and even Apple fans won't pretend otherwise. Before there was iCloud, there were three previous attempts at syncing data across devices. There was iTools in 2000, then .Mac in 2002, and finally MobileMe in 2008. It was all **goofed up**, and everybody knew it. Steve Jobs knew it……*
>
> (http://readwrite.com/2012/10/22/cupertino-we-have-a-problem/). Retrieved on 2 Dec 2012.

> *One interesting story is about Steve Jobs' reaction to the **poor** MobileMe rollout in 2008 which was **marred** with **slow services** and **downtime**……*
>
> (http://www.macrumors.com/2011/05/07/steve-jobs-reaction-to-mobileme-launch-and-other-anecdotes/). Retrieved on 2 Dec 2012.

Fig. 8 Example of compliance between linked blog posts with negative sentiments on the feature (MobileMe) of the application aspect.

> *When iOS 6 was announced, Apple said that iOS 6 would be compatible with the iPhone 3GS, but the major features they announced seemed to be **left off** of the list for the older device. With the release of iOS 6 beta 3, MacRumors is reporting that the iPhone 3GS has received a couple more iOS features.*
>
> (http://www.iphoneincanada.ca/iphone-3gs/iphone-3gs-gets-more-ios-6-features/). Retrieved on 2 Dec 2012.

> *………But with yesterday's release of iOS 6 Beta 3 for developers, Apple has **improved feature support** for the iPhone 3GS, adding shared Photo Streams and VIP mail filtering. The change was noticed by a number of iPhone 3GS users, and Apple has confirmed the removal of those restrictions on its iOS 6 preview page.*
>
> (http://www.macrumors.com/2012/07/17/apple-adds-shared-photo-streams-and-vip-mail-to-iphone-3gs-with-ios-6-beta-3/). Retrieved on 2 Dec 2012.
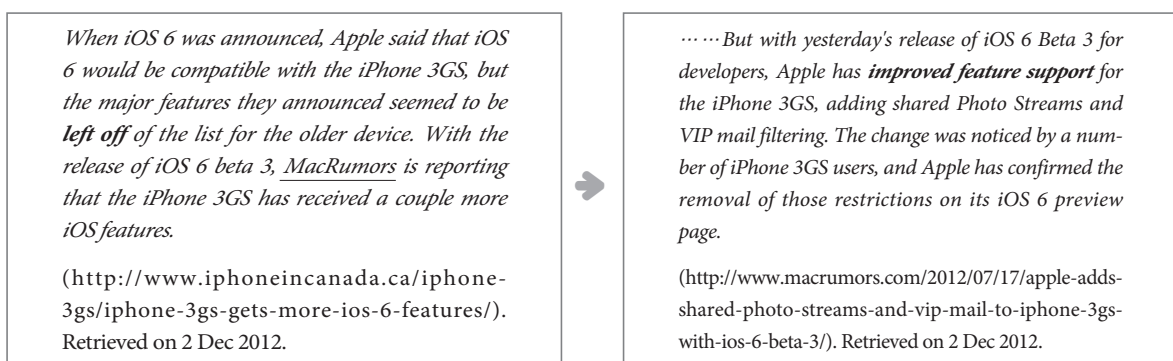
Fig. 9 Example of non-compliance in in-link blog post with negative sentiments on the feature (iOS 6) of the application aspect linked to target blog post with positive sentiments.

Table 3. Threshold Alues of Corresponding Influence Style Scores

Influence style	Low	Normal	High
Engagement Style			
• **Listening-Participating**			
- **Listening**	15% < Score <= 20%	20% < Score <= 50%	Score > 50%
- **Participating**	10% < Score <= 15%	5% < Score <= 10%	Score <= 5%
• **Sharing-Creating**			
- **Sharing**	15% < Score <= 20%	20% < Score <= 50%	Score > 50%
- **Creating**	10% < Score <= 15%	5% < Score <= 10%	Score <= 5%
• **Broad-Focused**			
- **Broad**			
- **Focused**		Comparative Evaluation	
• **Consistent-Casual**			
- **Consistent**			
- **Casual**			
Persuasion Style			
• **Subjective**	7% < Score <= 10%	10% < Score <= 12%	Score > 12%
• **Objective**	5% < Score <= 7%	3% < Score <= 5%	Score <= 3%
Persona Style			
• **Positive**	60% < Score <= 70%	70% < Score <= 80%	Score > 80%
• **Controversial**		40% < Score <= 60%	
• **Negative**	30% < Score <= 40%	20% < Score <= 30%	Score <= 20%

4. EXPERIMENTS AND EVALUATION RESULTS

We identified two influential blog sites (MacRumors, Technorati Authority = 742 and 9to5Mac, Technorati Authority = 779, retrieved on 1 July 2012) from Technorati.com based on their high Technorati Authority values. A total of 753 and 892 in-link posts dated from 1 July 2012 to 23 July 2012 were extracted for MacRumors and 9to5Mac respectively. From the in-link posts, we identified the target blog site posts with MacRumors having 425 posts and 9to5Mac having 391 posts. These numbers refer only to the linked target blog posts within the analyzed period, and not all the blog sites' posts. Within the target blog posts we extracted 1948 out-link posts for MacRumors and 2297 out-link posts for 9to5Mac. Further to that, we identified the authors of each blog post with the number of blog posts for the MacRumors bloggers and 9to5Mac bloggers shown in Table 4 and Table 5 respectively.

9to5Mac blog site had 10 bloggers as compared to Mac-Rumors with 6 bloggers. Though there were more bloggers in 9to5Mac, the number of target blog posts linked by the in-link posts was lower compared with MacRumors.

4.1. MacRumors Bloggers

4.1.1 Engagement Style: Listening-Participating Type

The Listening-Participating value was computed by counting the number of post titles that contained the target topic terms or related aspect terms, and normalized by the total number of target blog posts in each blog site. The aspect terms found in the blog posts were collated to the target topics to provide the aggregated result at target topic level. As seen from Table 6, the MacRumors bloggers had high participation rate (greater than 90%) on the target topic terms with respect to the other topics postings. This shows that

Table 4. MacRumors bloggers and their respective blog post count

Blogger Name	Number of target blog post
MacRumors_A	255
MacRumors_B	87
MacRumors_C	78
MacRumors_D	3
MacRumors_E	1
MacRumors_F	1

Table 5. 9to5Mac bloggers and their Respective Blog Post Count

Blogger Name	Number of target blog post
9to5Mac_A	105
9to5Mac_B	88
9to5Mac_C	60
9to5Mac_D	43
9to5Mac_E	39
9to5Mac_F	25
9to5Mac_G	17
9to5Mac_H	11
9to5Mac_I	2
9to5Mac_J	1

Table 6. Listening-Participating Values

Blogger	Listening-Participating values of target topics				Other topics
	iPad	iPhone	Mac	Total	
MacRumors_C	51(32.7%)	64(41.0%)	27 (17.3%)	142(91.0%)	14(**9.0%**)
MacRumors_D	7(41.2%)	7(41.2%)	2(11.8%)	16(94.1%)	1(**5.88%**)
MacRumors_A	231(34.4%)	238(35.5%)	176(26.2%)	645(96.1%)	26(3.9%)
MacRumors_B	82(40.4%)	80(39.4%)	34(16.7%)	196(96.5%)	7(3.5%)
MacRumors_E	0(0.0%)	0(0.0%)	1(100.0%)	1(100.0%)	0(0.0%)
MacRumors_F	1(50.0%)	1(50.0%)	0(0.0%)	2(100.0%)	0(0.0%)

the MacRumors bloggers were active participants towards Apple products related discussions. However, there was a slight deviation to other topics discussion by MacRumors_C (other topics' Listening-Participating score = 9.0%) and MacRumors_D (other topics' Listening-Participating score = 5.88%), indicating more participation in other non-Apple related products. We profile MacRumors_C and MacRumors_D as normal participating, with the rest of the bloggers (MacRumors_A, MacRumors_B, MacRumors_E, and MacRumors_F) as high participating type based on their other topics' Listening-Participating scores being less than 5%. Though MacRumors_E and MacRumors_F had only few postings, but their postings were specific to Apple related products leading to a low other topics' Listening-Participating score and subsequent high participating profile.

4.1.2 Engagement Style: Broad-Focused Type

The Broad-Focused type was evaluated by analyzing the unique target topics and the related features contained in the blog post titles for the MacRumors bloggers. MacRumors_A, MacRumors_B, and MacRumors_C were wide in their scope by covering comprehensively the target topics and related aspect features pertaining to Apple brand products. On the other hand, MacRumors_D was restricted in the scope of topics discussed. Only iPhone, iPad, Siri, iOS, iTunes, iCloud, and Mac topics were mentioned in MacRumors_D's postings. Further to that, MacRumors_E had a limited posting on the Mac topic, while MacRumors_F discussed only on applications pertaining to iPhone and iPad, indicating a focused or limited scope of topics. Nonetheless, the coverage of target topics and aspect features may be limited as the target topic and related

aspect-feature list were manually created based on prior knowledge of the Apple brand product domain. We aim to perform automatic topic discovery in our future work to improve the coverage of target topics, aspects, and features.

4.1.3 Engagement Style: Sharing-Creating Type

The Sharing-Creating type was analyzed through the similarity between the target blog posts content and its out-link posts content based on the Jaccard coefficient. Table 7 shows the maximum, minimum, and average Jaccard coefficient values with computed standard deviations. MacRumors_D had a low average self-links engagement style score (4.1%) indicating low content similarity between the self-linked posts. This meant that the blog post content was highly original even

mid-term (Jan 2011 to Dec 2011), and long-term (until year 2010) time period. Figure 10 and Figure 11 show a consistent number of links between the in-link blog posts and target blog posts for MacRumors_A, MacRumors_B, and MacRumors_C in the near-term and mid-term duration respectively. Similarly, there were links found in the long-term duration for the three top bloggers, indicating a consistent style. On the other hand, MacRumors_D had two links in the near-term duration, and one link in the long-term duration. Both MacRumors_E and MacRumors_F had only one link in the long-term duration. The results show MacRumors_D, MacRumors_E, and MacRumors_F to be causal bloggers with inconsistent link history.

Table 7. Engagement Style (Jaccard Coefficient) Scores

Blogger	Self-links (%)			Non-self-links (%)		
	Max	Min	Avg (Std dev)	Max	Min	Avg (Std dev)
MacRumors_E	15.4	0.7	9.1 (6.6)	NIL	NIL	NIL
MacRumors_C	79.5	0.2	8.2 (8.2)	19.3	1.2	7.8 (4.0)
MacRumors_B	96.1	0.3	7.9 (6.1)	15.3	0.3	7.2 (3.9)
MacRumors_A	98.1	0.3	7.8 (5.8)	16.4	0.5	7.6 (4.1)
MacRumors_F	6.5	6.5	6.5 (0.0)	7.5	7.5	7.5 (0.0)
MacRumors_D	14.8	0.5	**4.1** (4.6)	NIL	NIL	NIL

within self-linked posts, which shows a high creating style. The other MacRumors bloggers had Sharing-Creating scores between 5% and 10%, which classify them as normal creating type. It is noted that MacRumors bloggers are creative and original in its blog posts content even between self-linked content as seen from the similar engagement style values between the self-linked posts and non-self-linked posts.

4.1.4 Engagement Style: Consistent-Casual Type

The Consistent-Casual type measures the consistency of the target blog site to garner in-links over a period of time with respect to the number of target blog posts that are linked by the in-link posts within the analyzed period. Through registering the dates for the linked target blog posts, we grouped the blog posts according to the near-term (Jan 2012 to Jul 2012),

4.1.5 Persuasion Style

The persuasion style was measured based on the amount of subjectivity expressed in the target blog post content. From Table 8, MacRumors_A, MacRumors_B, MacRumors_C, MacRumors_D, and MacRumors_F were highly subjective in persuasion with their persuasion style scores greater than 12%. On the other hand, MacRumors_E had a mildly objective persuasion style as seen from the low persuasion style score of 6.9%. From the high subjectivity values, we could infer that MacRumors bloggers tend to be subjective rather than objective in their persuasion style, which indicates MacRumors to be a subjective product review blog site instead of being objective and reporting in style.

4.1.6 Persona Style
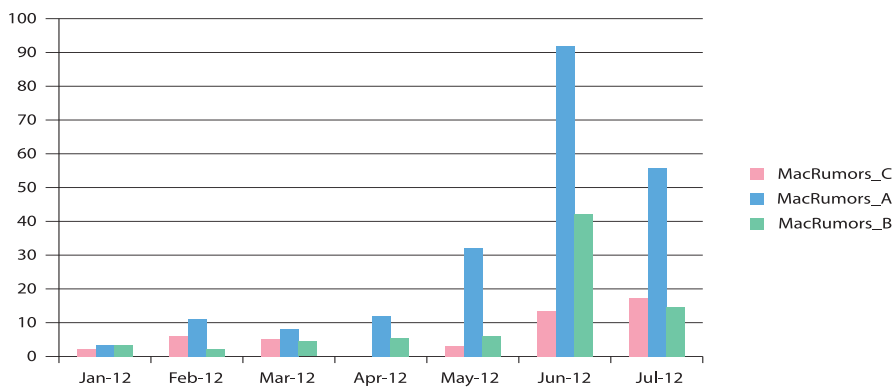
In the persona analysis, the similar sentiments and

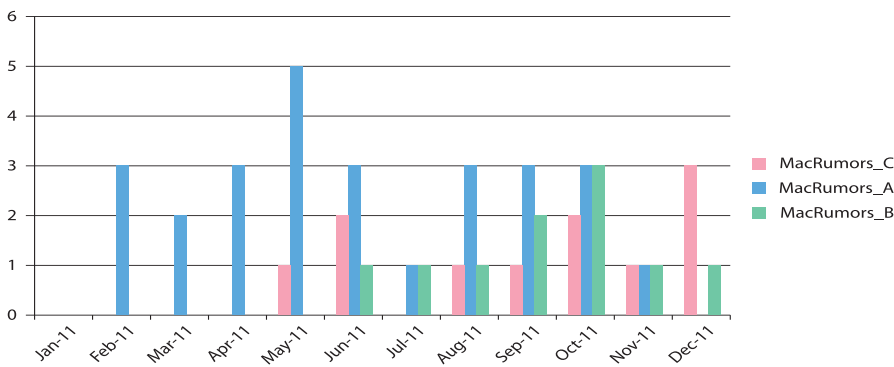**Fig. 10** Near-term linked posts count of top 3 MacRumors bloggers

Fig. 11 Mid-term linked posts count of top 3 MacRumors bloggers

Table 8. Subjectivity Values

Blogger	Subjectivity values (%)		
	Max	Min	Avg (Std dev)
MacRumors_D	15.6	13.8	15.2 (0.5)
MacRumors_F	13.7	13.7	13.7 (0.0)
MacRumors_B	13.9	10.6	12.7 (0.8)
MacRumors_C	14.2	9.2	12.5 (1.2)
MacRumors_A	15.7	10.5	12.2 (1.5)
MacRumors_E	6.9	6.9	**6.9 (0.0)**

dissimilar sentiments ratios between the in-link blog posts and the target blog posts towards the respective target topics were evaluated. From Table 9, it can be seen that MacRumors_A, MacRumors_C, MacRumors

_F have Persona Style scores greater than 80% indicating a high positive persona style. The other MacRumors blogger MacRumors_B is also of a normal positive persona style, inferring that MacRumors

Table 9. Similar Sentiment and Dissimilar Sentiment Values of MacRumors Bloggers

	Target							
	iPad		iPhone		Mac		Average	
	Similar Sentiment	Dissimilar Sentiment	Similar Sentiment	Dissimilar Sentiment	Similar Sentiment	Dissimilar Sentiment	Similar Sentiment	Dissimilar Sentiment
MacRumors_F	NIL	NIL	100.0%	0.0%	NIL	NIL	**100.0%**	0.0%
MacRumors_A	82.4%	17.6%	83.6%	16.4%	91.0%	89.0%	**89.3%**	10.7%
MacRumors_C	96.2%	3.8%	85.2%	14.8%	81.0%	19.0%	**87.5%**	12.5%
MacRumors_B	81.8%	18.2%	76.5%	23.5%	77.1%	22.9%	78.5%	21.5%
MacRumors_D	100.0%	0.0%	20.0%	80.0%	50.0%	50.0%	**56.7%**	43.3%
MacRumors_E	NIL	NIL	NIL	NIL	NIL	NIL	NIL	NIL

is a blog site that was generally positively received by its readers. However, MacRumors_D has a Persona Style score of 56.7% showing a controversial style within the positively received MacRumors blog site. This means that MacRumors_D blog posts readers expressed both agreement and disagreement on her postings. The evaluation result shows that bloggers from within a blog site could differ in their persona style. The persona style evaluation is not applicable for MacRumors_E because none of the blog postings were related to the target topics.

4.1.7 Macrumors Bloggers' Influence Profiles

Table 10 shows the evaluated profiles of the Macrumors bloggers. The results show the MacRumors bloggers to be participating and focused on the target topics as evaluated by the influence style

model. However, there were varying levels of participation and focus amongst the bloggers. In addition, the bloggers were evaluated to be creating in style, with MacRumors_D being more creative and original in blog posts content. MacRumors_A, MacRumors_B, and MacRumors_C had consistent influence over the duration of time compared to the other bloggers who were casual in their influence style. The Macrumors bloggers expressed high degree of subjectivity in their blog posts, except for MacRumors_E who is objective and reporting in nature. Likewise, the bloggers are positive in their persona, while MacRumors_D exhibits a controversial style. These differences in influence style show that bloggers within a common blog site could still exhibit varying style of influence, reflecting the need for further in-depth analysis to differentiate the bloggers' influence style. By providing an in-depth

Table 10. Profiles of MacRumors Bloggers Influence

Blogger	Engagement style				Persuasion style	Persona style
	Listening-Participating type	Broad-Focused type	Sharing-Creating type	Consistent-Casual type		
MacRumors_A	High Participating	Focused	Creating	High Consistent	High Subjective	High Positive
MacRumors_B	High Participating	Focused	Creating	Consistent	High Subjective	Positive
MacRumors_C	Participating	Focused	Creating	High Consistent	High Subjective	High Positive
MacRumors_D	Participating	Low Focused	High Creating	Casual	High Subjective	Controversial
MacRumors_E	High Participating	Low Focused	Creating	Low Casual	Low Objective	N.A.
MacRumors_F	High Participating	Low Focused	Creating	Low Casual	High Subjective	High Positive

analysis on the bloggers' influence style, we could further describe the bloggers' influence in details.

4.2. 9to5Mac Bloggers

9to5Mac blog site had 10 bloggers as compared to MacRumors with 6 bloggers. The greater number of individual bloggers in 9to5Mac would create more diversity in influence style amongst the bloggers. This is seen from the Listening-Participating evaluation results in Table 11, where the high Listening-Participating scores of 9to5Mac_J and 9to5Mac_D, and varying low to high scores of the other bloggers indicate differing engagement styles. In addition for the Broad-Focused type evaluation, 9to5Mac_A, 9to5Mac_B, 9to5Mac_C, and 9to5Mac_E were evaluated to be loosely focused on the Apple related target topics with more other non-target topics discussed, while the other 9to5Mac bloggers were restricted to the target topics range in their postings.

From Table 12, the duplication of content from non-self-link posts is minimal for the 9to5Mac bloggers, possibly due to copyrights constraints on usage of

Table 11. Listening-Participating Values

| Blogger | Listening-Participating values of target topics | | | | Other topics |
	iPad	iPhone	Mac	Total	
9to5Mac_J	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**1(100.0%)**
9to5Mac_D	27 (36.5%)	29 (39.1%)	6 (8.1%)	62 (83.7%)	**12 (16.3%)**
9to5Mac_B	88 (40.9%)	77 (35.8%)	21 (9.7%)	186 (86.4%)	29 (13.6%)
9to5Mac_E	39 (42.4%)	39 (42.4%)	7 (7.6%)	85 (92.4%)	7 (7.6%)
9to5Mac_C	31 (32.6%)	45 (47.3%)	13 (13.6%)	89 (93.5%)	6 (6.5%)
9to5Mac_H	13 (37.1%)	17 (48.5%)	3 (8.5%)	33 (94.1%)	2 (5.9%)
9to5Mac_A	114 (40.1%)	126 (44.3%)	30 (10.5%)	270 (94.9%)	14 (5.1%)
9to5Mac_F	18 (40.0%)	21 (46.7%)	4 (8.8%)	43 (95.5%)	2 (4.5%)
9to5Mac_G	13 (36.1%)	14 (38.9%)	9 (25.0%)	36 (100.0%)	0 (0.0%)
9to5Mac_I	4 (44.4%)	4 (44.4%)	1 (11.1%)	9 (100.0%)	0 (0.0%)

Table 12. Sharing-Creating (Jaccard Coefficient) Scores

| Blogger | Self-links | | | Non-self-links | | |
	Max	Min	Avg (Std dev)	Max	Min	Avg (Std dev)
9to5Mac_D	98.6%	0.4%	18.9% (28.5%)	93.6%	0.4%	10.6% (15.9%)
9to5Mac_B	96.3%	0.4%	32.7% (36.1%)	85.4%	0.7%	9.5% (6.7%)
9to5Mac_C	97.9%	0.2%	28.8% (29.2%)	93.8%	0.1%	8.5% (7.3%)
9to5Mac_E	95.4%	0.7%	32.2% (31.6%)	21.0%	0.6%	7.9% (3.4%)
9to5Mac_A	98.1%	0.5%	32.6% (33.5%)	85.9%	0.8%	7.8% (5.9%)
9to5Mac_I	95.3%	13.5%	**67.3% (46.5%)**	84.0%	3.2%	7.2% (9.6%)
9to5Mac_F	95.3%	1.9%	31.8% (31.0%)	39.7%	1.4%	7.2% (4.3%)
9to5Mac_H	98.0%	0.7%	31.5% (32.8%)	13.3%	2.3%	6.8% (2.2%)
9to5Mac_J	NIL	NIL	NIL	8.3%	4.0%	6.8% (1.8%)
9to5Mac_G	96.9%	0.5%	**43.5% (43.0%)**	12.1%	1.2%	**6.0% (2.0%)**

other sites' content. However, content similarity between self-links blog posts was high with 9to5Mac_I and 9to5Mac_G having Sharing-Creating scores greater than 40%. This shows that though 9to5Mac bloggers do observe the copyrights of other blogs' content, the tendency is to share their own blog content through the linked blog posts, indicating a sharing type in the influence style. The variations in the sharing-creating style amongst the bloggers are seen in the wide range of average similarity values (9to5Mac_I=67.3% to 9to5Mac_G=6.0%) between the 9to5Mac bloggers. It is noted that the similarity measure is affected by writing style as seen in short blog posts which tended to give higher Jaccard Coefficient Scores due to the higher similarity between short texts. In the Consistent-Causal type evaluation, it is observed that 9to5Mac_A, 9to5Mac_B, 9to5Mac_C, 9to5Mac_D, and 9to5Mac_E had high number of target posts linked by in-link posts in the near-term time period as seen in Figure 12. However, only 9to5Mac_A, 9to5Mac_B, and 9to5Mac_E had limited number of links (less than 6 for each blogger) in the mid-term duration. In the long-term duration, only 9to5Mac_A and 9to5Mac_F had one linked posting each. This shows that the influence consistency of the 9to5Mac bloggers is limited and differing across the bloggers, for example with 9to5Mac_A being more consistent

than 9to5Mac_J, who had limited number of linked posts in the near-term duration.

The persuasion style scores of the 9to5Mac bloggers are within the range of 10% to 12%, reflecting a subjective persuasion style type for the bloggers with similar level of subjectivity expressed in their blog posts. Only 9to5Mac_D and 9to5Mac_E have a highly subjective persuasion style with a score greater than 12%. Likewise, the 9to5Mac bloggers are positive in their Persona style, with the Persona style scores well above 70%, except for 9to5Mac_E who is controversial in style, and 9to5Mac_J, who did not post on the related target topics. From Table 13, the influence style profiles of the 9to5Mac bloggers show a participating and focused style with the exception of 9to5Mac_D being less active in the target related postings. The 9to5Mac bloggers, other than 9to5Mac_J, are sharing rather than creating in style. There are differing degrees of consistency amongst the bloggers with slightly more than half of them are causal in their influence style. On the whole, the 9to5Mac bloggers are positive in persona style, indicating general consensus from their readers, and they are also subjective in their persuasion style.

4.3. Evaluation Results

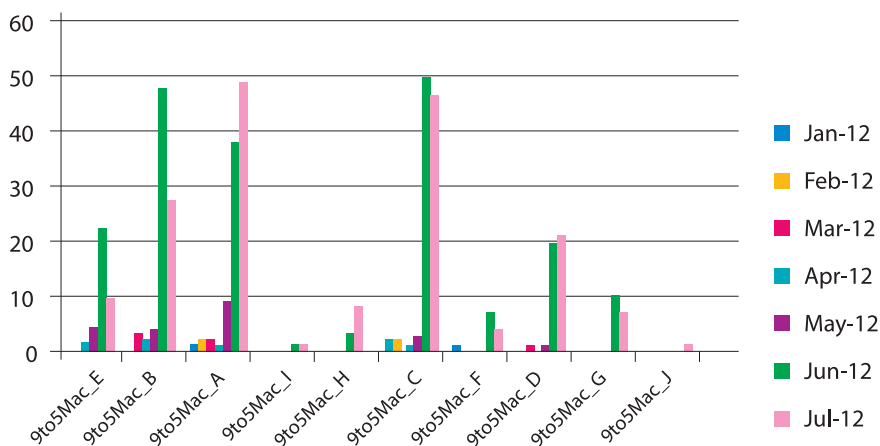A validation of the methods was performed to verify

Fig. 12 Near-term linked posts count of 9to5Mac bloggers.

Table 13. Profiles of 9to5Mac Bloggers Influence.

Blogger	Engagement style				Persuasion style	Persona style
	Listening-Participating type	Broad-Focused type	Sharing-Creating type	Consistent-Casual type		
9to5Mac_A	High Participating	Low Focused	Sharing	Consistent	Subjective	Positive
9to5Mac_B	Low Participating	Low Focused	Sharing	Consistent	Subjective	High Positive
9to5Mac_C	Participating	Low Focused	Low Sharing	Low Consistent	Subjective	Positive
9to5Mac_D	Low Listening	Focused	Low Creating	Casual	High Subjective	Positive
9to5Mac_E	Participating	Low Focused	Sharing	Consistent	High Subjective	Controversial
9to5Mac_F	High Participating	Focused	Low Sharing	Casual	Subjective	High Positive
9to5Mac_G	High Participating	High Focused	Sharing	Casual	Subjective	High Positive
9to5Mac_H	Participating	Focused	Low Creating	High Casual	Subjective	High Positive
9to5Mac_I	High Participating	Focused	Sharing	High Casual	Subjective	High Positive
9to5Mac_J	High Listening	High Focused	Creating	High Casual	Subjective	N.A.

the performance of the influence style model and the threshold values used. Approximately 12% (100 records out of 816) of the total target blog post records were extracted for the inter-coder reliability testing in the Persona evaluation, with a balanced distribution for the positive and negative tagged samples. A further 100 records were used for the subjectivity inter-coder reliability testing, where the samples were distributed evenly between the subjective and objective thresholds. A sample of 100 records (Approximately 2.5% of the total number of out-links) was used in the reliability testing for the Sharing-Creating evaluation with a balanced distribution between creating records and sharing records. The threshold values used in the validation are Sharing-Creating score (15%), Persuasion style score (7%), and Persona score (Positive>60%, Negative<=40%). The computed Kappa values for the two person coders' inter-reliability testing are as follow: Sharing-Creating evaluation (Kappa value=0.97), Persuasion evaluation (Kappa value=0.69), and Persona evaluation (Kappa value=0.61), which are in

the acceptable range (Cohen, 1960). The conflicting tags by the two coders were reviewed and manually re-classified and used as the answer keys. The F1 score defined as (F1 score = $(2*Precision*Recall)/(Precision+Recall)$) is used to measure the performance of the influence style model in relation to the precision and recall. The F1 score for both Sharing type (90.9%) and Creating type (90.1%) classification are high. This shows the Sharing-Creating type predictions made by the model to be reliable as validated by the coders.

The model's F1 scores for Persuasion evaluation were relatively lower (Objective type F1 score=7.1%, Subjective type F1 score=85.4%). Our threshold value (7%) for Objective type classification was too stringent resulting in the poor performance. Rather, a threshold value of 10% improves performance (Objective type F1 score=65.8%, Subjective type F1 score=78.7%), while a threshold value of 12% gives F1 score results of Objective type=36.9%, and Subjective type=49.1%. The results also show that persuasion style could vary even within individual blogger posts as seen from the varia-

tion between the maximum, minimum, and average subjective scores of the bloggers.

The F1 scores for positive Persona (74.6%) are higher than the scores for negative Persona (54.1%), because of the assumption that neutrality is considered as agreement, resulting in the higher scores for predicting similar sentiments between blog posts. Performance of the Persona evaluation is dependable on the sentiment analysis; therefore any improvement in the sentiment analysis process would result in better performance in the persona evaluation. In general, bloggers from popular sites have positive persona where in-link bloggers express agreement towards them. Nevertheless, there were bloggers who displayed different persona from the other bloggers as seen in MacRumors_D from Macrumors, and 9to5Mac_E from 9to5Mac, who had controversial persona style.

5. DISCUSSION

Our study analyzes blog posts at the blogger-level to give an in-depth evaluation of the individual blogger influence style. Manual observations and description of the blog sites in Technorati.com revealed both MacRumors and 9to5Mac to be specialized blog sites focusing on Apple brand related products, which are indicated by the Participating and Focused types of their bloggers in our evaluation results. It is observed that individual bloggers could differ from the overall blog site and other bloggers in influence style. For example, the overall MacRumors blog site was highly participating towards the main target topics, while individually MacRumors_C and MacRumors_D were less participating on the main target topics, with more diverse range of topics discussed. Similarly, the 9to5Mac bloggers had differing Participating and Focus type level. The level of participation and focus on topics could also differ between the individual bloggers. Being able to differentiate the participation level and topic focus of individual bloggers allows blog readers to determine the bloggers' depth and scope of knowledge in respective topics. Though bloggers could differ in influence style, there would be certain rules, such as the non-infringement of copyrights commonly observed by bloggers resulting in a similar Sharing-

Creating influence style for the bloggers, especially within non-self-link posts. Influence consistency is also observed to differ between bloggers. For example, MacRumors_D, MacRumors_F, and MacRumors_E were casual in their influence style compared with the other bloggers from MacRumors, while 9to5Mac_E, 9to5Mac_B, and 9to5Mac_A were more consistent in the more casual 9to5Mac blog site. Both the MacRumors and 9to5Mac bloggers, except for MacRumors_E were evaluated to be subjective in their persuasion style based on their subjectivity score. The homogeneity in persuasion style is indicative of a product review blog site, which is generally subjective in nature as compared to an objective news reporting blog site. We would attempt to improve subjectivity detection by taking into account the grammatical relations between words in our future work in consideration of the linguistic constraints in matching subjectivity terms for semantic orientation. For example, the subjectivity value of the phrase *"very mad"* would increase by considering the adverb modifier relationship "ADVMOD(mad, very)", instead of just matching *"mad"* as a subjective term.

We attempt to derive deterministic thresholds for the scores to provide an objective measurement to the respective influence style evaluation. Through this, we are able to provide a detailed breakdown of the bloggers' influence styles. Future studies could involve a formal validation of the threshold values through human coders testing, and by analyzing a sizable number of bloggers to improve the test quality. Influence style of a blogger is time dependent, where the bloggers influence on specific topics is based on the analyzed period as seen from the Consistent-Casual influence style result. A different period of analysis could have identified the influence style differently since the measurement values would have changed. In this study, the target topic and related aspects list were manually created based on prior knowledge of the Apple-product domain. We aim to perform automatic topic discovery in our future work to improve the coverage of target topics and aspects. An in-depth topic analysis would provide a more comprehensive coverage of the specific topics discussed.

The results from our study show that influence style is specific to individual bloggers, and the blog site

influence style is not representative of all bloggers. By analyzing the influence style of individual bloggers, we are able to provide more fine-grained descriptions of the bloggers' influence. Previous studies which assumed a monolithic definition of influence would not be able to differentiate the varying influence styles of the individual bloggers. The knowledge on individual blogger's influence style would be used in our future work to detect influence in the blogosphere. For example, we could use the Participating, Focused, Creating, Consistent, Low Subjective, and Highly Positive influence style of MacRumors_C as features to predict the possibility of influence propagation. Future studies could explore the combination of influence styles that are indicative of high influence, which could determine whether a mixture of influence styles would increase the influence level of a blog site.

6. CONCLUSION

In contrast to previous studies which used a monolithic definition of influence for blogs, our study further describes the influence of bloggers in terms of engagement, persuasion, and persona style. The detailed analysis of influence could differentiate the influence style of bloggers, and provide a more comprehensive description of the influence exerted. Our evaluation results show there are differences in influence style between individual bloggers and the overall blog site as well as amongst the bloggers. This indicates a need to analyze individual blogger's influence style to provide further in-sights into the blog influence profile. The provision of the bloggers' influence style would enhance the influential blog search experience through the detailed description of the manner in which the bloggers exert influence. Further to that, a more comprehensive and accurate portrayal of the blog influence through the evaluated influence profiles would in turn improve influence detection within the blogosphere.

REFERENCES

Adar, E., & Adamic, L. A. (2005). Tracking information epidemics in blogspace. In A. Skowron, R. Agrawal, M. Luck, T. Yamaguchi, P. Morizet-Mahoudeaux, J. Liu, & N. Zhong (Eds.), *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 207-214). Washington, DC: IEEE Computer Society.

Agarwal, N., & Liu, H. (2008). Blogosphere: Research issues, tools, and applications, *SIGKDD Explorations Newsletter, 10*(1), 18-31.

Agarwal, N., Liu, H., Tang, L., & Yu, P.S. (2008). Identifying the influential bloggers in a community. In M. Najork, A. Broder, & S. Chakrabarti (Eds.), *Proceedings of the International Conference on Web Search and Web Data Mining* (pp. 207-218). New York, NY: ACM Press.

Agarwal, N., & Liu, H. (2009). *Modeling and Data Mining in Blogosphere*. San Rafael, CA: Morgan & Claypool.

Brehm, J. W. (1966). *A theory of psychological reactance.* Academic Press.

Cai, K. K., Bao, S. H., Yang, Z., Tang, J., Ma, R., Zhang, L., & Su, Z. (2011). OOLAM: An Opinion Oriented Link Analysis Model for influence persona discovery. In I. King, W. Nejdl, & H. Li (Eds.), *Proceedings of the 4th International Conference on Web Search and Data Mining,* (pp. 645-654). New York, NY: ACM Press.

Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591-621.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D.T. Gilbert, & S.T. Fiskel (Eds.). *The handbook of social psychology* (4th ed.), 2, 151-192. New York: McGraw-Hill.

Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Costa, P. T., Jr., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment, 4,*

5-13.

Ghosh, R., & Lerman, K. (2008). Community detection using a measure of global influence. In L. Giles, J. Yen, H. Foley, M. Smith, & H. Zhang (Eds.), *Proceedings of the 2nd SNA-KDD Workshop on Social Network Mining and Analysis* (pp. 20-35). Berlin, Heidelberg: Springer-Verlag.

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology, 78*(6), 1360-1380.

Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Information diffusion through blogspace. In S. Feldman, M. Uretsky, M. Najork, & C. Wills (Eds.), *Proceedings of the 13th International Conference on World Wide Web* (pp. 491-501). New York, NY: ACM Press.

Guadagno, R. E., Okdie, B. M., & Eno, C. A. (2008). Who blogs? Personality predictors of blogging. *Computers in Human Behavior, 24*(5), 1993-2004.

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). Bridging the gap: A genre analysis of weblogs. In Sprague, R.H. (Ed.), *Proceedings of the 37th Hawaii International Conference on System Sciences*. Los Alamitos: IEEE Computer Society Press.

Herring, S. C., Kouper, I., Paolillo, J. C., Scheidt, L. A., Tyworth, M., Welsch, P., Wright, E., & Yu, N. (2005). Conversations in the blogosphere: An analysis "from the bottom up." In Sprague, R. H. (Ed.), *Proceedings of the Thirty-Eighth Hawaii International Conference on System Sciences,* (pp. 107b). Los Alamitos: IEEE Press.

Hogg, M. A., & Vaughan, G. M. (2005). Social *psychology.* Harlow: Pearson/Prentice Hall.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution,* 2(1), 51-60.

Krackhardt, D. (1992). The strength of strong ties: *The importance of philos in organizations,* 216-239. Boston, MA: Harvard Business School Press.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). Predicting positive and negative links in online social networks. In M. Rappa, P. Jones, J. Freire, & S. Chakrabarti (Eds.), *Proceedings of the 19th International Conference on World Wide Web* (pp. 641-650). New York, NY: ACM Press.

Li, H., Bhowmick, S. S., & Sun, A. (2011). CASINO: Towards conformity-aware social influence analysis in online social networks. In B. Berendt, A. Vires, W. Fan, C. Macdonald, I. Ounis, I. Ruthven (Eds.), *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (pp. 1007-1012). New York, NY: ACM Press.

Mackie, D. M. (1987). Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology,* 53, 41-52.

Matsumura, N., Yamamoto, H. & Tomozawa, D. (2008). Finding influencers and consumer insights in the blogosphere. In E. Adar, M. Hurst, T. Finin, N. Glance, N. Nicolov, & B. Tseng (Eds.), *Proceedings of International Conference on Weblogs and Social Media* (pp. 76-83). California CA: AAAI Press.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371-378.

Somasundaran, S., & Wiebe, J. (2010), Recognizing stances in ideological on-line debates. In D. Inkpen, & C. Strapparava (Eds.), Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text (pp. 116-124). Stroudsburg, PA: ACL.

Song, X., Chi, Y., Hino, K., & Tseng, B. (2007), Identifying opinion leaders in the blogosphere. In A.H.F. Laender, A. O. Falcão, Ø. H.Olsen, M. J. Silva, R. Baeza-Yates, D. L. McGuiness, & B. Olstad (Eds.), Proceedings of the 16th ACM Conference on Information and Knowledge Management (pp. 971-974). New York, NY: ACM Press.

Tan, L. K. W., Na, J.-C., & Theng, Y.-L. (2011). Influence detection between blog posts through blog features, content analysis, and community identity, *Online Information Review, 35*(3), 425-442.

Tan, L. K. W., Na, J.-C., Theng, Y.-L., & Chang, K. Y. (2012a). Phrase-level sentiment polarity classification using rule-based typed dependencies and additional complex phrases consideration. *Journal of Computer Sciences and Technology, 27*(3), 650-666.

Tan, L. K. W., Na, J.-C., Theng, Y.-L., & Chang, K. Y. (2012b). *Profiling Blog Sites Through an Influence*

Style (INFUSE) Model (Manuscript submitted for publication).

Wiebe, J., Wilson, T., Bruce, R., Bell, M., & Martin, M. (2004). Learning subjective language. *Computational Linguistics, 30,* 277-308.

Wilson, T., Wiebe J., & Hoffmann P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In R.J. Mooney, C. Brew, L-F. Chien, K. Kirchhoff (Eds.), *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing* (pp. 347-354). Stroudsburg, PA: ACL.