

Affect in social media: the role of audience and the presence of contempt in cyberbullying

Mihaela Cocea

University of Portsmouth, School of Computing, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth, PO1 3HE, United Kingdom

mihaela.cocea@port.ac.uk, <http://coceam.myweb.port.ac.uk/>

Abstract. Gervais & Fessler’s Attitude-Scenario-Emotion (ASE) model is a useful tool for the detection of affect in social media. An addition to the model is proposed – the audience – and its role in the manifestation of affect is discussed using a cyberbullying scenario. The presence of contempt in cyberbullying is also discussed.

Social media is now a standard way of interacting with other people – friends, acquaintances and a wider audience of unknown people. Given the wealth of online activity and its relevance to our “real” lives, analysis of social media has gained a lot of interest. In particular, affect detection from text attracted many researchers (e.g., Paltoglou & Thelwall, 2012; Altrabsheh et al., 2015; Tromp & Pechenizkiy, 2015), where affect is broadly understood as anything to do with human emotion and related phenomena. In this area, there is confusing terminology about what is expressed and what can be detected from social media text. A recent article (Munezero et al., 2014b) sought to clarify this terminology and it is largely aligned with Gervais & Fessler’s article in the sense that emotions can be observed and that sentiments can be inferred from different emotion expressions.

Research in this area focuses on broad categories such as subjective vs. objective text and polarity, i.e., positive/negative/neutral (e.g., Liu & Zhang, 2012; Gaber et al., 2015). Some attempts were made to detect specific emotions, mostly based on Ekman’s basic emotions model (Ekman, 1992) or Russell’s dimensions (Russell & Mehrabian, 1977) (e.g., Balahur, 2013; Paltoglou & Thelwall, 2013). In the last five years, detection of more complex affective phenomena has emerged, such as: humour and irony (Reyes et al., 2012), nastiness (Justo et al., 2014), and sarcasm (Justo et al., 2014; Altrabsheh et al., 2015). Moreover, there is a growing interest in the detection of online activity such as anti-social behaviour (Munezero et al., 2014a) and cyberbullying (Dinakar et al., 2012; Zhao et al., 2016) through the manifestation of affective text (e.g., insults).

In this context, the ASE model proposed by Gervais & Fessler would be a useful tool for sentiment detection. Given the characteristics of the medium of communication, i.e., social media, the addition of another dimension is proposed: the audience. The audience is implicit in social media; one may not be aware of the entire audience – in fact there is a tendency to underestimate the size of the audience (Bernstein et al., 2011), but one accepts that there is an audience and one’s mental model of the audience influences one’s online activity (Marwick, 2011; Litt, 2012).

To illustrate the ASE model with the addition of the audience a cyberbullying scenario is used, where cyberbullying is the repeated communication through digital media of hostile/aggressive messages

intended to harm/discomfort others (Tokunaga, 2010). The influence of attitudes, audience and emotions on behaviour are discussed for several cases (Table 1) – the cases and their explanations are not exhaustive; they are meant as an illustration of how the audience has an influence on one’s emotions, and, consequently, on one’s behaviour.

Scenario: X is repeatedly making nasty comments about Y’s appearance on a social network. The last comment is more hurtful and is accompanied by an unflattering picture.

	Case 1	Case 2	Case 3	Case 4	Case 5
Attitude	Appearance is indicative of one’s worth		One’s worth is independent of one’s appearance		
Audience	Acquaintances; friend with X	Friends; friend with X	Friends and acquaintances; friends with X	Friends	Friends and acquaintances; friends with Y
Emotion	Lack of compassion	Mirth	Pity for Y	Indignation	Fear
Behaviour	Indifference	Laughter at Y	Indifference	Defend Y	Indifference

Table 1. Cases of the Attitude-Scenario-Emotion+Audience (ASE+A) model for a cyberbullying scenario.

Case 1 and 2 illustrate the presence of contempt with “cold” and “hot” phenomenology respectively. In Case 1, the actor lacks compassion for Y, but the uncertainty over the audience (i.e., perceived risk of social negative evaluation/less social reward), prevents a reaction resulting in indifference. Case 2, on the other hand, with an audience of friends involved in the bullying (i.e., opportunity to gain social reward or fear of losing social approval if not joining) activates hostility in the form of laughter.

Case 3 shows a conflict between one’s beliefs and a friends’ behaviour, which could result in compassion for Y, but a reluctance to risk the friendship with X, leading to no reaction. In case 4, defending Y is an opportunity to gain social rewards from an audience of friends (presumed to have similar beliefs). In case 5, the actor is fearful that defending Y would lead to being targeted as well.

Thus, the audience has a major role in one’s reactions on social media, with two main dimensions: (a) relationship to main actors (X and Y) and (b) the social group (friends/acquaintances/unknown). Social reward depending on the audience of friends vs. acquaintances was explored in previous research on cyberbullying (Jones et al., 2011; Bastiaensens et al., 2014) – these cost-reward and power dynamic aspects are in line with the bookkeeping and commitment functions in social relationships, as discussed in Gervais & Fessler’s article.

Gervais & Fessler define contempt as the lack of caring, but with the lack of intrinsic motivation to harm others. In cyberbullying the intent to harm is present, without a clear distinction if this harm is the main goal or a means for other ends like social standing.

In face-to-face bullying and, possibly even more so, in cyberbullying the element of power or social standing may be the main goal. This may explain why some people join in the bullying when they do not know the victim or the initiating bully. Bertolotti & Magnani (2013) describes this behaviour as gratuitously humiliating another person in public and compares it to “sociopathy”. In this sense, it is

similar to the “hate crimes” mentioned by Gervais & Fessler and contempt rather than hate is more likely to justify such behaviour.

Using the ASE+A model for affect detection offers the potential to better understand affect in the context of social dynamics by integrating elements of the context (i.e., scenario), the audience (the network of relations) and attitudes. Of these, perhaps the most challenging is the detection of attitudes; in the case of cyberbullying, stereotypes about one’s appearance, intelligence and other personal characteristics are often involved (Dinakar et al., 2012) and could be used as a proxy for attitudes.

References

- Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2015). Detecting sarcasm from students’ feedback in Twitter. In *Design for Teaching and Learning in a Networked World* (pp. 551-555). Springer International Publishing.
- Balahur, A. (2013). Sentiment analysis in social media texts. In *4th workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 120-128).
- Bastiaensens, S., Vandebosch, H., Poels, K., Van Cleemput, K., DeSmet, A., & De Bourdeaudhuij, I. (2014). Cyberbullying on social network sites. An experimental study into bystanders’ behavioural intentions to help the victim or reinforce the bully. *Computers in Human Behavior*, 31, 259-271.
- Bernstein, M. S., Bakshy, E., Burke, M., & Karrer, B. (2013). Quantifying the invisible audience in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 21-30). ACM.
- Bertolotti, T., & Magnani, L. (2013). A philosophical and evolutionary approach to cyber-bullying: social networks and the disruption of sub-moralities. *Ethics and information technology*, 15(4), 285-299.
- Dinakar, K., Jones, B., Havasi, C., Lieberman, H., & Picard, R. (2012). Common sense reasoning for detection, prevention, and mitigation of cyberbullying. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(3), 18:1-30.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99 (3), 550–553.
- Gaber, M. M., Cocea, M., Wiratunga, N., & Goker, A. (Eds.). (2015). *Advances in social media analysis* (Vol. 602). Springer.
- Jones, S. E., Manstead, A. S. R., & Livingstone, A. G. (2011). Ganging up or sticking together? Group processes and children’s responses to text-message bullying. *British Journal of Psychology*, 102, 71–96.
- Justo, R., Corcoran, T., Lukin, S. M., Walker, M., & Torres, M. I. (2014). Extracting relevant knowledge for the detection of sarcasm and nastiness in the social web. *Knowledge-Based Systems*, 69, 124-133.
- Litt, E. (2012). Knock, knock. Who's there? The imagined audience. *Journal of Broadcasting & Electronic Media*, 56(3), 330-345.

- Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In C. C. Aggarwal & C. Zhai (Eds.), *Mining text data* (pp. 415-463). Springer US.
- Marwick, A. E. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, *13*(1), 114-133.
- Munezero, M., Montero, C. S., Kakkonen, T., Sutinen, E., Mozgovoy, M., & Klyuev, V. (2014a). Automatic detection of antisocial behaviour in texts. *Informatica*, *38*(1), 3 -10.
- Munezero, M., Montero, C. S., Sutinen, E., & Pajunen, J. (2014b). Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing*, *5*(2), 101-111.
- Paltoglou, G., & Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *3*(4), 66.
- Paltoglou, G., & Thelwall, M. (2013). Seeing stars of valence and arousal in blog posts. *IEEE Transactions on Affective Computing*, *4*(1), 116-123.
- Reyes, A., Rosso, P., & Buscaldi, D. (2012). From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering*, *74*, 1-12.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of research in Personality*, *11*(3), 273-294.
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in human behavior*, *26*(3), 277-287.
- Tromp, E., & Pechenizkiy, M. (2015). Pattern-Based Emotion Classification on Social Media. In Gaber, M. M., Cocea, M., Wiratunga, N., & Goker, A. (Eds.), *Advances in Social Media Analysis* (pp. 1-20). Springer.
- Zhao, R., Zhou, A., & Mao, K. (2016). Automatic detection of cyberbullying on social networks based on bullying features. In *Proceedings of the 17th International Conference on Distributed Computing and Networking* (43:1-6). ACM.