AAAI Technical Report WS-13-01
Computational Personality Recognition (Shared Task)

Predicting Conscientiousness through Semantic Analysis of Facebook Posts

Marc T. Tomlinson, David Hinote and **David B. Bracewell**
Language Computer Corporation
2435 N. Central Expy, Suite 1200
Richardson, Texas 75080
marc, dhinote, david@languagecomputer.com

Abstract

Here we present a method for detecting an individual's level of conscientiousness based on an analysis of the content of their Facebook status updates. Our model is based on the identification of semantic evidence of facets related to conscientiousness; an individual's belief of their control over events around them and their goal orientation. The model achieves a correlation of r=.27 on a subset of the Facebook data published for the myPersonality workshop, with an accuracy of 58.13% for detecting if an individual is above or below the median and 68.03% for those outside of one standard deviation. While we take a narrow approach and identify only one personality trait, the general methodology of directly looking for evidence of traits in an individual's utterances is applicable to discovering models for all of the personality traits.

Personality traits are enduring factors of an individual that are correlated with broad classes of behavior. They are generally manually assessed by a questionnaire which identifies an individual's likelihood of exhibiting relevant behaviors. Typically, automatic means of predicting personality analyze written communications by the individual based on the frequency with which they use words from certain classes. In contrast, we show that an analysis of the psychological implicatures encoded in an individual's unique way of expressing events in their surroundings and their relationship to those events can be used to identify their personality.

Currently, the most popular theory of personality is the big-five model. The big-five model suggests that an individual's long-term behavior is best characterized using five dimensions: openness to experience; conscientiousness; extraversion; agreeableness; and emotional stability (also called neuroticism). An individual can be characterized by the amount to which the individual exhibits each of these traits as well as facets along those dimensions (See McCrae and Costa 1999, for a review). Personality traits are meant to be enduring representations of an individual's behavior which persist over time. They are as predictive at long-term outcomes (e.g. mortality, divorce, and occupational attainment) as socio-economic status and intelligence (Roberts et al. 2007). In addition, individual dimensions and facets of those dimensions (sub-categories) are predictive of other outcomes, such as being in a leadership role. Many automated approaches to personality identification have taken a generic approach, linking word cooccurrence statistics or network metrics to all five personality dimensions. In contrast, we investigate a direct approach, which looks for only a single dimension, conscientiousness, based on an analysis of the behaviors expected by people exhibiting characteristics along that dimension.

Conscientiousness is related to an individual's "tendency to control behavior in pursuit of goals" (Chang, Connelly, and Geeza 2012). The facets making up conscientiousness include terms such as orderliness, achievement-striving, deliberate, and self-disciplined. Thus, conscientious individuals should be differentiable based on how their language encodes these facets.

We detect goal orientation and perceived control by examining differences in the way in which individuals express event structures that they or other individuals participate in. When reporting on an event, individuals choose to highlight various parts of the event based on internal factors (personality, cognitive state, goals, age, group identity, etc.) and external factors (principals of communication, the other individuals in the conversation, etc.). For example, when describing a recent victory in a competition, an individual can express that victory in many different ways, such as "I won", "I crushed the competition" or maybe "John was unable to withstand my attack". The first two sentences suggest that the individual has more control in the situation whereas the last sentence suggests that the protagonist's victory was more attributable to the actions taken by the antagonist. The way in which different individual's describe the same, or similar, situations can be used to reveal important things about the cognitive state of the individual and infer their motivation, goals, perceived control over situations, or even long-term enduring personality traits.

Related Work

Many of the approaches to identifying personality have been through an analysis of the words an individual uses to express themselves. Researchers in automated personality identification have generally used the term linguistic style analysis to refer to analyzing an individual through analysis of their communications. In particular, these types of analyses look at distributions of different types of words. The Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2007) dictionary is the most current set of categories. The dictionary provides lists of words that convey various psychological dimensions, such as words with positive emotional content or topics related to psychologically interesting phenomena. These types of analyses have demonstrated that it is possible to predict personality traits by looking at the distribution of the words used by individuals, for example the use of social language is correlated with extraversion (Mehl, Gosling, and Pennebaker 2006). Classifiers built using these approaches have achieved results 3-10% above chance on a large corpus of narratives across the big-five factors (Mairesse et al. 2007).

Alternatively, automated analysis can also be done by looking at an individual's behavior. This type of analysis has mostly focused around network metrics on twitter or Facebook. Users on these social-media sites are able to follow or friend other individuals (they automatically receive their news feeds). Individuals can also rebroadcast messages provided by their friends with or without their own comments. For example, Quercia et al. (2011) analyzed the predictive accuracy of the personality of Twitter users based on network metrics. Quercia et al. showed a significant correlation between emotional stability and several of the network metrics, in contrast to expert analysis of Facebook profiles which were not able to reveal those dimensions as well (Gosling, Gaddis, and Vazire 2007).

Model

Our model is based on extracting the nuances within an individual's communication that provide insights into their goal orientation and perceived control over situations. We do this by investigating the semantic content of their communications based on an analysis of the event-based verbs that they use to describe particular situations and the thematic roles that the individual uses to express their own behaviors or another individual's behaviors.

We are primarily interested in two different thematic roles, the agent role and the patient role. We base our definitions on Dowty's definition of proto-agents which demonstrate (Dowty 1991, pg. 572):

- Volitional involvement in the event or state
- Sentience
- Causal connectedness to an event or change in state of another participant
- Movement (relative to position of another participant)

Proto-patients, in contrast:

- Undergo change of state
- Incremental Theme
- Causally affected by another participant
- Stationary relative to movement of another participant

Thus, in the expression "I crushed John", I is the agent and John is the patient. These relations are encoded in the Propbank resource (Palmer, Gildea, and Kingsbury 2005), which provides a source of annotated sentences for learning to assign thematic roles for events.

Unfortunately, most annotations of agents and patients are done using a binary classification, instead of a graded scale. Therefore, an analysis of just these factors wouldn't give us much differentiation across individuals without a large amount of data. Thus, we also look for other signals within the predicate which specify control. Verbs which encode manner generally require animate agents (see Beavers and Koontz-Garboden 2012 for some more recent discussions on this), thus increasing the percieved agency of the subject. For example, when someone says, "I walloped him", that indicates greater control than would be indicated by using the phrase, "I won". We generalize this suggesting that verbs which encode greater specificity suggest more control on the part of the agent of the event. We calculate specificity using WordNet (Fellbaum 1998).

WordNet is a lexical resource which hierarchically encodes words and the semantic relations between those words. Each word is broken out into senses which convey differences in the usage of the word; for example, the word shoot has two noun senses and 20 difference verb senses. Verbs within WordNet are encoded hierarchically along troponymy relationships. Troponyms express the same high-level concept. For example, according to the WordNet verbal hierarchy, the verb gun down is a more specific way to say shoot, which is more specific than hurt, which in turn is a more specific way to say indispose, and finally, is a more specific way to say change. Because of gun down's depth in the WordNet hierarchy we can say that it is a very specific verb indicating that the statement is attributing a greater control over the event to the agent of the sentence. The top-level hiearchy is encoded based on psychological models of language and perception (Miller and Johnson-Laird 1976).

There are some issues with this approach. The top-level hierarchy in WordNet is not ideal. It is too restrictive in that each word-sense generally only has one top-level category. This ignores the multi-faceted nature of many verbal expressions which often entail multiple different top-level concepts. For example, when someone "guns down" another individual, this has a high likelihood of also entailing that they won a victory. In addition, there are some uncertainties in exactly what constitutes a verbal hyponym; this has led to at least some of the verb hyponym relations within WordNet include causal relations (Richens 2008). However, WordNet still provides a good proof-of-concept for representing the specificity of a verb.

To discover an individual's goal orientation we look at the objectivity of the predicate that is used to encode an event. Work on goal perception indicates that actions with negative social acceptability are more strongly percieved as goal of the agent (Knobe 2003). We extend this looking at predicates with less objectivity in general, suggesting that they are more goal oriented. For example, compare "I gave the money to the campaign fund" to "I donated the money to the campaign fund" to. Both statements entail the same basic event, but donate (which is less objective) seems to more strongly implies a goal.

Senti-WordNet (Baccianella, Esuli, and Sebastiani 2010) is a resource that identifies the positivity, negativity, and objectivity of word senses in WordNet. Each synset in WordNet is given a score along each of the three dimensions with the caveat that the sum of the scores for each of the three dimensions is 1 for each synset. Utilizing this encoding, objectivity can be thought of as representing the left-overs after a word's positive and negative qualities have been decided. For example, the synset representing the first sense of donate has a positivity of .625, a negativity of 0, and thus an objectivity of .375. While a sysnset containing the word give (cause to have) carries a positivity of 0, negativity of .125, thus it has an objectivity of .875. Accordingly, individuals should be more likely to infer that someone had the goal of donating rather than giving.

Methods

For our analysis, we looked at the data released by the MyPersonality project for the Workshop on Computational Personality Recognition (Celli et al. 2013). The MyPersonality project runs a Facebook application that collects data on individuals that agree to participate. Individuals can answer questionnaires leading to analysis of their personality according to the big-five dimensions, facets of the big-five, their IQ, and many others. The authors report that they have data on over 6 million tests from 4 million individuals. We used the subset of the data which was released for this workshop on personality detection. The subset contained profiles for 250 users with 9800 status updates across all of the individuals. Each individual is rated on a 4-point scale from 1-5 on each of the big-five dimensions.

We performed an automatic analysis of each user's status updates. Parsing was performed using an in-house part-of-speech tagger, event recognition software, and rule-based semantic parser. We identified between 0 and 17 event-based predicates for each poster, approximately 10% of posts. We removed 88 subjects from the analysis because an event-based verb could not be found in their posts. Our results are calculated only on this subset of individuals. Each verb within the sentence was graded according to its specificity and to its objectivity. In addition, each agent was marked for its person (first or other, i.e. second or third, "I am having a bad day" vs. "Rain is ruining my plans") as was the patient. Sentences that were missing an agent, but utilized a verb which requires an agent, were marked as having first person agents. Facebook users have a habit of dropping the sentence initial subject when they are referring to them-selves.

Analysis and Results

First, we ensured that the individuals dropped did not affect our results. We found no significant difference in the actual conscientiousness between those individuals that were not analyzed ($\mu = 3.53$) and those that were analyzed ($\mu = 3.52$), $t < 1$, nor did they differ significantly on any of the other personality dimensions.

For each individual we calculated an average depth and an average objectivity for predicates which had a first person agent, predicates which had a non-first person agent, predicates which had first-person patients, and finally, predicates with non-first person patients. Average values for each of these parameters are shown in Table .

	1^{st} Person		2^{nd} - 3^{rd} Person	
	Agent	Patient	Agent	Patient
Mean Depth	.99	.17	.73	1.23
Mean Objectivity	.95	.95	.99	.93

Table 1: Mean values for specificity and objectivity of verbal predicates associated with the given semantic argument

Interestingly, all means are significantly different, $p < .001$, except the difference in depth between first person agents and other agents ($p = .06$), and the difference between first person agent objectivity and first-person patient objectivity ($p = .117$). Objectivity scores are very close to 1; the predicates on the posts are fairly mundane. The general pattern of the means supports the idea that more specific verbs are used for greater agency (1st person pronouns are the most agentive). The 2nd and 3rd patient category does need to be interpreted with care though, as Facebook posts are often in the third person even when referring to oneself - this seems to be a property of the medium and not a reflection of distancing by the individual. The effects of this will need to be separated with further research.

The eight features shown in Table were analyzed using a linear regression model which contained each of the features plus the interaction terms between the objectivity and specificity for each of the dimensions (first-person agent, non-first person agent, first-person patient, non-first person patient). The linear model showed a significant fit to the data, $F(12, 149) = 2.18$, $p = .02$. The R^2 is .14, the adjusted R^2 is .08, which corresponds to a correlation of .28. Utilizing a similar analysis with a 10-fold cross-validation approach showed an average raw correlation of .26 to the held out test set and a root mean squared error (RMSE) of .74. These correlations bracket the .27 correlation obtained by human experts for conscientiousness (Gosling, Gaddis, and Vazire 2007). A plot of the predictions vs. the actual values is shown in Figure 1. The linear model showed a significant effect for the interaction term between verbal specificity and objectivity and its predictive accuracy on conscientiousness.

In terms of raw accuracy, a ten-fold cross-validation approach and predicting whether or not the individual's level of consciousness was above or below the median (3.5) obtained an accuracy of 58.13%. The data set was skewed slightly towards above (51.2%).

Discussion

The analysis above provides an initial test of the idea that the different ways in which individuals express events involving themselves and others can be used to reveal their personality. This approach relied on identifying the specificity and objectivity of verbs individual's use to describe an event that they participated in. There still exists considerable room for improvement and further analysis. For example, we need better detection of multiword event-based predicates. We also need to more fully investigate the interaction found

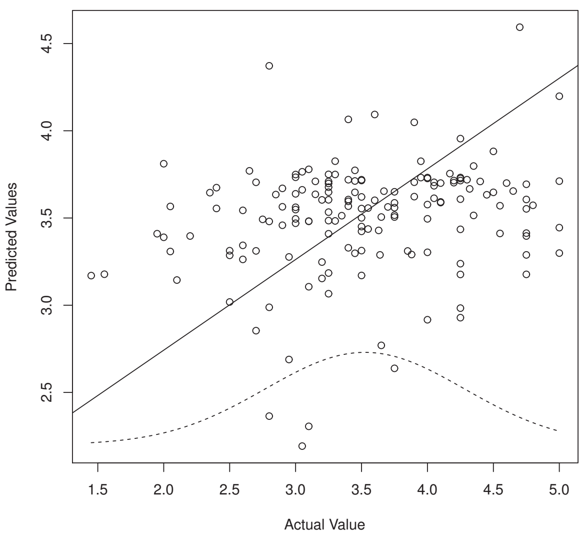

Figure 1: Plot of predictions vs. actual values for individuals and the resultant regression line. The dotted line shows the normal distribution of the actual conscientiousness scores.

within the model between specificity, objectivity, and conscientiousness and its exact bearing on goal orientation and perceived control.

One issue we encountered in our analysis is that conscientiousness (and most psychologically interesting variables) is normally distributed (see Figure 1), thus most individuals have a rating around the mean with relatively few outliers. This creates a problem with regards to framing a binary classification task, trying to detect if an individual is above or below the mean is always going to be a very difficult problem, because most people are very close to the mean. There is even good reason to say that separating individuals close to the mean will lower later prediction accuracy because these individuals should act the same. An approach that grades accuracy based on the detection of outliers, for instance those that are more than 1 standard deviation above or below the mean, might allow for higher accuracies and better downstream predictions. For reference, our approach achieved an accuracy of 68.03% when only classifying those individuals one standard deviation or more from the median.

Acknowledgments

This research was funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), and through the U.S. Army Research Lab. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI or the U.S. Government.

References

Baccianella, S.; Esuli, A.; and Sebastiani, F. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the 7th conference on International Language Resources and Evaluation (LREC10)*.

Beavers, J., and Koontz-Garboden, A. 2012. Manner and Result in the Roots of Verbal Meaning. *Linguistic Inquiry* 43(3):331–369.

Celli, F.; Pianesi, F.; Stillwell, D.; and Kosinski, M. 2013. Workshop on Computational Personality Recognition (Shared Task). In *WCPR13, in Conjunction with ICWSM13*.

Chang, L.; Connelly, B. S.; and Geeza, A. A. 2012. Separating method factors and higher order traits of the Big Five: A meta-analytic multitrait–multimethod approach. *Journal of personality and social psychology* 102(2).

Dowty, D. 1991. Thematic Proto-Roles and Argument Selection. *Linguistic Society of America* 67(3):547–619.

Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database.* Cambridge, MA: MIT Press.

Gosling, S.; Gaddis, S.; and Vazire, S. 2007. Personality impressions based on Facebook profiles. In *ICWSM*, 1–4.

Knobe, J. 2003. Intentional action and side effects in ordinary language. *Analysis* 63(3):190–194.

Mairesse, F.; Walker, M.; Mehl, M.; and Moore, R. 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research* 30(1):457–500.

McCrae, R. R., and Costa, P. T. 1999. The Five-Factor Theory of Personality. In *Handbook of personality: Theory and research*. 159–181.

Mehl, M. R.; Gosling, S. D.; and Pennebaker, J. W. 2006. Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *Journal of personality and social psychology* 90(5):862–77.

Miller, G. A., and Johnson-Laird, P. N. 1976. *Language and perception*. Cambridge university press Cambridge.

Palmer, M.; Gildea, D.; and Kingsbury, P. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics* 31(1):71–106.

Pennebaker, J. W.; Chung, C. K.; Ireland, M.; Gonzales, A.; and Booth, R. J. 2007. The Development and Psychometric Properties of LIWC2007 The University of Texas at Austin. *LIWC2007 Manual* 1–22.

Quercia, D.; Kosinski, M.; Stillwell, D.; and Crowcroft, J. 2011. Our Twitter Profiles, Our Selves: Predicting Personality with Twitter. *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing* 180–185.

Richens, T. 2008. Anomalies in the WordNet Verb Hierarchy. In *Proceedings of the 22nd International Conference on Computational Linguistics*, number August, 729–736.

Roberts, B. W.; Kuncel, N. R.; Shiner, R.; Caspi, A.; and Goldberg, L. R. 2007. The Power of Personality. *Perspectives on Psychological Science* 2(4).