

Show me how you write on social media and I'll tell you your age and sex

November 23 2015

Researchers at the Universitat Politècnica de Valencia (Polytechnic University of Valencia, UPV) have developed a new tool that can detect the sex and age range of the authors behind posts and other comments on social networks. Potential applications include its use in delinquent profiling and detection of pedophile cases. It is also a valuable tool for companies, offering a window onto their customer base and informing more focused marketing actions.

"Information about the age and sex of [social media](#) users is not always known or explicitly-stated. And even when it is, it might not always be true. Our tool decodes this information through the application of computational linguistic analytic techniques," explains Paolo Rosso, a researcher at the UPV's Pattern Recognition and Human Language Technology research group.

How does it work?

The tool developed at the UPV, together with Autoritas Consulting, applies [graph theory](#) to analyse the language used by social media users. It analyses verb tenses, the most repeated grammatical categories, discourse structure, type of expressions used and the affective content. From this data, it has proven possible to identify whether the person behind an anonymous text is male or female, and whether they are a teenager, a young person or an adult.

"We take a text and extract the grammatical categories to construct an initial graph. This graph is then enriched with information about the emotions expressed, the polarity of the words, the types of verb and types of noun used. We then apply graph theory to calculate the weight or importance of each element within the overall discourse structure. For each new case, we use machine learning algorithms to extract the graph and make a prediction", explains Francisco Rangel, CTO at Autoritas Consulting.

Their [tool](#) has already been used in police investigations into bomb threats. "In these cases, monitoring related accounts can be useful, not only to see what individuals are talking about, but also to profile their authors. The tools are also able to detect false profiles", the authors conclude.

The work was published last June in the Information Processing & Management journal. In it they approach the problem of gender and age identification using style-based and emotion-labelled [graph](#) features. Their study is carried out on Spanish social media posts, though the techniques can be applied to other languages.

More information: Francisco Rangel et al. On the impact of emotions on author profiling, *Information Processing & Management* (2015). [DOI: 10.1016/j.ipm.2015.06.003](https://doi.org/10.1016/j.ipm.2015.06.003)

Provided by Asociacion RUVID

Citation: Show me how you write on social media and I'll tell you your age and sex (2015, November 23) retrieved 23 May 2024 from <https://phys.org/news/2015-11-social-media-ill-age-sex.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.