

# Learning machines scour Twitter in service of bullying research

August 1 2012

---

Hundreds of millions of daily posts on the social networking service Twitter are providing a new window into bullying — a tough nut to crack for researchers.

"Kids are pretty savvy about keeping [bullying](#) outside of adult supervision, and bullying victims are very reluctant to tell adults about it happening to them for a host of reasons," says Amy Bellmore, a University of Wisconsin–Madison educational psychology professor. "They don't want to look like a tattletale, or they think an adult might not do anything about it."

Yet typical bullying research methods rely on the kids — victims and bullies alike — to describe their experiences in self-reporting surveys.

"For a standard study we may get access to students from one grade in one school," Bellmore says. "And then we get a one-time shot at it. We get one data collection point in a school year from these kids. It's very labor- and time-intensive."

Time and labor doesn't mean much to a computer, though, and Bellmore and graduate students Junming Sui and Kwang-Sung Jun have been helping Jerry Zhu, a UW–Madison computer sciences professor who studies [machine learning](#), teach computers to scour the endless feed of posts on [Twitter](#) for mentions of bullying events.

"What we found, very importantly, was that quite often the victim and

the bully and even bystanders talk about a real-world bullying incident on social media," Zhu says. "The computers are seeing the aftermath, the discussion of a real-world bullying episode."

Zhu fed the Twitter-monitoring computer two sets of tweets hand-selected by Bellmore's research group.

"The computer gets a set about bullying and a set definitely not about bullying," Zhu says. "In machine learning, the algorithm reads each tweet as a short text document, and it goes about analyzing the word usage to find the important words that mark bullying events."

Sufficiently trained, the computer went to work on samples of the 250 million publicly visible messages posted on Twitter on a daily basis. It wasn't long before the machine learning approach was identifying more than 15,000 bullying-related tweets per day. The traffic ebbs and flows on a weekly schedule — more active Monday through Thursday, presumably because the school-aged subjects see less of each other on the weekends.

Volume isn't the machine learning computer's only advantage over traditional research methods. As Bellmore and Zhu stepped up its training, the computer developed an eye for the roles played by the Twitter users wrapped up in bullying events.

"We taught it ways to identify bullies, victims, accusers and defenders," Bellmore says.

As the researchers dug into the tweets selected by the computer, they identified a new role: the reporter.

"The other roles were identified in the early '90s in the bullying literature," Bellmore says. "But the reporter role is new. It's just like it

sounds, a child who witnessed or found out about, but wasn't participating in, a bullying encounter. That role emerged out of studying the social media roles."

Data from social media has also thrown in the progression of time, a variable often left beyond the reach of student surveys. Bellmore and Zhu hope to follow groups of individual users through multiple bullying experiences.

"Paper surveys are not as dynamic as the social media tracks," Zhu says. "You just get one snapshot in time. You don't see the evolution of bullying events. You don't see the relationships evolving."

While the researchers are collecting data largely disconnected from the individual users, the machine learning technique could be used to identify children in need of an intervention.

"We want to add sentiment analysis, an assessment of the emotion behind a social media message, to our program," Zhu says. "The idea is that if someone is powerfully affected by the event, if they are feeling extreme anger or sadness, that's when they could be a danger to themselves or others. Those are the ones that would need immediate attention."

Using the data to show the bullied that they are not alone — the researchers have considered mapping social media mentions of bullying events — could also help children deal with their feelings.

"A way victims often make sense of their bullying is by internalizing it. They decide that there's something bad about themselves — not that these other people are jerks," Bellmore says. "When they're exposed to the idea that other people are bullied, actually it has some benefit. It doesn't completely eliminate the depression or humiliation or

embarrassment they might be feeling, but it can decrease it."

New insights will help the researchers supply policy-makers with a better understanding of bullying issues, Bellmore and Zhu say, which may result in more effective prevention methods.

Future work may include tapping other [social media](#) sites. For example, China's Weibo service claims 300 million users — maybe twice Twitter's count. Other services, like Facebook, may be even richer data sources.

**More information:** The group's work was presented at the North American Chapter of the Association for Computational Linguistics conference this summer, and is on the agenda at a Beijing sentiment analysis workshop in August.

Provided by University of Wisconsin-Madison

Citation: Learning machines scour Twitter in service of bullying research (2012, August 1)  
retrieved 9 May 2024 from <https://phys.org/news/2012-08-machines-scour-twitter-bullying.html>

<p>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.</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------