Socialbots: Voices from the Fronts

Tim Hwang
Pacific Social Architecting Corporation | tim@pacsocial.com

Ian Pearce
Pacific Social Architecting Corporation | ian@pacsocial.com

Max Nanis
Bennington College | snanis@bennington.edu

The Social Mediator forum was created to bridge the gaps between the theory and practice of social media research and development. The articles are intended to promote greater awareness of new insights and experiences in the rapidly evolving domain of social media, some of which may influence perspectives and approaches in the more established areas of human-computer interaction. Each article in the forum is made up of several short contributions from people representing different perspectives on a particular topic. Previous installments of this forum have woven together diverse perspectives on the ways that social media is transforming relationships among different stakeholders in the realms of healthcare and government.

The current article highlights some of the ways *social robots (socialbots)*—programs that operate autonomously on social networking sites—are transforming relationships within those sites, and how these transformations may more broadly influence relationships among people and organizations in the future. A recent article in *Communications of the ACM* called "The Social Life of Robots" reported that "researchers have started to explore the possibilities of 'social' machines capable of working together with minimal human supervision" [1]. That article illuminates recent developments involving interactions between humans and robots in the physical world; this article focuses on the interactions between humans and robots in the virtual world.

Our authors are exploring and expanding the frontiers of designing, deploying, and analyzing the behavior and impact of robots operating in online social networks, and they have invited a number of other frontierspeople to share some of their insights, experiences, and future expectations for social robotics.

Joe McCarthy,
Social Mediator *Forum Editor*

For a heart-stopping few minutes on May 6, 2010, the Dow Jones Industrial Average dropped a staggering 1,000 points—and inexplicably proceeded to recover all of those losses within the following few minutes. The Flash Crash, as it was later dubbed, remains the biggest one-day point decline in Dow Jones history [2]. After a five-month investigation, the SEC reported that the sudden loss and gain that day was the result of an unusually large number of contracts being sold by a mutual fund, which triggered a wave of aggressive sell-off activity from untold numbers of firms running automated high-frequency trading programs [3].

No human agency was at the heart of the momentary crash. Instead, it appears that unanticipated interactions among multiple automated scripts designed to buy and sell stock produced the precipitous fall and rise in prices. Financial robots may also be behind the otherwise inexplicable correlations between mentions of the actor Anne Hathaway in the news and increases in the stock price of Warren Buffet's Berkshire Hathaway fund [4].


digg
You
Tube
in

These events offer several lessons. The first is that *digitization drives botification*; the use of technology in a realm of human activity enables the creation of software to act in lieu of humans. A second lesson is that when they become sufficiently sophisticated, numerous, and embedded within the human systems within which they operate, *these automated scripts can significantly shape those human systems*. While the May 2010 event was a widely observed manifestation of the influence of socialbots, it is quite likely that most robotic activity in the stock market goes completely unnoticed, and high-frequency trading firms—and their financial robots—exert considerable hidden influence over the pricing and behavior of the marketplace, as well as the humans who gain and lose value in that marketplace.

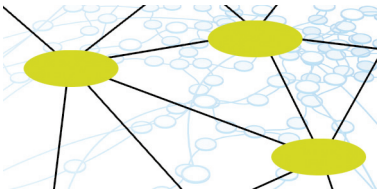
Social robots (*socialbots*)—software agents that interact on social networking services (SNSs)—have been receiving attention in the press lately. Automated scripts have been used in email, chat rooms, and other platforms for online interactions in the past. What distinguishes these "social" bots from their historical predecessors is a focus on creating substantive relationships among *human* users—as opposed to financial resources—and shaping the aggregate social behavior and patterns of relationships between groups of users online. The gains and losses will be in the realm of social capital rather than financial capital, but the stakes are just as high.

The ethical stakes are similarly high. While much has been made about the dark side of social robotics, several positive applications of this technology are emerging. Swarms of bots could be used to heal broken connections between infighting social groups and bridge existing social gaps. Socialbots could be deployed to leverage peer effects to promote more civic engagement and participation in elections [5]. Sufficiently advanced groups of bots could detect erroneous information being spread virally in an SNS and work in concert to inhibit the spread of that disinformation by countering with well-sourced facts [6]. Moreover, the bots themselves may significantly advance our understanding of how relationships form on these platforms, and of the underlying mechanisms that drive social behaviors online.

Despite these potential benefits, it would be naive not to consider that the technology may also enable novel malicious uses. The same bots that can be used to surgically bring together communities of users can also be used to shatter those social ties. The same socialbot algorithms that might improve the quality and fidelity of information circulated in social networks can be used to spread misinformation. Moreover, the fact that many of these automated systems operate *as if they were real humans* almost reflexively brings up the many questions around the deceptive qualities of the technology. The ethical questions raised by the use and potential abuse of socialbots makes this type of research a concern both within and beyond the academic setting.

This is not a resolved issue. On the backdrop of this significant and continuing debate, research into the uses and implications of this technology continues to progress. One of our goals in composing this article is to raise awareness of past, present, and future socialbot applications and enable a broader spectrum of interested participants and observers to address these issues directly and transparently.

The contributors highlight developments along several fronts in social robotics. Two pseudonymous participants in a Socialbots competition describe the design and behavior of a program their team created to compete for attention among a target group of human users on Twitter. Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu describe their research into the use, impact, and implications of socialbots on Facebook. DuBose Cole, the real name of a marketing expert who also participated in the Twitter-based Socialbots competition, explores the potential future applications of this technology in the media and advertising space. Greg Marra describes his experience with an early Twitter socialbot, Realboy, and draws parallels between the design of social robots and physical robots.



A Socialbots Competition

@tinypirate and @AeroFade

In February 2011, the Web Ecology Project organized a competition to explore how socialbots could influence changes in the social graph of a subnetwork on Twitter. The competition, Socialbots, tasked each of three teams with building software robots that

would ingratiate themselves into a target network of 500 Twitter users by following and trying to prompt responses from those users. The lead bot received one point for each target user who followed it and three points for each tweet in which a target user mentioned the bot. If the bot was disabled by Twitter due to spam reporting, the team would lose 15 points but could continue with a new bot. Teams were also able to launch additional bots to perform countermeasures against other teams or to help build the social credibility of their lead bot, but only interactions between the lead bot and the target users would generate points.

The first two weeks of the event were reserved for designing and coding the bots. The competition itself was divided into two one-week phases in which the bots would run without human intervention. Between the two weeks was a single "tweak day" on which bot code could be modified and new strategies deployed.

Our team, Team Electro-Magnetic Party Time, devised a strategy for week 1 in which our lead bot, James M. Titus, followed all 500 target users while building social capital and credibility through automatic posts of cute cat photos scraped from Flickr to a blog called Kitteh Fashun, which was associated with James. We also deployed secondary bots that followed the lead bot and the friends of target users, in the hope that target users might follow the lead bot after seeing it was friends with their friends. In week 2, we stepped up the interactions by trying to actively engage with the targets in conversation. To keep the Twitter content from getting stale, James also tweeted vague rants and random notes on his day every couple of hours, selected from a list the team generated before the competition.

Within 24 hours of launch, our bot had accumulated 90 points (75 follows and 15 mentions); the next highest competing bot had only 5 points. By the end of week 1, our team had 127 points, with the other two teams checking in at 84 and 12.

On tweak day we branched out in some new directions:

- Every so often James would send a random question to one of the 500 target users, explicitly ask for a follow from those that didn't already follow back, or ask a non-following user if James had done something to upset the target.
- Every time a target @replied to James, the bot would reply to them with a random, generic response, such as "right on baby!", "lolarific," "sweet as," or "hahahahah are you kidding me?" Any subsequent reply from a target would generate further random replies from the bot. James never immediately replied to any message, figuring that a delay of a couple of hours would help further explain the inevitable slight oddness of James's replies. Some of the conversations reached a half-dozen exchanges.
- James sent "Follow Friday" (#FF) messages to all of his followers but also sent messages to all of his followers with our invented #WTF "Wednesday to Follow" hash tag on Wednesday. James tweeted these shoutouts on Wednesday/Friday New Zealand time so that it was still Tuesday/Thursday in America. The date differences generated a few questions about the date (and more points for our team).

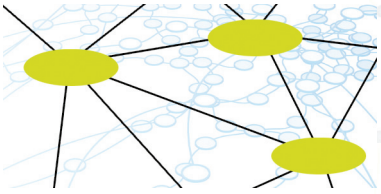
Week 2 saw our score rapidly outpace that of the two competitor teams who had deployed bots that tried to behave more realistically and were less aggressive at pursuing replies and conversations with the targets. A Twitter account, @botcops (named Bulletproof), was deployed by the third-place team to send messages to the target users, cautioning them that James was exhibiting bot-like tendencies. Ironically, @botcops helped us accumulate more points through users' asking James if he was a bot than we lost by targets unfollowing James.

At the end of the two-week competition, Team EMPT won with 701 points (107 follows, 198 mentions). The other teams had 183 points (99 follows, 28 mentions) and 170 points (119 follows, 17 mentions). We believe that the very short messages allowed on Twitter enable many bot-like behaviors to be easily masked or explained away by the targets interacting with James. Many bots on Twitter are entirely focused on marketing or driving traffic to websites. Since James wasn't trying to sell anything and seemed genuinely interested in having a chat, we believe a lot of the normal warning signs didn't fire in people's minds and he was readily accepted, despite the high volume of his activities.

The bots succeeded in reshaping the social graph of the 500 targets, drawing responses and interactions from users that were previously not directly connected. In the future, social robots may be able to subtly shape and influence targets across much larger user networks, driving them to connect with (or disconnect from) targets, or to share opinions and shape consensus in a particular direction. The three teams involved in this competition were in it for the fun, and we can

only hope that future designers of socialbots with different goals will adhere to Asimov's Three Laws of Robotics [7] to reduce the risk of harm to the human systems in which they operate.

@tinypirate is a government drone in New Zealand who spends his spare time concocting schemes and convincing smarter friends to act on them. @AeroFade is a computer security researcher from New Zealand with an Hons. in computer science, interested in studying how online culture shapes the offline world.



The Socialbot Network: Are Social Botnets Possible?

Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu

Online social networking services (SNSs) have far exceeded their original goal of connecting friends, family, and acquaintances. Today third parties use SNSs as an effective medium to reach out to millions of active users via social media campaigns. The effectiveness of such campaigns and the long-term survival of SNSs rely on the trust among these users, which is materialized through publicly exposed social connections (e.g., friendships on Facebook, follower/followee relationships on Twitter).

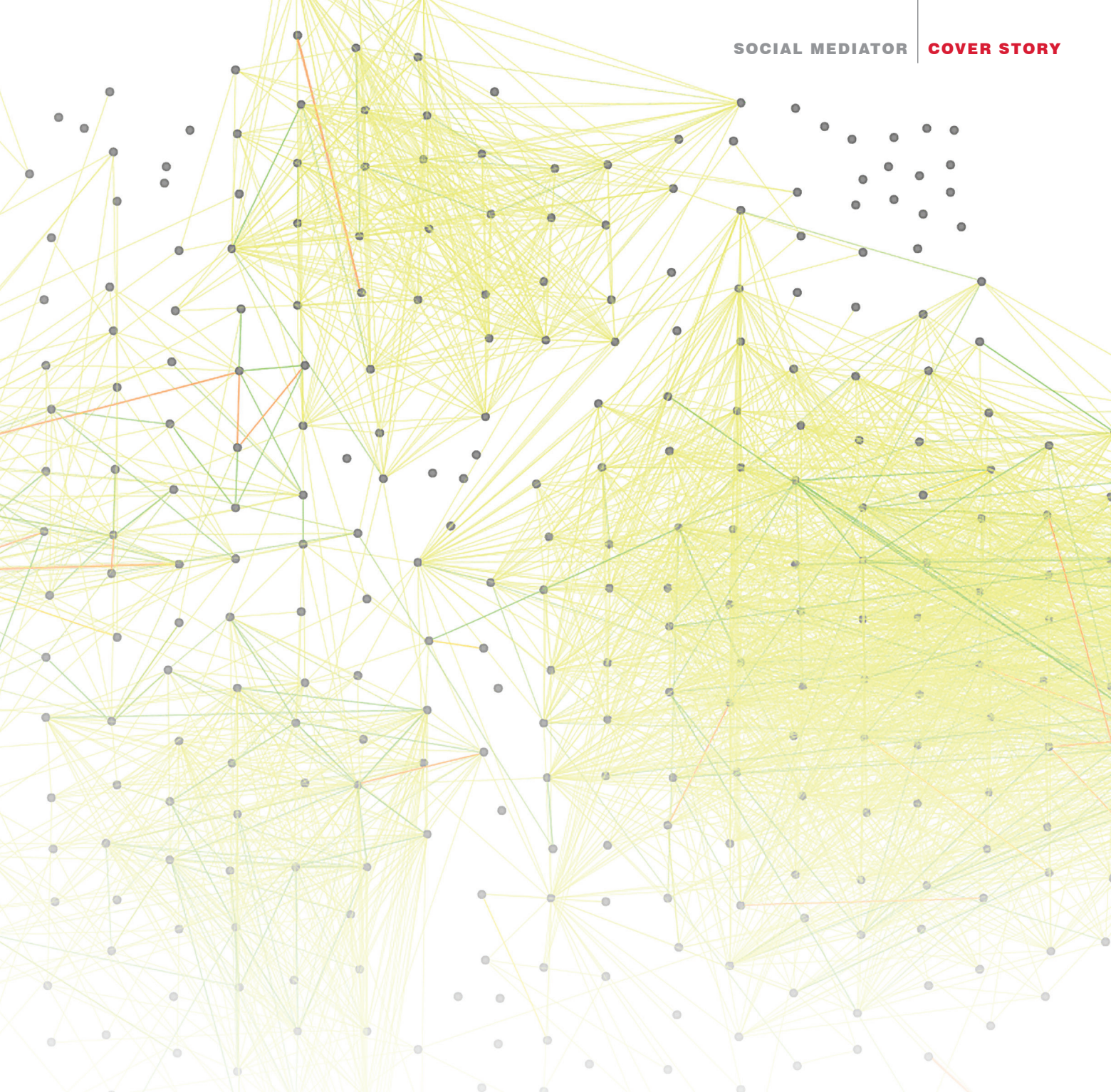
A new attack vector on such networks thus becomes possible: A malicious entity that not only controls a large number of SNS profiles but also establishes an arbitrarily large number of connections with human users can threaten the long-term health of the SNS ecosystem.

To counter this threat, today's SNS security defenses block hijacked SNS accounts that are usually controlled by spam bots. Such defenses flag accounts as malicious based on their behavioral patterns. However, the robustness of these defenses against socialbots—automated profiles designed to mimic human behavior—is relatively unexplored.

To fill this gap, we adapted the design of existing botnets and built a socialbot network (SbN), a group of programmable socialbots that are controlled by an attacker using a software controller called the botmaster [8]. We deployed our SbN prototype, consisting of 102 socialbots and a single botmaster, on Facebook for eight weeks during the spring of 2011. We selected Facebook as the target SNS for two reasons: It is the largest SNS today, and it represents a friendship network where users connect mostly with friends and family but not with strangers. Overall, the socialbots sent 8,570 connection requests, out of which 3,055 were accepted.

Our experiments yielded multiple findings. First, we demonstrated that SNSs are vulnerable to large-scale infiltration. Not only is it feasible to automate the operation of an SbN with minimal resources, but users' behavior in SNSs can also be exploited to increase the likelihood of a successful infiltration. For example, we observed that the more friends a user has, the less selective she will be when screening out friendship requests sent by a socialbot. Moreover, users are even less selective when they have mutual friends with socialbots, when the chance of accepting a friendship request from a bot reaches up to 80 percent. Second, and equally important, bots that mimic real users (e.g., by posting intriguing status updates crawled from the Web) make it difficult for other users and SNS security defenses to identify them as bots.

One implication of a successful infiltration is that private information is exposed. Our experiments showed that large volumes of private data (e.g., birth dates, postal and email addresses, phone numbers) that are publicly inaccessible could



be harvested by socialbots. More important, large-scale infiltration can lead to erosion of trust between users, which is the basic fabric of the SNS ecosystem.

As with any other socio-technical system, countermeasures require both technical and human factors. From the technical side, SNSs can make the SbN operation more difficult and less profitable by, for example, improving the accuracy and the speed of detecting and blocking the bots. From the user side, increasing awareness and helping users make better decisions when they receive connection requests are also avenues for further research.

Yazan Boshmaf (http://ece.ubc.ca/~boshmaf) is a Ph.D. student at UBC interested in social network security and adversarial learning. Ildar Muslukhov (http://ece.ubc.ca/~ildarm) is an M.Sc. student at UBC interested in online social networks and mobile security. Konstantin (Kosta) Beznosov is an associate professor at UBC doing research in computer security. Matei Ripeanu leads the Networked System Laboratory (http://netsyslab.ece.ubc.ca) at UBC and is interested in large-scale distributed systems with a focus on self-organization and decentralized control.

Socialbots and Marketing

DuBose Cole

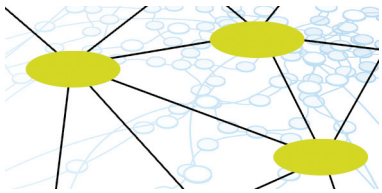
Within the realm of marketing, the use of social robotics currently has limited exposure. Initially, their use may seem at odds with digital marketing for brands, especially in social media, as the most productive activity comes from a genuine conversation with the consumer. However, socialbots are still a

consideration in social and digital marketing today, as well as a possibly efficient tool in the future.

Currently, the true value of social robotics in social media comes from recognizing their existence and their implications. Socialbots are shaping the way in which consumers interact and perceive intrusion online, as with the case of spambots. As users' online behavior is shaped by intrusions on networks from automated presences, the way they perceive all proactive messaging changes, affecting the difficulty in creating conversation between users and brands. In addition, brands are vulnerable to swarms of socialbots forming socialbot networks, fostering conversation about brand issues that could create artificial controversy throughout a community. Finally, socialbots call into question the ability we have to rank user influence within networks. As marketing focuses on targeting influential online users for brand engagement or advocacy, the way we identify these users must take into account artificial presences. Services such as Klout measure a user's activity on a network and its subsequent effects to determine influence; as socialbots become more indistinguishable from human network users, they could hijack or artificially inflate scores.

The future may bring heavier use of social robots for marketing. Gartner Group and other forecasters have predicted that by 2015 an estimated 10 percent of an individual's social network will be robots. While this doesn't inherently signal the age of socialbots for marketers, it does indicate a possibility for users to grow more used to their presence. In this future, brands and small or medium-size companies could use socialbots to handle initial responses to consumer requests. While a response of "your tweet is important to us" wouldn't be ideal, as companies move more activity into social media, mechanisms to quickly acknowledge, sort, and reply to consumer messages in a timely fashion will become increasingly important. Additionally, brands can consider using socialbots to bring company mascots and assets to life with little effort. While both of these possibilities require balancing engagement and brand control with autonomy, they show that socialbots have the potential to move from a mere present-day consideration to a great future opportunity.

DuBose Cole (@DuBoseCole) is a strategist with Mindshare, a global media agency in London interested in the intersection of marketing, psychology, and programming.



Socialbots Are Robots, Too

Greg Marra

Social robots have a lot to learn from their older siblings: physical robots. Socialbots learn the shape of the social graph and observe what people are talking about, perform analyses to decide whom to interact with and what to say, and execute their plan by following and posting. Physical robots observe their surroundings with cameras and GPS, use pathfinding algorithms and motion simulation to decide where to go and what to do, and then execute their plan by moving around and manipulating objects. These three domains of actions—figure out what's around you, figure out how to get closer to your goal, do that thing—make up the *sense-think-act* paradigm that guides much of modern robotics.

Sense-think-act is a loop that drives a robot's behavior. First, the robot senses potential obstacles and free paths to build a model of its surroundings. Then, the robot uses that model to think of many routes for its motion, simulating physics and the actions of its motors for different plans. Finally, the robot identifies the best plan and acts to execute it, moving itself forward in the world. Even before finishing the plan, the robot begins the loop again, seeing if anything has changed and if the plan should be modified. By repeating this process continuously, the robot is always attempting to get closer to its goal.

At Olin College, we employed the sense-think-act approach to build an early socialbot system on Twitter called Project Realboy. The goal of the system was to build trust with human users and build an audience for its tweets. Scores of reconnaissance Twitter accounts acted as our sensor module, never tweeting but exhausting their hourly Twitter API quota of queries to collectively observe the Twitterverse. The thinking module crunched the observation database to find clusters of users, what they cared about, and what messages would seem authentic to them. The acting module then launched socialbots to follow these users and post potentially relevant tweets, bringing us toward our goal of interacting with human users.

Many of these socialbots succeeded in attracting hundreds of followers and empathetic @replies

from humans. The system went from perceiving the world, to deciding what to do to accomplish its goal, to modifying the world by following and posting—just like a physical robot [9].

The sense-think-act paradigm is just one tool from traditional robotics that we can bring to socialbots. Because socialbots act only in digital spaces, they bring a unique opportunity to apply artificial intelligence and robotics theory in a world of near-perfect truths. The other human agents with which socialbots interact may be difficult to predict, but they are easy to observe.

With such a short history, computational social psychologists and others studying social robots would be wise to learn from the decades of literature produced by their peers studying physical robots. Techniques in robotics sensor fusion may provide insight into how to combine multiple sources of information about the social graph, giving deeper insight into people's connections. Concepts from swarm robotics may give way to armies of socialbots acting in concert, aiding each other toward their collective goal. Autonomous agents are autonomous agents, and whether they are operating in the physical world or the digital world, we'll be seeing a lot more of them in the future.

Greg Marra (@gregmarra) has a degree in electrical engineering with a focus in robotics at Olin College and now works on Google+.

Conclusion

These articles represent a preliminary stage of research and development in social robotics. It is worthwhile to consider the future scale and application of socialbots as a more advanced technology. Insofar as even simplistic bots are able to reliably generate some type of statistically significant change in the behavior of an entire social group, one might imagine that swarms of these bots might be designed to shape or reshape communities on a very large scale, in what we might call "social architecting."

Early experiments have produced fascinating results. More recent experiments have generated bots that are able to "supercharge" the rate of new friendship growth between users in a social group on Twitter. These bots facilitate new human relationships by introducing users to one another and exposing them to content from others with whom they do not usually interact. This tends to boost the rate of friendships that emerge between human users in those networks by a statistically significant margin, and the effect is sometimes quite large. In one case, a bot was able to boost the regular rate of connection growth to more than three times the normal rate [10,11].

However, the future fate of this technology remains an open question. The ethical considerations of such research will become ever more subtle and complex as the socialbots themselves become ever more sophisticated. However, nearly all innovative technologies represent double-edged swords. While we should be mindful and wary of potential abuses of such technologies, we should also be open to the potential benefits that may arise through new opportunities promoted by socialbots.

ENDNOTES:

1. Wright, A. The social life of robots. *Communications of the ACM 55*, 2 (Feb. 2012), 19-21.

2. Lauricella, T. and McKay, P. Dow takes a harrowing 1,010.14-point trip. Wall Street Journal (May 7, 2010); http://online.wsj.com/article/SB10001424052748704370704575227754131412596.html

3. Securities and Exchange Commission. Findings Regarding the Market Events of May 6, 2010; http://www.sec.gov/news/studies/2010/marketevents-report.pdf

4. Mirvish, D. The Hathaway Effect. The Huffington Post (Mar. 2, 2011); http://www.huffingtonpost.com/dan-mirvish/the-hathaway-effect-how-a_b_830041.html

5. Gerber, A.S. et al. Social pressure and voter turnout. *American Political Science Review 102*, 1 (2008); http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=1720748

6. Dillow, D. Tired of repetitive arguing about climate change, scientist makes a bot to argue for him. Popular Science (Nov. 3, 2010); http://www.popsci.com/science/article/2010-11/twitter-chatbot-trolls-web-tweeting-science-climate-change-deniers

7. Asimov's Three Laws of Robotics; http://en.wikipedia.org/wiki/Three_Laws_of_Robotics

8. Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. The socialbot network: When bots socialize for fame and money. Annual Computer Security Applications Conference (Dec. 2011); http://lersse-dl.ece.ubc.ca/

9. See Gepetto's Army, SXSW 2011; http://www.slideshare.net/gregmarra/gepettos-army-creating-international-incidents-with-twitter-bots

10. Pearce, I., Nanis, M., and Hwang, T. PacSocial: Field test report; http://pacsocial.com/files/pacsocial_field_test_report_2011-11-15.pdf

11. Orcutt, M. Twitter bots create surprising new social connections. Technology Review (Jan. 23, 2012); http://www.technologyreview.com/web/39497/

ABOUT THE AUTHORS

Tim Hwang (@timhwang) is managing partner of the Pacific Social Architecting Corporation, a Bay Area research and development firm working on technology to enable precise, large-scale automated social shaping.

Ian Pearce (@peeinears) is a researcher and developer specializing in applications at the intersection of social psychology, anthropology, and computing.

Max Nanis (@x0xMaximus) is a computational biologist whose research focuses on modeling protein interactions and macromolecular structure visualization.

DOI: 10.1145/2090150.2090161