2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing

Dirty Deeds Done Dirt Cheap

A Darker Side to Crowdsourcing

Christopher G. Harris
Informatics Program
The University of Iowa
Iowa City, IA 52242 USA
christopher-harris@uiowa.edu

***Abstract*— Several recent studies have examined the merits of crowdsourcing to aid in completing repetitive or complex tasks requiring human computation. In comparison, scant attention has been placed on the use of crowdsourcing for the purpose of meeting unethical objectives, which may or may not be known to the participants. In this paper, we explore the potential for which crowdsourcing may be used to bypass commonly-established ethical standards for personal or professional gain.**

Keywords*— *Crowdsourcing, Anonymity, Computer Ethics, Social Issues

I. INTRODUCTION

The merits of issuing an open call for participants to accomplish one or more prescribed tasks quickly, cheaply, and efficiently have been the subject of several studies, such as those by Snow [1] and Callison-Burch [2]. Some of these tasks are highly-repetitive in nature, such as relevance judgments or annotations of images, while others have focused on collaborative efforts, disintermediation of skills or knowledge, or assistance in obtaining diverse information. The overwhelmingly popular view is that crowdsourcing is on the positive end of the spectrum, and has benefitted many by more effectively allocating labor and resources, as pointed out by Howe in his frequently-mentioned book on crowdsourcing [3].

In comparison, there are relatively few criticisms of crowdsourcing to date. Some critics remark on the low pay a crowdsourcing participant may receive for a task, such as the discussion by Silberman in [4], but these criticisms rarely go beyond lamenting about the economic realities of a flexible global labor market. Other critics have remarked on issues of poor worker quality or the use of bots to perform tasks in an attempt to receive payment for vacuous submissions; however, most companies involved in crowdsourcing use complex rules or engage in well-known techniques to address these issues.

At the negative end of the spectrum, cybercriminal clearinghouses thrive where bulk sales of ill-gotten personal information are readily available for purchase. Many websites and forums offer passwords and cracks for popular software applications, clearly in violation of software piracy laws. Although an argument can be made that these transactions are an extreme form of crowdsourcing, such blatantly illegal activities are not the focus of our discussion here.

In this paper, we are interested in exploring the middle region of this spectrum, primarily tasks that are not illegal, but are regarded as unethical to many. Our purpose here is not to define what is ethical and what is not – that discussion has endured at least since the time of Socrates – though the examples used here illustrate violations of ethical behavior that are understandable to most readers. For instance, a manufacturer posting false positive reviews on their products or a jilted lover asking others to pose as potential (but phony) suitors to the former lover on a dating website in an effort to seek revenge may not be illegal, but many will find this action unethical. Likewise, providing expert assistance on student take-home exams, flowery testimonials for a poorly-constructed product or service, or surveillance on a stranger are not new phenomena – ploys such as these have existed for centuries. However, technology and the ease of sending micro-payments to areas of the world where netizens may be more willing to accept these tasks, have made these behaviors far easier to request and to fulfill than ever before.

The objective of this paper is to initiate a discussion of the role of ethics in crowdsourcing and is structured as follows. In the next section, we explore some of the challenges to ethical behavior in crowdsourcing activities. In Section III, we look at a few of the common techniques that could be applied in unethical tasks. Section IV discusses some examples of unethical crowdsourcing. Section V briefly examines steps that can be taken to possibly combat unethical behavior. We conclude our discussion in Section VI.

II. CHALLENGES IN POLICING ETHICAL BEHAVIOR ON THE INTERNET

Effective policing of computer ethics on the internet is difficult, particularly because nobody owns the internet. In crowdsourcing, one challenge in avoiding ethical issues is tied to a well-known strength – the diversity of crowd demographics. A number of classic studies on ethics have demonstrated that there is a substantial variety of what is regarded as unethical behavior between different demographic groups; for instance, a reprehensible deed conducted within one group may be considered perfectly ethical in another.

Other studies have concluded that the anonymity of the internet only emboldens behaviors that would not likely occur face-to-face, examined in detail in a 2004 study examining Generation Y behaviors in e-ethics by Freestone [5] and a 2007 computer ethics study by Johnson [6]. The chances of finding participants to perform an unethical task increases when

978-0-7695-4578-3/11 $26.00 © 2011 IEEE
DOI

everyone in the crowd, regardless of distance or demographic, is only a pseudo-anonymous internet connection away.

In addition, there are different motivating factors for crowd participants. Several studies have examined what motivates the crowd to work on low-paying tasks, e.g. [7, 8]; for example, the compensation for Wikipedia article edits and fact-checking, is ultimately personal satisfaction [9, 10]. However, for potential participants in regions where there may be few other employment opportunities, these tasks may represent a means to support an entire family [11]. With different motivations come different levels of acceptable behaviors; some may view every available task as an opportunity to earn additional compensation whereas others may see their reputation as a paramount consideration for future tasks and be far more reluctant to engage in potentially-dishonorable tasks.

A final challenge is that many task participants may be unaware that the task they have agreed to participate in is unethical. This may be related to how crowdsourcing can divide a task into several autonomous components – asking for help on a single question of a 100-question homework assignment may be considered ethically acceptable for some, but if 100 experts were each tasked a single question, it changes the task perspective from simply "helping out" to "blatant dishonesty".

III. Basic Techniques

In this section, we examine three techniques used by crowdsourcing task providers to address unethical objectives, either knowingly or unknowingly.

A. Social Engineering

Social engineering techniques are used to manipulate people into performing specific actions or divulging confidential information to be used for illicit financial gains, most commonly from identity theft. Social engineering for financial gain is explained in a classic study by Adams [12] and another more recent one by Thornburgh [13]. Given how frequently it appears in the news media, few people today would be surprised to learn that a global black-market exists for selling bulk personal data such as driver licenses, credit card numbers, and social security numbers. A common technique in social engineering is to find something in common with the potential victim, such as a mutual situation or demographic. Peer pressure can also be used to accomplish social engineering techniques – going "against the crowd" is inconsistent with the instincts of human nature. Using the crowd as artificial peers enables social engineering opportunities at a reduced cost.

B. Human Computation Tests

For malicious tasks requiring human computation, it is plausible that these could be accomplished using the crowd. CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart), which are screening devices for distinguishing virtual robots from humans on the web [14]: as a challenge, a user is required to decrypt a short sequence of distorted characters contained within an image in order to pursue a session. The premise is that computer programs, such as bots, cannot read distorted text as well as humans can.

CAPTCHA challenges remain a first line of defense against automated computer cracking techniques, although they are not undefeatable [15, 16]. Many of the password validation systems on free email services allow for a limited number of tries before a CAPTCHA challenge is presented. In recent years, a number of spammers have crowdsourced CAPTCHA-cracking tasks in exchange for access to pornography websites or at a rate as low as US$0.80 per 1000 CAPTCHAs cracked [17]. Likewise, enlisting the crowd to attempt attacks on passwords can be particularly advantageous when IP addresses are being tracked. This is particularly true when combined with social engineering techniques to limit the domain of potential passwords [18].

C. Attack and Run

Many forum and review websites fall along a continuum of identity requirements. *Identity-intensive* websites require and validate personal information before contributors can fully participate. This additional validation is done to ensure people are whom they claim. On the other end of this continuum are *identity-relaxed* websites, which maintain anonymity and thus require very little information upfront – the purpose is to attract a wider potential audience of readers and contributors. Many identity-intensive websites require a fee-based subscription for access, and use this as a screen of identity determination; on the other hand, free email services, Facebook, and many ratings websites fall on the relaxed end of the spectrum, which can be problematic as these are often used to validate user identities on other websites.

Identity-relaxed websites can become springboards for building false identities [19]. We call these "attack and run" techniques because they build a false sense of confidence based on weak or transient identities. If an audit was conducted on the reviews and an attempt was made to contact all contributors who raved about a particular product online, most should still be locatable; however, if after a short period of time few could be found, it might indicate a "attack and run" technique.

IV. Examples

In this section, we examine three examples illustrating the use of the techniques discussed in Section III to demonstrate how ethics can potentially be abused through crowdsourcing.

A. Review Manipulation

Websites that solicit user reviews are often identity-relaxed to encourage frequent contributions from a wide variety of reviewers. Consider TripAdvisor - the largest and most influential global travel review website, according to a 2008 examination by O'Connor [20]. A majority of the 50 million monthly TripAdvisor visitors, including 10 million unique visitors as of late March, 2011 [21], make the opinionated traveler reviews their first stop on planning a trip [22]. A positive review is so essential to the success of many travel-related businesses, such as hotels and restaurants, that submitting false reviews is a well-publicized concern. To combat this, TripAdvisor spends considerable effort to ferret out fake reviews, accomplished primarily by examining browser cookies, basic authentication information such as usernames, and the user's IP address. However, like many other websites that encourage unbiased user reviews,

TripAdvisor does little to validate the information contained within each review, according to a recent investigation on TripAdvisor's business model by Jurca [23]. Even if it chose to employ far more sophisticated methods, it would still be relatively easy for an outside company to crowdsource review submissions [24]. Indeed, in another publicized case, Belkin, a computer hardware manufacturer, was accused in 2009 of offering US$0.65 to crowdsourcing participants in exchange for writing positive reviews on their products on Amazon, among others, and marking other negative reviews as unhelpful [25]. These false reviews are generally made using the "attack and run" methods discussed in Section III.C. From a risk-return perspective, the cost to the hotel or restaurant is relatively little, the risk of detection is low, and the perceived value to a travel-related business is high. For review websites such as TripAdvisor, however, this could compromise their credibility – a core element of their business model [26].

B. Surveillance

The encouragement of participants from the crowd to engage in surveillance for compensation has been used for centuries, often as a technique in espionage or by private detectives. However, this technique is often used in business by "secret shoppers" and by quality inspectors to examine whether business standards are being met. As discussed in Section IV.A, when every customer is a potential reviewer of a business, there are more data points available to evaluate and the aggregate more closely represents the true experience of a customer. The use of crowdsourcing for surveillance is often beneficial to society: the use of the crowds to find missing and exploited children is a clear benefit; likewise, "Crimestopper" hotlines have been used by the police to obtain important tips on criminal activities for decades in many countries, and based on their increased use have been widely regarded as successful [27].

Using the crowd for surveillance can cross ethical boundaries: Consider a person who wishes to monitor the actions of his ex-spouse and is physically restricted from doing so, or a family member who wants to monitor their teenager's or parent's activities while away from home. He/she could hire a private detective to engage in surveillance at a considerable cost, or could solicit the crowd to observe the targeted person's activities and pay by the validated piece of information, making it harder for the person being monitored to identify who is following him or her. Recently Johnathan Zittrain illustrated how that it is possible to identify a single Iranian demonstrator from a photograph by crowdsourcing the comparison of the photograph with the ID cards of 72 million Iranians, four photos at a time [28]. He estimated the total cost would be about US$14,000 using a crowdsourcing platform like Mechanical Turk - but far less if this task was offered as a game to schoolchildren and played for no compensation. Even if the true motives were discovered by the crowd, it is still easy to find willing participants. Indeed, in societies where breaking social norms is lightly punished, a substantial subset of the populace will eagerly accept an opportunity to break social norms, especially when the right incentive is introduced [29, 30], providing a large pool of people to assist in these surveillance tasks.

C. Information Gathering

Closely related to surveillance is information gathering – the primary difference between the two is that information gathering is perceived as more invasive. Techniques such as password cracking and obtaining information through misrepresentation fall within the category of information gathering. In many jurisdictions, if the information gathered is not used for financial gain, it is not considered fraud. Often social engineering techniques discussed in Section III.A are utilized.

Consider a hypothetical "revenge"-oriented crowdsourcing platform, where the crowd is employed to conduct unethical information gathering tasks for various prices. Password cracking for a specific individual's email accounts might be listed for a certain fee (along with known information about that individual to make the task quicker and easier to perform); creating a profile on an online dating website in order to attract an ex-spouse into a false sense of commitment would involve a different level of activity and list for a different fee [31]. This platform, like most other crowdsourcing platforms, would seek to bring providers and requesters together at a market clearing price – except in this situation, they would be for potentially unethical tasks. One concern is that far more severe (and illegal) activities could be conducted based on these mild information gathering techniques.

Information gathering can be used to expose potential fraud: In June 2011, Guo MeiMei, reportedly the general manager of the Chinese Red Cross, was exposed living a lavish lifestyle far beyond what her salary could support – clues exposing potential fraud put together by hundreds of normal citizens collectively witnessing her lifestyle (such as photos of her dining receipts, and photos of her shopping exploits), combining their observations, and posting to Chinese social networking websites [32].

V. Combatting Unethical Behavior

The unethical behavior described here, even in those jurisdictions where it is considered against the law, is often difficult to eliminate or police for several reasons. First, the pseudo-anonymity of the crowd makes this a challenge [33]. Second, the crowd, especially those given a small task, may not be aware of the overall objective and thus unaware that they had participated in unethical behavior. Third, the targeted victim may not be aware that they were victimized or how the information about them was obtained.

However, there are a few known techniques that could be applied to reduce the number of incidents of unethical crowdsourcing behavior. First, laws could be enacted to make such acts illegal, but such laws would be difficult to enforce globally. A voluntary education effort on ethics would likely reduce the unethical tasks on crowdsourcing websites, but they will remain voluntary efforts without repercussions and would have little effect in parts of the globe where earning compensation trumps ethics. Additionally, reputation-based identification systems have shown an ability to discourage aberrant behavior while still preserving some degree of anonymity [34]. Other methods of uniquely identifying a computer that examine installed software, cookies, and usage

patterns are used in computer forensics [35]; these may eventually evolve into a more mainstream method of user identification while still preserving some level of anonymity. Finally, the "no honor among thieves" scenario presents an additional risk to the requester, particularly for surveillance tasks: although the requester remains anonymous online, it is easy to reverse engineer their identity; it might not be difficult to identify the limited pool of requesters desiring information on "Jane Doe" at "123 Maple Drive", which subjects each of these potential requesters to exploitation by others in the crowd wishing to expose their disreputable activities.

VI. Conclusion

Our intent is to initiate a discussion of ethics in crowdsourcing task design. We made a preliminary examination of the darker side of crowdsourcing, particularly the notion that crowdsourcing could be used for intentionally unethical objectives. We then examined several techniques and examples illustrating the role of crowdsourcing in these types of activities.

Like crowdsourcing, computer ethics is an evolving area of computer science, so there are several directions to extend the ideas introduced in this paper. One is to explore current ethical boundaries in crowdsourcing, and see how different crowdsourcing communities are able to police their own citizens. Another direction involves the introduction of some of the methods to combat unethical behavior introduced in Section V. A third future direction is to perform some experimental analysis of participant behavior when presented with unethical tasks.

VII. References

[1] Snow, R., O'Connor, B., Jurafsky, D. and Ng, A. Y. Cheap and fast---but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Honolulu, Hawaii, 2008). Association for Computational Linguistics, 2008.

[2] Callison-Burch, C. Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk. In *Proceedings of the 2009 Conference on Empirical Methods in NLP: Volume 1* (Singapore, 2009). Association for Computational Linguistics, 2009.

[3] Howe, J. Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business. Crown Publishing Group, New York, NY, 2008.

[4] Silberman, M., Irani, L. and Ross, J. Ethics and tactics of professional crowdwork. *XRDS: Crossroads, The ACM Magazine for Students*, 17:2 2010, pp 39-43.

[5] Freestone, O. and Mitchell, V. Generation Y attitudes towards e-ethics and internet-related misbehaviours. *Journal of Business Ethics*, 54:2 2004, pp. 121-128.

[6] Johnson, D. G. Computer ethics. Prentice Hall, New York, 2007.

[7] Mason, W. and Watts, D. J. Financial incentives and the performance of crowds. *ACM SIGKDD Explorations Newsletter*, 11:2 2010, pp. 100-108.

[8] Horton, J. J. and Chilton, L. B. *The labor economics of paid crowdsourcing*. ACM, 2010.

[9] Forte, A. and Bruckman, A. Why do people write for Wikipedia? incentives to contribute to open-content publishing, In *Proceedings GROUP 05 Workshop: Sustaining Community: The Role and Design of Incentive Mechanisms in Online Systems,* Sanibel Island, FL, 2005.

[10] Johnson, B. K. Incentives to contribute in online collaboration: Wikipedia as collective action. 2008.

[11] Paritosh, P., Ipeirotis, P., Cooper, M. and Suri, S. *The computer is the new sewing machine: benefits and perils of crowdsourcing*. WWW'11 Hyderabad, India, ACM, 2011.

[12] Adams, A. and Sasse, M. A. Users are not the enemy. *Communications of the ACM*, 42:12 1999. pp 40-46.

[13] Thornburgh, T. *Social engineering: the dark art*. ACM, 2004.

[14] Von Ahn, L., Blum, M., Hopper, N. and Langford, J. CAPTCHA: Using hard AI problems for security. *Advances in Cryptology—EUROCRYPT* 2003. p 646.

[15] Motoyama, M., Levchenko, K., Kanich, C., McCoy, D., Voelker, G. M. and Savage, S. Re: CAPTCHAs–Understanding CAPTCHA-solving services in an economic context. In *Proceedings of the 19th USENIX Security Symposium, USESEC'10*, August 11-13, 2010, Washington, DC.

[16] Yan, J. and El Ahmad, A. S. A Low-cost Attack on a Microsoft CAPTCHA. ACM, 2008.

[17] Bajaj, V. Spammers pay others to answer security tests. *NY Times,* April 25, 2010.

[18] Bonneau, J., Just, M. and Matthews, G. *Evaluating Statistical Attacks on Personal Knowledge Questions*. Springer-Verlag , New York, 2010.

[19] Levchenko, M. M. D. M. C. K. and Voelker, S. S. G. M. Dirty Jobs: The Role of Freelance Labor in Web Service Abuse. In *Proceedings of the 20th USENIX Security Symposium, USESEC'11*, August 8-12, 2011, San Francisco, CA.

[20] O'Connor, P. User-generated content and travel: A case study on Tripadvisor. com. *Information and Communication Technologies in Tourism* 2008. pp 47-58.

[21] Flosi, S. L. comScore Media Metrix Ranks Top 50 U.S. Web Properties for March 2011. comScore, Reston, VA, 2011.

[22] Keates, N. Deconstructing TripAdvisor. *The Wall Street Journal, October 13,* 2007. p W1.

[23] Jurca, R. and Faltings, B. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, 34:1 2009. pp 209-253.

[24] Ott, M., Choi, Y., Cardie, C. and Hancock, J. T. Finding deceptive opinion spam by any stretch of the imagination. *Association for Computational Linguistics*, 2011.

[25] Lai, C., Xu, K., Lau, R. Y. K., Li, Y. and Song, D. High-Order Concept Associations Mining and Inferential Language Modeling for Online Review Spam Detection. In *Proceedings of ICDM Workshops' 2010*. pp.1120-1127.

[26] Litvin, S. W., Goldsmith, R. E. and Pan, B. Electronic word-of-mouth in hospitality and tourism management. *Tourism management*, 29:3 2008. pp 458-468.

[27] Lippert, R. Policing property and moral risk through promotions, anonymization and rewards: Crime stoppers revisited. *Social & Legal Studies*, 11, 4 2002. pp 475-502.

[28] Zittrain, J. Minds for sale. Berkman Center for Internet and Society. 2009.

[29] Kandori, M. Social norms and community enforcement. *The Review of Economic Studies*, 59:1 1992. pp 63-80.

[30] Funk, P. Governmental action, social norms and criminal behavior. *Jour Inst & Theor Econ*, 161 3 2005. pp 522-535.

[31] Gray, T. E. Internet dating websites: A refuge for internet fraud. *Fl. Coastal L. Rev.*, Winter 2011, pp 389-407.

[32] ChinaSmack *Guo Meimei Red Cross Controversy Pissing Off Chinese Netizens*. Beijing, China, June 29, 2011. Available at: http://www.chinasmack.com/2011/stories/guo-meimei-red-cross-controversy-pissing-off-chinese-netizens.html

[33] Lim, D., Zo, H. and Lee, D. The Value of Anonymity on the Internet. *Service-Oriented Perspectives in Design Science Research.* 2011. pp 452-464.

[34] Ma, M. and Agarwal, R. Through a glass darkly: IT design, identity verification, and knowledge contribution in online communities. *Info Sys Research*, 18:1 2007. pp 42-67.

[35] James, J. I., Gladyshev, P. and Zhu, Y. Signature Based Detection of User Events for Post-mortem Forensic Analysis. *Digital Forensics and Cyber Crime* 2011, pp 96-109.