Understanding the Rogue User

Rich Gazan
Library & Information Science Program
University of Denver

Introduction

Visions of a harmonious online community are usually crushed quite quickly once human beings start to participate. A totem of user-centered design is that people use technologies in ways never intended by their designers (see, for example, Nielsen 1993), sometimes emotionally, but distinguishing destructive and creative interactions with an information system is often difficult. This distinction is addressed in this chapter via the concept of a rogue user, an active participant in an online community who violates the community's rules or spirit. Evidence of rogue behaviors in the Answerbag (www.answerbag.com) online question answering community was obtained through user postings and site logs, and analyzed through the lens of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR 2000), to suggest ways in which rogue behaviors can be understood and mitigated in the design of future online communities.

Background and theoretical framework

While the majority of affective information behavior research relies on direct observation of users, or the design of tools that are responsive to affective input, the emotional aspects of interaction within online communities have been comparatively understudied. Though evidence of affective interaction is usually restricted to text on a screen and transaction log data, it can indicate patterns of emotional behavior. Since online question-answering communities are both relatively recent phenomena, and access to data that does exist is usually tightly controlled by the owning organization, this study fills a gap in the literature by analyzing transcripts of online interactions in a naturalistic environment.

Increasingly, a person's online life is one of their many contexts (Turkle 1995), and an online community can serve as a 'small world' within which standards of behavior are defined and shaped (Chatman 1999). Solomon (2005) extends Chatman's work and conceptualizes a kind of exchange: an individual accepts the norms of a small world in exchange for making life manageable. Since one's online identity is often a vehicle for forms of expression and interaction that aren't possible elsewhere, some individuals become emotionally invested in both the online community and their identity within it. In an analysis of deviant behavior in cyberspace, Suler and Phillips (1998) challenge the concept of online disinhibition; they argue that though anonymity might remove some behavioral inhibitions, online community participants also seek the reward of recognition.

Baym (1997), studying Usenet, writes that online interactions allow participants to "co-construct the values, identities and conventions that make a group feel like community." Burnetta and Bonnici (2003) make a distinction between an online community's explicit norms and its implicit norms, such as who is entitled to post, what constitutes appropriate content, and community etiquette. Rogue users bend or break the rules in order to challenge or even attempt to control the implicit norms of a community.

The term "rogue user" has been used somewhat loosely in the computer science literature, to describe individuals who are not full-fledged hackers, but who have gained access to files or functions beyond their permission level. However, McNee et al. (2002, p. 118) specifically employ the term to describe users who undermine the rating system in a collaborative filtering environment. This chapter extends this definition to include any violation of an online community's rules or spirit by an active participant.

Site and method

Answerbag is an online question answering community that supports both anonymity and recognition. Users submit frequently-asked questions in a nearly limitless variety of categories, submit answers under a screen name, and also rate answers as useful (100%), somewhat useful (75%), or incorrect/not useful (50%). Multiple answers to a question are permitted, and the highest-rated answers are listed first, providing collaborative filtering while still allowing users to see the range of different answers. Participants with the highest percentage of useful answers gain status in the community; their screen name and statistics are displayed on the site, viewed by thousands of people per day. While this competitive aspect of recognition increases site traffic, it also motivates some users to bend the rules.

Answerbag is both a public Website and a research testbed, and administrator-level access to all site data was readily available. Initial research questions included:

- Why do some frequent online community participants engage in rogue behaviors?
- How can rogue behaviors be mitigated?

While some rogue behaviors can be measured statistically, many were initially reported by site moderators, who edit and categorize user submissions, monitor categories for inappropriate content, and interact with users via feedback forms. Data collection took place throughout 2005 and is ongoing. Instances of inappropriate activity collected by moderators were aggregated and analyzed, to develop a list of common rogue behaviors, including:

- *Vindictive rating*: e.g. downgrading all answers submitted by a given user, to advance one's own position.
- *Abusive language*: in answers or rating comments.
- *Flooding*: submitting multiple questions or answers in an attempt to dominate a category with a single point of view, or to increase submission statistics.

- *Excessive contact with administrators*: lobbying to get an offensive question, answer or rating removed, or to make self-serving design or policy suggestions. More than ten contacts is deemed excessive.
- *Creating 'sock puppets'*: separately registered identities under which rogue behavior is undertaken, to shield the user from identification or retaliation.
- *Requesting special privileges*: asking to be exempt from certain site rules and policies.

While this list was compiled to train moderators, these behaviors were observed so often that a more structured analysis was undertaken. Instances of rating irregularities, flooding and excessive contact were obtained from usage statistics. Examples of abusive language were flagged by moderators. Though 'sock puppets' are not always obvious, they were triangulated from login times, content analysis, unusual submission patterns and reports from other users.

Of roughly 40,000 registered users, only forty-six were identified as having engaged in one or more rogue behaviors. However, of these forty-six individuals, six ranked among the top 50 site contributors in terms of questions submitted, answers submitted, and useful rating percentage. Three ranked among the top 10. Understanding why such high participators in an online community would engage—often repeatedly—in behaviors that undermine the community's integrity motivated a deeper affective analysis.

Data analysis and Narcissistic Personality Disorder

The field of psychology has institutionalized personality diagnosis in DSM-IV-TR (2000). While DSM has been critiqued for taking an over-categorical approach to symptoms of mental illness, it does provide common dimensions of personality that can be compared and aggregated to form a diagnosis, with more or less confidence depending on how many criteria are observed. According to DSM, Narcissistic Personality Disorder (NPD) is an Axis II (diagnosed in adulthood), Cluster B (dramatic, emotional or erratic behavior) personality disorder, defined as a "pervasive pattern of grandiosity (in fantasy or behavior), need for admiration, and lack of empathy, beginning by early adulthood and present in a variety of contexts." Five of nine criteria must be met to diagnose NPD, which is estimated to occur in 0.7 to 1% of the population:

1. self-importance
2. fantasies of unlimited success, power, etc.
3. belief that he or she is special and unique
4. requires excessive admiration
5. unreasonable sense of entitlement
6. interpersonally exploitative
7. lacks empathy
8. envious of others, or believes that others are envious of him or her
9. arrogance

Since indicators overlap, it is both difficult and potentially misleading to categorize rogue behaviors as instances of one NPD dimension and not another, so a numerical breakdown of frequency data is not included here. Also, evidence of online interaction is certainly not sufficient for a clinical diagnosis. However, Table 1 demonstrates that many rogue behaviors observed in Answerbag users could be coded as positive dimensions of NPD:

Table 1: Examples of rogue behaviors and corresponding NPD dimensions

NPD dimension	Rogue behavior	Example [comment]
Self-importance	Excessive contact with administrators	*"I cannot sit idly by and let them ruin my ratings... It takes courage to stand alone and defend against this kind. When a person gives truthful responses and are rated so low, then the site itself is failing...I wonder if they are being allowed to personally attack me because it boosts interest in Answerbag?"* [To date, this user has made over 200 contacts with administrators, usually over 'unfair' ratings]
Fantasies of unlimited success, power, etc.	Requesting special privileges	*"I think you guys have the greatest site in the world, but you need someone to get the word out better. I would be happy to discuss a consulting arrangement...[that would] make us all rich."* [If this user's profile information and question submissions are any indication, he is a recent college graduate who is seeking employment]
Belief that he or she is special and unique	Requesting special privileges	*"I'm probably your best contributor, I given [sic] my heart and soul to AB, I think I deserve not to have my questions edited anymore."* [Neither the quality nor the quantity of this user's submissions was remarkable, save for a one-week submission binge immediately prior to this message]
Requires excessive	Excessive contact with	*"Can AB give acclaim to users whose*

admiration	administrators	*question submissions appeared on the front page by having their user profile show a medal?"* [This user's 80+ site suggestions usually include a 'look at me' element, and his aggressive defense of his submissions recalls a quote by Napoleon Bonaparte, circa 1807: "I have made the most wonderful discovery. I have discovered men will risk their lives, even die, for ribbons!"]
Unreasonable sense of entitlement	Flooding	*"I hope to have asked the most questions so whenever Answerbag launches a tangible rewards program that awards those who have the "bests" and "mosts" of anything, that I get rewarded dearly for having asked the most questions."* [In the Name Origins category, this user had recently submitted over 400 variants of the question "What is the origin of the name X?," to reach the coveted "Most Inquisitive" title.]
Interpersonally exploitative	Creating sock puppets	*"If you cannot stop [user] from calling me a liar and a satan, I'll have to call in my own troops."* [Soon after this communiqué, two new "members" registered on Answerbag, whose sole purpose seemed to be giving 100% ratings to every one of the above user's answers, and 50% to those of the alleged name-caller. Whether 'sock puppets' or not, this can be interpreted as exploiting another online identity to serve one's own ends.]
Lacks empathy	Abusive language	*"Who is monitoring your home, [username], to make sure you're not abusing children? Have you stopped molesting children yet?"* [This was posted in response to a

		question regarding secrecy in the practices of a certain religion, and how accusations of child abuse are handled.]
Envious of others, or believes that others are envious of him or her	Vindictive rating	*"I think it's wise to react to paranoia and stop the actions of anyone when they are destructive in a personal way. Do you agree? Does the crime of stalking carry the same penalty for the perpetrator if he or she were stalking someone online rather than doing so in person?"* [This user insisted that another user was envious of her position as a category expert, and was 'stalking' her across Answerbag by down-rating her answers. This message was her explanation for why she had vindictively rated the answers of the user in question.]
Arrogance	Abusive language	*"This has to rank as the most ignorant answer on Answerbag, maybe on the entire Internet…"* [A now-deleted rating comment by a frequent contributor of quality content, who nonetheless seems to go on jags of belittling others via insulting rating comments.]

When a rogue behavior was identified, the transaction data surrounding the question, answers, answer ratings, users, and the category in which the rogue behavior appeared were flagged for closer content and traffic analysis. When users elsewhere in the site referenced a rogue user, question or answer (such as in the 'Answerblog,' a bulletin board for free-form discussion about the site), any positive or negative comments were also factored into the analysis, to get a sense of the wider effects of the rogue behaviors.

Results and Conclusion

Though rogue users tend to be high participators themselves, when their roguish behaviors are apparent to other users, overall participation is reduced. There may be an initial flood of negative ratings and complaints in a category in which rogue behavior has taken place, but subsequently the number of participants tends to fall. This plea from one

non-rogue user, responding to abusive language in a Religion subcategory, indicates the chilling effect of rogue behavior on community participation:

> *"…curb the ongoing negativity…some of us haven't ventured onto the killing fields and don't plan to."*

The results suggest that using DSM to interpret rogue behavior through the lens of NPD yielded both greater understanding of why the behavior occurred, and several mitigation strategies. Attempting to reason with rogue users by first expressing appreciation for their contributions (appealing to their sense of self-importance), then requesting that they curb rogue behaviors for the good of the site, were usually met with resistance, defensiveness or defiance (NPD dimensions: lacks empathy, arrogance), and no change in rogue behaviors. However, in a parallel development, Answerbag began to allow users to construct enhanced personal profile pages on the site. Along with detailed usage statistics, users now had an 'About me' section where they could post personal information, links, quotes or anything else. Profiles are linked from the user's screen name, and accessible from any page on which the user has contributed content. Of those users who have visited the site since the upgrade, less than 10% have created enhanced profiles, while 77% of rogue users have. In the first four months after enhanced personal profiles were made available, instances of rogue behavior fell from an average of twelve per month to five. Providing an area of the community where one is by definition 'special and unique,' and where no one can edit or rate their content, seems to have served as an outlet for emotional expression, and less rogue behavior now occurs on the site.

Virtual identities can exhibit the same range of emotions as the people behind them. Rogue users in online communities can be viewed not simply as destructive miscreants, but as individuals with emotional needs that information systems might be better designed to address. Though this is a small initial study, the results suggest that diagnostic personality tools such as DSM can provide one avenue to understanding rogue behaviors, and suggest productive ways to channel the emotional needs behind them, to keep online communities useful and sustainable for all.

References

Baym, N. (1997). Interpreting soap operas and creating community: Inside an electronic fan culture. In: Keisler, S. (ed.) Culture of the Internet, Lawrence Erlbaum, Mahwah, NJ.

Burnetta, G., and Bonnici, L. (2003). Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library & Information Science Research* 25(3), 333-351.

Chatman, E.A. (1999). A theory of life in the round. *Journal of the American Society for Information Science* 50(3), 207-217.

DSM-IV-TR (2000). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association.

McNee, S.M., et al. (2002). On the recommending of citations for research papers. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work. New York: ACM Press, 116-125.

Nielsen, J. (1993). Usability engineering. San Francisco: Morgan Kaufmann.

Solomon, P. (2005). Rounding and dissonant grounds. In: Fisher, K.E., S. Erdelez and L. McKechnie (eds.). Theories of information behavior. Medford, NJ: Information Today, 308-312.

Suler, J.R. and Phillips, W. (1998). The bad boys of cyberspace: Deviant behavior in multimedia chat communities. *CyberPsychology and Behavior*, 1, 275-294.

Turkle, S. (1995). Life on the screen: Identity in the age of the Internet. New York: Simon & Schuster.