# What Ought A Program Committee To Do?

Mark Allman
*International Computer Science Institute*

## 1 Introduction

When the Internet was initially constructed it was a de-facto playground for researchers. The research community built the Internet and refined it by monitoring traffic and tinkering with protocols and applications. The days of the Internet being a research sandbox are clearly over. Today, society at-large depends on the Internet as a crucial piece of infrastructure for communication and commerce. As such, experimenting with and monitoring today's Internet is a much more thorny endeavor than in the past because researchers must consider the impact of their experiments on users. The community has started to face questions about the propriety of certain experiments and will no doubt continue to struggle with such issues going forward. In this note, we ponder the role of a program committee when faced with submissions that involve a potentially unacceptable experimental methodology.

We note that the community already charges program committees with dealing with a number of ethical issues such as plagiarism, simultaneous submissions and conflicts of interest. Should dealing with questions about the ethics of a particular experimental methodology also be on the PC's plate? Certainly we are aware of PCs that have taken it upon themselves to include such issues in the debate over particular submissions. Are such considerations "in bounds" for PC consideration? Or, should PCs stick to the technical aspects of submissions?

To give the reader a concrete flavor of the sorts of questions that might arise during PC deliberations we sketch possible reactions to several recently published paper as examples.

**Ex. 1:** When do active probing measurements cross the line and become attacks? For instance, in [11] the authors first collected a list of addresses that were believed to represent residential broadband-connected hosts. These hosts were then probed at various rates, which top out at 10 Mbps for 10 seconds. These "floods" (as [11] refers to them) were explicitly meant to congest the residential network to gather insights into their characteristics. Is this a reasonable amount of traffic over a short time interval? Or, is this an attack? If the latter, then when does probing constitute an attack and when is it benign? What distinguishes acceptable probing from an attack? Intent? Rate? Kind? We have heard various answers to these questions and so it seems clear to us that the community does not have a shared value in this space.

**Ex. 2:** How should shared measurement data be treated? In [10] the authors de-anonymize packet traces made freely available by the Lawrence Berkeley National Laboratory (LBNL) [3] and published a mapping between eight IP addresses found in the anonymized LBNL traces and the presumed real addresses.[1] Is this reasonable treatment of data that is offered as a resource to the community? Even though a strawman answer to that question has been offered [5] the question remains and the very presence of this pair of papers illustrates our lack of a common view on the topic across the research community.

**Ex. 3:** How should community resources be treated? PlanetLab [2] is a collection of several hundred machines that are provided and hosted by a variety of organizations around the world. Access to these hosts is granted to researchers as a way to enable large-scale systems research. While there is no doubt that PlanetLab has enabled much good research, it has also shown that such testbeds can expose unwitting participants to potential problems. For instance, [7] details an experiment whereby the researchers join the gnutella peer-to-peer network and exchange what they believe is copyrighted music with others in the network. On the one hand, such an experiment intends no harm and simply

tries to mimic user behavior to gain insight into the peer-to-peer system[2]. On the other hand, illegal exchange of music has caused legal problems for a large variety of people and organizations. Therefore, is such an experiment a legitimate use of a community resource? Or, is such an experiment exposing organizations supporting PlanetLab by hosting nodes to potential legal trouble? This and similar questions will grow increasingly important as testbeds like GENI [1] are realized as resources for researchers' experiments.

**Ex. 4:** How should users be treated during an experiment? The research community has done many studies that incorporate human behavior since such behavior often drives network dynamics. Traditionally, the researchers in our community have not sought *user consent* for such studies. Should we? Other sciences have guidelines for how to treat human subjects. Should those apply to us? Is there a distinction between watching user behavior at the border of some organization and watching user behavior by installing a monitor on the user's laptop (as was the methodology used in the study presented in [12])? Can we be expected to obtain consent from a large community of users to install a packet tap? Or, is institutional consent—whereby an organization's management approves the monitoring of the organization's network users—enough?

The above questions are all items that the research community will have to (formally or informally) puzzle through in some fashion. These questions are not posed to lay blame against the papers cited above. Nor are we posing these questions in an attempt to answer them in this note. Rather, we pose the above questions and cite the papers to give a flavor for the work that program committees *have grappled with recently*. The question we ask in this short note is: What is a program committee's role when faced with submissions that touch on the thorny issues of whether a particular experimental methodology is acceptable or a breach of etiquette or ethics?

## 2 Related Work

The discussion in this paper is in terms of "etiquette" and "ethics". Another aspect of the questions posed, however, is that of "legal" concerns. We largely ignore this final aspect, but no not wish to diminish its importance. We do not consider legal aspects here because ($i$) PCs for systems venues are not comprised of lawyers and therefore applying legal constraints to submissions is not likely to be accurate or useful and ($ii$) the global nature of our conferences and workshops mean that the submitted work is subject to myriad laws and precedents that PCs cannot be expected to understand or cope with. Researchers are encouraged to take steps to ensure they understand the legal implications of their work. Two recent papers written by legal scholars provide legal background within United States law (however, these are likely not a replacement for consulting a lawyer about specific projects) [13, 9].

The ACM Code of Ethics and Professional Conduct [6] can be read to address some of the issues we discuss in this note.[3] For instance, the Code indicates that professionals should "avoid harm" and "respect the privacy of others". As a broad framework the Code is reasonable and should provide researchers with a basis for thinking through their experimental methodologies. However, the lack of specifics means that the Code can be interpreted in a multitude of ways and therefore offers little help to PCs when considering thorny ethical issues (except, perhaps, in particularly egregious cases). In addition, reading the Code from a reviewer's standpoint might cause the reviewer to think that "avoid harm" means that a paper should be rejected to provide a disincentive to some particular behavior. Given these ambiguities it seems difficult to strongly lean on the Code for guidance on the sorts of specific questions that often face program committees.

Finally, while admitting similarity, we draw a distinction between research that is focused on users and networks and researched that is focused on algorithms. Therefore, we do not consider research such as outlined in [8], which discusses the security problems in 802.11 networks protected by WEP (and in fact how to compromise WEP). While such work can have concrete implications for users the research is fundamentally focused on the algorithms.

## 3 Decisions, Decisions

A program committee's overall task is obviously to decide which papers to include in a conference program. How much should these issues of etiquette and ethics play into the decision to accept, reject or shepherd a paper?

A simple approach would be for PCs to not bother with issues of etiquette and ethics at all and consider only the technical contribution of a particular paper. This option might be seen as reasonable because PCs are generally comprised of people with technical expertise, but not necessarily a broad grasp of the potential ethical issues involved in conducting various research projects.[4] This would leave judgments about the acceptability of particular techniques to the overall community's public scrutiny. For instance, the use of the LBNL data discussed in § 1 and published in [10] led to a rebuttal and call for more explicit guidance from data providers [5]. Perhaps a sys-

tem whereby the community-at-large polices behavior in such a fashion is best.

On the other hand, there are a number of reasons why a PC might want to consider the etiquette involved in a particular submission as part of its decision process, such as:

- PCs are in a unique position in that they see papers before the community at-large does and can take concrete steps to combat unacceptable behavior. Does this unique position carry a responsibility to the community to reject papers with inappropriate methodology?
- Not taking the etiquette of experimental techniques in account could be viewed as approval of the particular behavior, which could encourage researchers to not worry about these sorts of ethical issues when conducting their experiments. In turn, this could cause backlash against the community as outsiders grow weary of our methods and operators and administrators become less helpful.
- Similar to the above, behavior that violates etiquette could harm a particular venue's reputation or that of its sponsoring organization (e.g., ACM, USENIX, IEEE, etc.).
- Identifying unacceptable experiments involving the community's resources (e.g., PlanetLab, available data, etc.) could be viewed as protecting those resources for the benefit of the overall community.

While there are reasons a PC may want to consider breaches of etiquette in their decision-making process, such a path is not without problems. First, in the absence of a set of community norms each PC will have to reach its own consensus about the acceptability of a particular experimental technique. This will ultimately lead to uneven results and an unfairness to authors across venues. Further, rejecting a paper does not necessarily discourage what a PC considers to be inappropriate behavior since such decisions are not public. Finally, by rejecting questionable papers the community may lose out on some key new understanding that came from the research—which in turn begs the question as to whether the ends justify the means.

## 4 External Notification

In the normal course of business submissions are expected to be held in confidence by a PC. However, another question that has come up is to what degree a PC should violate this expectation when the committee finds the experiments presented in a submission to violate some form of etiquette. We know that in cases whereby reviewers and PC members have alerted PC chairs about a possible simultaneous submission that the chairs of the two venues have violated the expectations and shared the two papers to investigate the claims. Further, in cases of suspected plagiarism the ACM has a well-established policy for investigating such allegations that is beyond the scope of normal PC process and includes additional people [4]. Thus, there is precedent for PCs to involve external parties under exceptional circumstances.[5]

A broad question about whether there should be a body charged with investigatory power for certain unacceptable experimental behavior—as for plagiarism—is certainly something the community could puzzle through. However, that requires a set of ethical standards as a first step. Further, this question is somewhat beyond the scope of this note (i.e., what a *PC* can or should do when encountering inappropriate experimental techniques).

A more near-term question pertains to the use of a shared community infrastructure such as PlanetLab or some released dataset. When a PC encounters an experiment it ultimately considers inappropriate, is the PC in some way duty-bound to share the submission and the concerns with stewards of the community resource in the name of protecting the resource for the good of the entire community? As discussed above, PCs have a unique vantage point and therefore can raise concerns before a particular paper reaches the public. This can be important in cases such as de-anonymization of data whereby the results of the research may, for instance, have an impact on a network's security posture. In addition, if the administrators of some platform concur that unacceptable behavior is occurring they can sever a user's access. On the other hand, this violates the expectation of a PC holding submissions in confidence. Is protecting our community's resources a big enough concern to violate this expectation?

A similar situation would arise if a PC thought the privacy of a groups of users was being undermined in a submission. Does the PC have a duty to report such activity to the group of users (if possible) or the organization providing the data used in the submission?

## 5 Institutional Review Boards

Many organizations have processes for doing research that involves human subjects and this often involves an Institutional Review Board (IRB). Computer scientists have started using their institution's IRB processes for studies involving traffic monitoring. Anecdotally we find that some of this is driven by institutions becoming more aware of the implications of networking research and some is driven by researchers seeking to "cover them-

selves". Whatever the impetus for using the IRB process, a natural question is how it pertains to a PC's deliberations. If a submission notes that the experimental methodology has been vetted by the submitting institution's IRB is that enough to allay any concerns a PC might have about etiquette? While our intention is not to dissuade the use of IRBs, they are not a panacea. We note several issues:

- First, IRBs are generally setup to review research that involves human subjects. Examples 1–3 in § 1 do not involve human subjects and so seemingly would not be subject to traditional IRB oversight. Therefore, even in the best case scenario of a IRB approval being an additional aspect that a PC might leverage the systems community is left with thorny issues.
- Second, IRBs have traditionally been setup to deal with medical, biological, psychological, etc. research projects. It is not clear that IRB members will have the expertise needed to understand, for instance, the safe-guards (and their limitations) that systems researchers put into place to protect individual's identity. Can we really expect an IRB setup to vet medical research to understand that a packet-trace anonymization scheme is vulnerable to a side-channel attack relating to the clock drift present in TCP packets with the timestamp option?
- Finally, since ($i$) our community does not have a set of shared values (as discussed in § 1) and ($ii$) we cannot expect IRB boards to have the necessary domain expertise, as sketched above, it seems clear that in a global sense the IRB process will produce wildly varying results. Therefore, PCs are likely to continue to get papers that some members feel are problematic—even though they have been vetted by an IRB. Should we simply table these complaints as irrelevant? Or, is there a role for PC oversight to protect some greater global good (e.g., to protect users even if an institution will not)?

All that said, it seems clear that in some cases IRB approval can be used by PCs as an indication that an experiment is acceptable. For instance, if an experiment involves monitoring a campus wireless network and the appropriate IRB approves of the procedures for monitoring, storing the data, ensuring user privacy, etc. then shouldn't a PC respect that boards findings for the given setting described in a submission?

## 6 Discussion

In many ways this note contributes nothing to finding solutions to what PCs ought to do in cases where they discover what they believe to be inappropriate experimental behavior. Our goal in writing this is to not speculate on an answer to this question, but rather to start a discussion within the community about these issues. One avenue that the community may pursue is to develop a set of standard practices and/or a set of practices that are considered out-of-bounds.[6] If a set of community norms were to be developed, what should a PC's role be in enforcing those norms? In the absence of such a set of community standards, what role should a PC take? Our hope is that rather than positing what a PC's role should be we can first start with a community discussion of these issues.

## Acknowledgments

This note benefits from discussions with a great many people including Ethan Blanton, Aaron Burstein, kc claffy, Ted Faber, Mark Claypool, Shawn Ostermann, Vern Paxson, Colleen Shannon and the anonymous WOWCS reviewers. My thanks to all!

## References

[1] Global environment for network innovations (geni). http://www.geni.net/.

[2] Planetlab. http://www.planet-lab.org/.

[3] LBNL Enterprise Network Traces, Oct. 2005. http://www.icir.org/enterprise-tracing/.

[4] ACM Policy and Procedures on Plagiarism, Oct. 2006. http://www.acm.org/pubs/plagiarism

[5] ALLMAN, M., AND PAXSON, V. Issues and Etiquette Concerning Use of Shared Measurement Data. In *ACM SIGCOMM/USENIX Internet Measurement Conference* (Oct. 2007).

[6] ASSOCIATION OF COMPUTING MACHINERY. Code Of Ethics and Professional Conduct, Oct. 1992. http://www.acm.org/about/code-of-ethics.

[7] BANERJEE, A., FALOUTSOS, M., AND BHUYAN, L. Is Someone Tracking P2P Users? In *IFIP Networking* (2007).

[8] BORISOV, N., GOLDBERG, I., AND WAGNER, D. Intercepting Mobile Communications: The Insecurity of 802.11. In *ACM Conference on Mobile Computing and Networking* (July 2001).

[9] BURSTEIN, A. Conducting Cybersecurity Research Legally and Ethically. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)* (Apr. 2008).

[10] COULL, S., WRIGHT, C., MONROSE, F., COLLINS, M., AND REITER, M. Playing Devil's Advocate: Inferring Sensitive Information from Anonymized Network Traces . In *Proceedings of the Network and Distributed System Security Symposium* (Feb. 2007).

[11] DISCHINGER, M., HAEBERLEN, A., GUMMADI, K. P., AND SAROIU, S. Characterizing Residential Broadband Networks. In *ACM SIGCOMM/USENIX Internet Measurement Conference* (Oct. 2007).

[12] GIROIRE, F., CHANDRASHEKAR, J., IANNACCONE, G., PAPAGIANNAKI, D., SCHOOLER, E., AND TAFT, N. The Cubicle vs. The Coffee Shop: Behavioral Modes in Enterprise End-Users. In *Passive and Active Measurement Conference* (Apr. 2008). To appear.

[13] OHM, P., SICKER, D., AND GRUNWALD, D. Legal Issues Surrounding Monitoring During Network Research. In *ACM SIGCOMM/USENIX Internet Measurement Conference* (Oct. 2007).

## Notes

[1] In fact, seven of the eight address mappings given for the LBNL traces in [10] are wrong [5].

[2] [7] even notes that all music files were deleted immediately after the experiment.

[3] For brevity we only address the ACM Code in this note, but other professional societies have similar codes.

[4] This is different from PC members being unethical. It is possible for a given researcher to understand well the issues involved in their own work, but have little or no understanding about the sensitivities involved in work on a different topic. E.g., someone may understand the sensitivities involved in passive measurements, but not appreciate the issues involved with active probing.

[5] In addition, outside of computer science there is history in revealing otherwise private conversations in extreme cases. E.g., a lawyer is required to reveal any knowledge of a planned crime by a client—even though conversations between attorneys and clients are generally private and not revealed.

[6] A BOF was held at last year's Internet Measurement Conference as an initial discussion of whether such a set of norms would be useful. A possible workshop adjunct with the Passive and Active Measurement Conference in April 2008 may attempt to take the next step.