usenix ;login:
DECEMBER 2012 VOL. 37, NO. 6
ELECTRONIC SUPPLEMENT

Conference Reports

5th Workshop on Cyber Security Experimentation and Test (CSET '12)

Bellevue, WA
August 6, 2012

Anonymity and Privacy

Summarized by Aaron Alva (aalva@uw.edu)

Conducting an Ethical Study of Web Traffic

John F. Duncan and L. Jean Camp, Indiana University

John Duncan presented on the ethical aspects of network research conducted with university dorm students as subjects. He detailed the ethical design of the study and his experiences with the Institutional Review Board (IRB) at Indiana University. John also reviewed the state of ethical research standards while providing a charge for improved collaboration between ethics boards and researchers. Up front, John offered his conclusion highlights: that network researchers now have little incentive to follow the rules, that the rules are poorly defined to begin with, and that IRBs are neither fully equipped nor aware of what network research is, let alone how to deal with it.

John said that the main goal of the paper was to detail the ethical challenges in the process of conducting network research. The network research study on Web browsing habits of students used a server mirroring TCP port 80 traffic from the dorm students. There was no attempt to fetch content, no HTTPS packets were observed, session info was removed, and some post-processing was done on the data. While there was a process in place to secure the data, there was an unsecured copy the researchers were not aware of that was used in a publication by a colleague without consulting the original research team. There was no process set by the IRB for additional uses of the data, as happened in this case.

John stated that governing frameworks do matter, and the IRB needs to be involved in the process whether at the beginning, middle, or end. The IRB at Indiana University helped shape the research process and, together with the researchers, went through nine drafts of the study notification flyer. There were several areas in which the IRB did not allow the researchers to take action. For instance, Tor was originally suggested to students as an option for opting out. When using Tor was ruled to be against university computer use policy, a VPN was offered by the researchers instead. To protect student privacy, the researchers de-identified data, and addressed the concerns of inquiring students. As an ethical-design failure, though, opt-out was characterized as inherently more problematic than opt-in. Opt-in was deemed infeasible because viewing the packets to determine eligibility would have involved as much surveillance as the opt-out model.

John discussed another privacy case in which a student was using a poorly written anonymization service that was spoofing requests from uninvolved clients, with the name of the site communicating a certain sexual orientation. The researchers options were (1) to adjust collection to identify machine, individual, and notify the individual, directing them to better resources; or (2) to do nothing. In this ethical quandary, the answer was to do nothing. What if the researchers found potentially criminal activity, or activity that might imply criminal intent? John said that in the absence of actual physical threat, pre-crime information is protected against disclosure.

The second portion of the presentation provided a state-of-the-field in research ethics. John reviewed the human subjects process in the Belmont/Common Rule Update, an ANPR that included explicit questions for the research community and was motivated by new categories of risk such as bioinformatics research; and the Menlo Report (DHS-2011-0074), which suggests researchers must apply to IRBs to conduct network research. John then offered an alternative: draw from the IEEE and ACM (both of which filed comments to the ANPR) in order to ensure research protocols compliance. Additionally, since "if it doesn't get published, it isn't done," publishers could potentially bear responsibility by not publishing research unless it was cleared by an IRB. John again ends by inviting all types of network researchers to participate in the process to ensure the "research we do protects others and protects us."

Questioning began with a conjecture: suppose you have a study you want to do. You would float material on a dozen IRBs across the country and see what happens. For example, if you want to get something through, you would then know the right IRB to go to. Another participant noted his IRB had moved to Web forms, decreasing interaction between the IRB and the PI. There seems to be a fundamental problem with communicating, and this widens that gap. John rebutted the format change as an inherent problem, saying that the difference is whether the IRB is willing to sit down and talk to researchers. Communication comes down to the willingness to converse.

Steve Schwab, the program chair, asked whether we needed to talk about "future-perfect anonomyzation" when protecting privacy. John responded that network researchers must consider the implications of crypto users with regards to

"future-perfect crypto." The question there is how long will it be before the crypto's cracked, since today's gold standard is no guarantee against attacks in the infinite future. Steve then asked about why the researchers were not permitted by the IRB to conduct an educational panel on security in the dorm. John said that they tried to conduct a basic educational seminar to provide immediate benefit, but were not allowed to. The IRB said they could not approve it, but did not state why. When you're going through the IRB, the feeling is "denied until proven accepted."

Finally, Lucas Reber (University of Washington) pointed out a new university plan that exempted student work in classroom from going through IRB provided there was faculty oversight. John said this was another example showing the lack of uniform standards across IRBs.

Methodically Modeling the Tor Network

Rob Jansen, US Naval Research Laboratory; Kevin Bauer, University of Waterloo; Nicholas Hopper, University of Minnesota; Roger Dingledine, The Tor Project

Rob Jansen presented on the use of real measurements and metrics for modeling the Tor network as best as possible. Rob began the talk with an overview of why anonymity is important and the "Google herpes" example in which anonymity was breached. By using Tor, Rob explained, one can use telescope communications through three encrypted relays to provide anonymity. The Tor network, in terms of research, can benefit from experimentation. This can be done by (1) using the live Tor network itself, (2) using a distributed system, or (3) simulation. Using the live Tor network would be realistic, although it is difficult to manage for research purposes and there are associated privacy risks. The use of a distributed system such as PanetLab is also realistic, but it is difficult to manage, not scalable, and inaccurate due to lack of control over specific characteristics. With any of these choices, there are levels of complexity to understand including latency, jitter, bandwidth, location of nodes, behaviors, and more.

Rob said that because of these challenges, specific experimentation tools were created. Shadow and ExperimenTor both run real Tor software and were used for direct comparisons in this research. Shadow (https://shadow.cs.umn.edu/) uses discrete-event simulation of the network layer and simulates Internet traffic by using non-deterministic jitter. ExperimenTor (http://crysp.uwaterloo.ca/software/exptor/), similarly, allows researchers to run the actual Tor software within the isolated, simulated environment. When using models for Tor, how do we know what we have is like reality? Rob presented metrics used for accuracy of the model for network performance and network load. These metrics were validated through the use of Shadow and ExperimenTor to simulate real Tor networks well. Rob ended with an open question for further consideration: how can we tell when experiments have "enough rigor" to produce meaningful results?

Eric Eide took the open question and suggested that meaningful results would be measured as whether the results of the experiment have improved the real Tor network. Rob suggested that this experiment had sufficient rigor demonstrated through the discovery of a bug in the Tor software that was found and fixed. John Duncan asked what Rob's feelings were about the more limited past experiences and whether available experimentation options were better than nothing. Rob said that in the past, alternatives were limited and now experimentation is better than nothing. Network models and Tor-specific experimentation tools are available and configurable.

Collaborative Red Teaming for Anonymity System Evaluation

Sandy Clark, University of Pennsylvania; Chris Wacek, Georgetown University; Matt Blaze and Boon Thau Loo, University of Pennsylvania; Micah Sherr and Clay Shields, Georgetown University; Jonathan Smith, University of Pennsylvania

Sandy Clark presented a method of red teaming that was employed to conduct penetration testing on the SAFEST evaluation framework. Sandy began by describing Tor in general and the SAFEST evaluation framework—a DARPA-funded coordinate system on top of Tor. This coordinate system adds tunable link selection policies and enables a set of relay policies that precisely describes the characteristics desired, then tunes them properly. Sandy noted the importance of the coordinate system, as it could potentially be used to get access to the topology of the network. The SAFEST framework, then, is independently tested by an external red team.

Sandy said that there are two ways to penetration test. One way is to conduct pen tests covertly with a secret attack team. This is particularly good with IT staff or an implementation you want to test; however, because Sandy and her team were actually testing design, covert red teams were not necessary. The choice was made to conduct overt red teaming. The methodology for overt pen testing was to have the blue team get together with the red team and set up the scope. Then, the blue team also participated in the information gathering phase. Source code and a walkthrough of how the coordinate system works was provided, and design assumptions were also handed over to the red team. Once questions had been answered, the red team went away for several months. Following this, the actual red team attack occurred. A meeting occurred again following the attack phase to discuss successful and unsuccessful attacks, with the blue team attempting to repeat each attack.

Sandy said the result of this process was important for the security of this high-assurance software. The "School of Fish Attack" was a compelling example that stemmed from this process. The blue team did everything they could in the design of the system to brainstorm all that could be broken or attacked. As a result of the brainstorming, an important vulnerability in which a malicious user could make a (malicious) relay more enticing was discovered. This could not be changed, but the blue team built in an effective strategy where the safest nodes would keep a histogram that would skip a new and attractive node, thus mitigating the vulnerability. The red team launched an attack and failed to exploit this vulnerability because of the protection mechanisms put in place. As with normal procedure in the overt pen test process, the red team took note of the failed attack then engaged in an interesting question and answer session with the blue team. The discussion revolved around what happens if the attack was refined to exploit the protection mechanism itself. When the malicious node moves further out, other nodes migrate likewise. But then the malicious node jumps back into attractive position leaving the other nodes stuck too far away to be chosen.

Sandy emphasized that each vulnerability matters in this case. The back and forth discussion between blue and red teams resulted in the discovery of this vulnerability—and the protection of lives. Lessons learned from the process were that the red team gains better insight into the design assumptions; and the blue team learns new skills while experiencing the "attacker mindset," which is notably missing from the academic community.

Roy Maxion (Carnegie Mellon University) asked why the attack mind-set is missing from academia. Sandy responded that computer security students are bright, but they need to play by the rules to get through the academic system. To attack, you have to violate the assumptions on which the system is based. Cynthia Irving (Naval Postgraduate School) said it depends on the context in academia; she's from a military university. Roy noted an attitude in academia of "I've tried this and it doesn't work." Cynthia said the flaw hypothesis methodology was published a while ago and should be taught. Sandy observed that we unfortunately push the hacker mind-set aside in academia, but on the opposite side see lots of talks about breaking systems. The academic examples of attacker mind-set include DARPA funding for a hacker academy, hacker spaces, and the new game Yoshi Kohno developed. Dennis McCloy agreed that this is clearly a required component of teaching security.