DIANA SMETTERS, BRINDA DALAL, LES NELSON, NATHANIEL GOOD, AND AME ELLIOTT

ad hoc guesting: when exceptions are the rule

Diana K. Smetters is a senior security researcher at the Palo Alto Research Center. Her research interests include usability of security and applications of cryptography, particularly in social, mobile, and ubiquitous computing.

smetters@parc.com

Brinda Dalal is an anthropologist in the Ubiquitous Computing Area at the Palo Alto Research Center, focusing on sociotechnical design, usable security, and green technology innovation. Brinda received a PhD in Social Anthropology from the University of Cambridge, where she conducted research on nomadic identities and trade and barter in the Himalayas.

Brinda.Dalal@parc.com

Les Nelson is a senior research scientist at the Palo Alto Research Center. His research interests include social computing, mobile computing, ubiquitous computing, and in general the adaptation of sociotechnical systems in response to changing circumstances.

lesnelson@acm.org

Nathan Good is a research scientist at Palo Alto Research Center. His research interests include recommender systems, usable security, and mobile computing.

nathaniel.good@parc.com

Ame Elliott is a Senior Human Factors Researcher at IDEO, where she designs innovative technical solutions grounded in human need. She has a PhD in Architecture from the University of California, Berkeley, where she studied human-computer interaction. Prior to joining IDEO, she was a research scientist at PARC and Ricoh Innovations.

aelliott@ideo.com

PEOPLE INCREASINGLY RELY ON THEIR ability to access and share data in order to get their jobs done and to enrich their personal lives, yet corporate security policies around sharing are rarely effective in enabling their users to achieve their goals. We offer our observations on how policy and practice often clash, as well as suggestions for improving the security of file sharing.

We wanted to understand how users are sharing information and how their needs are or are not met by current tools, policies, and practices. We performed an ethnographic field study, interviewing a selected group of subjects about their practices around access control, security, and file sharing. Our intent was to understand three things: (1) Under what circumstances do people or companies share or restrict access to files? (2) What tools or behavioral practices do they use to accomplish that? (3) How are people's experiences, problems, and needs changing in regard to secure file sharing and access control, especially as they deal with geographically dispersed colleagues, clients, friends, and family members?

Background

Our research builds upon a growing body of literature on file sharing and access control. Previous studies have focused on personal file sharing, specifically, in the domains of photographs [1, 2] or music [3, 4], or professional collaborations in corporations [5, 6]. Overall, these studies concluded that current tools for managing sharing policies available in each domain were inadequate to meet their users' complex requirements [1, 2, 5, 6, 7].

In the corporate setting, email was routinely chosen as the preferred means of sharing files, even in the presence of other alternatives [5, 6]. However, both the studies of Voida et al. [5] and Whalen et al. [6] considered subjects selected from and predominantly sharing within single organizations, all of whom shared access to established file-sharing mechanisms (e.g., file servers). They did not consider the effect that these preexisting options had on the challenges users would face, or the choices they would make when sharing across organizational boundaries or operating in the absence of pre-existing shared infrastructure.

Our study focuses on the question of how users share content in the absence of existing shared infrastructure—for example, when sharing across organizational boundaries. Based on interviews of users across various domains, we were able to explore access control and sharing issues across different types of organizations, such as those with stricter or more lax regulations and compliance policies. We examined in some depth how file-sharing and access controls were used, not used, or circumvented in order to get work done. From this analysis we identified key challenges faced by those using and choosing among current file-sharing technologies, including email.

Our Study

We conducted two-hour, in-depth interviews with ten subjects in places where they did at least some of their work—homes, home offices, or cafes. These interviews consisted of semi-structured and open-ended questions about file sharing and access control. We selected participants with file-sharing and access-control challenges, such as having to work with multiple clients from different organizations or having to share data with geographically dispersed teams or with those who need access to confidential data. Altogether we interviewed six men and four women between 23 and 53 years of age, working in a variety of fields. All used multiple digital devices and traveled frequently. Subjects were asked to describe examples of their professional and personal practices around security, privacy, and file sharing.

Key Findings

From our data, we identified a number of key problems users face in sharing data:

Sharing with myself: Users are their own most common sharing partner, effortfully moving data among their own machines, accounts, and devices to ensure continued access.

Oversharing: Users grant more access than necessary when it is difficult to limit who has access to content or how much to share with others, or when pressed for time to extract information from larger data sets.

Transient data: Users often need to hold data only briefly while transporting it from one place or another, and that data may linger, be lost, and get forgotten.

Transient access: Users need to access data for only short periods of time—they intend only one-time access or to make data available in certain situations.

Impedance matching: Users spend considerable time and effort tailoring content for sharing based on their understanding of recipient needs or the demands of the sharing mechanisms in use.

Ad hoc sharing: Users often share content with groups of recipients they have not shared with before and may not again.

Based on these insights, we propose that the general nature of the problem faced by users is what we term *ad hoc guesting*: Users need to share data securely with unplanned sets of people with whom they have not previously shared. They may belong to another organization and thus cannot be “named” by traditional access control. These interactions are transitory and lightweight, often making it not worth the effort required to set up new sharing mechanisms or change administrative state.

In what follows, we quote directly from our respondents. Explanatory material (words garbled in transcription or context missing from a quote) is provided in square brackets ([]).

General Properties of Sharing

Our study highlights distinctions between personal and professional sharing. Professionally, 80% of respondents shared files with overseas collaborators or clients in Europe and the Asia-Pacific region, and 100% exchanged data with colleagues across the United States. When working from home, consultants and employees in larger corporations often shared files via distributed corporate servers, and, in three cases, on protected FTP sites. Predominantly, the data shared in professional settings revolved around project work: Shared documents included technical specifications, meeting minutes and action items, and proposals. One of the primary affordances of using a shared server within a company was the ability to reuse documents from one project to another. At the same time, people found it time-consuming to browse different versions of documents to find the proposal they wished to reuse and resorted instead to telephoning or emailing their colleagues to obtain the appropriate copy. On the whole, we found that individuals are deeply aware of and attempt to comply with security stipulations and privacy requirements for their clients and companies. However, compliance breaks down the instant that people perceive that they are unable to follow policy without compromising their accountability to clients or colleagues or their ability to complete a task.

In contrast, people's personal file sharing practices focused on ways experiences could be shared with others. The content being shared in this case—primarily multi-media—was relational in nature, such as sharing photographs of events with family members who live overseas. We also found that a surprising number of people shared the same personal account. For instance, relatives scattered across the United States used a photo sharing account that had a single login and password to ensure privacy. Another set of parents set up a "family email account" and used email messages within the same account to discuss homework with their children in the evening.

All respondents used email to share files. Fully 90% of subjects mentioned that they had multiple email accounts (largely personal accounts) and 80% said that they used personal email accounts for business.

A total of 80% of respondents, regardless of their demographics, also used a wide variety of social software, including wikis, blogs, social networking sites (including MySpace and Facebook), hosted services (such as Yahoo! Briefcase), public Web sites for sharing images and multimedia files (including Flickr and YouTube), and online forums and games.

People Are Their Own Best Friends

People are their own most common sharing partners. File sharing with oneself allows one to synchronize activities regardless of location (at work, while traveling, or at home) or what devices or network resources are accessible. For example, interviewees who did not have a printer at home often uploaded files to Yahoo! Briefcase, then downloaded and printed files out at their office. Eighty percent of the respondents used USB drives (rather than laptops) to download content at client sites, especially when policies required that they contact IT administrators before accessing electronic files.

Email is a convenient and preferred mechanism for sharing files with oneself. Often quicker and easier than accessing files via a VPN, it is often used to bridge home and work. Although this served a short-term need, people said they later ran into trouble trying to track source documents and different versions across their accounts: "I'll go home and look for a file and have to go through all the emails with no subjects and [ask], 'Oh, when did I email myself this file?'" Most respondents had multiple email accounts (some up to 12 or 15). Different types of content were filtered into different accounts—work, friends, dating services, rental businesses, family photographs, or spam. Professional and personal accounts bled into one another, opening avenues for significant security lapses. When email or corporate servers were inaccessible, people readily sent files to consultants using their personal email accounts.

Oversharing

Our subjects often found themselves forced to overshare—to share too much or to share inappropriate information with others or themselves. Oversharing often occurs when it is simply too difficult to share only the information needed. For example, a healthcare consultant noted that when she visits a client site, she lacks sufficient time to go through the database (which she is not permitted to access remotely) and extract only the records she needs. Instead she ends up downloading entire files, including social security numbers, onto USB drives. She remarked, "There are a lot of rules trying to get permission from state agencies [to access confidential data]. A lot of data really is protected, so a lot of times the only effective way for me to do the work really disturbs me. Like I can't get permissions, but I can dump huge amounts of data on flash drives that I can then [in theory] lose."

Our subjects also reported that although the initial decision of whether or not to share data was often well-considered and even heavily regulated, once that decision was made "everything relaxes"—in other words, sharing decisions are often all or nothing. For example, one respondent noted, "They say there's no way we can provide this information, HIPPA won't allow it—it's research, it's clearly protected, but then you get past that point and you have a data sharing agreement and they'll dump a bunch of stuff [people's social security numbers, date of birth] on a disk and mail it to a name they've never heard at an address on 29th Street, which strikes me as weird."

Managing Transient Data

Users frequently handle "transient data"—data useful for a single task or short period of time. Transient data is often copies created as a side effect of transporting data from one location to another (e.g., copies on USB flash drives or in emails to oneself or others). For example, one individual remarked that she had a shoebox full of USB drives. Other respondents reported having anywhere between 2 and 15 active or inactive flash drives stored in their cars or briefcases, at work, or at home. The "throwaway nature" of temporary storage and devices makes it difficult to remember where sensitive data has gone. Not only is it hard to find the most recent version when it is needed, but afterward it often languishes, unremembered, on such devices forever.

Users dealt with such "throwaway data" at different levels of granularity, up to and including entire accounts or identities. Respondents increasingly lacked the time to manage their many email accounts, and they tended to shed entire accounts and open new ones rather than sort, archive, or destroy private data.

Transient Access

A number of individuals noted a need for transient access to data. Consultants, for example, were only supposed to have access to client data during the period of their contracts or while working in a certain environment. People also wanted to grant temporary access in the personal domain—for example, a landlord wanted to make rental property photos accessible only when the property was available, and some users shared their passwords for photo-sharing sites with others to whom they really only wanted to grant one-time access to pictures.

Users often made data available when temporary access was needed, even at the cost of security. Most commonly, respondents made data temporarily available via personal email accounts, when unable to access their work email because of VPN difficulties. One financial analyst noted, "You're not supposed to use 'unprotected email addresses,' like Yahoo!, Gmail, or whatever, but it's just a fact of life. Even our CEO has a Gmail account. There are work requirements but there is also an undeniable fact that people will be attracted to whatever is out in the market." A geotechnical engineer commented, "There are policies against sending clients electronic documents of any sort. But it's just so backassward that no one can possibly adhere to it."

Unfortunately, it can be difficult to go back and "fix" unwanted lingering access, as with one respondent: "But that pretty much is just a few phone calls, desperate phone calls saying, 'Delete from your servers; delete from your company; make sure it's completely clean.' You're at the mercy of hoping they follow your request."

Impedance Matching

Possibly the most interesting and significant challenge faced by our subjects was impedance matching—they were expending tremendous effort figuring out how, rather than what, to share.

People have varying degrees of technical skills required to use systems; consequently, there is a disparity in their ability to master the details involved in moving data around. Often those with greater need were forced to shoulder the work to obtain or share files. As one subject reported about a newly installed Web-based repository, "I think we have folks with very limited technical comfort. So for that reason I always have to upload my files [to the repository] and then email them around, so it's sort of another step rather than saving a step."

A major concern among respondents was preventing data sharing failures. The majority of our subjects spent time anticipating their own and their recipients' current state—the speed of their network connection, the sharing mechanisms available to all peers, and familiarity of the individuals involved with those systems—and changed their actions according to the (assumed) result. Sometimes this anticipatory work had to do with what the recipient was explicitly allowed to access: "I need to know other people's permissions, and I need to know because there's usually a lag between the permissions and the actual access; even if I have permission, it might take me six weeks to get my approval for the system, so I need to know where people are on the sort of really ridiculous timeline."

More often, users were concerned about working around constraints in bandwidth, network availability, or storage. Fully half of our interviewees expressed frustration in sending or receiving large files. Some specifically mentioned that their personal accounts or corporate email could not handle files over 10 MB. A design consultant who provides audio-visual material to his clients was exasperated by the effort it took to reformat content for clients: "It's absolutely absurd in this networked economy that we can't share [large] files without going to some extreme effort." People spent considerable time reformatting data for others, based on two parameters. First, they anticipated the constraints of their own or a recipient's system (such as capacity or bandwidth); second, they anticipated the recipient's sociotechnical knowledge regarding his or her ability to receive data. A software engineer explained the reasons he compressed photographs for his relatives: "A lot of my relatives are not very techie, so I'll just put photographs in an email attachment. I try to compress them so they are small jpeg sizes and then all people have to do is just click [on the images]." Another respondent drew a similar distinction: "When you're trying to share with family or friends the speed of the network really decides whether you can share five photos or just one."

The need for impedance matching means users are forced to decide between sharing modalities based on whether the sharing mechanisms will work with a particular user (do they have X or are they on Y?) or piece of content (is the file too big?) or what sharing mechanisms work best with that user (can they be counted on to log onto a separate system?). Equally importantly, how well can you gauge the accuracy of your assumptions about another's state (can they even receive your files?)? It is clear that the onus of work currently resides on users rather than on the systems they use.

The result of this impedance matching work is an overwhelming fear of failure. Users select the simplest, "safest" mode of sharing—email—because it is most likely to work in all settings. They only move beyond it when some constraint, such as file size or cultural pressure, forces them to.

Ad Hoc Sharing

We can divide the data sharing performed by our subjects into two types: repeated sharing, which occurs multiple times with the same set of participants, and ad hoc sharing, in effect sharing with strangers, which is sharing with people you haven't shared with before and whom you don't know whether you will share with again. If you are going to share repeatedly with a particular group of people, it might be worth doing some up-front work to improve the sharing experience—setting up a server or making sure everyone involved learns how to use a particular service. But it turns out to be harder than you might expect to know when sharing is going to be repeated: except when sharing with oneself or with a stable work group, or possibly with family or friends, sharing tends to be ad hoc. Even closeness or stability of real-world relationships does not serve as a good predictor of future sharing. Digital sharing among real-world connections is still not universal, so although you may expect your family and friends to continue to be connected to you, you may not know how often you are likely to share documents or media with them in the future.

For example, consultants in our subject pool were forced to establish new sharing mechanisms for each new customer engagement. Often prevented from using email by the file sizes involved, they were often expected by their clients to provide secure but provisional electronic sites on which to store interim data or final reports.

Implications for Design

Our findings led us to identify a number of common sharing tasks that are undersupported by the current tools available to users: sharing with themselves, managing transient data, providing transient access, and a general class of sharing problems we term *ad hoc guesting*. The latter refers to the problem commonly faced by our subjects of sharing data with new and unplanned sets of people (unplanned either by themselves or their respective organizations), often without assurance that they would ever share with that group of people again.

Currently, users preferentially and almost overwhelmingly turn to email to solve this problem, except when their impedance matching processes indicate that email is unlikely to be successful. To be successful, alternatives to email must reduce this impedance matching burden—they must be so universal and easy to use that it is worth using them even for ad hoc or one-time sharing.

Organizations wishing to move their users from email onto more secure forms of sharing might find the most effective approach is to enable them to easily share more things, rather than "locking things down" in an effort to get them to share less. In current and future work, we are focused on designing new technologies that effectively balance these tensions between usability and security.

REFERENCES

[1] S. Ahern, D. Eckles, N.S. Good, S. King, M. Naaman, and R. Nair, "Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* pp. 357–366, 2007.

[2] A.D. Miller and W.E. Edwards, "Give and Take: A Study of Consumer Photo-Sharing Culture and Practice," *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* pp. 347–356, 2007.

[3] B. Brown, A.J. Sellen, and E. Geelhoed, "Music Sharing as a Computer Supported Collaborative Application," *Proceedings of the Seventh European Conference on Computer Supported Cooperative Work,* pp. 179–198, 2001.

[4] A. Voida, R.E. Grinter, N. Ducheneaut, W.K. Edwards, and M.W. Newman, "Listening In: Practices Surrounding iTunes Music Sharing," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* pp. 191–200, 2005.

[5] S. Voida, W. Edwards, M.W. Newman, R.E. Grinter, and N. Ducheneaut, "Share and Share Alike: Exploring the User Interface Affordances of File Sharing," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* pp. 221–230, 2006.

[6] T. Whalen, D. Smetters, and E.F. Churchill, "User Experiences with Sharing and Access Control," *CHI '06 Extended Abstracts on Human Factors in Computing Systems,* pp. 1517–1522, 2006.

[7] J.S. Olson, J. Grudin, and E. Horvitz, "A Study of Preferences for Sharing and Privacy," *CHI '05 Extended Abstracts on Human Factors in Computing Systems,* pp. 1985–1988, 2005.