Scaling Measurement Experiments to Planet-Scale: Ethical, Regulatory and Cultural Considerations

Tristan Henderson
School of Computer Science
University of St Andrews
St Andrews, Fife, UK
tristan@cs.st-andrews.ac.uk

Fehmi Ben Abdesslem
School of Computer Science
University of St Andrews
St Andrews, Fife, UK
fehmi@cs.st-andrews.ac.uk

ABSTRACT

Conducting planet-scale mobility experiments and measurements is of great interest to network researchers for building the next generation of wireless networking technologies, or for studying interdisciplinary problems in complex networks. There are many technical challenges that need to be addressed before such experiments can take place. But at the same time, there are many non-technical issues that need to be tackled in order to preserve the welfare of participants in these studies. While some of these issues have been addressed in previous small-scale studies, they become increasingly complex when differences between countries need to be taken into account.

This position paper highlights some of these issues and argues that they need to be addressed before planet-scale measurement experiments can be conducted. We discuss ethical, regulatory, cultural and privacy issues, and consider how to design measurement systems that will scale up to planet-wide experiments. We motivate our approach by discussing work in measurement of mobile and online social networks.

1. INTRODUCTION

Planet-scale mobility measurements may be useful for many research topics in several disciplines, such as opportunistic and delay-tolerant networks, epidemiological studies and urban planning. Collecting, analysing and interpreting such measurements involves many challenges, some of which are technical and can be addressed by the networking research community. But many others involve sociological, legal or ethical issues which must be considered before the technical challenges can be addressed. This position paper argues that we must engage experts from other research communities to help tackle these problems so that we can conduct planet-scale measurement in an efficient manner. In particular, we argue that the differences between countries in terms of ethics, regulation and culture mean that designing experiments which can be conducted successfully across countries requires careful planning.

This paper is outlined as follows. In the next section, we describe some privacy threats that impact planet-scale measurement, and how this impact may differ between countries. We then look at differences in ethics, regulation and culture between countries, and discuss how these might affect how we enable participation in planet-scale measurement studies. In Section 7 we outline some design principles which might help us engineer planet-scale systems that take these differences into account, and provide examples of some ongoing work. Finally we conclude in Section 8.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ACM HotPlanet '09 Krakow, Poland
Copyright ACM ...$10.00.

2. PRIVACY CONSIDERATIONS

Any planet-scale mobility measurement studies must ensure the privacy of those being measured. But privacy is a broad term, and the specific privacy threats must be characterised in order to be understood and addressed. Solove presents a comprehensive taxonomy of privacy violations in [31]. The main threats that apply to planet-scale measurement are those due to information collection: specifically, *surveillance* and *interrogation*. Surveillance can cause discomfort, and can also cause people to alter their behaviour. This could potentially impact the quality of the data collected through planet-scale measurement: if participants behave differently because they are being measured, then the data being collected and any resulting models or analysis may not be representative of normal behaviour. Planet-scale measurement data could also be used for interrogation; e.g., a mobility trace being used as evidence of a particular person's whereabouts. Solove also highlights several information processing threats which apply to measurement data: such data could be used to identify particular users, or aggregated with other collected data to violate privacy in manners unforeseen by those conducting the measurement study.

We should therefore design planet-scale measurement systems with privacy in mind. But attitudes to, and definitions of, privacy may differ between countries. The annual Privacy International survey [14] indicates significant differences between countries in Europe and beyond, in terms of constitutional protection of rights, implementation of identity cards, the level of government communication interception and access to data, and surveillance such as CCTV and workplace monitoring. A measurement study that is considered acceptable in one country may incur the wrath of participants and the media in another [30, 21].

3. REGULATORY CONSIDERATIONS

Conducting planet-scale measurement studies, and in particular the data being collected through such studies, may need to take legal requirements into account. Designing a study that is legal in an individual country may be simple, but is it possible to do so on a planet-wide scale?

The main area of legislation which applies to network and mobility measurement studies is that of data protection. Such legislation governs the handling of personal data and provides and protects the rights of those individuals whose data are being col-

lected. Data protection legislation, like most legislation, varies between countries. The EU Data Protection Directive (DPD) is implemented by individual EU states, for instance the UK's Data Protection Act [33]. The DPD outlines three principles for data protection which apply to the collection of personal data, where personal data are defined as "any information relating to an identified or identifiable natural person". These principles may apply to planet-scale measurement studies:

- *Transparency*: the persons whose data are being collected or accessed have the right to be informed when such data processing is taking place. Participants in a measurement study should therefore be aware, for instance, when their mobile phones are collecting data and what data are being collected.
- *Legitimate purpose*: data can only be collected for specific purposes. Studies should be careful to only collect those data that are needed — for instance, a study that is only interested in encounter patterns might only need to collect Bluetooth addresses and not Bluetooth names, since the latter might expose personal information.
- *Proportionality*: data should be processed in a fashion that is not excessive beyond the purposes for which they were collected. This principle has particular implications for the reuse of measurement data — do data archives such as CRAWDAD [19] need to monitor how their provided datasets are used?

While the DPD principles must apply to studies conducted within the EU, there may be cases where a researcher outside the EU wishes to conduct a study that does not obey these principles, for instance a study that collects data without consent [29]. The researcher would have to redesign their study and any measurement systems to scale such a study to planet-scale and include the EU.

Another way in which the DPD has an effect beyond the EU is through its stipulations on the transfer of data to states outside the EU ("third countries" in EU terminology). The Safe Harbour agreement enables the transfer of data to the US, if the US institution is able and willing to sign such an agreement. Even this agreement is no guarantee that data will be protected; some institutions who participate in Safe Harbour do not fulfil all of the requirements [12].

Moreover, while there has been much work studying EU-US data transfer [32, 27], what if no data transfer agreement exists for a particular country? For instance, Ligertwood studies the implications of transferring data between the EU and Australia [22]. How can we share measurement data on a planet-wide scale? Or will a planet-wide measurement study comprise a number of jurisdiction-specific smaller studies, with data never being exchanged between the various researchers?

One technical solution might be an equivalent of IEEE 802.11d [17].[1] Measurement systems could signal users of the regulatory requirements, so that mobile devices can be reconfigured accordingly to prevent the capture of illegal data. This might be useful in the cases where the users being measured move between jurisdictions.

Another solution might simply be to anonymise and sanitise any collected data at the point of collection, such that any personal identifying data are removed. This of course introduces new technical considerations, such as the impact on energy consumption, but also larger research considerations in terms of the tradeoff between the utility of the data and the level of sanitisation [26, 8]. It might also be impossible to sanitise data sufficiently — for instance it is difficult to "anonymise" location since most transforms can be reverse-engineered.

[1]The 802.11d standard is responsible for broadcasting regulatory information to 802.11 network users so that they can configure their NICs accordingly.

4. ETHICAL CONSIDERATIONS

Planet-scale mobility measurement clearly involves the study of human subjects, and thus any studies must take care to protect the welfare of these subjects. Traditionally such work must be approved by a research institution's ethics committee or IRB (Institutional Review Board). But how can we do this on a planet-wide scale?

There are again country-specific differences in ethical considerations for research. In the US, ethical approval tends to be considered in a utilitarian setting: the need to conduct the research and the outcomes of the research are considered most important. In Europe, on the other hand, a deontological approach predominates: the rights of the participants are more important than the rights of the researcher [2]. Therefore a study that is approved by an IRB in the US may not be approved in Europe, and vice versa.

Is it possible to achieve a compromise between these two systems, and indeed the many other systems that may exist across the planet? Can we decide on a common ethical system or framework under which to conduct planet-wide studies? Is this a desired outcome? Or again, as with regulatory requirements, perhaps different types of measurement might occur in different jurisdictions, so as to comply with jurisdiction-specific ethical systems.

5. CULTURAL CONSIDERATIONS

Even once all of the regulatory considerations have been taken into account, there may be differences between countries that are less tangible. Societal and cultural attitudes may differ between countries that have an impact on measurement systems.

The way in which mobile technologies are used varies across the globe. For instance, in Indonesia it is common to carry more than one mobile phone, while in Malaysia, mobile phones are shared amongst family members [5]. Measurement studies which assume that a single device represents an individual user would have very different results in these two countries. Finns are significantly more likely than Spaniards to turn away from nearby people when taking a phone call [16] — this may impact mobility patterns if some users move whenever they receive a call.

Selecting hardware for a measurement system may also be affected by cultural differences. Using a separate measurement device is fraught with difficulty [15, 3] and so we may choose to leverage users' existing devices such as mobile phones. But which mobile phone platform would be best to target for a planet-wide study? The iPhone may be popular in North America and so a useful platform for studies there, but not so in Europe or Japan [28, 11].

Cultural differences may also manifest themselves in areas besides human behaviour. The Google Street View measurement system, which comprises cars with roof-mounted cameras, was first deployed in North America and was able to successfully take photographs of public places. The system design did not take into account the differences in architecture between America and the rest of the world, however. When this system was deployed in Japan, it was found that many photographs were being taken of private residences: the cameras were mounted such that they could take photographs over the lower Japanese walls. As a result, Google had to repeat its Japanese measurement — costing time and money, as well as being a public relations embarassment [4].

6. PARTICIPATION CONSIDERATIONS

One major challenge to the deployment of a planet-wide measurement infrastructure is collecting high-fidelity and representative data. The data may be lossy if the devices used for measurement fail to capture data accurately. The data may also be inaccurate if the devices cannot be used by participants, for instance due to complicated user interfaces or insufficient battery life. The data may be unrepresentative if groups or sections of society are omitted: a study that relies on existing users' mobile phones will fail to capture data from anyone who does not own a mobile phone, while a study that leverages a social network application may omit any users who are not part of that particular social network.

So how can we create the appropriate incentives to enable planet-scale monitoring and experiments? There need to be appropriate incentives for participants, as well as for the data collectors themselves. If, for instance, mobile phone operators are to enable data collection by providing network or subscriber access, then they need sufficient motivation to do so. Similarly, users somehow need to be motivated to participate in measurement studies, for instance through payment in cash or in kind, or access to new applications. But what ethical and regulatory considerations are introduced by creating such incentives? Will we have accurate data if people know that they are being monitored, or if they have financial incentive to behave in particular ways?

There may be additional ethical considerations if we attempt to measure particular sectors of society. For instance, a researcher may choose to provide mobile phones to users who have never used such mobile technologies before. Such users may be willing to provide personal data in exchange for access to new technology, but is such an exchange ethical? And will the collected data be accurate?

7. DISCUSSION

The previous sections have discussed many inter-country differences which may complicate planet-scale measurement studies: both in terms of designing and conducting the studies, but also in interpreting and generalising the data collected through such studies. Many of these issues need to be addressed in an interdisciplinary fashion, since the solutions may lie outside of the computer science or networking research community. But in the interim, what can our research community do?

We propose that any planet-scale measurement studies should consider two simple design principles:

- *Scalability*. By this we do not mean that our measurement systems should be able to scale to millions of users. While this is clearly necessary for planet-scale measurement, another measure of scalability is that measurement systems should scale to *multiple countries and jurisdictions*. A measurement system that works in the researchers' home country may not necessarily work in other countries because of assumptions about ethics, regulation or culture. These considerations need to be placed at the forefront of experimental system design and any systems should be flexible enough to cope in different countries. If possible, systems should be tested in more than one country before deployment.

- *User-centered design.* User-centered design [25] holds that systems should be designed based on user needs, rather than other metrics such as aesthetics or performance. We believe that the same holds true for planet-scale measurement systems. If we take a deontological ethical approach to measurement, then user needs must be paramount. Systems should only measure what users want to be measured, and users should be able to reconfigure what is being measured in order to reflect their personal preferences. One mechanism for doing this might be to use "web 2.0" technologies to allow users to signal their privacy preferences and share experiences with each other [9].

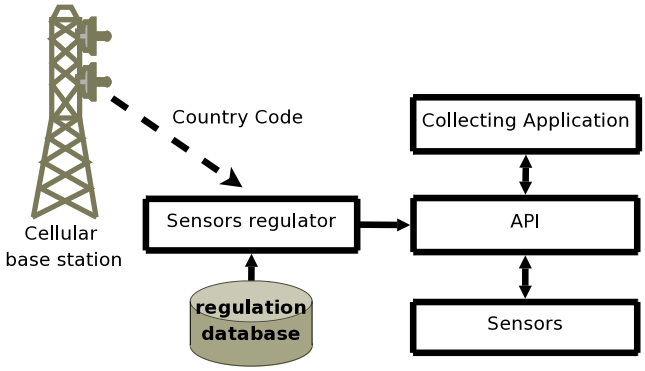

Figure 1: Local regulation-aware architecture for mobile phones to use in planet-scale experiments.

We have been attempting to apply these design principles to our current research, which we now discuss.

7.1 Case study: mobile/participatory sensing

Mobile phones have been increasingly used in the last few years to conduct research involving data collection from human subjects [13, 24, 6, 23, 18]. Two main reasons for using mobile phones instead of special sensor devices are that (i) mobile phones are now common in everyday use almost all over the world with over 2 billion devices [1]; (ii) more and more sensors are now embedded in mobile phones.

Instead of specific sensor devices such as the iMotes used to collect mobility and measurement data in previous experiments [10, 20], mobile phones are already used by people, even outside experimental contexts. Participants are therefore familiar with these devices, making their usage easier and limiting any experimental impact on user behavior, since many people already carry and use such devices everyday. Deploying a testbed of mobile phones for a planet-scale experiment is thus more convenient: participants can download software prepared by the researchers on their own mobile phone, and start collecting data. At the end of the experiment, the data are then uploaded to researchers' servers for analysis, for instance by using data transmission through a cellular network.

The second reason for using mobile phones in planet-scale sensing experiments is mainly due to advances in electronic engineering and to customer needs of advanced features in their device. This results in the presence of multiple sensors embedded in off-the-shelf mobile phones: in addition to a microphone, such phones may include one or more cameras, accelerometers, GPS receivers, 802.11 and Bluetooth radios. Exploiting these sensors opens up new opportunities to collect data on the users carrying them all day long, and studying their behaviours and characteristics, such as mobility patterns, proximity to other people, usage of the devices, privacy concerns, or location.

As discussed in Section 3, deploying these mobile phones in different countries needs to take local regulations into account when sensing and collecting data. For our mobile phone experiments, we are developing a "regulation controller" as part of our experimen-

Table 1: User information accessible through APIs of top 5 social network sites (SNS).

SNS	*User info (e.g., name)*	*Status*	*Friends*	*Friends of friends*	*Photos*
Facebook	√	√	√	×	√
Windows Live	×	×	×	×	√
MySpace	√	√	√	√	√
Hi5	√	√	√	√	√
Twitter	√	√	√	√	N/A

tal measurement system (Figure 1). This controller retrieves the country code, as broadcasted by cellular base stations, and queries a local database to determine regulation limitations in the current country. To retrieve data from the sensors, the actual collecting application uses an API customized by the regulation controller, that blocks the appropriate sensor functionalities to prevent infringing local regulations when collecting data. By implementing this controller, we hope that our measurement system should scale to more than one regulatory jurisdiction.

7.2 Case study: mobile social networks

Another active research area is the study of social network sites (SNS).[2] SNS such as Facebook are used by hundreds of millions of users around the world and as such might be leveraged as an off-the-shelf platform for conducting planet-scale measurements. But the open access policies and REST APIs provided by these SNS leave users open to the privacy threats described in Section 2. Applications installed by SNS users rely on REST APIs that provide the application owner with personal information about these users. Table 1 shows the ease by which a researcher can surveil SNS users for the top 5 SNS.[3] Facebook's API can only provide a list of friends for the current user (using the application), but cannot access the list of friends of these friends, even if these friends of friends can be found when browsing the website. Windows Live Spaces also provide a REST API, but only to access photos. Hi5 and MySpace both implement the OpenSocial[4] APIs, and hence are interoperable. Both can provide access to most data about the user. Since there are no photo features in Twitter, the API does not provide access to such information.

Such APIs to collect data planet-wide can be used with mobile phones (as in the previous case study) in a planet-scale testbed. For instance, we are currently developing *SpyMe*, a testbed for studying privacy in mobile social networks. As depicted in Figure 2, SpyMe first receives the regulation-compliant data from the phones and stores them in a central database. An activity-inferencing module infers the activity of users according to these sensed data. Part of the data are then displayed on a social network system such as Facebook or Twitter using the REST APIs, and retrieves other data from these SNS.

Our system is reconfigurable via SMS messages which can determine what data are collected and when. To ensure that users have control over the measured data, we allow the measurement system to be configured by the users. We are also deliberately cautious in the amount of data that we collect: for instance, we only collect data from a participant's immediate ("1-hop") social network, and participants have to request permission from these friends before data are transferred. If we want to collect data from additional hops (i.e., "friends of friends") then permission must be sought. Since social networks change over time, we check social network makeup every 24 hours to ensure that permission has been obtained from all of the users being measured.

[2]Some researchers use the term online social networks (OSNs); we prefer to use SNS following [7].

[3]We chose the top 5 SNS according to alexa.com, as retrieved on 24 May 2009.

[4]http://www.opensocial.org/.

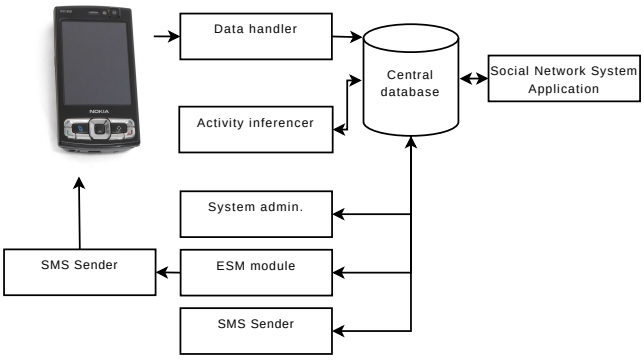

Figure 2: SpyMe testbed server architecture.

8. CONCLUSION

This position paper has described some of the differences between countries which may make it difficult to scale a measurement study from a single jurisdiction to a planet-wide scale. These include differences in regulation, ethics and culture. We believe that designing solutions that can deal with these differences requires an interdisciplinary research effort involving lawyers, sociologists, ethnographers and others. We should also attempt to learn lessons from other data collection systems which share data across countries, such as air traffic control or telecommunications networks. But in the interim, we propose that measurement system designers should consider how their systems will scale to beyond their particular measurement venue, and that they should enable participants to control the measurement to reflect personal preferences or jurisdiction-specific requirements.

9. ACKNOWLEDGEMENTS

This work is supported by the EPSRC *Privacy Value Networks* project, EP/G002606/1. The issues raised in this paper benefited from discussions with our colleagues Saleem Bhatti, Martin Bateman and Devan Rehunathan. We thank the reviewers for their useful comments.

10. REFERENCES

[1] C. I. A. *The World Factbook*. Central Intelligence Agency, Washington, DC, 2008.

[2] G. Allen, D. Burk, and C. Ess. Ethical approaches to robotic data gathering in academic research. *International Journal of Internet Research Ethics*, 1(1):9–36, Jan. 2008.

[3] I. Anderson, J. Maitland, S. Sherwood, L. Barkhuus, M. Chalmers, M. Hall, B. Brown, and H. Muller. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Networks and Applications*, 12(2):185–199, June 2007.

[4] Associated Press. Google to reshoot street views of Japanese cities, 14 May 2009. http://www.japantoday.com/category/technology/view/google-to-reshoot-street-views-of-japanese-cities, accessed 2009-05-19.

[5] G. Bell. The age of the thumb: A cultural reading of mobile technologies from Asia. *Knowledge, Technology, and Policy*, 19(2):41–57, June 2006.

[6] F. Ben Abdesslem, A. Phillips, and T. Henderson. Less is more: Energy-efficient mobile sensing with SenseLess. In *Proceedings of ACM SIGCOMM Mobiheld workshop*, Barcelona, Spain, Aug. 2009.

[7] D. Boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, Oct. 2007.

[8] J. Brickell and V. Shmatikov. The cost of privacy: destruction of data-mining utility in anonymized data publishing. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp 70–78, Las Vegas, Nevada, USA, Aug. 2008.

[9] E. Buchmann, K. Böhm, and O. Raabe. Privacy2.0: Towards collaborative data-privacy protection. In *Trust Management II*, volume 263 of *IFIP International Federation for Information Processing*, pp 247–262. Springer-Verlag, Boston, MA, USA, 2008.

[10] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6(6):606–620, June 2007.

[11] B. X. Chen. Why the Japanese hate the iPhone, 26 Feb. 2009. http://blog.wired.com/gadgets/2009/02/why-the-iphone.html, accessed 2009-04-11.

[12] C. Connolly. The US safe harbor - fact or fiction? *Privacy Laws & Business International*, (96), Dec. 2008.

[13] N. Eagle and A. S. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, May 2006.

[14] Electronic Privacy Information Center and Privacy International. *Privacy and Human Rights 2006*. Electronic Privacy Information Center, Dec. 2007. http://www.privacyinternational.org/article.shtml?cmd[347]=x-347-559458, accessed 2009-05-21.

[15] T. Henderson, D. Anthony, and D. Kotz. Measuring wireless network usage with the experience sampling method. In *Proceedings of the First Workshop on Wireless Network Measurements*, Trentino, Italy, Apr. 2005.

[16] J. Höflich. The mobile phone and the dynamic between private and public communication: Results of an international exploratory study. *Knowledge, Technology & Policy*, 19(2):58–68, June 2006.

[17] IEEE-SA Standards Board. IEEE Std 802.11d-2001 Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications amendment 3: Specification for operation in additional regulatory domains. Tech. rep., IEEE, July 2001.

[18] A. Kansal and F. Zhao. Location and mobility in a sensor network of mobile phones. In *Proceedings of the 17th International workshop on Network and Operating Systems Support for Digital Audio & Video (NOSSDAV)*. June 2007.

[19] D. Kotz and T. Henderson. CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth. *IEEE Pervasive Computing*, 4(4):12–14, Oct.-Dec. 2005.

[20] J. Leguay, A. Lindgren, J. Scott, T. Friedman, and J. Crowcroft. Opportunistic content distribution in an urban setting. In *Proceedings of the ACM SIGCOMM Workshop on Challenged Networks (CHANTS 2006)*, Pisa, Italy, Sept. 2006.

[21] P. Lewis. Bluetooth is watching: secret study gives Bath a flavour of Big Brother. *The Guardian*, 21 July 2008. http://www.guardian.co.uk/uk/2008/jul/21/civilliberties.privacy, accessed 2009-04-12.

[22] J. Ligertwood and M. Jackson. Transborder data protection and the effects on business and government. In *Usability and Internationalization. Global and Local User Interfaces*, pp 140–149, Beijing, China, July 2007.

[23] N. Marmasse, C. Schmandt, and D. Spectre. WatchMe: Communication and awareness between members of a closely-knit group. In *Proceedings of Ubicomp*, pp 214–231, Nottingham, UK, Sept. 2004.

[24] E. Miluzzo, N. Lane, K. Fodor, R. Peterson, S. Eisenman, H. Lu, M. Musolesi, X. Zheng, and A. Campbell. Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. In *Proceedings of ACM SenSys 2008*, Raleigh, NC, USA, Nov. 2008.

[25] D. A. Norman and S. W. Draper. *User-Centered System Design*. CRC Press, Boca Raton, FL, USA, 1986.

[26] V. Rastogi, D. Suciu, and S. Hong. The boundary between privacy and utility in data publishing. In *VLDB '07: Proceedings of the 33rd international conference on Very large data bases*, pp 531–542, Vienna, Austria, Sept. 2007.

[27] P. M. Regan. Safe harbors or free frontiers? Privacy and transborder data flows. *Journal of Social Issues*, 59(2):263–282, July 2003.

[28] J. L. Schenker. The iPhone in Europe - lost in translation. *Business Week*, 16 Apr. 2008. http://www.businessweek.com/magazine/content/08_17/b4081000500950.htm, accessed 2009-04-11.

[29] A. Singer. Conference password sniffing: Legal and ethical issues. *;login:*, 30(4):5–9, Aug. 2005.

[30] C. Soghoian. Researchers could face legal risks for network snooping, 24 July 2008. http://news.cnet.com/8301-13739_3-9997273-46.html, accessed 2009-04-12.

[31] D. J. Solove. A taxonomy of privacy. *University of Pennsylvania Law Review*, 154(3):477–560, Jan. 2006.

[32] R. M. Walczuch and L. Steeghs. Implications of the new EU directive on data protection for multinational corporations. *Information Technology and People*, pp 142–162, 2001.

[33] A. Warren and J. Dearnley. Data protection legislation in the United Kingdom. *Information, Communication and Society*, 8(2):238–263, June 2005.