Risk-Utility Paradigms for Statistical Disclosure Limitation: How to Think, But Not How to Act - Discussion: A Science of Statistical Disclosure Limitation?

Josep Domingo-Ferrer
Universitat Rovira i Virgili, Dept. of Computer Engineering and Mathematics,
UNESCO Chair in Data Privacy, Av Països Catalans 26,
E-43007 Tarragona, Catalonia. E-mail josep.domingo@urv.cat

Statistical disclosure limitation models

The final question raised in Cox *et al.* (2011) "Can there be a science of data confidentiality?" is an extremely honest and intriguing one, especially coming from authors who have long contributed to this discipline. It is evocative of Socrates's "I know that I know nothing". The history of data confidentiality, a.k.a. statistical disclosure limitation (SDL), statistical disclosure control (SDC) or statistical database privacy, started back in 1977 with the seminal paper by the Swedish statistician Tore Dalenius (Dalenius, 1977). As usual in pioneering work, Dalenius was not especially bothered by the scientific standing of the newborn discipline: he just stated a very ambitious privacy requirement, namely that anything that can be learned about a respondent having contributed to a statistical database should be learnable without access to the database.

For a number of years, new methods and results partially addressing the above goal were obtained, both in the statistical and the database communities. In the former community, Dalenius's ideas were partly formalized by Paass (1985) and Duncan and Lambert (1986) by leaning on probability theory. Also, a lot of practical progress in tabular data protection was made possible by the solid theory of tabular disclosure limitation established in the 1980s by Cox and other authors; this theory did not, however, deal with microdata, because microdata release was not regarded as an objective worth pursuing by statistical agencies of that time. Arisen in the database

community, the k-anonymity model (Samarati and Sweeney, 1998; Samarati, 2001; Sweeney, 2002) brought some conceptual order into anonymization of microdata by providing a simple yet powerful reference framework. The model made its way into the statistical community in Domingo-Ferrer and Torra (2005), where it was shown to be timewise and conceptwise parallel to multivariate microaggregation (Domingo-Ferrer and Mateo-Sanz, 1998 and 2002). However, k-anonymity did not aim to become a comprehensive theory of statistical database privacy: for one thing, it does not deal with tabular data and it has some privacy insufficiencies (Domingo-Ferrer and Torra, 2008).

Differential privacy (Dwork, 2006; Dwork, 2011) has been proposed with the ambition of bringing to statistical disclosure control the theoretical soundness of present day cryptography. Hence, if successful, differential privacy would allow answering Cox *et al.*'s final question in the affirmative. However, serious objections to the utility of differentially private numeric data have recently been brought forth (Muralidhar and Sarathy, 2010; Sarathy and Muralidhar, 2010; Sarathy and Muralidhar, 2011). In its current formulation, differential privacy seems to focus more on controlling risk than on preserving utility. Also, the data structures it considers are rather simplistic and idealized, compared to the real-life data encountered by the practitioner statistician (longitudinal data, movement data, administrative data, weighted data, linked tables, etc.), some of which are discussed in Cox *et al.* (2011).

Furthermore, both k-anonymity and differential privacy adopt a minimax approach to risk-utility conciliation: subject to a minimum privacy, utility is to be maximized. Beyond this minimax approach making utility a secondary goal at best, Cox *et al.* (2011) point out that it can lead to unstable solutions if privacy cannot be accurately measured.

Transparency

A major contribution of the paper under discussion is to shed light on transparency (that is, disclosure of SDL procedures used in anonymization), if the reader tolerates such a pun. Kerckhoff's assumption usual in cryptography, whereby encryption algoritms are public and only the keys are secret, is much thornier in SDL. Indeed, whereas in cryptography the legitimate receiver of a message and the intruder eavesdropping the channel are two different entities (with the intruder just seeing encrypted data), this is no longer true in SDL: the legitimate user/receiver of SDL-protected data is

a potential intruder too, perhaps interested in making illegitimate inferences about respondents using the *non-encrypted* anonymized data and any available side knowledge. To overcome this confusion between users and intruders, Cox *et al.* (2011) try to differentiate the typical analytical goals of a legitimate user and an intruder (integration vs maximum computation), but, as they acknowledge, theirs is more a conceptual distinction than an operational one.

As noted in Sebé (2003), in addition to SDL, there are other data protection disciplines, like steganography and watermarking, facing the same transparency debate. A common feature of these disciplines and SDL is the confusion between legitimate users and potential intruders. A digital content (*e.g.* image, tune, etc.) with an embedded watermark for copyright protection is in this respect parallel to an anonymized dataset; if the watermarking method and parameters are released, it might be easier for the legitimate user of the watermarked content to remove the watermark and obtain a version of the content without any copyright mark, ready to be unlawfully redistributed without any liability. The watermarking community attempts to satisfy Kerckhoff's assumption by developing tamper-resistant watermarking methods (see Cox *et al.* (2008) for background), in which only the embedding key (but not the method) remains secret and an intruder cannot hope to alter the embedded watermark without substantially damaging the perceptual quality of the content. How to import this tamper-resistance concept from watermarking methods to SDL methods is an open research issue, partly explored by the paper under discussion, where a partial classification is given between some methods which are unaffected by transparency (rounding and controlled tabular adjustment) and some which are weakened by it (noise addition and cell suppression).

Conclusions

SDL is a discipline which was born from daily statistical practice and any theory trying to bestow a unified scientific standing to it should be flexible enough to deal with the risk-utility trade-off in a rather vast number of situations (different data structures, different contexts of previously released information and intruders' side knowledge, etc.). How to accurately measure risk and utility, how to distinguish legitimate users from intruders and how far to implement transparency are some of the open challenges SDL faces today, most of which are insightfully highlighted in the paper under discussion.

It is our belief that cross-fertilization with related disciplines is a promising methodology to tackle some of those challenges. Just as transparency is a problem common to SDL and watermarking, dealing with longitudinal data is faced by official statisticians and by organizations interested in releasing anonymized trajectories of mobile objects in the contexts of location-based services and geographical information systems (*e.g.* see Bertino and Damiani (2009)). Increased communication and synergy across the various privacy research communities seems a sensible way to go.

References

Bertino, E., and Damiani, M. L. (2009). Foreword for the special issue of selected papers from the 1st ACM SIGSPATIAL Workshop on Security and Privacy in GIS and LBS. *Transactions on Data Privacy* 2(1):1-2.

Cox, L. H., Karr, A. F., and Kinney, S. K. (2011). Risk-utility paradigms for statistical disclosure limitation: how to think, but not how to act. *International Statistical Review*, this issue.

Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., and Kalker, T. (2008). *Digital Watermarking and Steganography*, Morgan Kaufmann, Burlington MA, 2008.

Dalenius, T. (1977). Towards a methodology for statistical disclosure control. *Statistisk Tidskrift* 5:429-444.

Domingo-Ferrer, J., and Mateo-Sanz, J. M. (1998). A method for data-oriented multivariate microaggregation. In *Proc. of Statistical Data Protection'98*. Office for Official Publications of the European Communities, Luxemburg, 1999, pages 89-99.

Domingo-Ferrer, J., and Mateo-Sanz, J. M. (2002). Practical data-oriented microaggregation for statistical disclosure control. *IEEE Transactions on Knowledge and Data Engineering* 14(1):189-201.

Domingo-Ferrer, J., and Torra, V. (2005). Ordinal, continuous and heterogeneous k-anonymity through microaggregation. *Data Mining and Knowledge Discovery* 11(2):195-212.

Domingo-Ferrer, J., and Torra, V. (2008). A critique of k-anonymity and some of its enhancements. In *Third International Conference on Availability, Reliability and Security-ARES 08*, pp. 990-993.

Duncan, G. T., and Lambert, D. (1986). Disclosure-limited data dissemination. *Journal of the American Statistical Association*, 81(393):10-18.

Dwork, C. (2006). Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP) (2)*, pages 1-12.

Dwork, C. (2011). A firm foundation for private data analysis. *Communications of the ACM* 54(1): 86-95.

Muralidhar, K., and Sarathy, R. (2010). Does differential privacy protect Terry Gross' privacy? In Domingo-Ferrer and Saygın, editors, *Privacy in Statistical Databases 2010*, volume 6344 of *Lecture Notes in Computer Science*, pages 200-209. Springer-Verlag, Berlin/Heidelberg.

Paass, G. (1985). Disclosure risk and disclosure avoidance for microdata. *Journal of Business and Economic Statistics* 6(4):487-500.

Samarati, P. (2001). Protecting respondents' identities in microdata release. *IEEE Transactions on Knowledge and Data Engineering* 13(6):1010-1027.

Samarati, P., and Sweeney, L. (1998). *Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression.* Technical Report, SRI International.

Sarathy, R., and Muralidhar, K. (2010). Some additional insights on applying differential privacy for numeric data. In Domingo-Ferrer and Saygın, editors, *Privacy in Statistical Databases 2010*, volume 6344 of *Lecture Notes in Computer Science*, pages 210-219. Springer-Verlag, Berlin/Heidelberg.

Sarathy, R., and Muralidhar, K. (2011). Evaluating Laplace noise addition to satisfy differential privacy for numeric data. *Transactions on Data Privacy* (in press).

Sebé, F. (2003). *Transparent Data Protection.* Ph. D. Thesis, Universitat Politècnica de Catalunya.

Sweeney, L. (2002). k-Anonymity: a model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, 10(5):557-570.