Privacy Preserving Data Mining:
A Process Centric View from a European Perspective[1]

Martin Meints and Jan Möller
Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein
Holstenstr. 98
24103 Kiel
{meints | moeller}@datenschutzzentrum.de

Abstract

Privacy Preserving Data Mining (PPDM) in a broad sense has been an area of research since 1991[2] both in the public and private[3] sector and has also been discussed at numerous workshops and international conferences[4]. Currently the research is mainly directed towards development of technical methods, such as application of cryptography or the development of specialised algorithms to meet security and privacy requirements for different data mining methods, such as classification or categorisation.

So far PPDM has found application in only a few cases. One example is documented in medical research to protect patients' privacy[5]. In all cases when data mining is applied in the context of personal data, basic data and data mining results have to be collected, stored and processed in compliance with data protection legislation. This results in responsibilities for data controllers, technical operators and others involved in those business or governmental processes where data mining plays a role.

In this article a brief overview of the state-of-the-art in PPDM and some current suggestions for proceeding towards standardisation in PPDM are summarised. This is followed by considerations of how PPDM could be improved based on the European Directive 95/46/EC, additionally taking into account procedural and process-related considerations. To illustrate these considerations, scoring practice in the financial sector is used as an example. Though this example certainly does not demonstrate all aspects possibly relevant in the area of data mining, it has been analysed from the perspective of recent data protection developments. In addition, with process chains containing providers for basic data, service providers for calculation of scoring values and banks using the mining results, the paper analyses the requirements that data controllers have to meet.

[1] This work is based on research carried out within the Project "Future of Identity in the Information Society" (FIDIS, http://www.fidis.net), which is funded by the European Union.
[2] See overview of articles by S. Oliveira at http://www.cs.ualberta.ca/%7Eoliveira/psdm/pub_by_year.html or K. Liu at http://www.cs.umbc.edu/~kunliu1/research/privacy_review.html
[3] For example research carried out by IBM, see http://www.almaden.ibm.com/software/disciplines/iis/
[4] See for example overview up to 2004 at http://www.cs.ualberta.ca/%7Eoliveira/psdm/workshop.html
[5] See for example http://www.lustat.ch/ms_Datenschutzkonzept_2001.pdf and http://e-hrc.net/media/ExtHealthNetworksMuscle02Feb2005.htm

1. State-of-the-art in PPDM

Oliveira and Zaïane (2004) define PPDM as data mining methods which have to meet two targets: (1) meeting privacy requirements and (2) providing valid data mining results. These targets are in some cases at odds with each other, depending on the type of data mining results and the attributes in basic data . In these cases the use of PPDM offers a compromise between the two targets mentioned.

Privacy preserving data mining typically uses various techniques to modify either the original data or the data generated (calculated, derived) using data mining methods. To achieve optimised results while preserving the privacy of the data subjects efficiently, five aspects or dimensions, have to be taken into account. These dimensions are (1) the distribution of the basic data, (2) how basic data are modified, (3) which mining method is being used, (4) if basic data or rules are to be hidden and (5) which additional methods for privacy preservation are used (Verykios et al. 2004). This overview shows from a technical perspective how many different methods and techniques in the context of PPDM can be used.

Though Oliveira and Zaïane (2004) observed a large and rapidly increasing variety of different methods and tools available to perform PPDM. In most cases these approaches seem to be limited to one data mining method or even a specialised algorithm. In addition there are no integrated PPDM solutions available on the market that allow for applications independent from the data distribution scheme, the mining method, the algorithm used or even type of attribute (Boolean, numerical etc.) used as basic data. Further development to achieve integrated solutions including PPDM is necessary. Moreover, they concluded that there is no common understanding of privacy in the context of PPDM. Some areas of PPDM application, for example, cover hiding of basic data or mined rules among organisations to protect trade secrets. In our understanding, this is an application of technical measures to meet confidentiality as one of the traditional three targets of IT-security. But PPDM also includes the technical implementation of data protection principles, for example the data minimisation principle, by anonymising personal data.

As a result of these different views on privacy there are currently no common metrics for (1) the quality of different methods and algorithms to meet privacy requirements and (2) the loss of quality of the data mining results as compared to today’s standard data mining systems.

Oliveira and Zaïane (2004) analysed the developments in PPDM to date and concluded that a standardisation process is needed to overcome the confusion being observed among developers, practitioners and others interested in this technology, caused by the excessive number of different PPDM techniques. For this process with respect to privacy they suggest using the OECD Privacy Guidelines[6] from 1980 which are accepted worldwide. From these guidelines they extracted eight principles and classified them with respect to their relevance for PPDM. This classification is focused on the view of the technical data operator within the mining process and is based on the understanding of privacy protection established in the USA:

1. Collection limitation principle – too general to be enforced in PPDM
2. Data quality principle – most of today's PPDM methods or algorithms assume that data are already prepared to an appropriate quality to be mined
3. Purpose specification principle – extremely relevant for PPDM
4. Use limitation principle – fundamental for PPDM
5. Security safeguard principle – unenforceable in the context of PPDM
6. Openness principle – relevant for PPDM
7. Individual participation principle - Oliveira and Zaïane suggest that the implications of this principle for PPDM should be carefully weighed in light of the ownership of the basic data otherwise the application could be too rigid in PPDM applications.
8. Accountability principle – too general for PPDM

Based on these principles, the authors suggest a set of four policies that should be defined when applying PPDM:

[6] See http://www.oecd.org/dataoecd/33/43/2096272.pdf

- Awareness Policy: The target is to define a policy of how the data subject is informed.
- Limit Retention Policy: The target of this policy is the deletion of data that are not up-to-date to avoid unnecessary risks.
- Forthcoming Policy: This policy contains the information regarding what data are processed for which purpose and how the results are to be used and with whom they are shared.
- Disclosure Policy: This policy defines that discovered knowledge is disclosed only for purposes which the data subject has given her or his consent.

According to Oliveira and Zaïane (2004), for the deployment of PPDM the following requirements have to be met. These are (1) identification of private information that is to be protected, (2) compliance with international instruments to state and enforce privacy rules, (3) logging of steps taken in PPDM in order to allow for transparency, (4) limitation of disclosure of the data subject's private information, and (5) matching the solution with privacy principles and policies especially in cases where policies or technical solutions (for example the data mining algorithm or its parameters) are updated.

2. Analysis and Discussion

Member countries of the EU have adopted the Data Protection Directive 95/46/EC[7] by implementing it into national data protection legislation. For the private sector and large parts of the public sector, excluding institutions and organisations that deal with state security, this Directive defines requirements that processes using data mining have to meet. The Directive states obligations of the data controller and data processor, thus in this context accountability (c.f. principle 8 based on the OECD Guidelines) is defined resulting in a need for compliance with national data protection legislation. As the Directive is widely applied across Europe, it is used for the following analysis.

2.1 Introduction of the Use Case: Scoring Practice in the Financial Sector

[7] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, download via http://ec.europa.eu/justice_home/fsj/privacy/law/index_en.htm

Typically, scoring is a directed data analysis to classify data subjects with respect to predefined risk categories for failure to pay back credit. One example is a classification in the three categories “normal”, “increased” and “high”. Such scoring is a type of profiling known as non-distributive group profiling (Hildebrandt, Backhouse 2005), since scoring values describe a certain *likeliness* of the data subject not to be able to pay a credit back, but they do not allow a precise determination of whether the credit really will be paid back or not.

For classification, a number of different data mining algorithms can be used such as (Schweizer 1999):

- Genetic algorithms to optimise parameter of regression algorithms or regression trees
- Regression trees
- Regression algorithms such as logistic regression
- Heuristics / neural networks

Typically the types of basic data used, the data mining algorithms applied and the parameters used by the algorithms are classed as trade secrets by the data controllers or data operators. That said, to the knowledge of the author, in the context of credit scoring heuristics are typically not being used on account of the problems of getting reproducible results. In many cases logistic regression seems to be used for that purpose. [8]

From a technical perspective we can discriminate two phases of the mining process:

- Selection and optimisation, in respect of the parameters of the mining algorithm; this includes business process understanding, data understanding and preparation, including selection of attributes; modelling and evaluation; and
- Deployment and application, which means that credit scoring values are calculated and used.

Of course the chosen example does not cover all possible organisational and technical aspects of data protection when using data mining methods. But with the market players we find today such as basic data providers, data mining service providers, and users of the mining

[8] See for example by the German “Schutzgemeinschaft für allgemeine Kreditsicherung“ (Schufa): Der Hessische Landtag, *Drucksache 16/1680 in der 16. Wahlperiode vom 11.12.2003*, p. 21, Wiesbaden 2003; Download: http://www.denic.de/media/pdf/dokumente/datenschutzbericht_2003.pdf

results in the financial sector (banks or insurance companies), this example allows for an analysis with respect to the requirements data controllers have to meet.

2.2 Differences in understanding of privacy and application of privacy protection

Compared with the European countries applying Directive 95/46/EC, in the USA there is a fundamentally different understanding of who owns the basic data that are mined by an organisation. While in the USA these data are considered to belong to the organisation, in the context of the Directive 95/46/EC data subjects have the right for self determination with respect to their own personal data. This leads to a different understanding of the implementation of the individual participation principle given the data subject is in a stronger position compared to the OECD Guidelines (Grimm, Roßnagel 2000).

In addition the understanding of legal processing of personal data is different. In the USA processing of personal data is allowed unless special legislation prohibits it. In the context of Directive 95/46/EC, processing of personal data is forbidden unless explicitly allowed by legislation or effective consent by the data subject.

The suggestions of Oliveira and Zaïane partially reflect these fundamental differences in the understanding of privacy protection. In addition they analyse the requirements PPDM has to meet with respect to privacy protection from a technical point of view in the core of the mining process. The application of Directive 95/46/EC requires a different perspective: the application of the legal norms of the Directive with respect to the entire business or governmental processes in which data mining is used. These processes include for example:

- Definition of target and purpose of the business or governmental process
- Selection and optimisation of the methods, algorithms and parameters of the mining process basing on reference data
- Application of the data mining including data collection and preparation
- Further use of the results in the process

With this perspective the scope of possible measures to enable data protection is much broader than the application of PPDM methods or algorithms. These cover, in addition, the organisational and technical measures of the whole process into which data mining is integrated.

Privacy Principles taken from the European Directive 95/46/EC

Taking these fundamental understandings into account and applying the principles used by Oliveira and Zaïane from the OECD Guidelines, Directive 95/46/EC can be summarised as follows (see for example Möller, Bizer 2006):

1. Legal basis for processing personal data
2. Data quality principle, including
    - Purpose binding principle (integrating the purpose specification and use limitation principles of the OECD Guidelines)
    - Data minimisation principle stating that data that are not needed or not needed any more for the originally defined purpose are to be deleted
3. Transparency principle
4. Security safeguard principle
5. Individual participation principle

These principles do not seem to be very different from the principles extracted from the OECD Guidelines by Oliveira and Zaïane, but their implementation will show differences.

With respect to the *legal basis for processing of personal data* the Directive 95/46/EC states a number of conditions under which processing can be lawful. In the private sector the most relevant conditions are:

- Unambiguous voluntary consent of the data subject
- Processing is necessary for the concluding and performance of a contract in which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract

Especially the second condition can be used to establish a legal basis for processing of personal data using data mining. Credit scoring in the financial sector can be taken as one example (Weichert 2005: 584), further example were investigated by Petri and Kieper (2003) and Weichert (2003).

Standards for Data Mining

In the context of data mining, standards such as CRISP-DM[9] or the semiotic model for knowledge discovery in databases (KDD, Hildebrandt, Backhouse 2005), developed based on the research of Stamper and Liu (see for example Xie, Liu 2003), are established. They support, in contrast to traditional PPDM methods, an integrated view on the business or governmental process where data mining is to be used. To reach an optimised quality they suggest a cyclic processes model based on the PDCA cycle (plan, do, check, act cycle, also called Deming cycle). In addition they suggest proper documentation for each step taken in the selection and optimisation process.

CRISP-DM as a significant standard for data mining lists six steps in the mining process:

1. Business understanding
2. Data understanding
3. Data preparation
4. Modelling
5. Evaluation
6. Deployment (including final report)

Each step taken in the mining process is to be reported and analysed carefully. In the case where compliance with data protection legislation is understood as a relevant business target, this general process model also can be used to optimise the level of privacy reached. In any case a higher level of data protection can be reached when organisational measures in the context of these standards based on the data protection principles are used. Organisational measures may be used alone or in combination with PPDM methods.

Use Case: Application of the Data Protection Principles in the Scoring Context

In the first two steps of CRISP-DM the business targets are defined and quality and origin of basic data are checked. From this the *legal basis* for the data processing can be checked. In

[9] See www.crisp-dm.org

the case where the legal basis and the business targets do not map, either the business target or the legal basis (e.g. by getting additional consent) should be adjusted.

The *data quality principle* has a number of consequences for the process in which data mining takes place. Getting back to our example (credit scoring) for each of the two general steps of data mining already introduced, the necessity of the collected data with respect to the defined purpose has to be checked and documented. In this case it can lead to differences in the attributes that are needed for the selection and optimisation of the mining algorithm, where anonymised data are to be used, and the deployment, where the link to a specific person is essential (Weichert 2005: 583). In cases where the anonymity set is small, for example when developing scoring algorithms for specific and small target groups, further PPDM methods such as perturbation can be applied to increase the reached level of privacy protection.

The data controller is responsible and liable for the data quality. To be able to demonstrate an appropriate quality of the basic data and the results of the mining process, we suggest applying the quality standards for data mining referred to. As changes in society, such as changes of income, employment, education, immigration etc. may have an impact on the classification, as well as the definition of the classes as the assignment of data subject to these classes, scoring algorithms should be regularly checked with respect to their quality, thus restarting the selection and optimisation cycles. As a result a new version of the scoring algorithm is deployed.

Attributes that show no significant impact on the scoring values in the optimisation phase or that are not up to date anymore are not to be used in the deployed version(s) of the algorithm (data minimisation). This legal requirement in some cases can make blocking or encryption of these attributes obsolete, as they are by law to be deleted anyway. The steps 1, 2, 3 and 4 of the CRISP-DM model can be used to identify attributes to be deleted. This covers the limit retention policy suggested by Oliveira and Zaïane (2004).

For the application of the deployed versions of the algorithms, logging has to be applied. In addition to aspects of data quality, logging also supports *transparency* (history of basic data and calculated scoring values) and *security* (access logs for personal data). In the case of scoring the logging should cover (Kamp, Weichert 2006: 89):

- Basic data (attributes), timestamp and the source they were taken from (e.g. questionnaire, interviewer, external data provider etc.)
- The mining result (which value, when, calculated by whom?)
- In cases where mining results are transferred to a user (which value, when and by whom?)

The implementation of the *transparency principles* is well defined, based on Directive 95/46/EC, describing the requirements for parts of the awareness, the forthcoming and the disclosure policy suggested by Oliveira and Zaïane (2004) precisely. To enable the participation of the individual, the data subject is to be informed about (Kamp, Weichert 2006: 92ff.):

- What is the purpose of the processing of personal data
- What personal data (attributes, mined results) are used
- How data are processed with respect to the governmental or business process, as well as the technical methods used. This means that the mining process has to be described at least in a general sense
- Who is doing the processing - data controller responsible, in this context mostly the bank or insurance company, and in cases where data are transferred to service providers, which data is transferred to whom, at least by category (e.g. service provider for scoring)
- In cases where basic data are stored and provided by external service providers it also has to be indicated what data (attributes) are used for the purpose

Security has to span the whole process from data collection, storage, transport, mining up to application and the documentation involved in the whole process. Directive 95/46/EC states "[The data] controller must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing."
In any case security only makes sense when applied to the whole business or governmental process. To assure this, internationally accepted information security management systems (ISMS) such as ISO/IEC 27001 or 17799 can be used to define, implement and document the organisational and technical measures needed to reach an appropriate security level. In cases

where service providers are used, the data controller has to ensure proper security measures for all parties involved in the process, e.g. via contracts including security service levels (SSLAs) and applying appropriate audit schemes.

With respect to the *individual participation principle* a number of rights of the data subject are also defined. This includes:

- The right to be informed about used basic data and calculated mining results
- The right to get "knowledge of the logic involved in any automatic processing of data concerning him […]" (Article 12, Directive 95/64/EC, see also for more details Kamp, Weichert 2006: 86ff.)
- The right to object against the data processing and to withdraw a given consent
- The right to get data corrected
- The right to get data deleted

The implementation of the *security safeguard* and the *individual participation principle* especially show that in some cases using additional organisational measures applied in the context of the overall business or governmental process, a higher level of privacy protection can be reached compared to using just PPDM methods and algorithms.

3. Conclusion

To implement effective privacy protection when applying data mining, it is not sufficient to focus on PPDM methods and algorithms. In addition to this the whole business or governmental process in which data mining is used has to be taken into account. As a result we can conclude that organisational and technical measures taken based on Directive 95/46/EC by applying national data protection legislation in combination with data mining standards allows for quite effective privacy protection. In some cases the strict application of these measures even can make the use of PPDM methods and algorithms obsolete, while in other cases PPDM potentially can enhance privacy protection further compared to the use of traditional data mining methods.

PPDM is still an area of research and not readily implemented on the market yet. However, first pilot implementations can already be observed, and integration of PPDM methods and algorithms in standard data mining tools will make them readily available as additional methods soon. For assurance of PPDM result quality, established standards such as CRISP-DM (in combination with comparison of results of traditional data mining) can be used as long as no general quality metric is available.

Bibliography

Grimm, R., Roßnagel, A., 'Datenschutz für das Internet in den USA', *Datenschutz und Datensicherheit* 8/2000, pp. 446-451, Wiesbaden 2000.

Hildebrandt, M., Backhouse, J. (Eds.), *FIDIS Deliverable 7.2 Descriptive Analysis and Inventory of Profiling Practice*, Frankfurt a. M. 2005. Download http://www.fidis.net/fidis_del.0.html

Kamp, M., Weichert, T., *Scoringsysteme zur Beurteilung der Kreditwürdigkeit – Chancen und Risiken für Verbraucher - Studie im Auftrag des Bundesministeriums für Verbraucherschutz, Ernährung und Landwirtschaft*, Berlin 2006. Download: http://www.bmelv.de/cln_045/nn_752314/SharedDocs/downloads/02-Verbraucherschutz/Finanzdienstleistungen/scoring,templateId=raw,property=publicationFile.pdf/scoring.pdf

Möller, J., Bizer, J., 'Datenschutzanforderungen and Digital Rights Management', *Datenschutz und Datensicherheit* 2/2006, pp. 80-84, Wiesbaden 2006. Download: http://www.datenschutzzentrum.de/drm/kurzfassungen.htm

Oliveira, S. R. M., Zaïane, O. R., *Towards Standardization in Privacy-Preserving Data Mining*, Edmonton 2004. Download: http://www.cs.ualberta.ca/%7Ezaiane/postscript/dm-ssp04.pdf

Petri, T. B., Kieper, M., 'Datenbevorratungs- und analysesysteme in der Privatwirtschaft', *Datenschutz und Datensicherheit*, 10/2003, pp. 609-613, Wiesbaden 2003.

Schweizer, A., *Data Mining Data Warehousing – Datenschutzrechtliche Orientierungshilfe für Privatunternehmen*, Orell Füssle Verlag AG, Zürich 1999.

Verykios, V. S., Bertino, E., Fovino, I. N., Provenza, L. P., Saygin, Y., Theodoridis, Y., 'State-of-the-art in Privacy Preserving Data Mining', *SIGMOD Record*, Vol. 33, No. 1, New York, March 2004. Download: http://dke.cti.gr/CODMINE/SIGREC_Verykios-et-al.pdf

Weichert, T., 'Datenschutzrechtliche Anforderungen an Data-Warehouse-Anwendungen bei Finanzdienstleistern', *Recht der Datenverarbeitung* 3/2003, Frechen-Königsdorf 2003.

Weichert, T., 'Datenschutzrechtliche Anforderungen and Verbraucher-Kredit-Scoring', *Datenschutz und Datensicherheit* 10/2005, pp. 582-587, Wiesbaden 2005.

Xie, Z., Liu, K., *Improving Business Modelling with Organisational Semiotics*, Reading 2003. Download: http://www.rug.nl/bdk/onderzoek/onderzoeksgroupen/castor/OS02/Xie.pdf