



DISSERTATION | DOCTORAL THESIS

Titel | Title

Soft-Computing-Based Privacy-Preserving Model for Distributed
E-Business Environments

verfasst von | submitted by

Fatemeh Amiri

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Doktorin der technischen Wissenschaften (Dr.techn.)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 786 880

Dissertationsgebiet lt. Studienblatt | Field of
study as it appears on the student record
sheet:

Informatik

Betreut von | Supervisor:

Univ.-Prof. Dipl.-Ing. Dr. Dr. Gerald Quirchmayr

Acknowledgments

I am deeply grateful to my family for their patience, understanding, and unwavering support during my doctoral studies. In particular, I am deeply grateful to my sister, Mahya, whose encouragement helped me overcome many challenges throughout this journey.

I would also like to express my sincere gratitude to my supervisor, Prof. Dr. Quirchmayr, for his valuable guidance, constructive feedback, and continuous support throughout this research. His expertise and encouragement greatly contributed to the development and completion of this thesis.

My sincere thanks also go to Prof. Dr. Edgar Weippl and the members of the research group at Secure Business Austria (SBA Research) for their valuable discussions, insights, and for providing a stimulating research environment during the preparation of this work.

During the preparation of this thesis, selected passages were language-edited with the assistance of an AI-based writing tool for clarity and style (OpenAI scientific GPT engine in overleaf). All technical content, analysis, interpretations, and conclusions are solely the responsibility of the author.

Abstract

Large-scale data mining and data sharing have become key drivers of innovation in modern online services. However, they also raise significant privacy concerns among users. While data-driven services provide substantial benefits, preserving user privacy remains a fundamental requirement when personal information is disclosed. Consequently, the central challenge is to enable efficient and accurate data analysis while maintaining strong privacy guarantees. This thesis addresses this challenge by investigating privacy-preserving techniques for data mining and recommender systems. The work focuses on identifying and mitigating privacy gaps that arise during the mining and sharing of user data. Potential privacy breaches are examined, privacy-preserving algorithms are proposed, and the trade-offs between data utility and privacy protection are analyzed for selected highly relevant algorithmic problems in the domains of big data analytics and recommender systems. To support these objectives, a model is developed and implemented that integrates efficient techniques for privacy-preserving computation in data mining. In particular, the thesis introduces improved soft-computing-based clustering methods designed for distributed environments. Privacy protection is achieved through the anonymization of sensitive user attributes and the non-deterministic removal of infrequent elements. The proposed methodologies are evaluated to demonstrate their effectiveness in supporting privacy-preserving data analysis. Furthermore, this thesis investigates the relationship between privacy loss and utility in collaborative recommender systems and provides a comparative analysis of privacy–utility trade-offs in recommender networks.

Kurzfassung

Data Mining und Datenaustausch spielen eine wichtige Rolle für Innovationen in modernen Online-Diensten. Gleichzeitig verursachen sie zunehmenden Bedenken hinsichtlich des Schutzes der Privatsphäre von Nutzerinnen und Nutzern. Obwohl datenbasierte Anwendungen erhebliche Vorteile bieten, bleibt der Schutz der Privatsphäre von Nutzerinnen und Nutzern eine grundlegende Voraussetzung, insbesondere wenn individuelle Nutzerdaten verarbeitet oder weitergegeben werden. Eine zentrale Herausforderung besteht daher darin, leistungsfähige und präzise Verfahren zur Datenanalyse zu ermöglichen und gleichzeitig verlässliche Garantien zum Schutz der Privatsphäre sicherzustellen. Diese Dissertation erforscht Methoden für Data Mining und Recommender Systems die die Privatsphäre schützen. Der Schwerpunkt der Arbeit liegt auf der Erkennung und Reduktion von Risiken, die bei der Verarbeitung und Weitergabe von Nutzerdaten entstehen können. In diesem Zusammenhang werden potenzielle Verletzungen der Privatsphäre analysiert, algorithmische Ansätze zum Schutz sensibler Informationen entwickelt sowie die Wechselwirkungen zwischen Analysequalität und Schutz persönlicher Daten in ausgewählten algorithmischen Fragestellungen der Analyse umfangreicher Daten und Recommender Systems untersucht. Darüber hinaus wird ein Modell entwickelt und implementiert, das effiziente Verfahren für datenschutzwahrende Berechnungen integriert. Insbesondere werden verbesserte Clustering-Verfahren auf Grundlage von Soft-Computing-Ansätzen vorgestellt, die speziell für verteilte Systemumgebungen konzipiert sind. Der Schutz sensibler Informationen wird dabei durch die Anonymisierung relevanter Nutzerattribute sowie durch die nichtdeterministische Entfernung seltener Elemente erreicht. Die vorgeschlagenen Methoden werden umfassend ausgewertet, um ihre Eignung für eine Analyse von Nutzerdaten unter strengen Anforderungen an den Schutz persönlicher Informationen aufzuzeigen. Darüber hinaus wird eine theoretische untere Grenze für den minimalen Verlust an Privatsphäre in kollaborativen Empfehlungssystemen mit hoher Leistungsfähigkeit hergeleitet. Ergänzend erfolgt eine vergleichende Untersuchung der Wechselwirkungen zwischen Analysequalität und Privatsphärenschutz in Empfehlungsnetzwerken.

Contents

Part I Major Building Blocks and Contributions

1	Privacy in Recommender Systems: Definitions and Motivation	3
1.1	Privacy Concerns in Recommender Systems	3
1.2	Definition of Privacy and Confidentiality in Recommender Systems	4
1.3	Known Mechanisms for Achieving Privacy in Recommender Systems	5
1.3.1	Methods	5
1.3.2	Models	6
1.4	A Threat Model Against Recommender Systems	10
1.5	Taxonomy of Privacy Attacks Targeting Recommender Systems	11
1.5.1	White-box Attack	11
1.5.2	Black-box Attack	12
1.6	Motivation: Why Is Privacy Preservation Important in Recommender Systems?	13
1.6.1	The Benefits of Training and Sharing User Data	13
1.6.2	Privacy Threats in Online Platforms: The Case of Recommender Systems	14
1.7	Summary	15
2	Structure and Contributions of the Thesis	17
2.1	Challenges and Research Questions	17
2.1.1	Privacy Challenges Associated with Sharing User Data During Training	17
2.1.2	Research Questions	18
2.1.3	Aim, Scope, and Limitations of the Thesis	19
2.2	Contributions	21
2.2.1	Protecting Privacy When Training and Sharing User Data	21
2.2.2	Assignment of Contributions and Research Questions (C & Q)	22
2.3	Thesis Structure and Focus Areas	24

2.3.1	Part II – A Need to Strike the Right Balance Between Utility and Privacy	24
2.3.2	Part III – Algorithms for Training and Sharing User Data Privately	24
2.3.3	Part IV – Quantifying the Privacy Level for Recommender Systems	25
2.4	Summary	25

Part II Addressing the Trade-Off Between Utility and Privacy

3	The Evolution of the Privacy Concept in the Digital World	29
3.1	The Emergence of the Concept of Privacy in the Digital World	29
3.1.1	The Concept of Privacy as a Human Right	29
3.1.2	Digital Privacy	30
3.2	Privacy Violation Examples in Online Environments	31
3.3	Privacy in Recommender Systems: Literature Review	34
3.3.1	Definitions	34
3.3.2	Survey of Recommender System Types	35
3.3.3	Privacy in Recommender Systems: The Trade-off Between Privacy and Accuracy	37
3.4	Summary	38
4	Conventional and Soft-Computing-Based PPDM Approaches	41
4.1	Introduction	41
4.1.1	Privacy in e-Business and Knowledge Discovery	42
4.1.2	Problem Definition	43
4.1.3	Methodology	44
4.1.4	Literature Selection Criteria	45
4.1.5	Keywords and Databases	46
4.1.6	Contributions	46
4.2	Conventional Privacy-Preserving Approaches in Knowledge Discovery	47
4.2.1	The k-Anonymity Model and l-Diversity	47
4.2.2	The Perturbation Approach	47
4.2.3	Cryptographic Distributed Techniques	48
4.2.4	Privacy-Preserving Outsourced Data Mining	48
4.2.5	Privacy-Preserving Algorithms in Data Mining	48
4.3	Soft-Computing Approaches to Privacy Preservation in Knowledge Discovery	49
4.3.1	Fuzzy Logic	49
4.3.2	Neural Networks	49
4.3.3	Rough Sets	49
4.3.4	Genetic Algorithms	50
4.4	Analysis and Results	50
4.4.1	Key Characteristics and Evaluation Metrics	50

4.4.2	Comparative Analysis of Conventional and Soft-Computing Approaches	51
4.5	Insights and Implications	55
4.5.1	Conceptual Framework	55
4.5.2	A Unified Taxonomy of Privacy-Preserving Methods	55
4.5.3	Research Gaps and Future Directions	56
4.6	Summary	57

Part III Training Algorithms and Sharing of User Data

5	Metaheuristic-Based Strategies for Data Anonymization in Data Mining	61
5.1	Introduction	61
5.1.1	Motivation	63
5.1.2	Distributed Approaches for Privacy-Preserving Data Mining (PPDM)	64
5.1.3	Contributions	64
5.2	Related Work	65
5.3	Theoretical Foundations and Problem Statement	66
5.3.1	Design and Formulation of the GAFSOM Method	66
5.3.2	GA Concepts	67
5.3.3	Fuzzy Set Concepts	68
5.3.4	Kohonen SOM	69
5.3.5	Privacy Attacks	71
5.3.6	Loss Metric	72
5.3.7	The Optimization Problem	72
5.4	Setting and Notation	73
5.5	GAFSOM Algorithm	73
5.6	Empirical Analysis	75
5.6.1	Datasets	76
5.6.2	Experimental Setup	77
5.6.3	Statistics Evaluation	78
5.6.4	Performance Evaluation and Analysis	94
5.7	Interpretation of Experimental Results	94
5.8	Possible Attack Strategies	95
5.9	Open Issues and Future Directions	96
5.10	Summary	96
6	A Differentially Private Self-Organizing Map	99
6.1	Introduction	99
6.1.1	Motivation	101
6.1.2	Contributions	101
6.2	Related Work	102
6.3	Theoretical Foundations and Problem Statement	103
6.3.1	SOM Clustering	103

6.3.2	Definition of Differential Privacy	105
6.3.3	Global Sensitivity	108
6.3.4	The Privacy Budget and Local Function Sensitivity	108
6.3.5	Mechanisms of Differential Privacy	108
6.3.6	Sequential Composition	110
6.3.7	Data Release Under Differential Privacy	110
6.3.8	Uncertainty of Outcome	111
6.3.9	Properties of Differential Privacy	111
6.3.10	Problem Statement	111
6.4	Method	112
6.4.1	Approach Overview	112
6.4.2	Initial Calculations and Transformations	113
6.4.3	DP2SOM Approach and Its Analysis	113
6.5	Experimental Evaluation	118
6.5.1	Datasets and Experimental Setup	118
6.5.2	Evaluation Metrics	119
6.5.3	Parameter Selection	119
6.6	Analysis and Results	120
6.6.1	Accuracy Analysis of the Iris Dataset Before and After DP	120
6.6.2	Accuracy Analysis of the MNIST Dataset Before and After DP	120
6.6.3	Experimental Dataset Analysis	122
6.6.4	Clustering Quality Measurement	123
6.6.5	Comparison Between DP2SOM and Similar Methods	124
6.7	Discussion	125
6.7.1	Challenges in Differential Privacy	127
6.8	Summary	128

Part IV Quantitative Analysis of Privacy in Recommender Systems

7	Case Study: Measuring Privacy Levels in Recommender Systems	133
7.1	Motivation	133
7.2	Contributions	134
7.3	Preliminaries	135
7.3.1	Recommender Systems	135
7.3.2	SOM Clustering	139
7.3.3	Differential Privacy	140
7.4	Related Work	140
7.5	Model Setting	141
7.5.1	General Steps of Collaborative Filtering with Privacy Considerations	141
7.5.2	Personalized Social Recommenders	143
7.5.3	Assumptions and Attacker Model	144
7.5.4	Privacy Model	144
7.6	Proposed Conceptual Framework	145

7.7	Case-Study Objectives	145
7.8	Experimental Study	146
7.8.1	Dataset	146
7.8.2	Experimental Setup	146
7.9	Results and Discussion	152
7.9.1	Analysis of the Privacy–Accuracy Trade-Off	152
7.9.2	Test Dataset for Prediction	156
7.9.3	How Does DP Protect Recommendations?	157
7.9.4	How Does DP Protect Against the Threat Model?	162
7.9.5	Complexity Analysis	163
7.9.6	Competing Algorithms	164
7.9.7	Challenges	168
7.10	Summary	168
8	Conclusion, Contributions, and Future Work	171
8.1	Review of Contributions	171
8.2	Results and Discussion	174
8.3	Open Problems	177
8.4	Future Work	178
8.5	Summary	179
	References	180

Part I
Major Building Blocks and Contributions

This section defines the conceptual framework and scope of the thesis. In addition, it introduces the key research questions and provides an overview of the thesis structure.

In Chapter 1, the adopted definition of privacy—selected from several existing definitions and adapted to the context of this research—is introduced. The chapter also presents the motivation of the study and reviews existing privacy-preserving models and methodologies. In addition, the threat model used to address the challenges considered in this thesis is described.

Chapter 2 outlines the key elements of the research, including the main challenges, the scope of the investigation, and the central research questions. The chapters in which each research question is addressed are clearly identified. The main scientific contributions of the thesis are also summarized. The chapter concludes with an overview of the structure of the thesis.

Chapter 1

Privacy in Recommender Systems: Definitions and Motivation

Chapter Overview. The purpose of this chapter is to introduce the fundamental concepts of privacy and to present the theoretical foundations relevant to this research. Privacy plays a crucial role in users' adoption of online services and digital platforms [Oliveira and Martins, 2010], which are widely used in modern information systems. In this context, users are often reluctant to disclose personal information; consequently, privacy has become a major concern that requires particular attention from both data holders and infrastructure providers [Bélanger and Crossler, 2011]. Privacy protection is the primary motivation of this thesis. Among the many open research topics in this area, particular attention is given to privacy in e-business environments that process large volumes of data and commonly rely on recommender systems. This chapter introduces the terminology used throughout the thesis and presents the relevant models, methods, attack vectors, and privacy concerns addressed in this research. The chapter also provides a detailed discussion of Research Questions 1 and 2.

1.1 Privacy Concerns in Recommender Systems

A privacy breach may involve a variety of parties, including fellow users and both internal and external service providers. Depending on the nature of the exposed information, substantial harm may result. A data breach may be intentional (for example, snooping or hacking) or inadvertent (for example, mismanagement or lingering data) [Jeckmans et al., 2013]. Several classifications of privacy concerns exist in recommender systems, including personal information leakage, contextual information exposure, preference leakage, social network identifier exposure, inferred sensitive information, evaluation metric leakage, and model inversion [Friedman et al., 2015, Jeckmans et al., 2013]. In recommender systems, privacy risks primarily arise from the extensive collection of user data by service providers and the potential exposure of sensitive information.

In one taxonomy, Shyong et al. [Shyong et al., 2006] presented a clear and comprehensive definition of privacy concerns in recommender systems, stated as follows:

- Data collection. This concern refers to the extent of information use because the gathering procedures do not match users' expectations.
- Data retention. Data removal from the internet is generally difficult; in other circumstances, the service provider may actively block or obstruct data removal. Data retention poses a risk because data may remain accessible longer than intended.
- Data sales. Third parties may find the large amount of information contained in online systems valuable, and in some cases such information may be sold. User ratings, preferences, and purchasing histories may all be useful for marketing purposes. Data sales often conflict with users' privacy preferences.
- Recommendations reveal information. Recommendations are inherently based on the data stored in the recommender system. Each recommendation may reveal a small portion of personal information. This issue is addressed in the approach presented in this thesis.

1.2 Definition of Privacy and Confidentiality in Recommender Systems

Information privacy has been defined in multiple ways in the literature. In order to provide the functionality required by a recommender system and to generate accurate recommendations, user information must be disclosed. However, the sharing of personal information may result in data breaches; therefore, a balance must be maintained between service value and user privacy.

In this section, the concept of privacy in recommender systems and the potential privacy concerns affecting users are examined. The wording of Kang's Information Infrastructure Task Force (IITF) definition [Kang, 1997] is adopted, as cited below:

Information privacy is an individual's claim to control the terms under which personal information—information identifiable to the individual—is acquired, disclosed, or used.

In the broader context of information security, the concept of information privacy is closely related to confidentiality. Privacy concerns the protection of individual pieces of information, the consequences of their disclosure for each person, and the role of control and consent. Confidentiality, by contrast, concerns the secrecy of individual pieces of information. The scope of privacy considered in this thesis is as follows:

This thesis focuses on the unintended disclosure and misuse of information rather than on the individual-level consequences of disclosure. In this thesis, privacy is understood as preserving information within its intended scope of use and disclosure. This scope is determined by a web application in which personal information, such as product ratings or comments, is contributed by users. According to the privacy assumptions adopted in this thesis, user data must be protected throughout all stages of collection, processing, and storage on the server side. It is further assumed that servers and communication channels cannot be fully trusted; therefore, data is transmitted to the server only in a controlled manner.

The size of the audience, the amount of authorized use, and the duration of storage all influence this scope. A privacy breach occurs when information is unintentionally or maliciously transferred beyond its intended scope. Such a breach may occur when information is used for a purpose other than the one originally intended, or when it is retained beyond its intended lifetime [Palen and Dourish, 2003].

1.3 Known Mechanisms for Achieving Privacy in Recommender Systems

This section reviews some of the most common privacy-preserving methods, models, and threat vectors.

1.3.1 Methods

According to [Willenborg and De Waal, 2012], privacy methods fall into two broad categories: suppressive and perturbative.

1.3.1.1 Suppressive Privacy Methods

In suppressive approaches, data values are not altered; instead, the amount of published data is restricted. The most common suppressive techniques are described below.

- **Cell suppression.** Individual attributes or combinations of data-point properties, as well as full rows or columns in tabular data, are hidden in order to protect privacy.

- **Query restriction.** Certain queries may be suppressed if their responses are likely to violate a given privacy-protection standard and if the data recipient accesses the database through a query mechanism.
- **Aggregation.** Data are grouped or clustered before publication in order to reduce identifiability.
- **Sampling.** A synthetic dataset is generated by sampling from a distribution estimated from the original data.

1.3.1.2 Perturbative Privacy Methods

Perturbative methods alter the values of data attributes [Stewart, 1990]. The main perturbative methods are as follows.

1. **Data swapping.** For numerical values, two forms are commonly used: (i) individual value swapping within a dataset and (ii) rank swapping. In the latter, the set of values is sorted and then swapped within a restricted rank range.
2. **Adding noise.** Random noise is added to the data values; this approach is generally applicable only to continuous numerical data.
3. **Rounding.** Original values are replaced with rounded values. This is similar to clustering in suppressive approaches, but here the data values themselves are modified.
4. **Resampling.** A synthetic dataset is generated by drawing samples from a distribution estimated from the original dataset.

1.3.2 Models

Over the past few decades, numerous privacy models have been established; however, this thesis does not attempt to cover all of them. Instead, it focuses on a selected subset regarded as most relevant to the research presented here. Traditional anonymization approaches are introduced first, followed by more recent developments, namely soft computing and differential privacy. The adaptation of anonymization in survey settings is also discussed.

1.3.2.1 Perturbative Anonymization

A significant proportion of privacy-preserving techniques use randomized data perturbation to protect sensitive information. This approach hides sensitive data by randomly modifying values, often through additive noise [Agrawal and Srikant, 2000, Shah and Gulati, 2016b, Datta, 2004, Liu et al., 2005, Jahan et al., 2014]. However, later studies argued that random objects in the spectral domain may still exhibit predictable structures, and extensive experiments showed that in many cases

random data distortion preserves relatively little privacy [Kargupta et al., 2003, Kargupta et al., 2005].

In the literature, three perturbation methods are commonly distinguished: additive, multiplicative, and geometric data perturbation. The boundaries between these categories are not always sharply defined. Shah et al. [Shah and Gulati, 2016a] investigated and compared the efficiency of these methods. Their simulation and experimental results indicate that additive and geometric data perturbation preserve the statistical utility of the original dataset. Multiplicative perturbation methods generate a perturbed copy with minimal record linkage but do not preserve the statistical properties of the original dataset. This finding further illustrates the trade-off between information loss and privacy. According to [Aggarwal and Philip, 2008], multiplicative perturbation may provide useful privacy protection while approximately preserving inter-record distances, allowing transformed records to be used in a variety of data-mining applications. This technique therefore forms the basis of privacy-preserving clustering and classification [Chen and Liu, 2005, Chen and Liu, 2011].

Despite the advantages of perturbation and randomization, a major issue is the trade-off between privacy and data utility. In this context, [Liu et al., 2008] designed a model that allows users to select their own level of privacy, and through simulation argued that this model may also be applicable in practice. However, the model should be extended to account for multiple attributes. Lohia et al. [Lohiya and Ragha, 2012] proposed a hybrid model based on perturbation and then applied generalization methods to protect privacy. The evidence suggests that perturbation alone is rarely sufficient and that hybrid models tend to be more convincing. Similar observations were reported in [Balasubramaniam and Kavitha, 2015, Vlachos et al., 2015, Kadampur and Somayajulu, 2008].

In summary, the limitations of perturbation-based methods demonstrate their weakness in the trade-off between privacy and information loss. Despite their ubiquity and ease of use, perturbation methods appear to be most useful when combined with other techniques. Although such methods are straightforward and may preserve a degree of accuracy, they are no longer sufficiently secure in the face of increasingly capable adversaries.

1.3.2.2 Soft Computing

In the previous sections, conventional methods of privacy protection were classified and reviewed. A common characteristic of these methods is the attempt to achieve a trade-off between the level of privacy for individuals and the performance of data-mining techniques in producing valid results. The tension between these two goals motivates an examination of another branch of computer science that has become increasingly productive in many applications. Soft computing is a combination of methodologies that provides flexible knowledge-processing tools for handling ambiguous real-world problems in a synergistic manner [Yager et al., 1994]. It is often

associated with robust and computationally efficient approximate solutions in terms of time and space complexity [Maimon and Rokach, 2007].

Soft-computing methods have been proposed as promising approaches for addressing complex data-analysis problems [Mitra et al., 2002]. Unlike conventional exact methods, soft computing emphasizes approximate reasoning and tolerance to uncertainty, partial truth, and imprecision. It provides robust, human-interpretable, and relatively low-cost approximate solutions compared with those developed using conventional techniques [Maimon and Rokach, 2007]. For this reason, it is worthwhile to provide an overview of state-of-the-art soft-computing methods in order to justify why they deserve attention in the context of privacy protection. Fuzzy logic, neural networks, and genetic algorithms are three prevalent soft-computing approaches investigated in this thesis and are described in the following chapters.

1.3.2.3 Differential Privacy

In 2006, Dwork refined earlier work on privacy-preserving data analysis and proposed the framework of differential privacy [Blum et al., 2005, Dwork, 2006, Dwork et al., 2006]. As a privacy criterion, differential privacy can be achieved through distortion mechanisms and does not require assumptions about the adversary’s background knowledge or computational power. In this technique, random noise is added to the data in order to protect personal privacy against inference from indirect information [Dwork, 2008, Dwork, 2006]. Dwork argued that, in a provable privacy model, the output of a query should not permit reliable inference about any particular record in the input.

An ϵ -differential privacy guarantee is satisfied when, for a privacy parameter $\epsilon \geq 0$, the following inequality holds for all neighboring databases D and D' that differ in at most one record, and for all measurable subsets $M \in \text{Range}(F)$.

Definition 1.1 (ϵ -Differential Privacy) Let D and D' be adjacent datasets that differ in at most one data record. Let F denote a randomized function with output space $\text{Range}(F)$. If F satisfies ϵ -differential privacy, then for all $M \in \text{Range}(F)$:

$$\Pr[F(D) \in M] \leq \exp(\epsilon) \Pr[F(D') \in M] \quad (1.1)$$

This condition ensures that the output of a query changes only slightly regardless of whether a single record in the database is modified [Dwork et al., 2006].

Two parameters are central to differential privacy. First, the privacy budget parameter ϵ determines the magnitude of the difference between two neighboring datasets after the addition of random noise. A lower value of ϵ provides a stronger privacy guarantee. Second, the amount of noise added depends on the local sensitivity Δf of the evaluated function, where f is the query function applied to dataset D . In general, the local function sensitivity is defined as:

$$\Delta f(D) = \max_{D'} \|f(D) - f(D')\| \quad (1.2)$$

where $\|\cdot\|$ denotes the chosen norm.

Among the various mechanisms used to add noise, the Laplace mechanism is adopted in this thesis, relying on the sensitivity of the function being protected [Dwork, 2008]. The formal function of the Laplace mechanism is defined as follows:

Definition 1.2 (Laplace noise) Let the location parameter be 0 and the scale parameter be $b = \Delta f / \epsilon$. The noisy function is defined as:

$$\text{Laplace}(b) = \exp\left(-\frac{|r|}{b}\right) = \exp\left(-\frac{|r| \times \epsilon_i}{\Delta f}\right) \quad (1.3)$$

Here, r is treated as a random parameter set during the initialization phase of SOM training. By adding noise, the availability of the data is altered. To simplify the problem, the sensitivity is assumed to be constant with $\Delta f = 1$. In this way, only one differing record is assumed to exist between neighboring databases, and the problem is regarded as an optimization problem in which ϵ is the sole variable.

Considering the limitations discussed in the previous section, a balance between privacy and availability is sought by determining acceptable values of the privacy budget ϵ through the following optimization function:

$$\min_{\omega} (F(D), \omega) \quad \epsilon \geq 0 \quad (1.4)$$

where D denotes the dataset, ω denotes the desired result, and $(F(D), \omega)$ denotes the objective function to be minimized.

Under the definition of differential privacy, a differentially private mechanism reflects the dataset as a whole and is not sensitive to the value of any individual entry [Friedman and Schuster, 2010]. Therefore, in this recommendation-system scheme, an adversary is unlikely to determine whether a particular person is contained in the database on the basis of any query response, since responses from neighboring databases remain indistinguishable within a given factor.

1.3.2.4 K-anonymity

K-anonymity [Samarati and Sweeney, 1998] is a well-known model for preventing privacy leaks. It requires that a sufficient number of indistinguishable individuals exist in published data, making it impossible for an adversary to identify the specific individual whose privacy is at risk [Ren and Yang, 2010]. In a k-anonymized dataset, with respect to a given set of identifying attributes, each record is indistinguishable from at least $k - 1$ other records [Machanavajjhala et al., 2007]. It is generally assumed that the dataset is tabular and that each record corresponds to an individual.

It has been shown that k-anonymity alone does not always ensure the privacy of a k-anonymized dataset, because it is associated with several subtle but serious privacy problems [Domingo-Ferrer and Torra, 2008]. Domingo analyzed related approaches such as p-sensitive k-anonymity, l-diversity, and t-closeness in order to identify their

shortcomings. According to this line of research, none of these methods is entirely convincing as a universally best choice.

Some studies have extended the original model by adding further constraints in order to improve the protection of sensitive information. For example, [Wong et al., 2006] proposed the α -k-anonymity model to protect both identifiers and relationships involving sensitive information. This method is NP-hard. It limits the confidence of implications from quasi-identifiers to raw attribute values within a threshold α , thereby reducing the risk that sensitive information can be inferred. Empirical results showed execution times that were approximately three times faster in local cases. However, more general cases still need to be evaluated because of the nature of diverse sensitive data.

Two important limitations of k-anonymity are the lack of protection for sensitive data when diversity is low and its weakness against attacks that exploit background knowledge. In these two situations, k-anonymity is vulnerable. [Machanavajjhala et al., 2007] therefore proposed the privacy criterion of l-diversity and demonstrated its effectiveness. They showed that k-anonymity and l-diversity are structurally similar enough that k-anonymity algorithms can be adapted to support l-diversity. Despite this improvement, l-diversity still needs to be extended to multiple sensitive attributes and further optimized in terms of algorithmic complexity. To minimize the disclosure risk of sensitive attributes, some researchers such as [Bhaladhare and Jinwala, 2016] proposed two approaches based on k-anonymity that create distinct combinations of quasi-identifiers and sensitive attributes. Their experiments, however, were limited to pure k-anonymity and did not consider different environments or settings. Similar work was presented by [El-Rashidy et al., 2010, Lin and Wei, 2009, Xiong and Zhu, 2012], who examined different aspects of sensitive attributes and the trade-off with optimization; however, none of these methods proved entirely convincing.

Because of its conceptual simplicity, k-anonymity has been extensively studied [Fung et al., 2005], but many research issues remain open, not only at the conceptual level but also in computational terms if an effective trade-off is to be achieved.

1.4 A Threat Model Against Recommender Systems

From a privacy perspective, a general threat model against recommender systems assumes that the adversary may gain access to the database through background knowledge. This includes both the training dataset and the trained model in the case of reconstruction attacks. The adversary's primary objective is to obtain information about data owners whose records are used to train the recommender system, whether the data were introduced through APIs or third-party services. From this perspective, the main goal of this thesis is to protect identifiable user data in a way that prevents the adversary from achieving this objective while simultaneously ensuring that the outcomes of model training remain practically useful in terms of utility and accuracy. The threat model addressed by the proposed solutions is illustrated in Fig. 1.1.

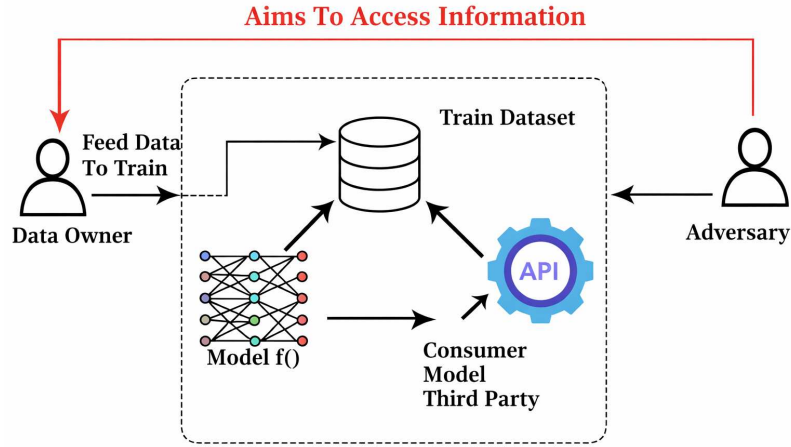


Fig. 1.1: Privacy threat model for recommender systems. The adversarial actions against the protected assets are indicated by red lines.

1.5 Taxonomy of Privacy Attacks Targeting Recommender Systems

Based on the amount of information available to the adversary, attacks are classified in this thesis as white-box and black-box attacks. In the former case, the adversary has access to the data and potentially even to trained models. In the latter case, the adversary does not have access to the training data or the trained model. The main types of privacy attacks considered in this thesis are described in the following sections and are used in later chapters to evaluate the effectiveness of the proposed algorithms. The attack vectors and details of each threat are illustrated in Fig. 1.2.

1.5.1 White-box Attack

A white-box attack refers to a situation in which the adversary has full access to the training dataset or even to the trained model. Such circumstances commonly arise in distributed training settings. This threat model is mainly understood as a background-knowledge attack, in which an association between one or more quasi-identifier attributes and a sensitive attribute is used to narrow the set of possible values of the sensitive attribute [Nasr et al., 2019]. White-box access is not necessarily the attacker's final objective; rather, such information may be used to accelerate the

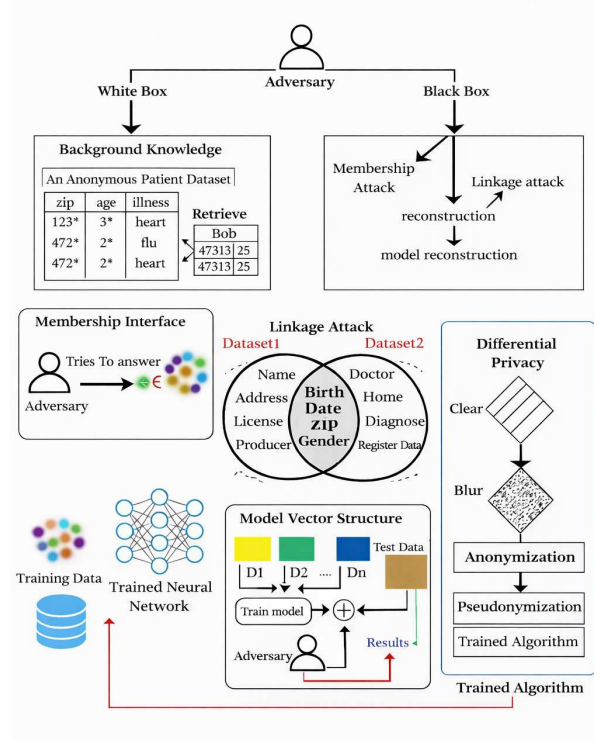


Fig. 1.2: Privacy attack vectors in recommender systems, defined according to the level of difficulty and the extent of data access available to the adversary.

search for identifiable user information. Consequently, the attacks classified in the black-box category may be regarded as potential subsequent actions.

1.5.2 Black-box Attack

Black-box attacks are attacks in which the adversary does not have access to the training data, model parameters, or the trained model itself. Strong security frameworks or controlled machine-learning applications may be responsible for this lack of access. Models used in machine-learning applications are generally not expected to reveal information directly about the data used for training. Nevertheless, the adversary still has opportunities to exploit the system and infer sensitive information. The major attacks in this group are shown in Fig. 1.2 and are described in the following sections.

1.5.2.1 Linkage Attack

A linkage attack consists of combining different datasets in such a way that the resulting information reveals facts that the subject did not disclose to the previously separate sources and, most importantly, did not intend to disclose [Solove, 2005]. One example is the increased risk of re-identifying anonymized data [Barbaro et al., 2006]. Linkage attacks in anonymized datasets occur when an adversary uses record identifiers or overlapping attributes to connect the target dataset to another dataset that still contains explicit identifiers.

1.5.2.2 Inference Attack

An inference attack occurs when a user is able to infer more sensitive information from seemingly trivial data in a database without direct access to the underlying protected information. This is effectively a data-mining process in which analysis is used illegitimately to obtain information about an individual in the database [Shokri et al., 2017]. If an adversary can infer the true value of sensitive information with high confidence, then that information can be considered leaked.

1.6 Motivation: Why Is Privacy Preservation Important in Recommender Systems?

This section explains why recommender systems were selected as the focus of this study and which privacy issues are central to the contributions of this thesis.

1.6.1 The Benefits of Training and Sharing User Data

The use of recommender systems has enabled personalized services in contexts such as web navigation and online shopping. The data are used by the recommender framework to improve the effectiveness of the platform and to adapt recommendations based on user preferences and contextual relevance. In other words, in recommender systems, the precision of recommendations and the extent to which users can share information about their preferences are closely related [Xin and Jaakkola, 2014].

In most real-world applications, recommender systems must address data scarcity. Owing to the lack of sufficient information, recommender systems struggle to generate recommendations for new products or users, a difficulty known as the cold-start problem [Gope and Jain, 2017]. Without sufficient data, model parameters and user preferences cannot be reliably estimated, and the integration of heterogeneous external knowledge becomes more difficult [Fernández-Tobías et al., 2016, Guo et al., 2020]. The use of a machine-learning framework for sparse-data training that main-

tains a balance between user utility and privacy is, therefore, a fundamental issue of this thesis, particularly in view of the risks associated with transmitting personal data. The corresponding relationship among these objectives is illustrated in Fig. 1.3. The next section examines the disadvantages of knowledge sharing and how they are addressed in this study.

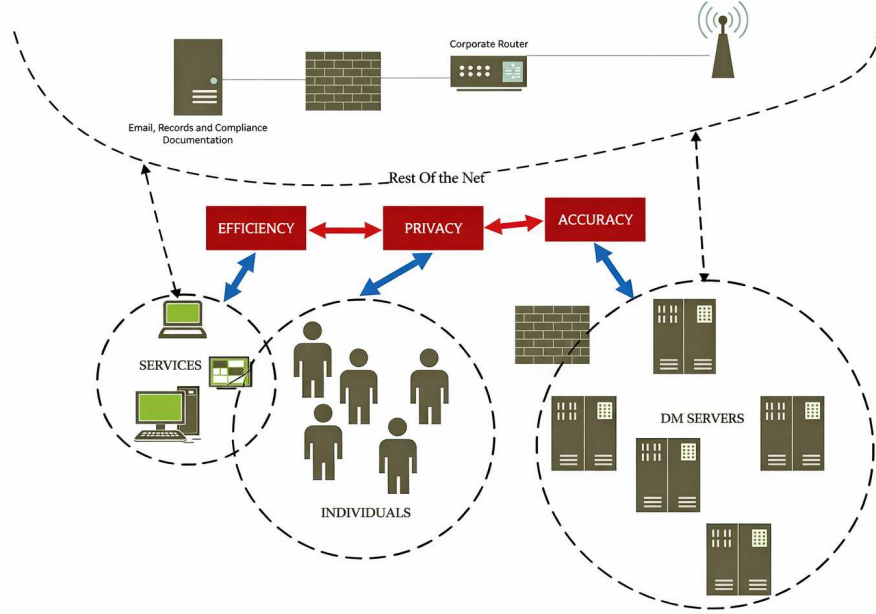


Fig. 1.3: Conceptual representation of the relationship between privacy preservation and core data-mining objectives—accuracy, computational efficiency, and output quality—illustrating the inherent trade-offs in privacy-preserving data analysis.

1.6.2 Privacy Threats in Online Platforms: The Case of Recommender Systems

With the rapid growth of online services, recommender systems have become an integral part of the digital economy. Recommendations are inherently based on mechanisms that rely on extensive personal data about users' interests in order to provide reliable and personalized results.

The privacy issues associated with collecting and processing such personal data are frequently overlooked or underestimated, and the sharing of personal infor-

mation can lead to data breaches. This situation is commonly referred to as the privacy–personalization trade-off [Chellappa and Sin, 2005, Awad and Krishnan, 2006, Kobsa, 2007, Zhang et al., 2014].

This thesis examines privacy in recommender systems in the context of personalization, with particular emphasis on user privacy, which is inherently implicated in the generation of personalized recommendations.

1.7 Summary

This chapter introduced the fundamental concepts relevant to this thesis, including the notion of privacy and its importance in recommender systems. It also reviewed key privacy challenges, common privacy-preserving approaches, and relevant threat models in order to establish the conceptual foundation of the study. The following chapter presents the structure of the thesis, formulates the research questions, and summarizes the main contributions.

Chapter 2

Structure and Contributions of the Thesis

Chapter Overview. This chapter presents the research objectives, problem statement, research questions, and scope of the thesis. It also outlines the main contributions of the study and explains how they relate to the corresponding research questions. Finally, the overall structure of the thesis is presented in order to guide the reader through the subsequent chapters.

2.1 Challenges and Research Questions

This section defines the scope of the study by focusing on privacy challenges in specific research areas and the associated objectives. It then explains the relationship between the research contributions and the corresponding research questions, followed by a statement of the research to be undertaken in this thesis.

2.1.1 Privacy Challenges Associated with Sharing User Data During Training

As discussed in Chapter 1, sharing personal information in recommender systems may lead to data breaches, while the implications of many information-disclosure practices remain unclear [Ramakrishnan et al., 2001]. This may expose sensitive information either about individual users or about the recommender system itself, potentially facilitating reverse engineering of the system. Consequently, recommender-system designers and service providers must find a balance that allows them to produce reliable recommendations while still protecting their users' privacy. The overall objective is therefore to preserve privacy while maintaining the accuracy and efficiency of the system (see Fig. 2.1).

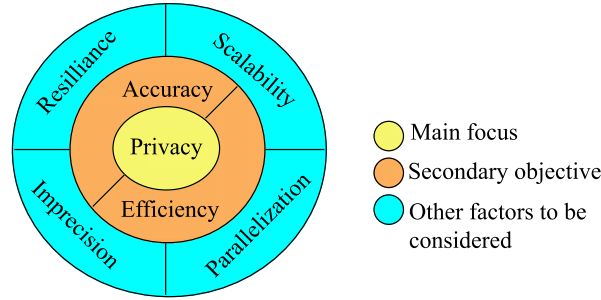


Fig. 2.1: Conceptual representation of the Privacy–Accuracy–Efficiency (PAE) trade-off in privacy-preserving data mining. Privacy represents the main objective of the system, while accuracy and efficiency act as secondary objectives that must be balanced. Additional factors, such as scalability, resilience, parallelization, and imprecision, further influence the overall performance and design of privacy-preserving systems.

2.1.2 Research Questions

This section presents the research questions that guide the investigation throughout the different phases of this thesis. To clarify how the research questions are addressed throughout the thesis, Table 2.1 summarizes the relationship between each research question and the chapter in which it is primarily investigated.

In this thesis, the research questions are examined from the following perspectives:

- **Application environment.** Each privacy-preserving method is evaluated within a clearly defined environment, characterized by its own requirements in terms of utility and privacy protection.
- **Data-mining task.** To limit the scope of the research, a selected set of data-mining tasks is considered for privacy improvement.
- **Evaluation expectations.** The expected outcomes of the proposed methods are defined in advance in order to establish thresholds for accuracy and utility in the final evaluation.
- **Methodological scope.** The set of privacy-preserving methods considered in this thesis is explicitly delimited in order to maintain a focused research scope.

Each chapter addresses a specific problem and explains how it contributes to the overall objective of the thesis. Table 2.1 summarizes the chapters in which the research questions are primarily addressed, while Chapter 8 revisits them in light of the overall findings of the thesis.

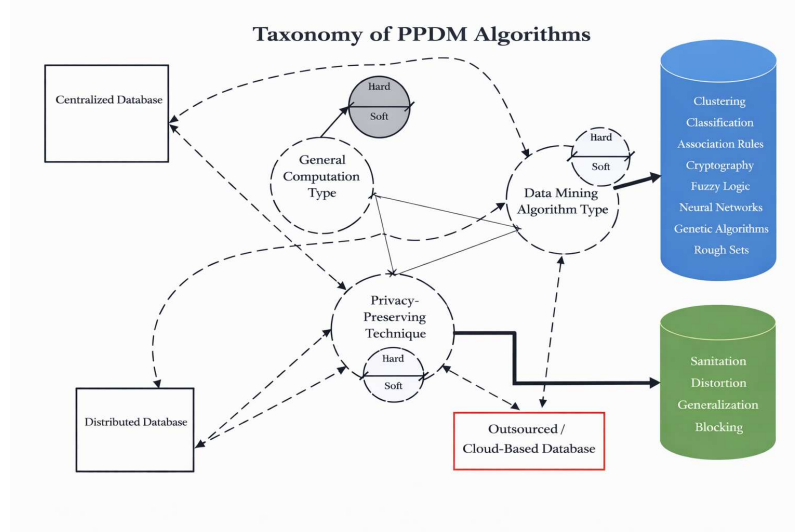


Fig. 2.2: Taxonomy of privacy-preserving data-mining algorithms. The framework classifies methods according to the application environment, computational paradigm, privacy-preserving technique, and data-mining task, thereby highlighting the relationships between conventional and soft-computing-based approaches in distributed and outsourced settings.

2.1.3 Aim, Scope, and Limitations of the Thesis

In this thesis, the challenge of privacy preservation in big data analysis is addressed in the context of data mining, motivated by the large volume of data processed in recommender systems.

Although numerous approaches have been proposed, the problem of preserving privacy in data mining remains open, particularly when privacy guarantees must be balanced against accuracy and reliability. Privacy concerns may reduce the accuracy of data-mining tasks and compromise the reliability of the overall process. Consequently, the problem can be understood in terms of three interrelated dimensions, which must be balanced by any effective solution.

A selected clustering approach, namely the self-organizing map (SOM), is used on the basis of the PAE variables shown in Fig. 2.2. The decision to use SOM was motivated by research showing that, in certain e-business applications, SOM outperformed hierarchical clustering techniques.

Mingoti and Lima [Mingoti and Lima, 2006] compared the clustering capabilities of Self-Organizing Maps (SOM) with those of other methods, including fuzzy c-means and k-means, across different data structures. Their findings indicate that, although SOM is not uniformly the most effective method, it performs well in several important respects, including topological preservation, interoperability, and suit-

ability for clustering high-dimensional data. In addition, SOM has been successfully applied in e-business and recommender-system contexts, which further supports its relevance for the present study. In [Roh et al., 2003], SOM clustering was applied to an e-business application, in which it was empirically demonstrated that the SOM-based recommender system provided higher-quality recommendations than other comparative models. Accordingly, SOM clustering is used in this thesis as a basis for improving privacy while also supporting the secondary goals of accuracy and utility.

To improve SOM, the approaches proposed in this thesis employ different soft-computing techniques. Some of these methods are novel and use techniques that have not previously been applied in this context. Others build on existing procedures and refine them further. The next chapters provide an overview of these approaches.

Several categorizations and methodological frameworks have been proposed in the literature. However, because new trends and objectives may alter the structure of such categorizations, a more focused framework is needed for the purposes of this thesis. The structure adopted here is defined in terms of the trade-off between complexity, correctness, and privacy level across conventional and soft-computing-based techniques. Its purpose is to clarify the state of the art in privacy-preserving data mining, establish conceptual relationships among the main approaches, and identify the associated challenges and research gaps. On this basis, a taxonomy of privacy-preserving data-mining (PPDM) methods combined with soft-computing techniques is proposed, building on the model introduced by [Wu et al., 2007].

The problem definition is presented in the following section, together with the perspective on the goals and limitations of the thesis.

2.1.3.1 Problem Statement

As stated in the previous section, no single privacy-preserving approach provides the best solution. This limitation is particularly evident in conventional PPDM methods, many of which are designed primarily for centralized environments and become less effective in distributed or outsourced settings. Another central issue is information loss, since many studies have shown that stronger privacy protection often comes at the expense of data utility. Better results are often achieved by combining two or more techniques.

Soft computing, due to its ability to handle uncertainty, ambiguity, partial truth, and imprecision, may provide more suitable solutions according to the findings of this study. However, these soft-computing approaches are not sufficient on their own and should be combined with reliable conventional techniques, such as association-rule-based methods and k-anonymity. Many other surveys discussed in Chapter 4 have focused either on developing new methods for stronger privacy preservation or on designing new algorithms for improved knowledge discovery. Accordingly, this thesis gives explicit consideration to combining soft-computing and conventional privacy-preserving methods in order to develop approaches that are both effective and practically applicable. Several reasons support this direction.

Soft-computing techniques are often computationally efficient, but when used in isolation they typically provide weaker privacy guarantees than conventional approaches. Some soft-computing approaches, such as rough sets, may offer stronger privacy-related capabilities; however, their mathematical complexity may make them more difficult for practitioners to implement and interpret. Conversely, most existing PPDM approaches perform best in centralized databases and are less effective in newer distributed settings, including cloud-based environments. Cryptographic techniques may provide strong protection in distributed databases, but their execution time and resource requirements can be prohibitive in practice. Soft-computing techniques may help address this limitation because of their flexibility and computational efficiency. Consequently, a novel combination of different approaches from both classes may provide a better solution to the three dimensions illustrated in Fig. 2.3. The problem addressed in this thesis is therefore formulated as follows.

A privacy-preserving approach for distributed systems with large volumes of data flow, such as recommender systems.

The main aim of the proposed approaches is to protect privacy against established privacy threat models. In addition to the primary privacy objective, a secondary aim is to maintain the utility and accuracy of the common processes within the system. The proposed methods are intended to combine modern privacy-preserving mechanisms with the functional strengths of conventional approaches.

2.2 Contributions

2.2.1 Protecting Privacy When Training and Sharing User Data

Our privacy-improvement methods are applied to recommender systems, which are considered a critical component of the e-commerce industry and are affected by insufficient privacy protection. Recommender systems are employed by e-commerce sites to assist customers in selecting appropriate goods, as the task of selecting suitable products becomes increasingly difficult as the number of available options grows [Ahn, 2008]. In such systems, ratings can be collected from a sufficiently large number of users so that accurate and reliable predictions can be generated. When access is granted to only a small amount of user data, the formation of sufficiently accurate and broad neighborhoods is hindered, resulting in low-quality recommender services. Furthermore, when the number of users is insufficient, a cold-start problem is encountered, in which recommendations can be predicted for only a limited number of products because adequate data are not available.

The lack of consistency in recommendations may result in a loss of customers [Bobadilla et al., 2012, Clifton et al., 2004]. Companies, on the other hand, may

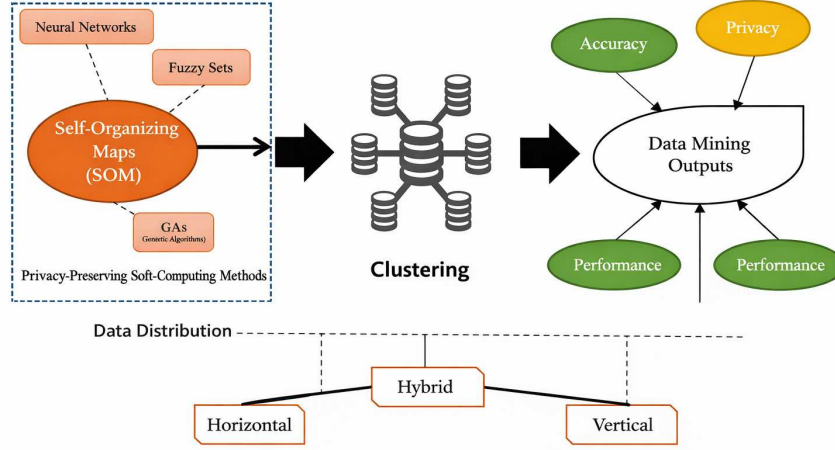


Fig. 2.3: Conceptual framework of the proposed approach. Clustering is applied in a distributed setting, with either horizontal or vertical partitioning, while privacy, accuracy, and performance are jointly considered through a hybrid soft-computing strategy.

be unwilling to exchange personal information because of concerns related to competitive intelligence and antitrust restrictions. Consequently, combining data in such applications is often unavoidable; however, this can only be achieved if privacy is maintained. Without adequate privacy protection, the participating parties may be reluctant to cooperate for legal, commercial, and ethical reasons.

To address this critical need, privacy-preserving methods for machine-learning algorithms in recommender systems are presented. With strong privacy protection and minimal impact on system utility, a self-organizing map technique combined with our privacy concept is applied to a recommendation system, as shown in Fig. 2.3. Previous studies have reported privacy and communication challenges in distributed clustering environments, which are also relevant to SOM-based methods [Bélanger and Crossler, 2011]. In our proposed approaches, these issues are addressed by incorporating soft-computing methods such as fuzzy sets, genetic algorithms, and differential privacy.

2.2.2 Assignment of Contributions and Research Questions (C & Q)

The primary objective of this research is to investigate privacy preservation in e-business environments. In this work, privacy gaps in PPDm are examined and explained, and a model is proposed that combines several modern techniques to reduce

Table 2.1: Mapping of research questions to the chapters in which they are addressed

RQ	Description	Chapter
RQ1	Definition of the objectives and scope of the privacy-preserving problem in data mining and recommender systems.	Chapter 2
RQ2	Identification and specification of relevant privacy attributes in e-business environments.	Chapter 2
RQ3	Determination of the criteria required to satisfy privacy preservation requirements in e-business applications.	Chapter 3
RQ4	Investigation of methods that can effectively address privacy-preserving challenges in data-mining systems.	Chapter 3
RQ5	Analysis of why soft-computing approaches may provide advantages when solving privacy-preserving problems.	Chapter 3
RQ6	Design of a suitable soft-computing algorithm aimed at reducing privacy gaps in recommender systems.	Chapter 5
RQ7	Evaluation of how soft-computing methods improve the privacy level of user data under potential privacy attacks.	Chapter 6
RQ8	Simulation and validation of the proposed model against privacy attacks such as linkage attacks.	Chapter 7

these gaps in specific data-mining tasks. This study makes several contributions to the state of the art in privacy-preserving data mining for distributed environments, with particular emphasis on recommender systems.

- A privacy-preserving clustering model based on soft computing is proposed. The model is designed to improve the Privacy–Accuracy–Efficiency (PAE) factors by combining several methods, with different soft-computing techniques employed to address specific objectives within the overall process.
- A set of algorithms is introduced to enhance privacy protection beyond conventional privacy-preserving data-mining (PPDM) approaches. In particular, soft-computing methods are used to improve privacy while reducing the limitations of traditional techniques. Within the scope of this thesis, a privacy breach is understood not only as the unauthorized disclosure of sensitive information, but also as the possibility of linking an individual to otherwise unrelated information, even with a low probability.
- Particular emphasis is placed on recommender systems. To this end, specific solutions are proposed for neural-network-based clustering in environments with both vertical and horizontal data distribution.
- The proposed algorithms are designed to preserve privacy while maintaining the accuracy and utility of the data-mining process.
- The integration of privacy-preserving components is expected not to impose substantial additional computational or communication costs, since the proposed methods are based on efficient modern techniques.
- The proposed methods aim to achieve a balanced trade-off among the PAE factors. In particular, privacy is strengthened while accuracy and performance are maintained or improved.

- The proposed soft-computing-based methods are intended to be flexible. Although they are applied in this thesis to clustering tasks, the underlying ideas may also be extended to other data-mining tasks, such as classification.

Chapter 8 revisits each research question and contribution in light of the overall findings of the thesis.

2.3 Thesis Structure and Focus Areas

This section outlines the structure of the chapters and the main focus of each part, which is divided into three parts.

2.3.1 Part II – A Need to Strike the Right Balance Between Utility and Privacy

In this part, the state of the art relevant to the thesis is examined from multiple perspectives. In Chapter 3, the importance of privacy protection in the field of e-business and recommender systems is discussed on the basis of real-world examples. In Chapter 4, a systematic comparative analysis of traditional and modern privacy-preserving approaches is presented, with particular attention to their respective strengths and limitations. Based on this analysis, the motivation for moving toward newer approaches is examined, and the research gaps to be addressed in the subsequent chapters are identified.

2.3.2 Part III – Algorithms for Training and Sharing User Data Privately

This part presents the two main models developed in this thesis to address the requirements of the problem statement. The first model, based on a metaheuristic approach, is described in Chapter 5, while the second model, grounded in differential privacy, is introduced in Chapter 6. Each chapter presents the corresponding model together with its analysis, experimental evaluation, and discussion.

2.3.3 Part IV – Quantifying the Privacy Level for Recommender Systems

This part presents a case study that builds on the contributions of the preceding sections. In Chapter 7, the proposed personalized recommender framework is applied to a real-world dataset and evaluated to demonstrate the effectiveness of the developed models. Chapter 8 concludes the thesis by summarizing the main arguments, identifying open problems, and outlining directions for future research.

2.4 Summary

This chapter presented the structure of the thesis, formulated the main research questions, and summarized the contributions of the study. It also outlined the scope of the investigation and explained the organization of the subsequent chapters. The following part begins with a review of the current state of research on privacy and utility trade-offs in digital environments.

Part II
Addressing the Trade-Off Between Utility
and Privacy

This section reviews the current state of research from several perspectives related to privacy-preserving data analysis. Chapter 3 discusses the evolution of privacy in the digital era and presents the state of the art in recommender system research, including selected real-world examples of privacy failures. Chapter 4 analyzes the two main categories of privacy protection approaches and compares their advantages and limitations. Based on this discussion, a research roadmap is presented that outlines how the subsequent chapters address the identified challenges.

Chapter 3

The Evolution of the Privacy Concept in the Digital World

Chapter Overview. This chapter examines the evolution of privacy in the digital era, with particular emphasis on online environments and recommender systems. It reviews major privacy failures, outlines the historical development of privacy as a concept, and discusses the state of the art in recommender-system privacy research. On this basis, the chapter identifies the research gap that motivates the subsequent contributions of this thesis. Research Questions 3, 4, and 5 are primarily addressed in this chapter.

3.1 The Emergence of the Concept of Privacy in the Digital World

Privacy has a long intellectual and legal history and has evolved in close connection with technological development. A key starting point in modern privacy literature is the seminal article by Warren and Brandeis [Warren and Brandeis, 1890], in which privacy was formulated as the right to be let alone. Since that time, successive advances in communication and information technologies have transformed both the meaning of privacy and the forms of threat against it. Historical analyses therefore emphasize that the development of privacy concerns is inseparable from technological progress and the expanding ability to collect, store, and process personal information [Warren and Brandeis, 1890, Holvast, 2007].

3.1.1 The Concept of Privacy as a Human Right

Early notions of privacy were associated with bodily integrity, the home, and personal autonomy, and later evolved toward the protection and control of personal information. Socrates and other Greek philosophers addressed the topic of privacy in their writings [Moore Jr, 1984]. In a famous paper published in 1890, Samuel

Warren and Louis Brandeis described the right to privacy as the right to be let alone [Warren and Brandeis, 1890]. With the publication of Alan Westin's *Privacy and Freedom* in 1967, a new milestone was reached when privacy was defined in terms of self-determination: privacy is the right of individuals, organizations, or institutions to choose when, how, and to what extent information about them is communicated to others [Westin, 1968]. However, the modern debate on privacy intensified after the Second World War, particularly in response to the growing use of computers and large-scale data storage. These publications focused on defining the notion of privacy and examining the development of privacy-invasive technologies, particularly computers, which were increasingly viewed as a major source of privacy risk. The project *The Impact of Science and Technology on Privacy* was initiated in 1962 in response to these concerns [Westin, 1966a, Westin, 1966b].

At the same time, a new dimension of privacy rights emerged, namely the use of computers as primary devices for data storage and processing. Several authors emphasized that computers and databases introduced unprecedented risks through the large-scale storage and processing of personal information. A number of publications raised serious warnings regarding large-scale intrusions into personal privacy, stressing that the central problem lay in the mass collection and electronic processing of personal data [Warner and Stone, 1970, Miller et al., 1971].

3.1.2 Digital Privacy

In the context of e-business, the term *digital privacy* is frequently used to describe concerns related to individual and consumer privacy rights [Rice and Sussan, 2016, Bergemann et al., 2019]. In 2001, research on internet use examined the growing popularity of online commerce together with users' concerns regarding privacy and security [Brown, 2001]. Recent studies have focused on how user data are voluntarily disclosed or inferred from user actions, such as purchasing behavior over time, the formation of social links, patterns of platform use, or data collected through monetary transactions [Fainmesser et al., 2019]. Changes in the disclosure levels of Facebook profile items between 2005 and 2011 suggest that users increasingly preferred to keep more information private over time [Scholz, 2012]. However, over seven years, Facebook generated profits of approximately \$100 billion while collecting and sharing user data with third-party marketers [Stutzman et al., 2013].

The following table, adapted from [Bergemann et al., 2019], summarizes the evolution of information privacy described in this section.

Table 3.1: Privacy evolution steps

Timeline	Details
Privacy emergence 1950s	Privacy is defined as a human right. Limited legal and governmental support due to the low level of information technology.
Contemporary privacy age 1960s–1970s	Privacy increasingly becomes a legal and social issue. Privacy-invasive technologies become more prevalent. Rapid advances in information technology. Growing recognition of privacy as a topic in computer science. Development of established frameworks and regulations such as the Privacy Act of 1974.
Third generation of privacy issues 1980s–1990s	Technological development accelerates further. Greater government regulation and institutional support emerge.
Fourth generation and the big data challenge 2000s–present	The rise of internet users, big data, and recommender systems. Reported privacy breaches become increasingly frequent.

3.2 Privacy Violation Examples in Online Environments

The following examples illustrate major privacy violations in online environments. They include both deliberate attacks and cases in which data were insufficiently protected and became publicly exposed.

Large-scale privacy breaches have become increasingly common. In some of the most significant incidents of the twenty-first century, hundreds of millions or even billions of records were affected.¹

Violation case (year). Netflix (2005)

Impact. 500,000 users

Details. In the Netflix case, the company released movie ratings of 500,000 subscribers in 2006 in an attempt to improve its recommendation algorithms. To preserve rating-to-rating consistency, Netflix anonymized the records by removing identifying information such as usernames while assigning each user a unique identifier [Narayanan and Shmatikov, 2006].

Although movie ratings are less sensitive than medical records, the release of large-scale behavioral data on individual Netflix users raised significant privacy concerns. Narayanan and Shmatikov showed that some Netflix and IMDb users were likely to rate the same films in similar ways and at similar times, which made it possible to link anonymized Netflix entries to identifiable users on IMDb [Narayanan and Shmatikov, 2008].

Violation case (year). AOL Search Log (2006)

Impact. 650,000 users

Details. In 2006, AOL publicly released a large number of user search queries. Although users were not explicitly identified, several queries contained personally identifying information. By cross-referencing the leaked and anonymized search data with phone book listings, the *New York Times* was able to identify an individual user.

¹ <https://www.csoonline.com/article/2130877/the-biggest-data-breaches-of-the-21st-century.html>

As a result, the case generated substantial debate regarding the ethical implications of using such data in research [Hafner, 2006, Hansell, 2006].

Violation case (year). Facebook Ads (2010)

Impact. Tens of millions of Facebook app users

Details. Facebook disclosed a flaw that allowed certain advertisers to view a user's name and the names of their Facebook friends, and potentially to monitor websites visited by the user. According to reports, this issue affected external applications accessed through Facebook rather than core platform features.²

Violation case (year). LinkedIn (2012 and 2016)

Impact. 165 million user accounts

Details. In 2012, LinkedIn reported that attackers had obtained user passwords. In 2016, it was revealed that the same hacker who sold MySpace data was also selling the email addresses and passwords of approximately 165 million LinkedIn members. LinkedIn stated that it had been informed of the breach and reset the passwords of all affected accounts.³

Violation case (year). Adobe (2013)

Impact. 153 million user records

Details. Adobe reported in October 2013 that attackers had obtained approximately 3 million encrypted customer credit card records, as well as login credentials for an unknown number of user accounts. Customers' names, IDs, passwords, and debit and credit card information were also exposed. In August 2015, Adobe agreed to pay \$2 million in legal fees to customers as part of a settlement.^{4,5}

Violation case (year). Yahoo (2013–2016)

Impact. 3 billion user accounts

Details. Yahoo was the target of one of the largest data breaches in history. In 2014, attackers gained access to 500 million real names, email addresses, birth dates, and phone numbers. A second breach occurred in 2016, when a separate attacker gained access to the names, dates of birth, email addresses, passwords, and security questions and answers associated with 1 billion user accounts. In October 2017, Yahoo revised its estimate to all 3 billion user accounts. These incidents reduced the company's value by an estimated US\$350 million.⁶

Violation case (year). eBay (2014)

Impact. 145 million users

Details. In May 2014, eBay reported that an attack had exposed the account information of 145 million users, including names, addresses, birth dates, and encrypted passwords. The company was criticized at the time for insufficient user communication and weaknesses in its password reset process.

² <http://abcnews.go.com/Technology/facebooks-privacy-controls-work/story?id=10760108>

³ <https://blog.linkedin.com/2016/05/18/protecting-our-members>

⁴ <https://krebsonsecurity.com/2013/10/adobe-breach-impacted-at-least-38-million-users/>

⁵ <https://www.csoonline.com/article/3268035/adobe-s-cso-talks-security-the-2013-breach-and-how-he-sets-priorities.html>

⁶ <https://www.csoonline.com/article/3180762/inside-the-russian-hack-of-yahoo-how-they-did-it.html>

Violation case (year). New York City Taxi Trips (2014)**Impact. 173 million individual taxi trips**

Details. New York City released data on 173 million individual taxi trips but inadvertently made it easy to identify drivers' personally identifiable information. Each record included pick-up and drop-off times and locations, as well as anonymized license and medallion numbers. Subsequent analysis showed that restoring the anonymized data to individually identifiable form was relatively straightforward [Douriez et al., 2016].⁷

Violation case (year). Adult Friend Finder (2016)**Impact. 412.2 million accounts**

Details. The stolen information included names, email addresses, and passwords and covered 20 years across six databases. As reported by CSO, a vulnerability in a module running on the company's production servers was being exploited at the time.⁹

Violation case (year). Equifax (2017)**Impact. 147.9 million consumers**

Details. A data breach affecting 147.9 million users was caused by an application vulnerability on one of Equifax's websites. The breach exposed Social Security numbers, birth dates, addresses, and, in some cases, driver's license numbers. Credit card information for approximately 209,000 individuals was also exposed. Equifax was widely criticized for broader security and incident-response failures.¹⁰

Violation case (year). Canva (2019)**Impact. 137 million user accounts**

Details. Canva, an Australian graphic design service, was hacked in May 2019, exposing 137 million users' email addresses, usernames, cities of residence, and salted and hashed passwords for users who did not rely on social logins. The attacker also claimed to have obtained OAuth login tokens for users who had signed in through Google.

Violation case (year). Microsoft (2020)**Impact. 280 million customer records**

Details. A customer support database containing approximately 280 million Microsoft customer records was left exposed on the web. The leaked data included email addresses, IP addresses, and support case information.¹¹

Violation case (year). 23andMe (2023)**Impact. Millions of user accounts / profile records**

Details. In 2023, 23andMe disclosed a security incident in which account compro-

⁷ <https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymized-data-researchers-warn>

⁸ <https://medium.com/@vijayp/of-taxis-and-rainbows-f6bc289679a1>

⁹ <http://www.csoonline.com/article/3132533/security/researcher-says-adult-friend-finder-vulnerable-to-file-inclusion-vulnerabilities.html>

¹⁰ <https://www.csoonline.com/article/3444488/equifax-data-breach-faq-what-happened-who-was-affected-what-was-the-impact.html>

¹¹ <https://www.zdnet.com/article/microsoft-discloses-security-breach-of-customer-support-database/>

mise exposed sensitive profile-related information. The case is particularly important because it involved highly sensitive personal and genetic-related data and demonstrated the privacy risks associated with credential-stuffing attacks.¹²

Violation case (year). BetterHelp (2023)

Impact. Sensitive health-related user data

Details. In 2023, the U.S. Federal Trade Commission took action against BetterHelp for sharing sensitive mental-health-related data with advertising platforms. This case shows that privacy violations may arise not only from external breaches but also from improper platform-side disclosure of personal data.¹³

Violation case (year). Change Healthcare (2024)

Impact. Approximately 190 million affected individuals

Details. In 2024, Change Healthcare was affected by a major cyberattack involving highly sensitive healthcare and personal information. The incident illustrates the large-scale privacy consequences of attacks on digital health infrastructures.¹⁴

Violation case (year). Ticketmaster (2024)

Impact. Customer data in a third-party cloud environment

Details. In 2024, Live Nation disclosed unauthorized activity in a third-party cloud database environment containing primarily Ticketmaster data. This case highlights the privacy risks associated with cloud-hosted infrastructures and third-party data management.¹⁵

3.3 Privacy in Recommender Systems: Literature Review

The first definitions of what are now known as recommender systems emerged approximately two decades ago [Goldberg et al., 1992]. Since then, numerous approaches have been introduced in both academia and industry [Schafer et al., 1999]. Early commercial recommender systems were studied extensively, and several of the systems examined are still in use. This section reviews how recommender systems operate, how they are applied, and which privacy issues arise in this domain.

3.3.1 Definitions

Collaborative filtering was first introduced in the form of a community-based system for annotating emails as interesting [Goldberg et al., 1992]. Various definitions have

¹² <https://www.sec.gov/Archives/edgar/data/1804591/000119312523287449/d242666d8ka.htm>

¹³ <https://www.ftc.gov/news-events/news/press-releases/2023/07/ftc-gives-final-approval-order-banning-betterhelp-sharing-sensitive-health-data-advertising>

¹⁴ <https://www.hhs.gov/hipaa/for-professionals/special-topics/change-healthcare-cybersecurity-incident-frequently-asked-questions/index.html>

¹⁵ <https://www.sec.gov/Archives/edgar/data/1335258/000133525824000081/lyv-20240520.htm>

been proposed in the literature; in this thesis, the definition by [Jannach et al., 2011] is adopted.

Let U be the set of users and let S be the set of all possible objects that can be recommended. A recommender system is an automated system that provides user u with recommendations drawn from the set S , including items with which u may be unfamiliar. Let F be a utility function that measures the usefulness of object s to user u . The goal of the recommender system is to maximize this utility for each user $u \in U$:

$$\forall u \in U, \quad s'_u = \arg \max_{s \in S} F(u, s) \quad (3.1)$$

Since the publication of the earliest studies on collaborative filtering, recommender systems have become a major research topic [Hill et al., 1995, Resnick et al., 1994, Shardanand and Maes, 1995]. More recently, the use of recommender systems in e-commerce has grown considerably, facilitating their application in various domains [Castro-Schez et al., 2011, Huang et al., 2007]. In recommender systems, data can be collected either directly, by gathering user ratings, or indirectly, by monitoring user activities [Lee et al., 2010, Núñez-Valdéz et al., 2012]. Pazzani [Pazzani, 1999] outlined three common types of filtering used in early recommender systems: collaborative, content-based, and demographic. Other types have also proven useful. A survey focusing on hybrid recommender systems was conducted by [Burke, 2002]. However, hybrid approaches often overlook the role of social filtering, a technique that has become increasingly relevant in recent e-business research. A collection of recommendations for constructing neighborhood-based prediction systems is provided by [Herlocker et al., 2002].

Recommender systems in the e-business domain are often classified into four main categories [Jannach et al., 2010, Li et al., 2011]: collaborative filtering, content-based filtering, knowledge-based techniques, and hybrid approaches. According to the survey presented by [Jannach et al., 2012], collaborative filtering is the most widely used strategy in recommender-system research, whereas other methods appear less frequently. The widespread use of collaborative filtering is due to the “wisdom of the crowd,” which is domain-independent and does not require precise, domain-specific knowledge of the recommended items. This classification was later supported by [Schafer et al., 2007, Candillier et al., 2007]. The next section discusses the main characteristics of these approaches.

3.3.2 Survey of Recommender System Types

A variety of methods have been proposed for the design of recommender systems [Deshpande and Karypis, 2004]. Among these, collaborative filtering has proven particularly effective in e-commerce applications [Van Meteren and Van Someren, 2000]. Before focusing on the recommender-system type adopted in this thesis, this section briefly reviews the most common classes of recommender systems

discussed in academia and industry. Several surveys have examined the state of the art in recommender systems and their extensions [Adomavicius and Tuzhilin, 2005, Bobadilla et al., 2013, Burke, 2002].

- **Content-based.**

Content-based recommender systems analyze item descriptions and recommend items that are similar to those a user has preferred in the past [Pazzani and Billsus, 2007]. In this approach, recommendation is treated as a search for items whose content is closely related to the user's interests. Newsweeder [Lang, 1995] is a classical example of a content-based recommender system. The main advantage of this approach is that it can generate recommendations based on item characteristics without relying on the preferences of other users [Van Meteren and Van Someren, 2000].

- **Knowledge-based.**

In knowledge-based recommender systems, users specify their preferences explicitly, and recommendations are generated on the basis of domain knowledge encoded in the system. An example is Entrée [Burke, 2000], which was developed to assist users in finding suitable restaurants. Case-based recommender systems also belong to this category and may incorporate user feedback to refine later recommendations [Lorenzi and Ricci, 2003]. A limitation of this approach is that it often requires substantial domain knowledge and may involve greater modeling effort and longer development time.

- **Random prediction.**

A random prediction algorithm selects items arbitrarily from the set of available alternatives and recommends them to the user. Owing to its lack of predictive accuracy, this approach is not considered a practical recommender-system method and is mainly used as a baseline for comparison with more advanced techniques [Papagelis and Plexousakis, 2005].

- **Demographic.**

When detailed information about a user's preferences is unavailable, demographic information can be used to generate partially personalized recommendations. One early example is Rich's Grundy system [Rich, 1979]. However, because this approach relies on generalized user stereotypes, the level of personalization it provides is limited.

- **Collaborative Filtering.**

Collaborative filtering relies on user ratings or interactions in order to identify patterns of shared preferences. In general, the more users evaluate items, the more accurate the resulting recommendations become [Sarwar et al., 2001]. One of the earliest collaborative filtering systems was introduced by Goldberg et al. [Goldberg et al., 1992] for filtering Usenet messages according to user interests. Another influential system was GroupLens [Resnick et al., 1994], which generated automated recommendations for Usenet postings based on the preferences of similar users. Collaborative filtering may also be implemented through item-based similarity, in which relationships between items are computed from the rating matrix.

In addition to the approaches listed above, other recommender-system techniques have also been proposed. However, collaborative filtering offers several important advantages over many alternative approaches. In particular, it can handle complex and difficult-to-represent notions such as taste and quality, and it does not depend on explicit machine analysis of item content, which may be error-prone [Herlocker et al., 2000].

According to [Papagelis and Plexousakis, 2005], collaborative filtering has been widely adopted in both research and e-commerce applications because it provides an effective mechanism for filtering large volumes of information and identifying items that users are likely to prefer based on their past interactions. At the same time, trust and privacy remain critical issues in recommender systems.

Because users generally do not know how recommendations are generated, they must trust the correctness and reliability of the recommender system [Shyong et al., 2006]. Most recommender systems require users to register before receiving personalized recommendations, and during this process they are often asked to provide personal information such as names, birth dates, postal codes, and email addresses. Combinations of such attributes may function as quasi-identifiers and may therefore increase the risk of re-identification.

In addition, personal preferences expressed in recommender systems may themselves become quasi-identifiers, especially when users exhibit unusual or highly specific preferences [Sen et al., 2006]. Such information may enable the re-identification of users across systems. Recommender systems also face broader security concerns, since they are frequently targeted by malicious users and adversarial manipulation attempts [Chirita et al., 2005]. Earlier research on recommender systems focused primarily on algorithmic performance [Herlocker et al., 2004, Adomavicius and Tuzhilin, 2005]. More recent studies, however, have shown that recommendation accuracy alone is not sufficient; user trust and privacy protection are equally important [Jeckmans et al., 2013]. The next section therefore examines privacy issues in recommender systems in greater detail.

3.3.3 Privacy in Recommender Systems: The Trade-off Between Privacy and Accuracy

Recommender-system designers and service providers must understand how reliable recommendations can be created while protecting user privacy. This section summarizes the main privacy risks affecting recommender-system users and reviews technological approaches proposed to mitigate them. Recommender systems may jeopardize user privacy during data collection [Ricci et al., 2011]. Some studies suggest that users are willing to disclose a certain amount of personal information for recommendation purposes; however, they may still be concerned that systems collect too much information, make inaccurate inferences, or use the data for inappropriate purposes [Regard, , Awad and Krishnan, 2006, Chellappa and Sin, 2005, Kobsa,

2007]. Knijnenburg et al. [Knijnenburg et al., 2017] distinguish three categories of privacy threats.

The first category arises from the recommender system itself and includes risks such as unsolicited data collection, targeted advertising, and discrimination [Knijnenburg and Kobsa, 2013, Mikians et al., 2012]. The second category involves malicious internal processes and adversarial users, who may exploit recommender outputs in order to infer private information about a target user [Knijnenburg et al., 2017]. The third category concerns external entities that may attempt to obtain acquired data or interfere with the interaction between the system and its users [Narayanan and Shmatikov, 2010].

Different mechanisms have been introduced to address privacy issues in recommender systems. Studies such as [Cissée and Albayrak, 2007, Heitmann et al., 2010] emphasize distributed architectures, portable user profiles, and client-side recommendation algorithms that restrict user-data access and reduce linkability. Anonymization, obfuscation, differential privacy, and homomorphic encryption are among the algorithmic strategies that have been investigated for data protection [Canny, 2002, Dwork and Naor, 2010, Kobsa and Schreck, 2003, Wang and Kobsa, 2006].

A third set of studies focuses on laws, regulations, and industry self-regulation mechanisms that may be introduced by governments, legislative bodies, or organizations themselves [Kobsa, 2007].

Another important issue concerns privacy decision support, that is, the extent to which users can control their privacy settings. Several paradigms have been proposed to empower users and regulate privacy concerns. For example, [Cate and Mayer-Schönberger, 2013, Taylor et al., 2009] investigated the notice-and-choice model, which aims to maximize user control but may also create misunderstandings regarding the consequences of privacy settings [Madejski et al., 2012]. The privacy nudging paradigm seeks to increase the likelihood that users make beneficial privacy decisions [Acquisti, 2009], but it does not work equally well for all users because privacy preferences vary considerably [Interactive, 2000].

User-tailored privacy is a third and more adaptive approach, in which a system measures users' privacy-related traits and actions, uses this information to model their privacy preferences, and then adjusts privacy settings accordingly [Sivakumar et al., , Xhafa, 2016, Knijnenburg, 2015]. Within the scope of this thesis, the privacy decision support factor is treated as fixed rather than variable. Instead, this thesis contributes to algorithmic approaches based on anonymization and differential privacy, in light of the privacy threats outlined above. The next chapter provides a more detailed comparative review of these approaches.

3.4 Summary

This chapter examined the historical development of privacy and its transformation in the digital era. It also reviewed privacy failures in online environments and

analyzed the main concepts, system types, and privacy challenges associated with recommender systems. On this basis, the chapter identified the main research gap addressed in the subsequent comparative analysis of privacy-preserving approaches. The following chapter builds on this foundation by presenting a systematic comparison of conventional and soft-computing-based approaches to privacy preservation in data mining and e-business environments.

Chapter 4

A Comparative Study of Conventional and Soft-Computing-Based Privacy-Preserving Approaches in e-Business

Chapter Overview. The growing scale of data and the increasing demand for knowledge extraction have made privacy preservation a central challenge in knowledge discovery and data mining. This challenge is particularly important in big data environments, where privacy protection must be achieved without substantially reducing analytical accuracy or computational efficiency. This chapter presents a comparative analysis of conventional and soft-computing-based privacy-preserving approaches in KDD and data mining, identifies their main strengths and limitations, and proposes a conceptual framework for highlighting research gaps and future directions.

4.1 Introduction

Knowledge discovery is the process of extracting useful hidden information from raw data. According to [Xhafa, 2016, Sumaryanti, L et al., 2021], KDD seeks to identify patterns in data that are valid, novel, and understandable. Relevant data sources include medical histories, criminal records, financial accounts, and shopping habits [Madan, 2022]. Current trends increasingly involve storing sparse data in heterogeneous and distributed databases [Aggarwal and Philip, 2008]. Consequently, methods used in KDD must perform effectively in such environments. As the core component of the KDD process, data mining enables practitioners to extract useful knowledge from raw data in order to gain insights and competitive advantages in serving their customers [Pramanik et al., 2021].

Data mining enables the extraction of patterns, rules, and useful information, which may subsequently be shared with different parties. One important issue in knowledge discovery and data mining is privacy preservation [Machanavajjhala et al., 2007]. Within this context, sensitive and confidential data must be protected while preserving the utility of the data-mining process [Temuujin et al., 2019]. Early studies on this topic primarily focused on protecting sensitive data intended for publication [Muttoo, 2023]. Numerous approaches have been proposed to address

this issue. Some attempt to hide sensitive data, whereas others modify the data or introduce perturbations in order to protect it. Some techniques rely on cryptographic mechanisms, while others employ additive perturbation approaches [Chamikara, 2018].

Some studies, such as [Devi et al., 2020], classify privacy-preserving methodologies in data mining into two main types. The first type, referred to as input privacy, manipulates the data before applying data-mining techniques, ideally with little or no effect on the final results. The second type, referred to as output privacy, seeks to hide sensitive rules and patterns while minimizing alterations to the underlying data. A central issue in this area is the need to balance the efficiency of the data-mining process with the protection of sensitive information. This challenge motivates the present study and leads to the following question: where can an appropriate balance be achieved between privacy protection, efficiency, and the validity of data-mining results? This challenge is particularly important because many data-mining methods are computationally expensive. In distributed environments, the problem becomes even more pronounced. In such cases, even methods that provide strong privacy protection may fail to deliver practically useful results if their computational cost is too high.

This challenge motivates the consideration of soft computing. Soft computing refers to a class of computational methods designed to handle uncertainty, imprecision, and partial truth in complex problem settings [Chaturvedi et al., 2008]. Such methods have proven useful for handling uncertainty, improving efficiency, and addressing certain NP-complete problems. The problem considered in this study, which involves privacy, efficiency, and correctness, appears well suited to this class of methods. Accordingly, this chapter provides a comparative assessment of conventional and soft-computing-based approaches. In addition, it proposes a conceptual framework for categorizing privacy-preserving methods and identifying existing research gaps and directions for future work.

4.1.1 Privacy in e-Business and Knowledge Discovery

Privacy in e-business and knowledge discovery concerns the protection of sensitive personal and transactional data during collection, storage, processing, and analysis [Verykios et al., 2004]. In e-business environments, this issue is particularly important because commercial platforms routinely process financial records, behavioral data, and user preferences, often in distributed or outsourced infrastructures. As a result, privacy preservation is not only a legal and ethical requirement but also a technical challenge that directly affects the design of data-driven systems [Sumaryanti, L et al., 2021].

Knowledge discovery aims to extract useful patterns, rules, and insights from large datasets. In this context, privacy becomes a central concern because the data used for mining and analysis may contain sensitive or identifiable information. Privacy-preserving knowledge discovery therefore seeks to enable data analysis

while reducing the risk of unauthorized disclosure, re-identification, or misuse of personal information [Narwaria, 2016, Xhafa, 2016].

These concerns are especially pronounced in distributed environments, where data are often shared across multiple sites, third-party services, or cloud-based platforms. Under such conditions, privacy-preserving data mining (PPDM) methods must achieve an appropriate balance between privacy protection, analytical accuracy, and computational efficiency. This challenge motivates the comparative analysis presented in this chapter, in which conventional and soft-computing-based approaches are examined with particular attention to their suitability for distributed knowledge discovery and e-business applications [Madan, 2022].

4.1.2 Problem Definition

The privacy-preserving problem in data mining remains an open topic that requires more effective methods for protecting sensitive data. Several PPDM strategies have been introduced over the past decade; however, methods that improve privacy often reduce the accuracy and efficiency of data-mining tasks. Most existing solutions therefore seek to balance privacy protection with the quality of the knowledge obtained through the mining process. This challenge can be viewed from three inter-related perspectives—privacy, accuracy, and efficiency—and an effective solution must balance all three dimensions, as illustrated in Figure 4.1:

- improving privacy,
- maintaining efficiency, and
- minimizing information loss.

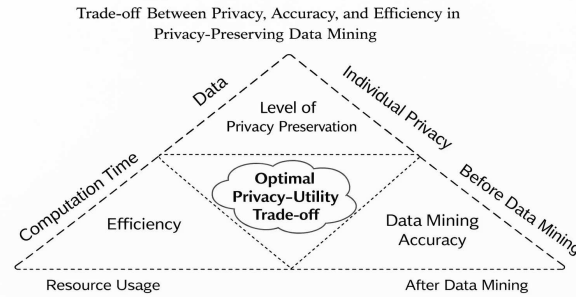


Fig. 4.1: Relationship between privacy preservation, accuracy, and efficiency in privacy-preserving data mining.

Figure 4.1 illustrates the interdependence between privacy preservation, data-mining accuracy, and computational efficiency. The goal is to approach an effective

balance among these three objectives. In e-business applications, additional factors, such as those shown in Figure 4.2, must also be taken into account. Although many models and algorithms have been proposed to reduce privacy risks, privacy is not the only relevant consideration. Accuracy, efficiency, and system-level constraints must likewise be considered. To the best of our knowledge, no existing method fully satisfies all Privacy–Accuracy–Efficiency (PAE) factors simultaneously.

Data-mining privacy concerns two main aspects: first, protection against the disclosure of individual or sensitive data, and second, prevention of the disclosure of sensitive patterns through the results of mining tasks [Madan, 2022]. In this study, the primary focus is on preserving the privacy of individuals. Based on this perspective, the following dimensions are emphasized:

1. **Application environment.**

This study focuses on distributed environments, where privacy and optimization concerns are particularly significant and relatively less explored. With the rapid growth of data and the increasing need for performance, identifying the challenges of privacy preservation in distributed settings has become especially important [Saxena et al., 2017].

2. **Privacy-preserving methodology.**

According to [Verykios et al., 2004], the privacy-preserving challenge may be NP-hard. Soft-computing algorithms are therefore considered promising because they can derive knowledge from complex data and uncover patterns and trends that may otherwise remain hidden [Malik et al., 2012].

3. **Selected data-mining task.**

Soft clustering is selected as the main application setting for the proposed privacy-preserving methods. This choice reflects the observation that different data-mining tasks exhibit different characteristics and behaviors, and it therefore helps to define the scope of the study more precisely.

SOM clustering is examined from several perspectives. SOM was selected on the basis of studies showing that, in certain e-business applications, SOM-based approaches can outperform hierarchical clustering methods. Within this thesis, SOM is extended through several methods based on different soft-computing approaches. Some of the proposed methods are entirely novel and employ techniques that have not previously been applied in this context, whereas others build on existing strategies and further refine them.

4.1.3 Methodology

This study adopts the design science research methodology, as the main objective is the development and evaluation of privacy-preserving algorithmic artifacts for knowledge discovery and data mining. This methodology is appropriate for the present work because it supports the systematic design, implementation, and assessment of technical solutions to a clearly defined problem. In particular, it provides a

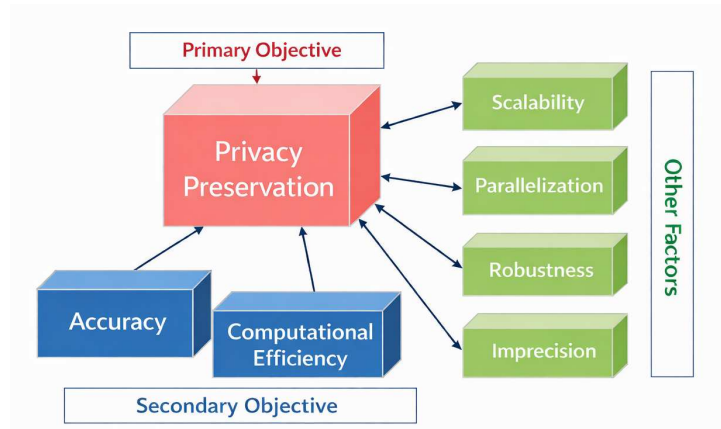


Fig. 4.2: Conceptual representation of the Privacy–Accuracy–Efficiency (PAE) trade-off.

suitable framework for developing and comparing privacy-preserving methods with respect to privacy, accuracy, and efficiency [Hevner et al., 2008]. The main steps of the adopted methodology are illustrated in Figure 4.3.

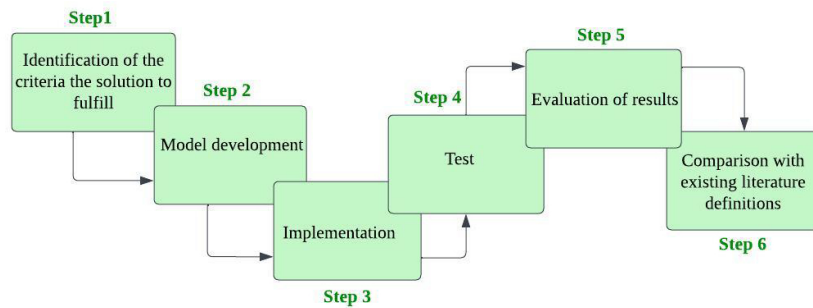


Fig. 4.3: Main steps of the design science research methodology adopted in this study.

4.1.4 Literature Selection Criteria

To perform an in-depth analysis of the literature relevant to the aims of this research, a content analysis methodology is applied. This methodology is commonly used in

Table 4.1: Applied framework of content analysis methodology [Zhao et al., 2022].

Step	Workflow
1	Selection of the study sample from the available population of documents. This step includes the specification of search criteria, the selection of the journal pool, the definition of the search string and time range, and the extraction of relevant research articles.
2	Determination of the unit of analysis and coding scheme. This review records the research methods, sample sizes, and descriptive information for each article.
3	Categorization of the content according to predefined rules.
4	Analysis of the data and identification of research gaps and issues.

rigorous review studies [Wafo Soh et al., 2020, Zhou et al., 2009]. It enabled the systematic identification, classification, and analysis of relevant studies. Accordingly, the literature review was conducted using the framework proposed by [Zhao et al., 2022]. This framework is shown in Table 4.1 and was applied throughout the present investigation.

Addressing the research questions defined earlier motivated a focused examination of soft-computing approaches and existing PPDM methods. To this end, more than 200 research publications available by the end of 2022 were analyzed, with particular attention given to studies addressing privacy-preserving techniques in data mining and knowledge discovery.

4.1.5 Keywords and Databases

The data were collected from the Scopus, Web of Science, ACM, Springer, ScienceDirect, and IEEE databases. The search was restricted by the following keywords: “privacy-preserving,” “PPDM,” “conventional methods,” “soft computing,” “review,” “survey,” “KDD,” “knowledge discovery,” “data mining,” “big data,” “distributed sites,” “distributed environment,” “scattered sites,” and the methods associated with each approach.

4.1.6 Contributions

The main contributions of this chapter can be summarized as follows:

- a comprehensive review of conventional and soft-computing-based methods for privacy preservation in knowledge discovery;
- the development of a conceptual framework for emerging PPDM classification trends; and

- the identification of research gaps and promising directions for future work.

4.2 Conventional Privacy-Preserving Approaches in Knowledge Discovery

Privacy-preserving data mining (PPDM) has become an important topic in the literature [Aggarwal and Philip, 2008]. The review of PPDM methods is necessary in order to identify their strengths, weaknesses, and research gaps. To provide a broader understanding of the state of the art, a number of studies have proposed different classifications of conventional PPDM methods [Pramanik et al., 2021]. Several conventional approaches can be considered for preserving privacy while still extracting useful information from data. The most common of these are discussed below.

4.2.1 The k-Anonymity Model and l-Diversity

The k-anonymity model aims to ensure that individuals cannot be re-identified from a dataset. A dataset is k-anonymous if each record is indistinguishable, with respect to selected quasi-identifiers, from at least $k - 1$ other records [Machanavajjhala et al., 2007]. In this way, the probability of re-identification is reduced.

The l-diversity model extends k-anonymity by requiring diversity in the sensitive attribute values within each equivalence class. It guarantees that there are at least l distinct sensitive attribute values within each k-anonymous group. By increasing the diversity of sensitive data within each group, this strategy seeks to limit information disclosure and reduce the risk of inference [Temuujin et al., 2019, Muttoo, 2023].

4.2.2 The Perturbation Approach

This technique introduces controlled random noise or perturbations into the original data in order to obscure sensitive information. The goal is to protect individual privacy without materially reducing the value of the data for analysis or decision-making. Depending on the data type and privacy requirements, several perturbation techniques may be applied, including transformation-based perturbation, noise addition, and randomized response [Chamikara, 2018, Devi et al., 2020].

4.2.3 Cryptographic Distributed Techniques

These methods rely on cryptographic algorithms and protocols in distributed computing environments in order to protect data privacy and enable secure cooperation among multiple parties. Such techniques allow secure processing and analysis while maintaining the confidentiality of sensitive data. Common examples include secure multiparty computation (SMPC), homomorphic encryption, secure function evaluation (SFE), and zero-knowledge proofs [Pinkas et al., 2002, Subramani et al., 2023].

4.2.4 Privacy-Preserving Outsourced Data Mining

This approach combines privacy preservation with outsourced data mining. It enables organizations to benefit from data mining while protecting the privacy and security of sensitive data transferred to third-party services. Since outsourced data may contain confidential information, privacy-preserving outsourced data-mining methods employ mechanisms such as data anonymization, secure data aggregation, secure computation protocols, and differential privacy to reduce the associated risks [Aldeen et al., 2015, Chauhan et al., 2020].

4.2.5 Privacy-Preserving Algorithms in Data Mining

Privacy-preserving algorithms in data mining are designed to protect sensitive information and preserve the privacy of individuals while extracting useful patterns and insights from data. Using different strategies, these algorithms seek to balance data utility with privacy protection. Typical examples include perturbation, differential privacy, data anonymization, and secure multi-party computation [Verykios et al., 2004].

4.2.5.1 Association Rule-Based PPDM

Association-rule-based privacy preservation refers to methods for protecting private information when mining associations or patterns from a dataset. Because association rule mining can reveal sensitive relationships among variables, privacy-preserving variants often incorporate mechanisms such as anonymization, differential privacy, or secure multi-party computation [Zhang et al., 2019].

4.2.5.2 Classification-Based PPDM

Classification-based PPDM comprises techniques designed to protect sensitive data while performing classification tasks. The objective is to build accurate predictive models while reducing the risk of disclosing private information. These methods typically preserve privacy during both the training and prediction phases, using mechanisms such as data anonymization, differential privacy, secure multi-party computation, and feature selection [Nasiri et al., 2020, Verykios et al., 2004].

4.3 Soft-Computing Approaches to Privacy Preservation in Knowledge Discovery

4.3.1 Fuzzy Logic

Fuzzy logic is a mathematical framework for handling uncertainty and imprecision in decision-making processes. In privacy-preserving knowledge discovery, fuzzy logic can be used in methods such as fuzzy clustering, privacy-preserving data perturbation, and fuzzy rule-based systems. These methods can help balance useful knowledge extraction with the protection of sensitive data [Afzali et al., 2018].

4.3.2 Neural Networks

Neural-network-based approaches to privacy-preserving knowledge discovery use machine learning models to perform data analysis while protecting sensitive information. Neural networks can identify complex patterns and relationships in data and can be combined with techniques such as secure multiparty computation, homomorphic encryption, and differential privacy in privacy-preserving settings [Barni et al., 2006, Li et al., 2017].

4.3.3 Rough Sets

Rough set theory provides a mathematical framework for data analysis under uncertainty and ambiguity. It is especially useful when the data are incomplete or imprecise, as it supports approximation-based reasoning. In privacy-preserving knowledge discovery, rough sets can be used for attribute reduction, privacy-preserving knowledge approximation, and rule generation [Wafu Soh et al., 2020, Zhou et al., 2009].

4.3.4 Genetic Algorithms

Genetic algorithms can optimize privacy-related parameters by iteratively evolving a population of candidate solutions. In knowledge discovery, they may be used to search for suitable trade-offs between privacy and utility [Zhao et al., 2022].

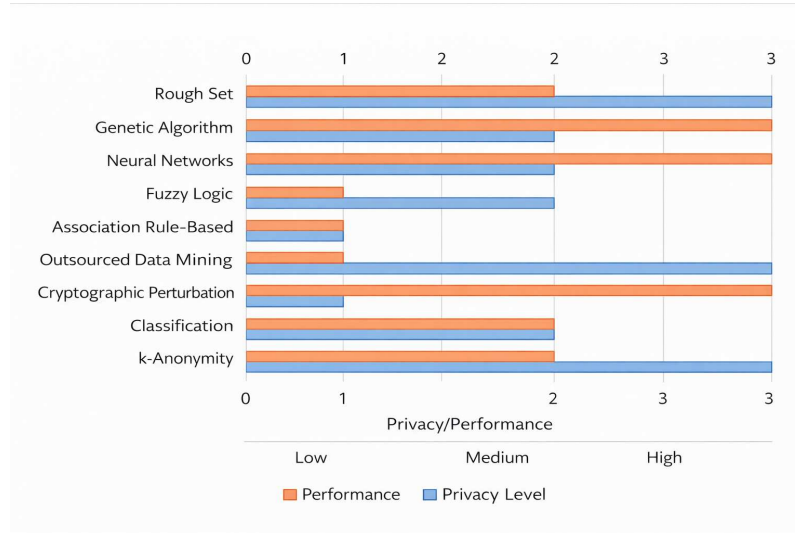


Fig. 4.4: Comparison of privacy level and performance across the evaluated methods, illustrating that neither conventional nor soft-computing approaches fully satisfy all PAE factors.

4.4 Analysis and Results

4.4.1 Key Characteristics and Evaluation Metrics

A detailed description of the main characteristics of the privacy-preserving methods identified in previous research is provided in Table 4.2. The purpose of this comparison is to clarify the main challenges involved in identifying an effective solution. The variables summarized in the table were derived from more than 200 articles published in conferences and peer-reviewed journals over recent years. Table 4.4 presents representative studies included in this review. This comparison provides a structured basis for evaluating the relative strengths and limitations of the considered approaches.

For each method, the following aspects are assessed: privacy level, correctness or information loss, time and resource performance, resilience to dynamic data insertion, parallelization, scalability, and tolerance to partial truth and imprecision. For example, traditional cryptographic methods provide strong privacy protection without removing data, but they often incur high computational cost, particularly in distributed settings. Their ability to support parallelization may nevertheless be advantageous in some dynamic scenarios.

As shown in Table 4.2, none of the methods is complete with respect to all evaluation criteria. Similarly, the comparison of privacy level and performance in Figure 4.4 indicates that neither conventional nor soft-computing approaches fully satisfy all PAE requirements. For this reason, hybrid soft-computing methodologies are considered promising candidates for the solutions developed in the subsequent chapters.

Table 4.2: Privacy-preserving method characteristic evaluation.

	PP Methods	Privacy level	Correctness / Performance		Resilience to dynamic data	Parallelization	Scalability	Tolerance to imprecision
			Information loss	(time/resource)				
Conventional methods of PPDM	k-anonymity	High	Yes	Medium	No	No	No	No
	classification	Medium	Yes	Medium	No	No	No	No
	perturbation	Low	No	High	No	Yes	No	No
	cryptographic-distributed	High	No	Low	No	Yes	Yes	No
	outsourced data mining	Low	Undefined	Low	No	Yes	Yes	No
Soft approaches	association-rule-based	Medium	Yes	Low	No	No	No	No
	fuzzy logic	Low	Yes	High	Yes	Yes	Yes	Yes
	neural net-	Medium	Yes	High	Yes	Yes	Yes	Yes
	works							
	genetic algo-	Medium	Yes	High	Yes	Yes	Yes	Yes
	rithms							
	rough sets	High	No	Medium	Yes	Yes	Yes	Yes

4.4.2 Comparative Analysis of Conventional and Soft-Computing Approaches

The previous sections reviewed and categorized conventional privacy-preserving data mining (PPDM) methods. Table 4.5 summarizes representative studies in this area and highlights the main advantages and limitations of traditional approaches. Despite their differences, these methods share a common objective: to achieve an appropriate balance between the protection of individual privacy and the ability of data-mining techniques to produce useful and reliable results. The tension between these two objectives motivates the investigation of alternative methodological directions.

Table 4.3: Representative peer studies included in this review.

Type	Method	References
Conventional techniques	k-anonymity & related topics	[Samarati and Sweeney, 1998, Ren and Yang, 2010, Samarati, 2001, Machanavajjhala et al., 2007, Fung et al., 2005, Domingo-Ferrer et al., 2005, Wong et al., 2006, Domingo-Ferrer and Torra, 2008, El-Rashidy et al., 2010, Lin et al., 2008, Xiong and Zhu, 2012]
	perturbation	[Jahan et al., 2014, Agrawal and Srikant, 2000, Aggarwal et al., 2008, Shah and Gulati, 2016a, Datta, 2004, Liu et al., 2005, Kargupta et al., 2003, Kargupta et al., 2005, Shah and Gulati, 2016b, Thuraishingham et al., 2006, Chen and Liu, 2011, Lohiya and Ragha, 2012, Balasubramaniam and Kavitha, 2015, Vlachos et al., 2015, Kadampur and Somaya-julu, 2008]
	cryptographic & distributed	[Verykios et al., 2004, Kargupta et al., 2005, Xu et al., 2014, Kerschbaum et al., 2014, Shah et al., 2014, McGregor et al., 2010, Sharma et al., 2010]
	outsourced data mining	[Chen and Liu, 2011, Tai et al., 2010, Giannotti et al., 2012, SANTHOSHI et al., 2015, Qiu et al., 2008, Nithin et al., 2017]
	association-rule-based PPDM	[Verykios et al., 2004, Xu et al., 2014, Wu et al., 2006, Wang et al., 2013, Moustakides et al., 2008, Dasseni et al., 2001, Verykios et al., 2008, Afzali et al., 2016, chandrakar et al., 2010, Jain et al., 2012, Rao et al., 2014]
	classification-based PPDM	[Yang et al., 2005, Zhang et al., 2005, Zhang et al., 2017, Delis et al., 2010, Gupta et al., 2011, Patal et al., 2014, Fung et al., 2007, Yu et al., 2006, Tripathy et al., 2012, Raghuram et al., 2012, Jia et al., 2016]
Soft computing	fuzzy logic	[Kumari et al., 2008, Poovammal et al., 2009, Mukkamala et al., 2011, Jahan et al., 2016, Ahmed et al., 2014, Torra et al., 2008, Kasugai et al., 2013, Kumar et al., 2011, Honda et al., 2012]
	neural networks	[Samet et al., 2012, Malik et al., 2015, Zadeh, 1996, Wimmer et al., 2014, Malik et al., 2016]
	genetic algorithms	[Grefenstette et al., 1993, Dehkordi et al., 2009, Matatov et al., 2010, Keshavamurthy et al., 2013, Mandapati et al., 2013, Zhu et al., 2009, Hong et al., 2010]
	rough sets	[Pawlak et al., 1995, Mitra et al., 2002, Hu et al., 2016, Ye et al., 2013, Zhou et al., 2009, Hu et al., 2008]

Table 4.4: Representative peer studies included in this review.

Type	Method	References
Conventional techniques	k-anonymity & related topics	[Samarati and Sweeney, 1998, Ren and Yang, 2010, Samarati, 2001, Machanavajjhala et al., 2007, Fung et al., 2005, Domingo-Ferrer et al., 2005, Wong et al., 2006, Domingo-Ferrer and Torra, 2008, El-Rashidy et al., 2010, Lin et al., 2008, Xiong and Zhu, 2012]
	perturbation	[Jahan et al., 2014, Agrawal and Srikant, 2000, Aggarwal et al., 2008, Shah and Gulati, 2016a, Datta, 2004, Liu et al., 2005, Kargupta et al., 2003, Kargupta et al., 2005, Shah and Gulati, 2016b, Thuraisingham et al., 2006, Chen and Liu, 2011, Lohiya and Ragha, 2012, Balasubramaniam and Kavitha, 2015, Vlachos et al., 2015, Kadampur and Somayajulu, 2008]
	cryptographic & distributed	[Verykios et al., 2004, Kargupta et al., 2005, Xu et al., 2014, Kerschbaum et al., 2014, Shah et al., 2014, McGregor et al., 2010, Sharma et al., 2010]
	outsourced data mining	[Chen and Liu, 2011, Tai et al., 2010, Giannotti et al., 2012, SANTHOSHI et al., 2015, Qiu et al., 2008, Nithin et al., 2017]
	association-rule-based PPDM	[Verykios et al., 2004, Xu et al., 2014, Wu et al., 2006, Wang et al., 2013, Moustakides et al., 2008, Dasseni et al., 2001, Verykios et al., 2008, Afzali et al., 2016, chandrakar et al., 2010, Jain et al., 2012, Rao et al., 2014]
	classification-based PPDM	[Yang et al., 2005, Zhang et al., 2005, Zhang et al., 2017, Delis et al., 2010, Gupta et al., 2011, Patal et al., 2014, Fung et al., 2007, Yu et al., 2006, Tripathy et al., 2012, Raghuram et al., 2012, Jia et al., 2016]
Soft computing	fuzzy logic	[Kumari et al., 2008, Poovammal et al., 2009, Mukkamala et al., 2011, Jahan et al., 2016, Ahmed et al., 2014, Torra et al., 2008, Kasugai et al., 2013, Kumar et al., 2011, Honda et al., 2012]
	neural networks	[Samet et al., 2012, Malik et al., 2015, Zadeh, 1996, Wimmer et al., 2014, Malik et al., 2016]
	genetic algorithms	[Grefenstette et al., 1993, Dehkordi et al., 2009, Matatov et al., 2010, Keshavamurthy et al., 2013, Mandapati et al., 2013, Zhu et al., 2009, Hong et al., 2010]
	rough sets	[Pawlak et al., 1995, Mitra et al., 2002, Hu et al., 2016, Ye et al., 2013, Zhou et al., 2009, Hu et al., 2008]

In this context, soft computing provides a promising complementary perspective. Unlike conventional approaches, soft-computing methods are characterized by their tolerance to uncertainty, partial truth, and imprecision [Malik et al., 2012]. As a

collection of methodologies, soft computing offers flexible knowledge-processing tools for dealing with complex real-world problems [Zadeh, 1996]. These properties make it particularly suitable for privacy-preserving data analysis, where robustness, tractability, and computational efficiency must be considered simultaneously [Maimon and Rokach, 2007, Mitra et al., 2002].

Recent data-mining research increasingly relies on algorithms capable of handling heterogeneous data sources and modeling complex relationships in practical settings [Malik et al., 2015]. In this regard, soft-computing approaches have attracted growing attention because they can provide workable solutions within limited computational resources. Moreover, many real-world datasets contain continuous and uncertain information that cannot be adequately represented by purely discrete or binary models. For this reason, the expanded set of characteristics and evaluation criteria presented in Table 4.2 and Figure 4.4 is used to illustrate why soft-computing methods are particularly relevant to privacy-preserving knowledge discovery in big data environments. On this basis, Figure 4.6 introduces a taxonomy that clarifies the relationships among the main privacy-preserving techniques considered in this chapter.

Table 4.5: Representative studies on conventional PPDM methods.

Study	Methods considered	Main findings
[Aggarwal et al., 2008]	Randomization, k-anonymity, l-diversity, distributed privacy preservation, data downgrading	Presents a broad classification of privacy-preserving methods and discusses several fundamental limitations of privacy preservation in data mining.
[Zhang et al., 2005]	K-anonymity, perturbation, cryptographic techniques, randomized response	Summarizes the advantages and limitations of several conventional techniques and introduces a condensation-based approach for constructing constrained clusters and generating pseudo-data.
[Narwaria, 2016]	Hiding sensitive information, secure multiparty computation, data distribution, data modification	Provides an overview of several well-known PPDM algorithms and their main categories.
[Natarajan et al., 2012]	Cryptographic techniques, generative techniques, data modification methods, data restriction techniques, data ownership techniques	Evaluates privacy-preserving methods with respect to privacy level, hiding failure, data quality, and complexity.
[Nayak et al., 2011]	Randomization, k-anonymization, distributed privacy preservation	Provides a detailed analysis of selected methods and discusses their strengths and weaknesses, although the survey covers a limited number of studies.

4.5 Insights and Implications

4.5.1 Conceptual Framework

In addition to evaluating the merits and limitations of existing techniques, this chapter proposes a conceptual framework for organizing the state of the art in privacy-preserving methods and identifying research gaps and future directions. Representative conventional PPDM and soft-computing approaches are shown in Figure 4.5. These techniques are discussed in detail in Sections 4.2 and 4.3. The review of more than 200 studies suggests that hybrid algorithms combining conventional and soft-computing methods often yield improved results. In Figure 4.5, this relationship is illustrated by the arrow connecting the two categories.

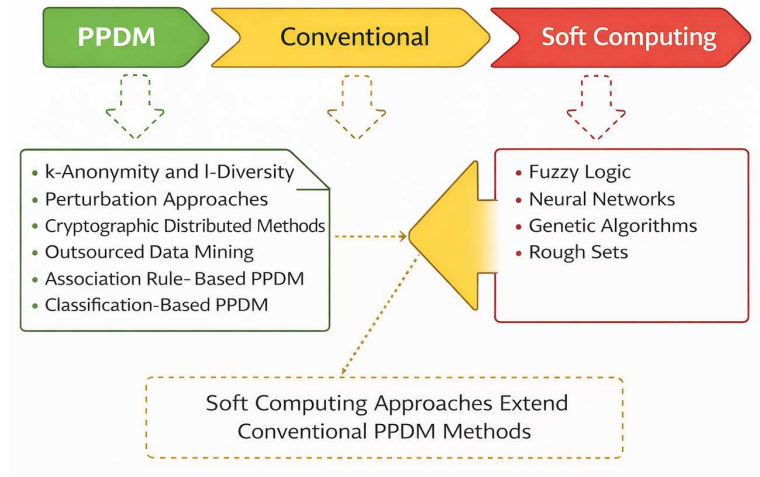


Fig. 4.5: Conventional and soft-computing-based privacy-preserving methods included in the comparative analysis.

This conceptual framework is particularly relevant to cloud-based outsourced environments, which are likely to become increasingly important as data volume and distribution continue to grow.

4.5.2 A Unified Taxonomy of Privacy-Preserving Methods

In addition to comparing the benefits and drawbacks of various privacy-preserving data-mining techniques, it is necessary to clarify how these techniques can be cat-

egorized. Such a taxonomy is particularly useful for researchers who wish to focus on specific branches of the field. Although many classifications and methodologies have been proposed in the literature, emerging approaches may alter the boundaries of existing categories. The framework presented here is intended to clarify the relationships among the state of the art, the main methodological branches, and the associated research gaps. The proposed taxonomy of PPDM methods is based on the model introduced by [Wu et al., 2007]. Figure 4.6 presents the suggested structure for privacy protection in knowledge discovery, with particular emphasis on outsourced environments.

Different algorithms may be appropriate for different scenarios and PP approaches, as shown in Figure 4.6. This taxonomy explicitly accounts for outsourced environments, such as cloud-based databases that manage large volumes of sparse and distributed data. Accordingly, careful attention must be paid to the core criteria of privacy protection when designing effective algorithms for such settings.

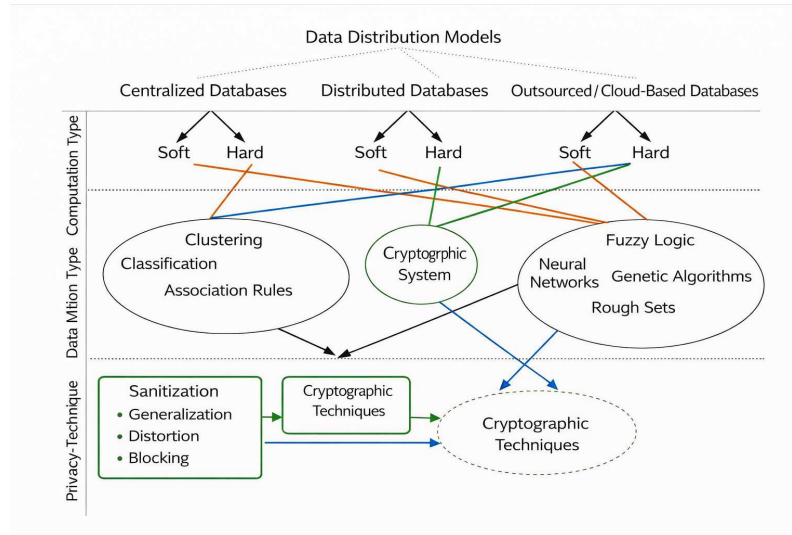


Fig. 4.6: Proposed taxonomy of privacy-preserving data-mining techniques based on data distribution and computation type.

4.5.3 Research Gaps and Future Directions

Based on the analysis of more than 200 publications, together with the proposed framework and the comparative results presented in the previous sections, it can be concluded that no single approach is sufficient to achieve an optimal balance

among privacy, accuracy, and efficiency. This limitation is particularly evident in conventional privacy-preserving data mining (PPDM) techniques, many of which are primarily designed for centralized environments and become less effective in distributed or cloud-based settings. In addition, information loss remains a central challenge, since stronger privacy protection is often achieved at the expense of data utility.

The comparative analysis further indicates that improved results are more likely to be achieved by combining multiple techniques rather than relying on a single method in isolation. Soft-computing approaches offer several attractive properties, including tolerance to uncertainty, partial truth, and imprecision, which make them promising candidates for privacy-preserving applications. However, these methods are generally not sufficient on their own and should be integrated with more established privacy-preserving techniques, such as association-rule-based methods and k-anonymity-based models.

The findings of this chapter therefore point to a clear research gap: the need for hybrid approaches that combine the strengths of conventional privacy-preserving methods with the flexibility and adaptability of soft-computing techniques. While soft-computing methods are often computationally efficient and well suited to complex and uncertain environments, they typically provide weaker privacy guarantees when used in isolation. Conversely, many conventional PPDM techniques offer stronger privacy protection but are less suitable for modern distributed and large-scale data environments. Cryptographic methods, for example, may provide strong protection in distributed settings, but their computational and communication costs can be prohibitive in practice.

Accordingly, this study suggests that future progress in privacy-preserving knowledge discovery is most likely to emerge from hybrid models that combine both classes of methods. Such combinations may offer a more effective way to address the three core dimensions of the problem—privacy, accuracy, and efficiency—as illustrated in Figure 4.2.

4.6 Summary

This chapter presented a comparative analysis of conventional and soft-computing-based approaches to privacy-preserving data mining. The review showed that no single method fully satisfies the joint requirements of privacy, accuracy, and efficiency. On this basis, the chapter proposed a conceptual framework for organizing existing approaches and identified hybrid approaches as a promising direction for addressing the limitations of individual techniques. These findings provide the conceptual foundation for the algorithmic developments introduced in the subsequent chapters.

Part III
Training Algorithms and Sharing of User
Data

This section presents the two main models developed in this thesis to address the defined problem statement. The first model, based on a metaheuristic approach, is described in Chapter 5. The second model, which is based on differential privacy, is introduced in Chapter 6. Each chapter presents the corresponding model together with its methodological analysis, experimental evaluation, and discussion.

Chapter 5

Metaheuristic-Based Strategies for Data Anonymization in Data Mining

Chapter Overview. Privacy-Preserving Data Mining (PPDM) requires methods that safeguard sensitive information while retaining the analytical value of data. This chapter presents GAFSOM (Genetic-Fuzzy Algorithm for Self-Organizing Maps), a metaheuristic anonymization framework that combines fuzzy sets with genetic optimization to balance privacy protection and clustering fidelity. By selectively applying fuzzification to high-risk attributes and using genetic search to minimize distortion, GAFSOM preserves the topological structure of the Self-Organizing Map (SOM), a property often overlooked in existing anonymization techniques. The approach is evaluated on two benchmark datasets, Adult and Bank Marketing, against a range of baselines including traditional SOM, k-anonymity, Fuzzy C-Means, genetic clustering, and differential-privacy-enhanced SOM. Experimental results show that GAFSOM achieves strong clustering accuracy, lower information loss, and reduced topographic error while maintaining competitive computational efficiency. In addition, structural analyses indicate that the method preserves SOM neighborhood relationships with only limited distortion under anonymization. These findings identify GAFSOM as an effective and scalable solution for privacy-preserving, topology-sensitive data-mining tasks. The methods proposed in this chapter are developed primarily to address Research Questions 6–8.

5.1 Introduction

Big-data approaches have been developed to address real-world problems and are applied across a variety of disciplines because of the large volume of data generated by many digital applications [Marx et al., 2013, Templ et al., 2022, Rodríguez et al., 2016]. At the same time, privacy concerns have become more complex, particularly in light of regulations such as the General Data Protection Regulation (GDPR), which aim to safeguard the fundamental value of collected datasets. In contemporary data environments, anonymization methods that preserve data utility while ensuring confidentiality are indispensable in sensitive domains such as financial transactions,

medical records, and census data. For example, preserving topological and semantic links between patient clusters while protecting individual privacy is critical when analyzing health records to detect disease spread [Rieke et al., 2020]. Similarly, socioeconomic data from government repositories, such as the Adult dataset, are essential for policy research but must comply with privacy regulations such as HIPAA and GDPR [Haverkamp et al., 2024, Amiri et al., 2019, Amiri et al., 2019]. This context highlights the dual challenge of extracting actionable insights from data while complying with privacy mandates.

Privacy-preserving data analysis has attracted considerable attention over the past two decades [Mendes et al., 2017]. Numerous approaches have been proposed for safeguarding privacy in association-rule mining, classification, and clustering tasks [Duggirala et al., 2016]. Conventional techniques such as perturbation, generalization, suppression, and k-anonymity provide only partial guarantees and often do so at the expense of utility [Mendes et al., 2017, Feng et al., 2014]. More recently, machine-learning-based methods have demonstrated strong potential for balancing these trade-offs, particularly in high-dimensional and complex data environments.

Our prior work [Amiri et al., 2019] categorized these approaches into conventional and soft-computing-based techniques and showed that methods rooted in evolutionary computation, fuzzy logic, and neural networks exhibit greater adaptability to privacy-utility trade-offs. Building on this foundation, the present study introduces a metaheuristic framework, GAFSOM (Genetic-Fuzzy Algorithm for Self-Organizing Map), which integrates genetic algorithms (GA) with fuzzy sets in order to preserve data confidentiality through anonymization while minimizing information loss. By embedding GA-driven subset selection and fuzzy-membership-based anonymization into the SOM clustering pipeline, GAFSOM aims to preserve structural fidelity and improve utility compared with traditional PPDM strategies.

The rationale for employing GA rather than alternative metaheuristics such as Particle Swarm Optimization (PSO) or the Grey Wolf Optimizer (GWO) is three-fold. First, the search space is mixed, consisting of discrete variables (attribute subset selection) and continuous parameters (fuzzy boundaries), a setting for which GA is particularly suitable [Amiri et al., 2018]. Second, the problem is inherently multi-objective, requiring a balance among accuracy, privacy, execution time, and topological preservation. GA's evolutionary diversity supports robust exploration of such rugged optimization landscapes and helps reduce premature convergence [Gad et al., 2022]. Third, GA has a demonstrated history in PPDM, where it is flexible enough to incorporate customized fitness functions, such as the hiding-failure metric used in this chapter.

Empirical evaluation on benchmark datasets shows that GA-based anonymization can provide strong clustering accuracy, reduced information loss, and competitive execution efficiency compared with classical SOM and swarm-based PPDM techniques. These properties make GAFSOM a promising foundation for privacy-preserving clustering in sensitive domains.

Beyond traditional privacy-preserving clustering, GAFSOM is relevant to current adversarial and security landscapes. Adversarial attacks on machine-learning models, including model inversion, membership inference, and reconstruction attacks, ex-

exploit structural weaknesses in data representations. By embedding fuzzy anonymization, GAFSOM introduces stochastic variation that may improve resilience against inference-based attacks. This is particularly relevant in healthcare analytics, where adversaries may attempt to reconstruct patient features from cluster outputs.

Similarly, the growth of Internet of Things (IoT) devices has generated unprecedented volumes of sensitive streaming data. IoT data are vulnerable to both passive background attacks, through the correlation of contextual metadata, and active probing attacks. GAFSOM's anonymization framework may be extended to IoT scenarios by dynamically fuzzifying sensitive streams, thereby supporting both compliance with privacy mandates and greater robustness against adversaries targeting edge or distributed systems.

Although validated here on the Adult and Bank Marketing datasets, the implications of GAFSOM extend to diverse real-world domains. In electronic health records, preserving topological relationships between patient subgroups while anonymizing sensitive features can support disease-trend analysis under GDPR/HIPAA constraints. In financial fraud detection, anonymization that preserves transaction topology may strengthen detection models while reducing re-identification risks. Likewise, in urban mobility and IoT sensor networks, GAFSOM could support compliance with emerging AI-governance frameworks while retaining the utility needed for smart-infrastructure optimization.

By combining metaheuristic optimization with privacy-preserving clustering, GAFSOM demonstrates that machine learning is not only feasible but also promising as a mechanism for balancing privacy and utility in real-world, high-stakes applications.

5.1.1 Motivation

In this chapter, a metaheuristic model is introduced for the protection of data confidentiality through anonymization. The minimization of information loss and the maximization of privacy protection are pursued through the combined use of genetic algorithms and fuzzy sets. Although a number of approaches have been proposed to mitigate privacy breaches in big-data processes, the accuracy and utility of data-mining results are often reduced in the process. In this chapter, utility is understood as both the accuracy and the efficiency (time and memory) of the overall process. Accordingly, a model is proposed in which data confidentiality is protected through anonymization. The resulting model is denoted GAFSOM and builds on the method presented in [Dua et al., 2017].

5.1.2 Distributed Approaches for Privacy-Preserving Data Mining (PPDM)

Although the current study focuses on static datasets, GAFSOM provides a basis for distributed PPDM in federated or streaming environments. Distributed anonymization frameworks, such as federated SOM or in-memory anonymization streams, have shown that parallel processing can reduce latency while scaling to very large datasets. GAFSOM could be integrated into such infrastructures by distributing GA population-based operations across multiple nodes.

A rough estimate suggests that distributing the GA population across p nodes can reduce optimization time by approximately $O(1/p)$ under balanced partitioning assumptions. For example, a clustering task on 10 million records that requires 200 minutes on a single machine could potentially be reduced to approximately 25–30 minutes on an 8-node cluster. In addition, distributed fuzzy anonymization can adapt to local data distributions, which may improve utility in heterogeneous environments such as smart cities or federated healthcare systems. This scalability aspect highlights GAFSOM's potential relevance for real-time, large-scale data anonymization.

5.1.3 Contributions

Although the method presented in this chapter shares certain conceptual similarities with traditional PPDM approaches such as differential privacy and k-anonymity, it is based on a fundamentally different paradigm, namely machine learning. The main contribution is the introduction of a new machine-learning-based privacy-preserving method. Current state-of-the-art research continues to indicate the need for approaches that combine strong protection with reduced information loss. The main contributions of this study are summarized as follows:

1. GAFSOM is introduced as a privacy-preserving algorithm in which genetic algorithms are integrated with fuzzy logic so that sensitive data can be anonymized while clustering utility is preserved.
2. GA-based subset selection is employed to handle mixed search spaces effectively, including both discrete attribute selection and continuous fuzzy parameters, thereby supporting a flexible balance between privacy and accuracy.
3. Through the incorporation of fuzzy sets for anonymization, the reliance on blanket generalization is reduced, while structural and topological consistency in SOM clustering is maintained.
4. A hiding-failure parameter is introduced as a fitness criterion in order to minimize information loss and improve the robustness of anonymization.
5. The model is evaluated on two real-world datasets, Adult and Bank Marketing, and its advantages over classical SOM and comparable metaheuristic approaches are demonstrated in terms of execution time, accuracy, topological preservation, and information loss.

6. The framework provides a practical trade-off between privacy and utility and positions GAFSOM as a suitable method for real-world applications in which both data protection and structural fidelity are important.

5.2 Related Work

Privacy-preserving data mining (PPDM) was introduced in [Agrawal and Srikant, 2000]. Applying privacy-improvement techniques may have side effects on running data-mining tasks and associated administrative processes. Smarter techniques, including machine-learning-based methods, may therefore help reduce such side effects. Machine learning, as a collection of methodologies, can provide flexible knowledge-processing facilities for handling ambiguous real-world problems in a synergistic way [Zadeh, 1996].

Genetic algorithms are often used to improve both privacy and accuracy in database-oriented tasks. For example, Jahan et al. [Jahan et al., 2014] used GAs because of their robustness with respect to local optima and their domain-independent suitability for large search spaces in exact or approximate optimization. In combination with fuzzy sets, the amount of processing time may be reduced relative to some of the best-known methods discussed in the introduction. The authors of [Jahan et al., 2016] used fuzzy methods for geometric data perturbation by perturbing multiple columns of a multivariate database through a single transformation. This transformation was designed to be invariant across different classifiers and clustering methods. Their results indicated that the hybrid method performed well across the membership functions considered.

Soft clustering is widely used in data mining for practical applications [Mingoti and Lima, 2006]. The empirical results in [Roh et al., 2003] show that an SOM-based recommender system in e-business applications can provide higher-quality recommendations than comparable models. Several methods have been proposed to reduce privacy gaps in SOM clustering, although their results did not fully satisfy utility-related criteria. Privacy-preserving SOM was introduced by [Han et al., 2007], and this was followed by a model proposed in [Malik et al., 2015] that sought to preserve privacy without sacrificing neural-network accuracy to a significant extent. The main advantage of this method is that it attempts to preserve both data relationships and privacy. In [Malik et al., 2016], the authors aimed to protect the privacy of individuals without significantly affecting neural-network prediction results while also reducing training time. In that model, privacy was preserved by generalizing and fuzzifying the data, and the neural networks trained on the transformed data performed better than those trained on the raw data.

There are other metaheuristic techniques, such as PSO and GWO [Shami et al., 2022, Makhadmeh et al., 2023]. Although these algorithms have been successfully applied in clustering and feature-selection tasks, their suitability for privacy-preserving anonymization is more limited. Swarm-based algorithms typically rely on continuous position updates, which makes them efficient for continuous optimization

but less adaptable to mixed-variable domains where both discrete (attribute subset selection) and continuous (fuzzy membership boundaries) parameters must be optimized simultaneously. By contrast, GA's flexible representation through bit strings and its evolutionary operators, such as crossover and mutation, are particularly advantageous for privacy-preserving problems that demand broad exploration of the solution space. Moreover, PSO and GWO often converge more quickly, but may also stagnate prematurely in high-dimensional multi-objective landscapes, which can weaken the balance between privacy and utility. GA, by maintaining population diversity through stochastic recombination, has a stronger ability to escape local optima and is therefore well suited to the trade-offs inherent in privacy-preserving SOM clustering.

Taken together, these studies support the view that machine-learning-based methods are promising for achieving a better trade-off between privacy and utility. However, most of the reported studies focus on only a single aspect of utility or privacy. In this chapter, SOM extensions are examined with the aim of addressing all three factors simultaneously: privacy, efficiency, and accuracy. The proposed method is conceptually related to [Amiri et al., 2018], where machine learning was used to strengthen privacy protection. However, in that work some sensitive data were removed in order to preserve efficiency, whereas the present approach seeks to pursue both privacy and efficiency without incurring substantial information loss.

5.3 Theoretical Foundations and Problem Statement

This section reviews the technical preliminaries underlying the designed algorithms.

5.3.1 Design and Formulation of the GAFSOM Method

In this chapter, a machine-learning-based algorithm is introduced that uses genetic algorithms and fuzzy sets to improve the privacy and utility of Self-Organizing Map clustering. The method is therefore referred to as GAFSOM. Its design builds on the method introduced in [Amiri et al., 2018].

Definition 1 (Input and Output). The input and output of the proposed GAFSOM model are defined as follows. Let F be the original database and let F_q ($1 \leq q \leq N$) denote a record in the database with a unique identifier FID and a set of sensitive data items to be hidden, denoted by $SX = \{SX_1, SX_2, \dots, SX_n\}$. Each SX_i is a field in the current record that satisfies user-defined criteria for being treated as sensitive. These parameters constitute the input values. Let F' denote the output of GAFSOM, namely the anonymized database. The transformed dataset F' is subsequently used as the input to SOM clustering.

5.3.2 GA Concepts

The proposed GA method seeks to minimize the extractability of transactions in dataset F that belong to the sensitive subset SX . To achieve this objective, a hiding-failure parameter denoted by α is used. This parameter represents the intersection between the original database, the sensitive subset S , and the anonymized database F' (see Figure 5.1).

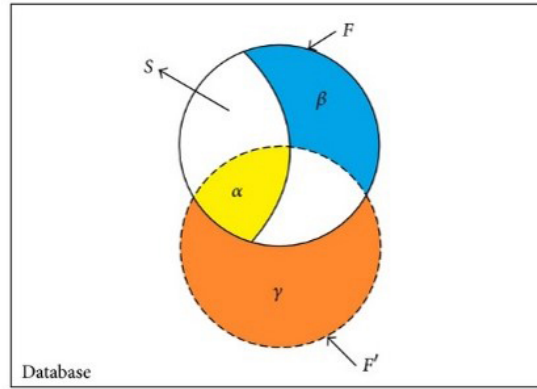


Fig. 5.1: Hiding-failure factor α used to minimize extractability from the main database [Wu et al., 2006].

Definition 2 (Hiding-failure value). To evaluate hiding failures for each processed transaction during sanitization, the α parameter is used. Figure 5.1 shows the relation between the main dataset and its intersection with reduced datasets.

When a processed transaction contains a sensitive item, the sum of the α values for the processed transaction T_j is calculated as:

$$\alpha^j(S_x) = \frac{MAX_{sx} - freq(S_x) + 1}{MAX_{sx} - \lceil |T| \times MST \rceil + 1} \quad (5.1)$$

where MAX_{sx} is the maximum number of sensitive records containing S_x , $|T|$ is the number of records, and $freq(S_x)$ is the frequency of S_x in the current record. MST (Minimum Support Threshold) is a predefined parameter that limits the number of records to be selected as sensitive. As a determinant of both the speed and quality of the GA procedure, MST is also employed as a termination condition. The total α value of each record is then defined as:

$$\alpha^j = \frac{1}{\sum_{i=1}^n \alpha^j(S_i) + 1} \quad (5.2)$$

Definition 3 (Fitness Function). The core of the GA method is the fitness function, which evaluates hiding failures for each processed transaction. To identify the optimal transactions containing sensitive items that should be transformed, the GA requires a dedicated fitness function. Based on [Hong et al., 2010], the fitness function is defined as:

$$Fitness = W_1\alpha + W_2\beta + W_3\gamma \quad (5.3)$$

where W_1 , W_2 , and W_3 are user-defined weighting parameters. The value α is computed using Equation (5.1). The factor β denotes the number of missing items, and γ denotes the number of artificial items. Given the strength of SOM clustering during the training phase and the desire to keep complexity low in distributed execution, $W_1 = 1$ is used in this chapter and the other factors are omitted.

Definition 4 (Threshold of Termination). The termination threshold is defined by the minimum-support-threshold ratio MST , which plays a central role in the GA process. This factor determines when the GA procedure should terminate. In other words, MST identifies when a candidate subset is regarded as sufficiently optimal.

Definition 5 (GA Operation Parameters). The efficiency and convergence rate of the algorithm depend on the population size, the number of generations, and the crossover and mutation operators. In the introduced algorithm, several parameter values are explored in order to determine acceptable settings for real-world databases. The details of these parameters are described below.

5.3.3 Fuzzy Set Concepts

In GAFSOM, the output of the GA phase is used as the input to a fuzzy function. The Triangular Membership Function anonymizes sensitive content using the *TriMF* relation.

Definition 6 (Fuzzy membership function).

$$TriMF(D, X, Y, Z) = \begin{cases} 0, & \text{when } D \leq X \\ \frac{D-X}{Y-X}, & \text{when } X < D \leq Y \\ \frac{Z-D}{Z-Y}, & \text{when } Y < D < Z \\ 0, & \text{when } D \geq Z \end{cases} \quad (5.4)$$

where D is the sensitive data value from the database, X is the lower boundary (the point at which membership begins to increase), Y is the peak (the generalized representative value with maximum membership, i.e., 1), and Z is the upper boundary (the point at which membership decreases back to 0). For example, the Age attribute may be generalized into ranges such as (20–30), (30–40), (40–50), and (50–70), while job categories may be anonymized into the broader difficulty domain (*high*, *medium*, *low*), as shown in Table 5.1.

Table 5.1: Example of fuzzy anonymization: original records and their anonymized counterparts. Sensitive attributes are generalized into broader fuzzy categories while preserving the clustering-relevant structure of the data.

Age	Job	Marital	Education	Housing	Age	Job	Marital	Education	Housing
56	Manager	Married	PHD	Yes	50–70	High	Married	*	Yes
25	Services	Single	Basic.6	Yes	20–30	Medium	Single	*	Yes
40	Admin	Married	High-school	No	30–40	High	Married	*	No
35	Technician	Married	Professiona	No	30–40	Low	Married	*	No
54	Retired	Divorced	Basic.9	No	50–70	Low	Divorced	*	No
37	Admin	Single	High-school	No	30–40	High	Single	*	No
25	Services	Single	University	Yes	20–30	Medium	Single	*	Yes

5.3.4 Kohonen SOM

A Kohonen map, also known as a Self-Organizing Map (SOM), is a neural-network model widely used in data-mining tasks, especially for clustering. SOM is an unsupervised learning model that can learn from complex multidimensional data and is well known for dimensionality reduction. It maps multidimensional data onto a two-dimensional grid. Its architecture consists of two fully connected layers: an input layer and a Kohonen layer, also referred to as a topology-preserving map [Yang et al., 2012]. The steps of the SOM clustering algorithm and the constants used in the algorithm are described in [Gan et al., 2020]. Figure 5.2 shows the topology of SOM clustering.

The two layers of neurons, an input layer and an output map layer, define the vector representation. Each output neuron is connected to each input neuron, which makes SOM a fully connected network. In privacy-preserving settings, SOM exhibits several communication-related privacy gaps:

- Securely identify the winning neuron from data privately held by two parties.
- Securely update the weight vectors of neurons.
- Securely determine the termination status of SOM [Han et al., 2007].

Definition 7 (SOM training). The SOM architecture comprises two fully connected layers: an input layer and a Kohonen layer, also called a topology-preserving map [Yang et al., 2012]. The steps of the SOM clustering algorithm and the constants used in the algorithm are described in [Gan et al., 2020]. Based on the constants defined by Haykin [Haykin 1998], in order to identify the winning neuron in the Kohonen layer, a random object x is selected from the input data X , and the winning neuron KN_i is determined by the minimum Euclidean distance between x and w_j , as follows. Here, w_j represents initial weights chosen randomly among objects in X for $j = 1, 2, \dots, T$, where T is the number of neurons in the Kohonen layer and s denotes the iteration:

$$KN_i^{(s)} = \arg \min_j \|x^{(s)} - w_j^{(s)}\| \quad (5.5)$$

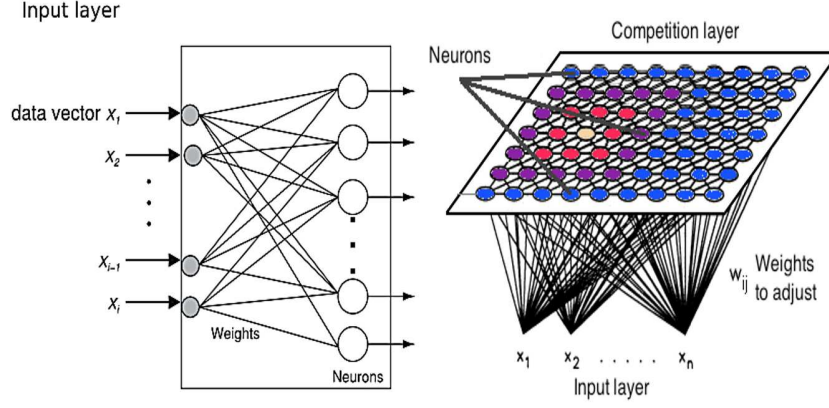


Fig. 5.2: SOM clustering topology [Amiri et al., 2018].

The weight vectors of all neurons are then updated as follows:

$$w_j^{(s+1)} = w_j^{(s)} + \eta(s) h_{j,i}(s) (x - w_j^{(s)}) \quad (5.6)$$

where $h_{j,i}(s)$ is the neighborhood function. The learning rate $\eta(s)$ and the neighborhood function $h_{j,i}(s)$ are computed using:

$$\eta(s) = \eta_0 \exp\left(-\frac{s}{\tau_2}\right), \quad s = 0, 1, 2, \dots \quad (5.7)$$

$$h_{j,i}(s) = \exp\left(-\frac{d_{i,j}^2}{2\sigma^2(s)}\right), \quad \sigma(s) = \sigma_0 \exp\left(-\frac{s}{\tau_1}\right) \quad (5.8)$$

These steps are repeated until no substantial change is observed in the resulting map.

Definition 8 (Evaluation of the Model). For measuring the results of SOM clustering, two factors are used:

- Quantization error (QE): the average distance between the current BMU and each data vector [Kohonen, 2001]:

$$QE = \frac{1}{N} \sum_{k=1}^N \|x_k - m_k\| \quad (5.9)$$

where N is the number of data vectors and m_k is the best-matching prototype corresponding to the data vector x_k . The smaller the value of QE , the better the algorithm.

- Topographic error (TE): describes how well SOM preserves the topology of the studied dataset:

$$TE = \frac{1}{N} \sum_{k=1}^N u(x_k) \quad (5.10)$$

where N is the number of input samples, and $u(x_k)$ is 1 if the first and second best-matching units of x_k are not adjacent; otherwise, $u(x_k)$ is 0. A smaller value of TE is more desirable [Kiviluoto, 1996].

The training process of SOM, as formulated by Haykin [Haykin 1998], is adopted in GAFSOM as the core clustering mechanism. In the proposed framework, SOM functions as the structural backbone, where best-matching units are iteratively updated according to Euclidean distance and neighborhood adaptation. This supports topological preservation and low quantization error, both of which are essential for reliable clustering. The integration with GAFSOM occurs before SOM training, when sensitive data items are anonymized through the GA-fuzzy module. Specifically, the genetic algorithm identifies optimal subsets of attributes to anonymize, and fuzzy membership functions transform these attributes into generalized representations. The anonymized dataset is then provided to the SOM training phase, ensuring that the clustering structure is learned from privacy-preserved data rather than raw sensitive inputs. In this way, Haykin's SOM methodology remains intact as the clustering core, while GAFSOM extends it through a preprocessing layer that protects privacy without substantially compromising the accuracy and topological properties of the resulting map.

5.3.5 Privacy Attacks

User data are considered effectively protected when an adversary cannot identify a particular user's data through linkages between a record owner and a sensitive feature in the published data [Tan et al., 2016]. Linkage attacks can be broadly classified into three categories: record linkage, attribute linkage, and table linkage [Fung et al., 2010]. The method proposed in this chapter is designed primarily to resist attacks of the table-linkage type.

Definition 9 (Privacy Attack Specification). In all attack scenarios, it is assumed that the adversary knows the quasi-identifiers of the victim. If an adversary can successfully link a record owner to the published data table, then the associated privacy risk is referred to as table linkage [Fung et al., 2010]. In this situation, the attacker attempts to determine whether the victim's record appears in the released table. To prevent table linkage, privacy models such as δ -presence, ϵ -differential privacy, (d, λ) -privacy, and distributional privacy may be employed. In the present chapter, the design strategy is to identify optimal subsets to be removed or transformed from the dataset in order to protect the data against this form of attack.

5.3.6 Loss Metric

Preventing the disclosure of sensitive items may reduce the utility of the data because it necessarily involves some form of data transformation [Gkoulalas-Divanis et al., 2014]. One way to quantify this effect is by measuring the difference between the original and transformed data using general-purpose metrics, such as Generalization Cost, Normalized Average Equivalence Class Size, Normalized Certainty Penalty, and Information Loss Metric. In this chapter, general-purpose metrics are used to evaluate information loss.

Definition 10. The information loss (IL) for a partitioned and refined GAFSOM protocol is defined as $\sum_i len(T_i^*) - len(T)$, where T^* is the optimized and reduced dataset of T , and $len(T)$ denotes the number of records contained in T .

Definition 11. The information loss for a sensitive item i is defined as $SL(i) = len(i, T^*)$, where $len(i, T^*)$ denotes the number of records in the transformed dataset that contain the sensitive item i . Accordingly, the information loss for an anonymized integrated dataset T is defined as:

$$IL(T) = \sum_{tran \in T} IL(tran) + \sum_{i \in sen} IL(i) \quad (5.11)$$

where sen is the set of items to be hidden, defined by the administrator. IL measures the information loss introduced by the item-hiding process using the number of sensitive items. The larger the value of IL , the less certain the results are with respect to the original information contained in the trajectories and transactions.

5.3.7 The Optimization Problem

The hiding-failure parameter α , defined earlier in this chapter, may be used to estimate the ability of an adversary to reconstruct the dataset. The optimization problem considered here is to minimize information loss (IL) while satisfying the specified anonymity constraint. There exists a large space of possible solutions, constrained by the threshold MST , that must be explored in solving this optimization problem. Given a large dataset $F = (N, E)$ with n nodes, the goal is to provide anonymization (SX) by clustering the nodes into artificial groups. An $n \times n$ adjacency matrix is used as the input. The problem is to determine the optimal partition set F^* such that the value of $IL(F, F^*)$ is minimized as far as possible for the given graph F . For this purpose, a genetic algorithm is employed. Subsequently, fuzzy anonymization is applied to F^* to produce F' . It is expected that:

$$SOM(F) \approx SOM(F') \quad (5.12)$$

That is, the result of SOM clustering on the original dataset F and on the anonymized dataset F' extracted by GAFSOM should remain approximately equivalent. In this way, both privacy and utility goals are pursued.

5.4 Setting and Notation

One of the most common techniques used to protect personal information from disclosure in data mining is the hiding of sensitive data [Lin et al., 2014]. In this section, a utility-aware hiding algorithm is designed to conceal sensitive items through the optimal deletion or transformation of records. To evaluate whether records should be removed in order to hide sensitive items, the hiding-failure parameter is considered. Records containing any of the sensitive items are first evaluated using the GA algorithm designed to identify minimal hiding-failure values. These records are then processed and the resulting reduced dataset is provided for SOM training on local servers.

The algorithm was implemented in MATLAB and executed on a Linux Ubuntu virtual machine with four vCPUs based on an Intel(R) Xeon(R) E5-2650 v4 processor and 4 GB of memory.

5.5 GAFSOM Algorithm

In the proposed method, a GA approach is used as a search heuristic to identify the optimal sensitive items to anonymize, based on the evolutionary principles of genetics for solving optimization problems [Holland, 1975]. The GA framework was chosen because of its flexible formulation and its ability to identify suitable solutions with manageable computational resources. Among the different techniques available for privacy protection, anonymization and generalization are selected and implemented using membership functions in a fuzzy set. Fuzzy sets are used to anonymize sensitive items associated with objects whose boundaries are not sharply defined. Figure 5.3 provides a simplified illustration of GAFSOM.

As Figure 5.3 shows, the process starts by importing the raw database F from the data provider, together with the sensitive items SX defined by the user. This database is sent to the GA component to extract the optimal sensitive subset, denoted F^* . Next, the fuzzy function anonymizes this subset and appends the result to the primary database. Finally, the data miner uses the anonymized database F' to derive useful patterns and information for decision-makers and other downstream consumers.

The GAFSOM steps are described in Algorithm 1.

The algorithm begins by deriving MST and SX , which are user-defined in this implementation of GAFSOM. Next, the GA loop evaluates the fitness function and applies crossover and mutation. The loop terminates either when the maximum number of generations is reached or when a satisfactory fitness level determined by MST is obtained. The set F^* denotes the optimal subset of sensitive records. The next step is to anonymize F^* using a fuzzy function. Finally, the fuzzified output is merged into the primary dataset to produce F' , which serves as the final output and should remain relatively similar to the original dataset F . The usefulness of GAFSOM is then evaluated through SOM clustering in the following sections.

Algorithm 1 GAFSOM algorithm

Input: A dataset F , a set of sensitive itemsets SX , a threshold of records to be hidden MST , and a population size n .

Output: An anonymized dataset of records F'
 {*// SOM results based on F' and F are temporary outputs used to compare the utility of GAFSOM*}

- 1: **Begin**
- 2: Derive the lower support threshold MST and scan the database to identify the sensitive itemsets SX .
- 3: Randomly generate a population of n individuals, where each gene represents the FID of a record to be hidden.
- 4: **Repeat**
- 5: Calculate the fitness value of each chromosome F_i in the population. {*// According to the selected fitness function using α* }
- 6: Execute crossover and mutation on the population.
- 7: Select individuals for the next generation based on the selection scheme (probabilistically).
- 8: **Until**
- 9: The termination criterion is satisfied. {*// MST* }
- 10: Create the output set of hidden records from the best chromosome as F^* .
- 11: Fuzzify F^* using the defined membership function.
- 12: Append the results to the main dataset F to obtain F' .
- 13: **End**

and, therefore, the higher the probability of including individuals with better fitness values. However, this also increases computational resource consumption [Holland, 1975]. In the present implementation, the population size for a 1000-record dataset is fixed at 50 individuals.

Crossover recombination is the process of generating offspring from parent solutions. It is applied under the expectation that the next generation will combine the beneficial parts of existing chromosomes and thus achieve a better fitness value. At the same time, retaining a small number of chromosome strings from the previous population is useful for survival into the next generation. In this chapter, arithmetic crossover with a rate of 80% is used. After crossover, individuals are subjected to mutation in order to reduce the risk of confinement to local minima. The mutation scheme randomly selects and swaps two nodes from different subnodes, thereby introducing diversity into the algorithm [Holland, 1975]. To avoid compromising the integrity of the raw data excessively, uniform mutation with a rate of 10% is applied.

5.6 Empirical Analysis

The designed GAFSOM algorithm is used to address the optimization problem of identifying a suitable trade-off between privacy and utility. The number of studies on privacy preservation through machine-learning-based clustering anonymization remains limited. Therefore, in order to evaluate different aspects of the algorithm, including quality, time, and cost, GAFSOM is compared with classical SOM cluster-

ing and GASOM. As noted earlier, GAFSOM was originally designed on the basis of the GASOM algorithm introduced in [Amiri et al., 2018]. In GASOM, the most traditional anonymization method is used, namely the deletion of sensitive data. A GA is applied in GASOM to identify a minimally sensitive subset, which is then removed from the database, thereby reducing the number of records that must be eliminated. In addition, the results of GAFSOM are compared with those of classical SOM clustering in order to validate the outputs.

In the proposed approach, the data are not deleted but anonymized. Both methods are compared with respect to the quality of the resulting clustering, measured through structural similarity and information preservation. In this section, utility is evaluated using two real datasets. Subsequently, privacy protection and information loss are examined in relation to SOM clustering outcomes. Most importantly, clustering results on the raw datasets are compared with those obtained by GASOM and GAFSOM so that the usefulness of the method can be assessed.

It should be noted that penetration testing of GAFSOM against privacy attacks such as membership inference lies beyond the scope of this section. Only the cost and correctness of GAFSOM are evaluated here, since these aspects are treated as the immediate priority. A privacy-preserving approach should first demonstrate acceptable utility and computational properties before stronger claims regarding resistance to known attacks are made. Nonetheless, practical privacy-protection techniques such as deletion and anonymization, which are implemented in GAFSOM, may already provide an initial level of protection for sensitive data.

Table 5.3 summarizes the cases considered in the experiments. The following sections describe the experimental procedure, the algorithms used, and the process through which the reported results were obtained.

Table 5.3: Dataset specifications for the testing phases.

Type	Description
Main dataset	Original dataset
GASOM	Main dataset – GA result
GAFSOM	GASOM + fuzzification(GASOM)

5.6.1 Datasets

The algorithms were implemented in MATLAB [Moler, 1980]. For the experimental evaluation, two real-world datasets from the UCI Machine Learning Repository were used: the Adult dataset and the Bank Marketing dataset [Asuncion, 2007, Moro et al., 2011]. These datasets were selected because they contain realistic socioeconomic and marketing information with attributes that can naturally be treated as sensitive. Their main characteristics are summarized in Table 5.4.

- **Adult.** The Adult dataset contains 48,842 records and 14 attributes and is commonly used in privacy-preserving data-mining research.
- **Bank Marketing.** This dataset was generated from direct marketing campaigns conducted by a Portuguese banking institution and contains 45,212 records and 17 attributes.

Different test cases were designed using these datasets and selected sensitive attributes. The utility and protection properties of the method are evaluated according to the way these test cases are defined. The details of the test cases are described in the next section.

Table 5.4: Datasets used in the experimental evaluation.

Dataset	Records	Attributes	Domain	Missing values
Adult	48,842	14	Social	Yes
Bank Marketing	45,212	17	Marketing	No

5.6.2 Experimental Setup

Two distinct test scenarios were designed in order to assess the proposed methods from multiple perspectives.

Test 1: Clustering results on the Adult dataset when the sensitive criteria are defined by the data miner as “black female who is post-graduate, works more than 20 hours per week, and is younger than 30 years old.”

Design 1: In this experiment, a test case consisting of 1000 tuples from the Adult dataset was used. The sensitive attributes were age, work class, gender, education, and race. The tuples were selected randomly from the main dataset.

Test 2: Clustering results on the Bank Marketing dataset when the sensitive criteria are defined by the data miner as “young employee who is married and works in a high-ranked position such as manager.”

Design 2: In this experiment, sensitive tuples were selected deliberately rather than sampled randomly. In order to assess the method under more demanding conditions, records containing higher levels of sensitive information were chosen as the test set. For the Bank Marketing dataset, a relatively large test set of 3000 tuples was used. The attributes treated as sensitive were age, work class, and marital status.

In both experiments, the selected tuples were processed by the fuzzification algorithm. The algorithm begins by identifying tuples that satisfy the predefined sensitivity criteria, so that these records can be hidden in an optimal way. The following sections present the parameters and results used to demonstrate the effectiveness of the proposed method.

5.6.3 Statistics Evaluation

5.6.3.1 Execution Time

Figure 5.4 compares the execution time of vectorized and non-vectorized cases for the Bank Marketing dataset. Despite the larger number of tuples, the elapsed time in part (a) is lower than in part (b). The same comparison was also performed for the Adult dataset (see Figure 5.5). In both databases, vectorization leads to a substantial reduction in execution time.

To evaluate the cost of preserving sensitive data and storing the transformed data in the anonymized database, the execution time of GAFSOM was measured in both serial and vectorized settings. Figure 5.6 summarizes the execution time of GAFSOM for these two implementations. Because GA-based methods can be time-consuming, their practical applicability depends strongly on computational efficiency. For this reason, the fitness function was optimized by employing the hiding-failure parameter in the designed GA. The execution times obtained under different minimum support thresholds were then compared while maintaining a fixed sensitivity rate of 5%, as defined by the user. The results are shown in Figure 5.7. In this experiment, the number of tuples was kept relatively similar; however, more complex conditions were used to define sensitive data in the Bank Marketing dataset. Consequently, a significant increase in runtime was observed there.

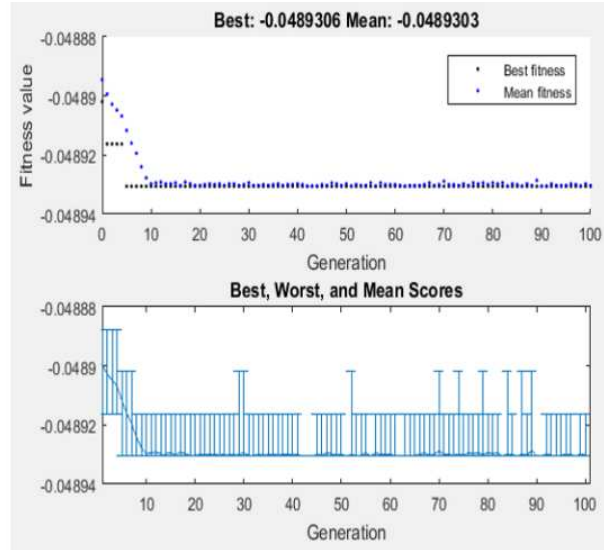
The GA designed for GAFSOM was tested in both serial and vectorized implementations. In serial execution, each generation evaluates the best-fit value record by record. In vectorized execution, the algorithm first extracts all candidate sensitive transactions as a subdatabase and then performs the optimization process. As Figure 5.7 shows, changes in the MST factor influence both runtime and the balance between privacy and accuracy. It was observed that MST values near 0.4 can support a reasonable compromise between utility and privacy.

5.6.3.2 Analysis of the Accuracy of GAFSOM

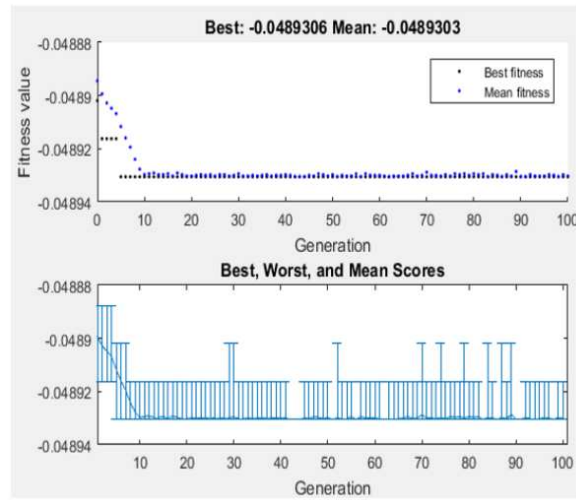
The experiments were carried out using SOM Toolbox v.2. The lattice radius was set to $3/2$, the network topology was defined as a hexagonal lattice, and the optimal number of groups was fixed at three [Kaleli et al., 2012]. To evaluate the precision of the introduced algorithm, its results were compared with those obtained from traditional SOM clustering.

For the first experiment, a test set from the Adult dataset was used, including 1000 tuples with 14 attributes. Age, work class, gender, education, and race were treated as sensitive attributes. A similar implementation was used for the Bank Marketing dataset, with 3000 tuples and 17 attributes. In this case, the sensitive features were age, occupation, and marital status.

The neural network was implemented in MATLAB using SOM Toolbox, and the attributes were represented numerically. The approach consisted of identifying the tuples that matched the sensitivity criteria and then hiding those records optimally

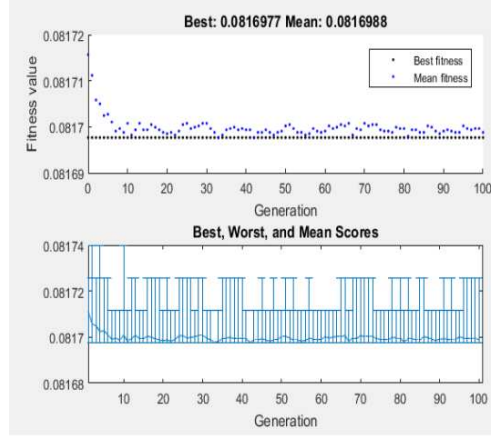


(a) Bank vectorized. The elapsed time is 42.992104 seconds.



(b) Bank non-vectorized. The elapsed time is 47.505873 seconds.

Fig. 5.4: Execution-time comparison of vectorized and non-vectorized cases for the Bank Marketing dataset.



(a) Adult vectorized. The elapsed time is 40.151578 seconds.



(b) Adult non-vectorized. The elapsed time is 53.473483 seconds.

Fig. 5.5: Execution-time comparison of vectorized and non-vectorized cases for the Adult dataset.

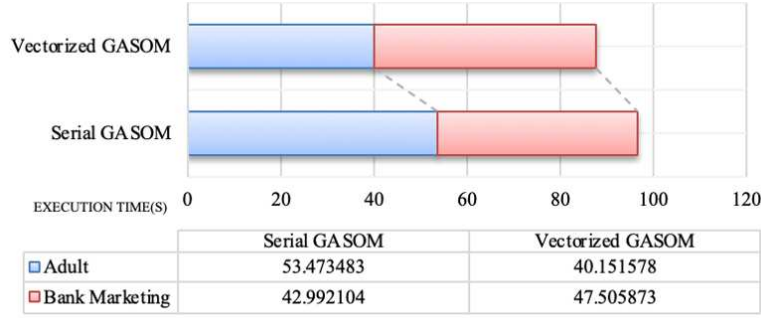


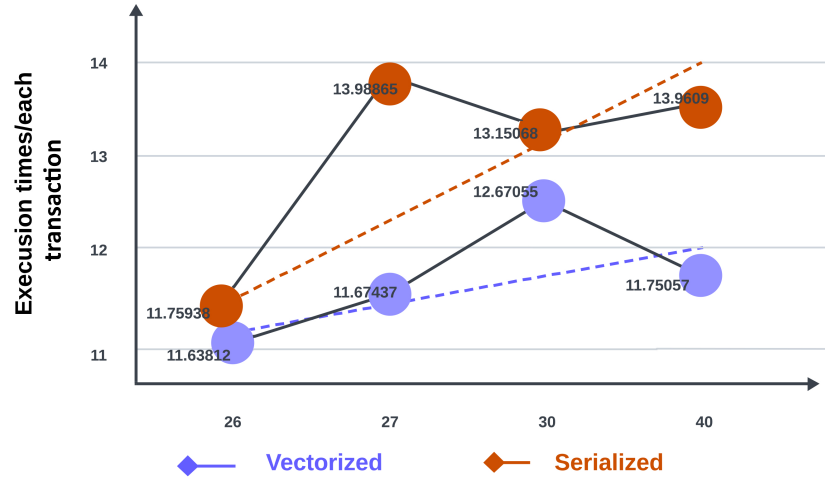
Fig. 5.6: Execution time of the GASOM algorithm under serial and vectorized implementations.

using the proposed GA. In the initial experiments, the datasets were clustered before and after anonymization. The time required to hide sensitive items is illustrated in Figure 5.7. Subsequently, the transformed datasets were used for neural-network training. The U-matrices of the clustering scheme for the two datasets, both before and after the data-hiding process, are shown in Figures 5.8 and 5.9.

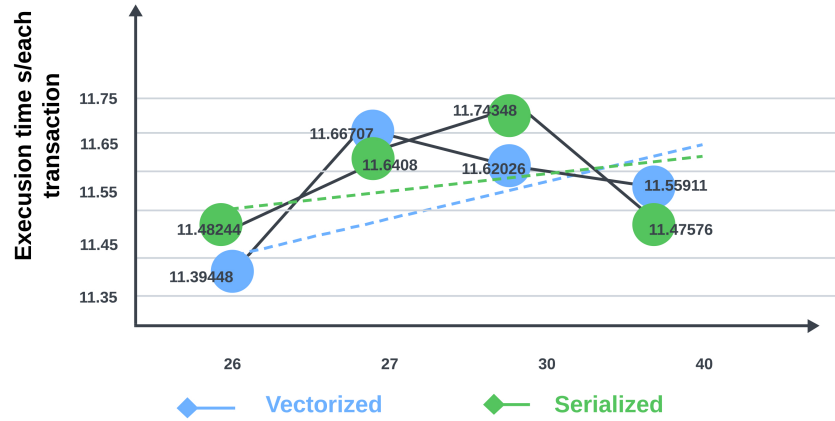
The figures show that the differences in weighting distances across clusters are limited, although they are influenced by dataset size. Figures 5.10 and 5.11 show the positions of the weight vectors, which illustrate how the trained maps are affected by anonymization.

The precise measurement of similarity between the raw database and the transformed database is affected by both dataset size and the level of anonymization. However, a visual comparison of the trained maps and U-matrices between classic SOM and the proposed approach in Figures 5.12 and 5.13 indicates that the model remains useful for the intended objectives. Figure 5.12 compares SOM clustering results for the Adult dataset before and after transformation. Figure 5.13 provides the corresponding comparison for the Bank Marketing dataset.

Figure 5.14 shows the U-matrices of the clustering schemes for all cases defined in Table 5.3 for the Adult dataset. Figure 5.15 provides the same comparison for the Bank Marketing dataset. As Figure 5.14 shows, the difference between the clustering results for the main dataset and GASOM is relatively small. It should be noted that the sensitive items defined for GASOM constitute only 5%, and their deletion does not substantially affect convergence. To allow a more detailed evaluation of GAFSOM, a larger number of sensitive items is defined. As noted earlier, in GAFSOM no sensitive items are removed; instead, anonymization is applied. As can be seen, the topological difference in part (c) of Figure 5.14 is more pronounced. This indicates that, despite the overall similarity, the level of anonymization still influences the resulting topology. However, the analysis of information loss in the next section further supports the usefulness of GAFSOM.



(a) Serialized and vectorized GA in the Bank Marketing dataset with different minimum support thresholds (MST).



(b) Serialized and vectorized GA in the Adult dataset with different minimum support thresholds (MST).

Fig. 5.7: Execution time for the Adult and Bank Marketing datasets under different MST values. Execution time was measured for both serial and vectorized GA implementations. The results show that vectorization is faster in both settings. In addition, the influence of MST on execution time is visible. In both datasets, the results suggest that an MST value around 0.3 may offer a useful compromise between utility and privacy.

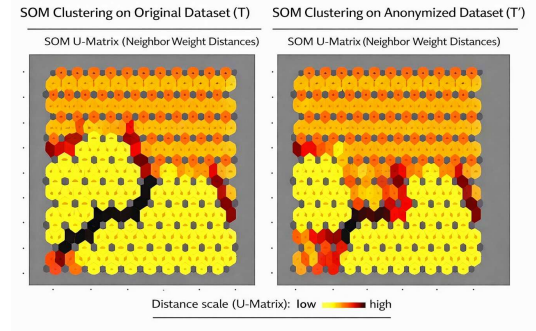


Fig. 5.8: U-matrices of traditional SOM and the proposed SOM-based approach on the Adult dataset, showing limited distortion after anonymization. The left plot shows clustering on the original dataset (T), while the right plot shows clustering on the anonymized dataset (T'). The similarity of the cluster patterns indicates that the anonymization process preserves the underlying data structure.

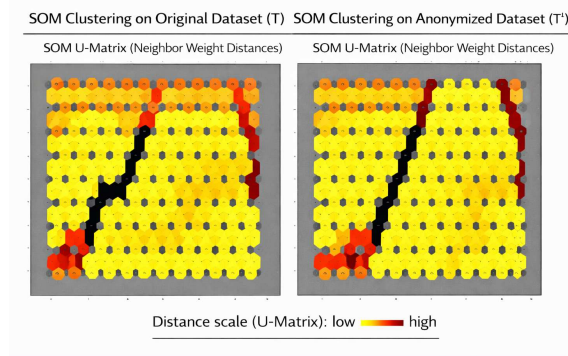


Fig. 5.9: U-matrices of traditional SOM and the proposed SOM-based approach on the Bank Marketing dataset, showing limited distortion after anonymization. The comparison of clustering results for the original dataset (T) and the anonymized dataset (T') indicates that the main topological and neighborhood-distance patterns remain largely preserved.

5.6.3.3 Analysis of Information Loss in GAFSOM

To validate GAFSOM, visual comparisons of trained maps and U-matrices between classic SOM and the proposed method were complemented by quantitative criteria, including quantization error (QE) and topographic error (TE). In the experiments, maps were initialized randomly and a batch SOM was used. For both the classic and proposed SOM models, the initial sigma was set to 2, the final sigma to 1, and the training phase to 1 epoch. The resulting values are shown in Figures 5.16 and 5.17.

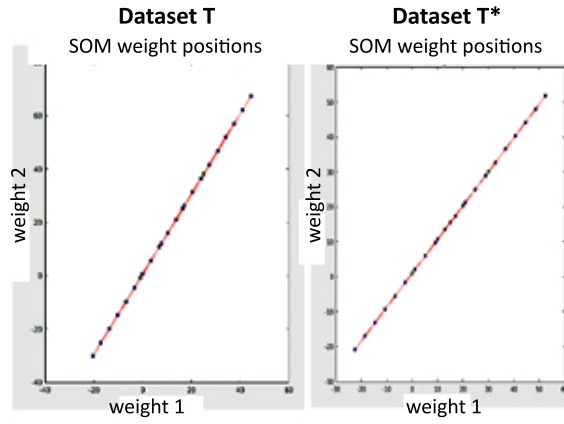


Fig. 5.10: Weight-position vectors of traditional SOM and the proposed SOM-based method on the Adult dataset. The results show a near-linear alignment of weight vectors before (T) and after anonymization (T^*), indicating that the proposed method preserves the original topological structure with limited distortion.

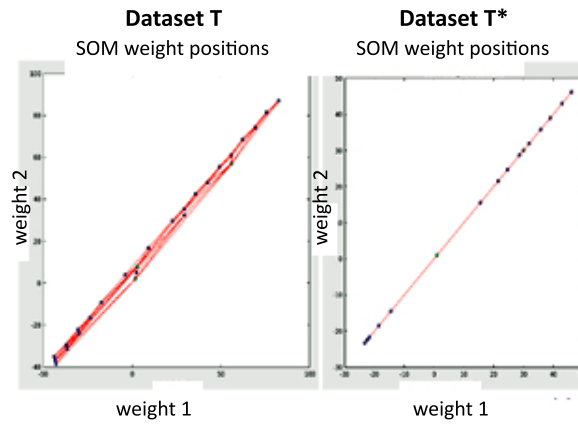
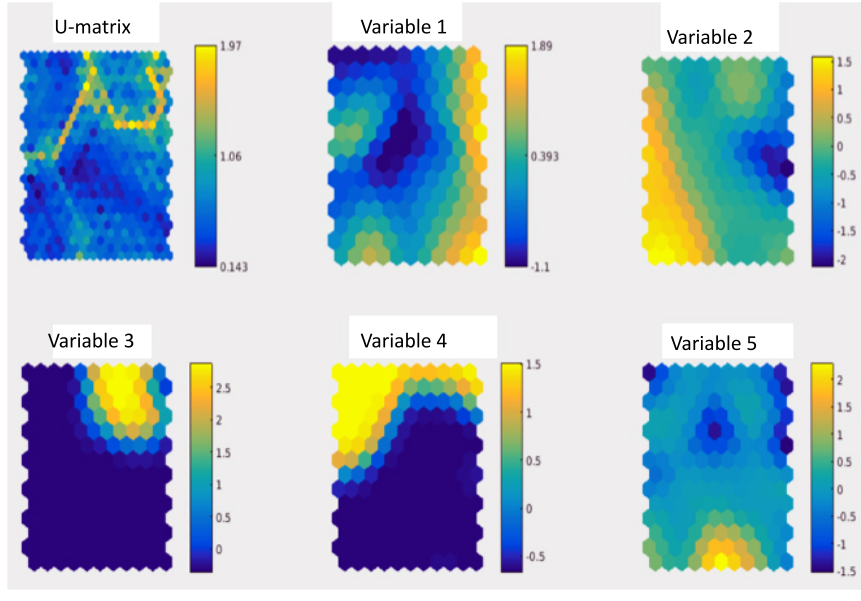
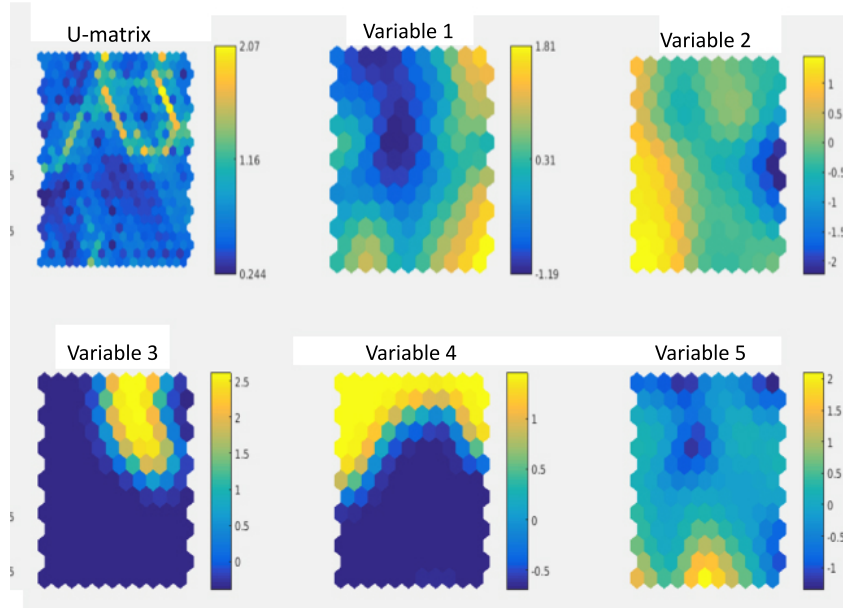


Fig. 5.11: Weight-position vectors of traditional SOM and the proposed SOM-based method on the Bank Marketing dataset. While some dispersion is observed in the anonymized version (T^*), the overall topology remains consistent, indicating that the method preserves structural similarity even in the more complex dataset.

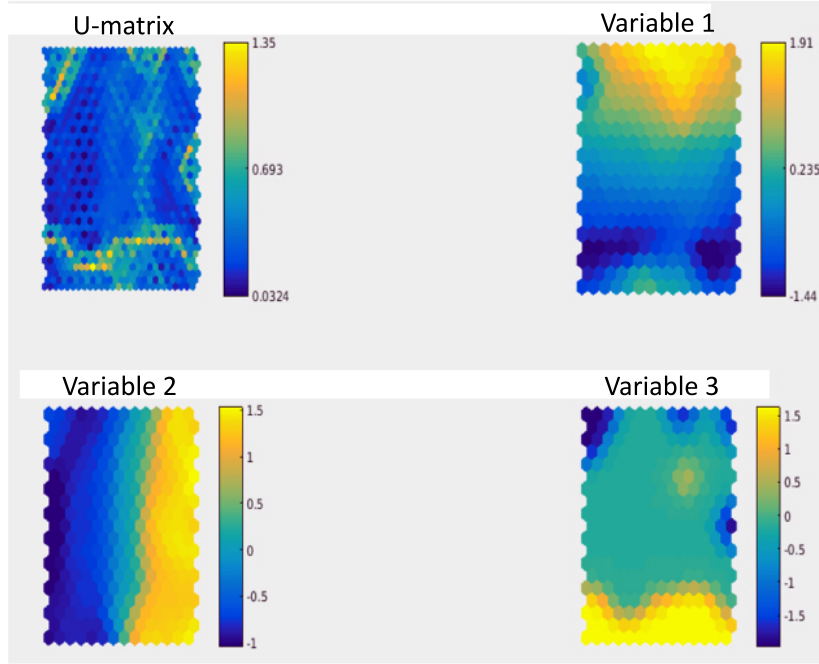


(a) Comparison of the U-matrix for the original Adult dataset with final quantization error 0.617 and final topographic error 0.037.

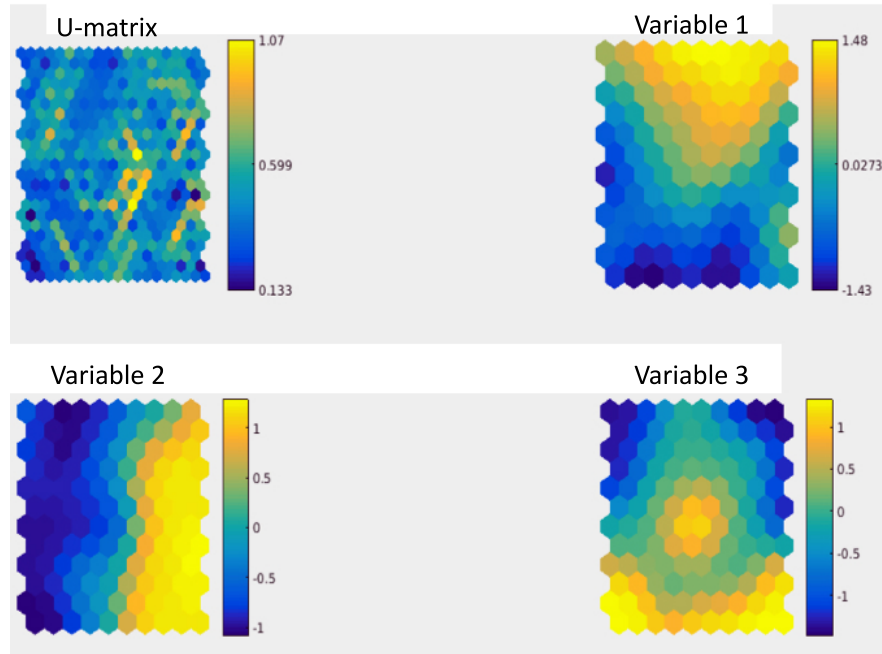


(b) Comparison of the U-matrix for the transformed Adult dataset with final quantization error 0.654 and final topographic error 0.055.

Fig. 5.12: Comparison of SOM clustering results for the original Adult dataset D and the transformed dataset D' .



(a) Comparison of the U-matrix for the original Bank Marketing dataset with final quantization error 0.193 and final topographic error 0.042.



(b) Comparison of the U-matrix for the transformed Bank Marketing dataset with final quantization error 0.372 and final topographic error 0.034.

Fig. 5.13: Comparison of SOM clustering results for the original Bank Marketing dataset D and the transformed dataset D' .

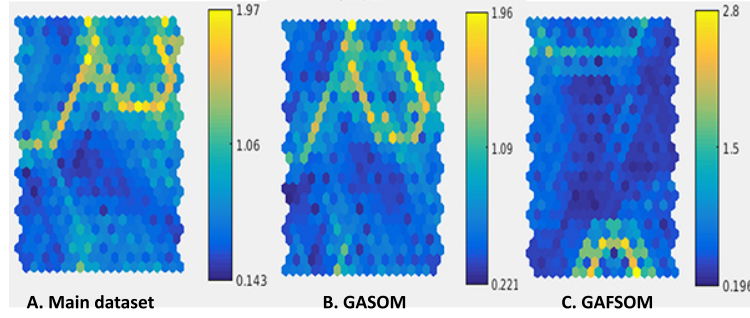


Fig. 5.14: Comparison of U-matrices for the Adult dataset. Part (a) shows classic SOM on the original dataset, part (b) shows GASOM on the transformed dataset, and part (c) shows GAFSOM. The results indicate that GAFSOM remains useful, although the clustering topology differs more visibly from the classic SOM baseline.

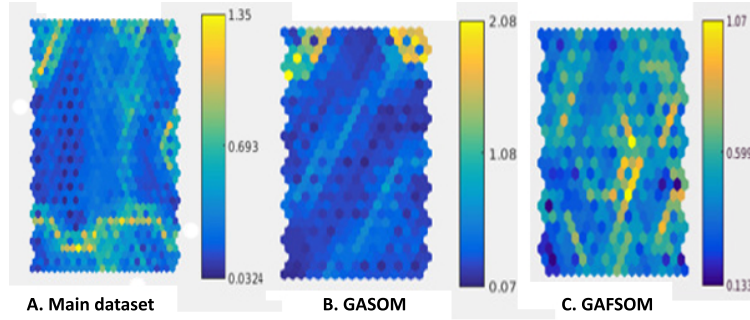


Fig. 5.15: Comparison of U-matrices for the Bank Marketing dataset: (a) classic SOM on the original dataset, (b) GASOM on the transformed dataset, and (c) GAFSOM on the anonymized dataset. Although the clustering experiments support the usefulness of GAFSOM, the topological similarity remains more distorted in this dataset.

These results show the usefulness of GAFSOM in maintaining cluster quality while improving privacy protection. In addition, in order to estimate the level of protection achieved for sensitive items, the difference between the number of sensitive items before and after the hiding process in the genetic algorithm was calculated using the ratio:

$$\frac{Senfreq(T^*)}{Senfreq(T)} \quad (5.13)$$

where T^* denotes the transformed dataset and T the original dataset before hiding. The resulting values are generally close to zero, which supports the effectiveness of the proposed method in reducing the visibility of sensitive items. Although the experimental results indicate that the introduced protocol is useful, further refine-

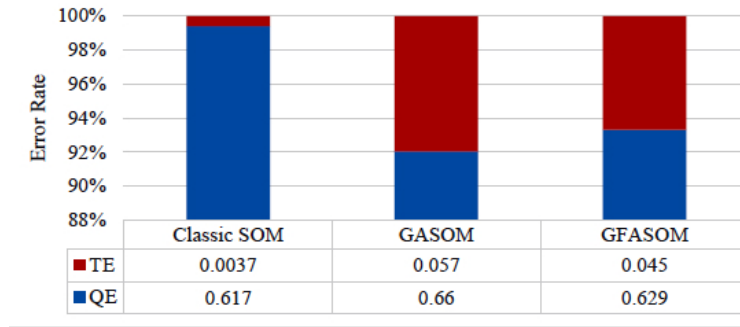


Fig. 5.16: Comparison of Classic SOM, GASOM, and GAFSOM in terms of quantization error (QE) and topographic error (TE), showing the usefulness of GAFSOM for preserving clustering quality while enhancing privacy protection.

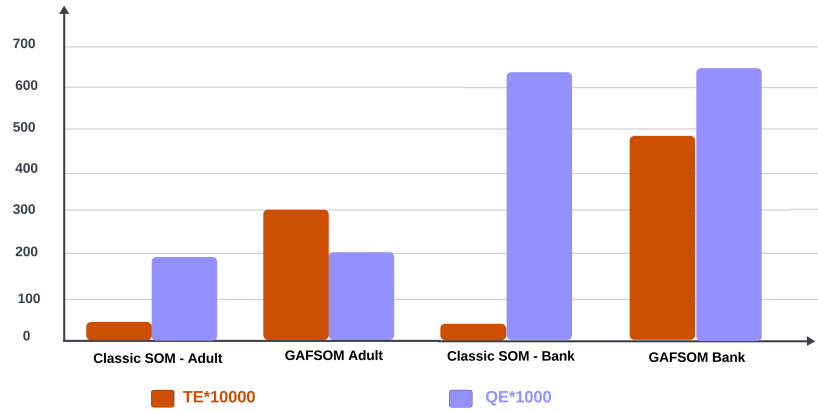


Fig. 5.17: Accuracy analysis of clustering results for the Adult and Bank Marketing datasets under both the classical SOM and GAFSOM models.

ment is still needed in order to improve clustering accuracy and runtime efficiency. Figure 5.18 illustrates the relationship between the hiding factor and the minimum support threshold (MST), thereby showing the trade-off between privacy protection and information loss.

The results indicate that information loss and privacy level are opposing factors and should be selected according to the requirements of the current database. Under the assumed model, the use of transformed quasi-identifiers reduces the feasibility of direct linkage attacks. In the present method, the quasi-identifier is defined as the subset of attributes that can uniquely identify a record. Since this subset is altered through the GA-based transformation process, its direct recovery by an attacker becomes more difficult. At the same time, the experiments were intended primarily

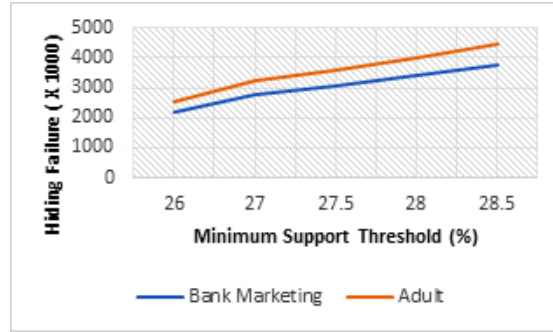


Fig. 5.18: Relationship between MST and hiding failure, showing the trade-off between information loss and privacy level.

to verify the effectiveness of the proposed GA method in identifying and transforming sensitive items. Although the accuracy of GAFSOM is relatively acceptable, further refinement, such as improved fuzzification of the optimal subset identified by the fitness function, may be useful in reducing the need to distort records more strongly.

5.6.3.4 Topological Analysis of GAFSOM

Some differences are visible in the topological comparison of clustering results. These differences indicate that, despite the general similarity, the level of anonymization influences the final clustering structure.

As shown in Figures 5.14 and 5.15, the difference between classical SOM clustering and GASOM is relatively small. However, improving the accuracy of GAFSOM remains necessary in order to achieve even greater precision in the clustering results. A substantial topological difference can be observed in Figure 5.15(c). This suggests that the fuzzy function used in the current implementation is not yet fully satisfactory and should be refined. Nonetheless, the mathematical evaluation criteria still support the usefulness of the proposed method.

5.6.3.5 Statistical Validation of Clustering Performance

To determine whether the observed differences among Classic SOM, GASOM, and GAFSOM in quantization error (QE) and topographic error (TE) were statistically significant, a Kruskal–Wallis H-test was employed. This non-parametric test is a suitable alternative to ANOVA when comparing multiple groups without assuming normality.

Null Hypothesis (H_0): There is no statistically significant difference in QE and TE distributions across Classic SOM, GASOM, and GAFSOM.

Alternative Hypothesis (H_1): At least one algorithm exhibits a significantly different QE or TE distribution.

The test results are shown in Table 5.5.

Table 5.5: Comparison of QE and TE across the three algorithms: Classic SOM, GASOM, and GAFSOM.

Metric	Test Statistic	p-value	Significance ($\alpha = 0.05$)
QE	68.80	1.15×10^{-15}	Significant
TE	74.12	8.03×10^{-17}	Significant

The results show that the differences in both QE and TE across the three algorithms are highly significant ($p < 0.001$). This indicates that the observed improvements of GAFSOM over GASOM, particularly in reducing topographic error while maintaining low quantization error, are unlikely to be due to random variation and instead reflect genuine performance differences. The statistical evidence therefore supports the claim that GAFSOM achieves a more favorable balance between privacy preservation and clustering accuracy.

5.6.3.6 Comparison Between GAFSOM and Similar Methods

Table 5.6 presents a comparative overview of the GAFSOM model and related privacy-preserving or clustering-based anonymization methods, using representative values and qualitative assessments drawn from the literature.

Table 5.6 indicates that GAFSOM achieves low execution time because of its lightweight fuzzy integration and its GA-optimized search process. Although DP-based techniques often provide stronger formal privacy guarantees, they usually incur a greater computational burden because of noise injection and additional post-processing. Evolutionary clustering methods and traditional k-anonymity-based approaches typically exhibit moderate execution times. With respect to the accuracy of data-mining outcomes, GAFSOM, GA-based clustering, and DP-based methods preserve important data structures required for downstream tasks. By contrast, fuzzy clustering methods such as FCM may lose precision because of over-generalization.

GAFSOM and DP-based methods may reduce information loss through selective anonymization and carefully calibrated privacy mechanisms. In contrast, FCM often exhibits substantial information loss because of its broad fuzzification strategy. K-anonymity-based techniques occupy an intermediate position, often sacrificing part of their utility in order to guarantee generalization. GAFSOM maintains a high degree of topological similarity, which is especially important for SOM performance and cluster quality. The topological similarity of DP- and GA-based methods is often moderate, whereas fuzzy approaches may reduce topological accuracy because of stronger generalization.

Table 5.6: GAFSOM compared with related PPDM and SOM-based methods.

Method	Execution Time	Precision of Data-Mining Task	Information Loss	Topological Similarity
GAFSOM	Low	High	Low	High
K-Anonymity + SOM [Kabir, M.E et al., 2011]	Medium	Medium	Medium	Medium
Differential Privacy + SOM [Pinot, R. et al., 2018]	High	High	Low	Medium
Fuzzy Clustering (FCM) [Bezdek, J.C., 1981]	Low	Medium	High	Low
Genetic Algorithm + Clustering [Maulik, U. et al., 2000]	Medium	High	Medium	Medium
DP + Genetic Clustering [Liu, J. et al., 2018]	High	High	Low	Medium
Federated SOM [Zhou et al., 2024]	Medium–High	High	Low–Medium	High
Noise-based DP + SOM (ϵ -bounded) [Zhu et al., 2023]	High	Medium–High	Medium–High	Medium
Synthetic Data Generation [Sarmin et al., 2024]	Medium	High	Low–Medium	Medium–High
In-Memory Anonymization [Shamsinejad et al., 2025]	Low	Medium–High	Low–Medium	Medium

Overall, the comparison suggests that GAFSOM improves several important dimensions relative to traditional anonymization techniques and provides a balanced compromise between privacy and utility through the combined use of controlled fuzzification and evolutionary optimization. The method appears especially well suited for sensitive clustering tasks with SOMs because of its low computational overhead, high precision, limited information loss, and preservation of topology. While differential privacy continues to offer stronger formal guarantees, GAFSOM may provide a computationally lighter alternative in applications where structural preservation and performance are as important as privacy [Amiri and Quirchmayr, 2025].

Table 5.6 also includes two recent PPDM methods: PPML-Omics [Zhou et al., 2024], which uses federated learning and differential privacy in bioinformatics settings, and DP-STTC [Zhu et al., 2023], which applies ϵ -differential privacy to spatio-temporal trajectory clustering. PPML-Omics tends to achieve high precision and low to medium information loss, but with higher execution time due to distributed computation and secure-aggregation overhead. DP-STTC also demonstrates strong privacy performance for small ϵ values, but often at the cost of higher information loss and degraded topological similarity. Compared with GAFSOM’s selective fuzzy anonymization and GA-based optimization, these methods are competitive in precision and information loss under suitable settings; however, GAFSOM may

retain an advantage in topological similarity for SOM-based clustering tasks and in implementation simplicity and efficiency.

Compared with these alternatives, GAFSOM appears to provide a favorable compromise between privacy and utility in the reported setting and may preserve SOM-related structural properties effectively. Thus, in applications where topology and utility are critical and centralized training is feasible, GAFSOM appears to be an effective alternative. Synthetic-data approaches [Sarmin et al., 2024] complement GAFSOM in scenarios where throughput and real-time anonymization are the dominant constraints, whereas GAFSOM remains preferable when topology preservation and selective fuzzification are central objectives.

5.6.3.7 Unified Comparison Across Methods and Datasets

Table 5.7 aggregates the principal performance metrics and qualitative assessments for GAFSOM and a selection of representative privacy-preserving and clustering methods from the literature. Where possible, quantitative metrics (QE and TE) from the experiments in this chapter on the Adult and Bank Marketing datasets are provided. For external methods reporting different evaluation criteria, qualitative assessments are used to enable approximate cross-method comparisons.

Three main observations arise from this unified comparison. First, GAFSOM achieves a favorable balance between utility and privacy: it attains a low quantization error ($QE = 0.629$) close to the Classic SOM baseline (0.617) while maintaining a topographic error ($TE = 0.045$) substantially lower than GASOM ($TE = 0.057$). This suggests that selective GA-driven anonymization combined with fuzzy representation preserves map fidelity and neighborhood structure more effectively than deletion-based GA approaches. Second, methods that offer formal privacy guarantees, such as ϵ -differential privacy variants and DP+GA combinations, typically involve higher computational cost and may degrade topology or utility under stricter privacy budgets. Third, alternative paradigms such as federated SOM and synthetic data generation provide useful trade-offs: federated methods preserve topology and precision while increasing communication overhead, whereas synthetic data pipelines can provide competitive downstream utility but require careful calibration of generative models and privacy budgets.

Overall, the unified comparison positions GAFSOM as a practical middle ground for centralized applications in which topology preservation and computational efficiency are both important. It achieves topology and quantization metrics close to the SOM baseline while reducing information loss through selective fuzzy anonymization. For scenarios requiring formal ϵ -DP guarantees or decentralized training, hybrid pipelines or federated/DP variants may remain preferable, albeit with higher runtime and implementation complexity.

Table 5.7: Unified comparison of GAFSOM and representative PPDM/SOM-based methods. QE = Quantization Error, TE = Topographic Error. Numeric QE/TE values are reported where available from this study; for external methods, qualitative assessments are used when comparable numeric values are not reported in the cited sources.

Method	Dataset(s)	QE	TE	Execution Time	Precision	Info Loss	Topo. Simi- larity	Notes/Citation
Classic SOM (baseline)	Adult, Bank Marketing	0.617	0.037	Medium	Medium	Medium	Medium	Baseline SOM (no privacy); values from this study.
GASOM (GA + deletion)	Adult, Bank Marketing	0.660	0.057	Medium	High	Medium	Medium	GA-based hiding via deletion; values from this study.
GAFSOM	Adult, Bank Marketing	0.629	0.045	Low	High	Low	High	Selective GA + fuzzy anonymization; values from this study.
K-Anonymity + SOM [Kabir, M.E et al., 2011]	Generic/Adult			Medium	Medium	Medium	Medium	Generalization/suppression approach; moderate information loss.
DP + SOM (noise) [Pinot, R. et al., 2018]	Generic/Adult			High	High	Low	Medium	Strong formal privacy; utility depends on the noise budget.
Fuzzy Clustering (FCM) [Bezdek, J.C., 1981]	Generic			Low	Medium	High	Low	Over-generalization often increases information loss.
GA + Clustering (gen-eral) [Maulik, U. et al., 2000]	Generic			Medium	High	Medium	Medium	Evolutionary clustering approaches with variable topology preservation.
DP + GA Clustering [Liu, J. et al., 2018]	Generic			High	High	Low	Medium	DP improves privacy but often increases runtime.
DP-STTC [Zhu et al., 2023]	Trajectories/LBSN			High	Medium–High	Low–Medium	Medium	ϵ -DP applied to spatio-temporal clustering; utility depends on ϵ .
PPML-Omics [Zhou et al., 2024]	Distributed/Omics			Medium–High	High	Low–Medium	High	Federated learning + DP; increased communication overhead.
Noise-based DP (ϵ -bounded) + SOM [Zhu et al., 2023]	Generic			High	Medium–High	Medium–High	Medium	Utility and topology depend strongly on noise level.
Synthetic Data Generation [Sarmin et al., 2024]	Various			Medium	High	Low–Medium	Medium–High	Synthetic-data pipelines such as PrivBayes and PATE-GAN.
In-memory Anonymization [jad et al., 2025]	Streaming/Big Data			Low–Medium	Medium–High	Low–Medium	Medium	Designed for throughput and stream-oriented anonymization.

5.6.4 Performance Evaluation and Analysis

A central issue in PPDM is the side effects introduced by privacy-preserving techniques on standard data-mining tasks. Conventional privacy-protection methods such as perturbation, suppression, k-anonymity, l-diversity, and t-closeness often cannot eliminate these side effects. Machine-learning-based methods appear to offer more adaptive solutions because of their flexibility, robustness, and speed. The use of machine learning in PPDM is still relatively recent. Consequently, multiple criteria must be defined and evaluated. The primary aim of this study is therefore a feasibility-oriented investigation of whether machine-learning techniques such as GA and fuzzy sets can be applied successfully to privacy protection. A secondary aim is to clarify the differences between this approach and more classical PPDM techniques, especially differential privacy and k-anonymity, in order to support future integration or comparison. The experiments indicate that an acceptable trade-off between privacy protection and utility can be achieved using machine learning, such that both improved privacy protection and useful clustering performance are maintained. The next sections interpret these findings and outline the remaining open problems.

5.7 Interpretation of Experimental Results

The evaluation of GAFSOM shows that the integration of GA-driven subset selection with fuzzy anonymization provides a robust framework for privacy-preserving clustering. The empirical results indicate that the approach achieves high precision, competitive execution efficiency, and limited information loss when compared with classical SOM and GASOM baselines. The introduction of fuzzy anonymization improves privacy protection; however, it also introduces challenges related to topological preservation. Specifically, structural fidelity in cluster outputs decreases as stronger anonymization is applied, with visible divergence from baseline SOM performance in some cases. This illustrates the central tension between privacy strength and clustering utility, which remains a recurring challenge in privacy-preserving data mining.

In the GAFSOM framework, termination of the genetic algorithm is governed by a dual condition: (i) reaching the maximum number of generations or (ii) achieving a satisfactory fitness level defined by the Minimum Support Threshold (MST). This design balances computational efficiency against the need to converge toward privacy-preserving yet utility-aware solutions. In the experiments, MST values around 0.3–0.4 produced the most balanced outcomes: execution time remained low, information loss was limited, and hiding-failure ratios were small. Lower thresholds offered stronger privacy but at the cost of excessive runtime and greater hiding failures, whereas higher thresholds favored efficiency but weakened protection. In this way, MST functions both as a convergence parameter and as a measurable representation of the privacy–utility trade-off.

A broader comparison with alternative privacy-preserving strategies provides further context for GAFSOM’s contribution. While differential privacy remains the gold standard in terms of theoretical guarantees, its reliance on random noise injection may compromise topological structure and introduce significant computational overhead. Likewise, fuzzy clustering offers efficient anonymization but may suffer from degraded precision and structural distortion. By contrast, GAFSOM, together with GASOM, provides an optimized balance: its selective anonymization preserves clustering utility while maintaining reasonable execution speed. These results suggest that genetically augmented frameworks offer a practical middle ground for applications where both structural fidelity and privacy are important, such as healthcare, finance, and IoT analytics.

Beyond its immediate evaluation, GAFSOM is also relevant to emerging adversarial concerns. The stochasticity introduced by fuzzy anonymization may improve resilience against inference-based attacks such as model inversion, membership inference, and reconstruction attacks, all of which exploit regularities in the data in order to re-identify individuals. However, the framework has not yet been systematically tested against linkage attacks in which anonymized outputs are combined with auxiliary datasets. Future work should therefore include a dedicated privacy penetration-testing framework in order to evaluate robustness against such more sophisticated attack scenarios.

Finally, although this study focused on static datasets, scalability remains an important consideration. The population-based nature of the GA lends itself naturally to distributed and federated environments, with potential speedups related to the number of available computational nodes. For example, tasks requiring hours on centralized architectures may be reduced substantially on clustered infrastructures. This adaptability makes GAFSOM a candidate for future extension to large-scale settings such as IoT ecosystems or federated healthcare platforms, where dynamic anonymization of streaming data is essential. These observations reinforce that the present phase of research is primarily intended to establish the feasibility of machine-learning-based privacy protection, while later work should address scalability, distributed implementation, and deeper adversarial evaluation.

5.8 Possible Attack Strategies

Systematic penetration testing of GAFSOM against common privacy attacks, such as membership inference, lies beyond the scope of this thesis. Here, the evaluation focuses on utility, accuracy, and computational cost. A privacy-preserving approach should first demonstrate acceptable performance in these respects before stronger claims about robustness against known attacks are made. Nevertheless, the practical privacy-protection mechanisms implemented in GAFSOM, especially transformation and anonymization, may already provide an initial level of defense for sensitive data.

5.9 Open Issues and Future Directions

The present study was intentionally limited in scope, with the primary goal of determining whether machine learning, and in particular an SOM-based approach, can serve as a feasible foundation for privacy-preserving clustering. Accordingly, issues such as scalability to streaming or real-time big data, benchmarking execution performance against industrial standards, and resilience against sophisticated adversarial models were not examined in full detail. Although highly relevant to long-term applicability, these aspects extend beyond the feasibility-driven objectives of the present phase.

The GAFSOM model still requires further refinement in order to improve clustering accuracy and minimize information loss. At the same time, the current evaluation does not yet address scalability in large-scale or distributed environments. This limitation arises because the current implementation was designed for controlled datasets, allowing the privacy-preserving capabilities of the model to be isolated and evaluated. Likewise, while GAFSOM may provide some resistance to active attacks, its robustness against more advanced or adaptive adversaries has not yet been thoroughly analyzed. A systematic examination of such vulnerabilities, including the scalability trade-offs that arise under high-volume or high-velocity data streams, should be prioritized in future work.

Future work can extend GAFSOM in three main directions. First, scalability and efficiency should be investigated in real-time and big-data contexts, including in-memory and federated architectures. Second, alternative randomized or adaptive fuzzification functions may be introduced to improve the balance between accuracy and anonymization and potentially increase resistance to passive attacks. Third, the integration of differential privacy and additional machine-learning techniques may be explored in order to provide stronger formal robustness guarantees. Through these extensions, a more complete evaluation of GAFSOM can be developed, moving beyond feasibility toward applicability in realistic deployment settings.

5.10 Summary

PPDM based on machine-learning techniques may provide more adaptive solutions not only for protecting against privacy breaches but also for maintaining the quality of data-mining results. In this chapter, the GAFSOM method was introduced, combining a genetic algorithm with fuzzy sets in order to achieve a trade-off between privacy and utility. The results suggest that selective transformation of sensitive data can be less destructive than broad anonymization through uniform fuzzification, and that topology-aware optimization is especially useful in SOM-based privacy-preserving clustering. The computational overhead of GAFSOM is limited, and its correctness is supported by topological and quantization-based evaluation criteria. However, the current form of general fuzzy anonymization still requires improvement.

Compared with standard fuzzy clustering, GAFSOM balances data distortion and privacy by using a genetic search to determine an appropriate degree of fuzzification. In contrast to previous clustering-based privacy methods that do not explicitly account for topological structure, GAFSOM directly optimizes for topological quality metrics, which makes it especially suitable for SOM-based pattern analysis. Differential-privacy-based clustering methods often suffer from structural fragmentation caused by injected noise, whereas GAFSOM's evolutionary selection of anonymization strategies allows comparable privacy effects to be pursued with reduced topological distortion.

On the basis of the discussion and results presented in this chapter, the next methodological step is to move toward a more formal privacy mechanism in order to respond more fully to Research Questions 6–8. The following chapter therefore introduces and evaluates a differential-privacy-based approach through experimental analysis.

Chapter 6

A Differentially Private Self-Organizing Map

Chapter Overview. Differential privacy (DP) provides a mathematically grounded framework for safeguarding sensitive information in large datasets [Dwork, 2006]; however, traditional clustering methods often suffer from reduced accuracy under strict privacy constraints. This chapter presents a local distance-based differential privacy model for Self-Organizing Maps (SOM), designed to preserve data utility while ensuring strong privacy protection. By introducing noise proportional to local distances within SOM’s neighborhood-update process, the proposed method reduces quantization error and maintains topological consistency. Unlike global DP methods such as DP k-means or conventional local DP models that perturb centroids independently, this approach supports stable convergence, preserves structural integrity, and flexibly adapts to varying privacy budgets (ϵ). Performance evaluation on several benchmark datasets emphasizes accuracy and topology preservation under DP constraints. Comparative analyses with state-of-the-art differentially private clustering techniques highlight three main benefits: (1) preservation of neighborhood and topological structure, (2) stable convergence with low quantization error, and (3) scalability to high-dimensional data without major loss of cluster coherence. Overall, the proposed framework demonstrates an improved privacy–utility trade-off compared with existing differentially private clustering methods. The proposed method and its contributions are intended to provide improved responses to Research Questions 6–8 while also supporting the requirements of Research Questions 3–4.

6.1 Introduction

The purpose of this chapter is to address concerns related to privacy preservation in clustering and machine-learning tasks, which significantly affect the reliability and real-world applicability of data-driven systems. As discussed in previous chapters, this challenge can be understood in terms of three dimensions: privacy level, accuracy of data-mining tasks, and computational efficiency, that is, the Privacy–Accuracy–Efficiency (PAE) trade-off [Amiri et al., 2017]. To address the limitations of standard

techniques, an efficient privacy-enhancement strategy with low overhead is required (see Chapter 4). In earlier chapters, a machine-learning-based method for protecting privacy in e-business applications was presented. For data-mining tasks, especially clustering and dimensionality reduction, one of the fundamental neural-network models employed is the SOM [Amiri et al., 2018] (see Chapter 5). Our previous findings showed that certain privacy gaps arise in SOM communication components, such as the inability to securely update neuron weight vectors, reliably determine the termination condition of SOM, or securely identify the winning neuron when confidential data are held by multiple parties [Amiri et al., 2018].

Fuzzy sets, genetic algorithms, and vectorized metaheuristic techniques were used in the complementary method described in the previous chapter (see Chapter 5). Two approaches, GASOM (Genetic Algorithm Self-Organizing Map) and GAFSOM (Genetic Algorithm and Fuzzy Self-Organizing Map), were developed for grouping large datasets [Amiri et al., 2019, Amiri et al., 2018]. To improve utility, GASOM combines a vectorized optimization algorithm with a genetic algorithm. It focuses on the Kohonen map and is motivated by privacy gaps in SOM training. A hiding-failure parameter is used in the design of the GA fitness function. In GAFSOM, the output of the GA phase is used as the input to a fuzzy function. Sensitive content is anonymized through a triangular membership function. The results reported in the previous chapter demonstrated the effectiveness of the designed GA-based technique [Amiri et al., 2019].

Nonetheless, a more robust method for protecting privacy while minimizing information loss is still needed. Our previous results showed that selective deletion of valuable data items is less destructive than broad anonymization through fuzzification; however, a stronger privacy mechanism is still desirable. The overhead of GAFSOM is limited, and its effectiveness was supported through topological-error analysis (see Chapter 5), but the current fuzzy anonymization mechanism still leaves room for improvement. In this chapter, differential privacy is applied to SOM clustering in order to compare its validity and utility with the previous approach.

Differential privacy [Qardaji et al., 2013] has increasingly become a standard approach for protecting against indirect privacy attacks when publishing or analyzing data. Figure 6.1 provides a high-level illustration of differential privacy. This chapter introduces DP2SOM, a novel protocol in which differential privacy is used to protect the privacy of sensitive data during SOM clustering. A distance-based form of differential privacy is employed so that relationships are protected through the addition of random noise, while a local convergence mechanism is used to adjust privacy and utility simultaneously. In this way, both privacy protection and clustering quality are pursued. The privacy protection and effectiveness of the proposed clustering method are then evaluated experimentally.

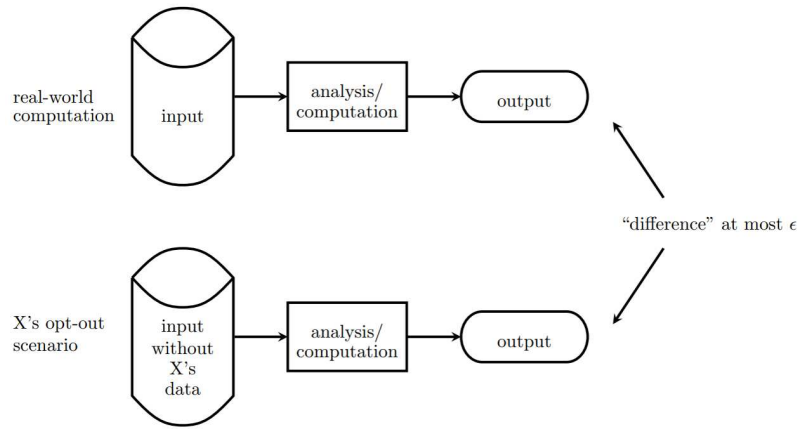


Fig. 6.1: High-level illustration of differential privacy.

6.1.1 Motivation

With the rapid growth of data collection, many organizations have accumulated large datasets containing private information. Data-mining analysis can enable the inference of private information, and privacy concerns have therefore increased as more personal data are collected and used [Perugini et al., 2004]. The critical challenge is to protect privacy by eliminating or minimizing the risk of personal-data disclosure without sacrificing the functionality of the information system [Amiri et al., 2019]. Existing privacy-preserving techniques include data distortion, encryption, and restricted data publication. Because of the required encryption and decryption steps, cryptographic approaches are often computationally expensive [Tex et al., 2018]; moreover, they do not in themselves eliminate all possibilities of information leakage about individuals [Van Dijk et al., 2010].

Restricted publication methods, such as k -anonymity, attempt to protect data through anonymization [Zhao et al., 2018]. Data-distortion methods, such as aggregation, modify sensitive data while attempting to preserve some of its relevant attributes [Leontiadis et al., 2015]. These traditional privacy-preserving methods defend mainly against direct attacks but are not always sufficient to prevent privacy disclosure caused by indirect inference [Dwork, 2006]. It is therefore necessary to develop a new privacy-preservation method that can be applied to data-mining techniques while balancing privacy and data availability.

6.1.2 Contributions

The main contributions of this chapter are as follows:

1. **Differentially private SOM clustering.** A distance-based Self-Organizing Map framework under differential privacy is presented. Unlike prior work that applies noise only to post hoc outputs, the proposed method perturbs centroid updates during SOM training, thereby protecting privacy at each interaction step while preserving SOM neighborhood behavior and topology.
2. **Local privacy-adjustment mechanism.** A novel per-centroid privacy-budgeting strategy is designed to provide locally adaptive privacy levels, so that a better balance is achieved between privacy protection and data availability. This mechanism injects calibrated noise into a local distance-based scheme while supporting centroid convergence in the selected region.
3. **Mathematical and empirical validation.** A mathematical discussion of sensitivity, privacy accounting across iterations, and convergence behavior is provided, together with an empirical evaluation of clustering quality and convergence.
4. **Extensive experimental study.** DP2SOM is evaluated on five real-world datasets (Cancer, Iris, OlivettiFaces, MNIST, and Covertypes) and compared with three representative clustering approaches: classical SOM, a differentially private k-means implementation based on IBM Diffprivlib, and the PrivGene model, a differentially private model fitted using genetic algorithms. Mean Squared Deviation (MSD) is used as the primary utility metric.
5. **Improved privacy-utility trade-off relative to the evaluated baselines.** When compared with representative differentially private clustering baselines, including DPLloyd, Hybrid DP k-means, DP k-means, local-DP k-means, Diffprivlib, and PrivGene, DP2SOM shows a more favorable privacy-utility trade-off in the evaluated settings. The method maintains clustering quality across a range of privacy budgets.
6. **Robustness across settings.** The results indicate that the local distance-based mechanism generalizes across algorithms and data scales, from small datasets with $k = 3$ to larger datasets with $k = 10$, which supports the broader applicability of DP2SOM.

6.2 Related Work

This section briefly summarizes prior work on differentially private data mining, particularly in clustering. Dwork [Dwork, 2006] introduced the notion of differential privacy and developed it further in a series of studies [Blum et al., 2005, Dinur et al., 2003, Dwork et al., 2006, Dwork et al., 2004, Friedman and Schuster, 2010]. In summary, several major mechanisms support differential privacy: the Laplace mechanism for numeric data [Dwork et al., 2006], the exponential mechanism for categorical data [McSherry et al., 2007], the sample-and-aggregate framework [Kasiviswanathan et al., 2011], and randomized response, originally developed in the social sciences to allow individuals to answer sensitive questions in a privacy-preserving way [Ma et al., 2021].

The authors of [Zhao et al., 2018] described a differentially private system for gradient-boosting decision trees in regression and binary classification. Rana et al. [Rana et al., 2015] proposed a differentially private decision-tree induction algorithm. Another important privacy-preserving method for multi-core DBSCAN clustering was suggested by [Ni et al., 2018]. Friedman et al. [Friedman and Schuster, 2010] addressed privacy and algorithmic requirements simultaneously, focusing on decision-tree induction as a sample application, and made the ID3 algorithm differentially private. The studies in [Chaudhuri et al., 2008, Chaudhuri et al., 2011] proposed a differentially private logistic-regression algorithm. In another study, an efficient and private solution to the horizontally federated k-means problem was introduced by using a computationally differentially private model [Diao et al., 2024].

Although extensive research has been conducted on both privacy in clustering and differential privacy, to the best of our knowledge no prior study has applied differential privacy directly to SOM clustering. However, several variants of differentially private k-means have been proposed, and k-means is sufficiently similar to SOM to provide a useful comparison. A group of methods is characterized by injecting Laplace noise into the iterations of Lloyd's algorithm. Our approach is compared with studies such as [Blum et al., 2005]. The main distinction between these methods lies in how the privacy budget is allocated across iterations. The sample-and-aggregate framework for k-means clustering was developed by Kasiviswanathan et al. [Kasiviswanathan et al., 2011] and later implemented in the GUPT system [Mohan et al., 2012]. In [Zhang et al., 2013], a general private model-fitting framework based on genetic algorithms, PrivGene, was proposed for k-means clustering. In PrivGene, candidates for the next iteration are selected through the exponential mechanism. When the dataset is relatively small, PrivGene can achieve acceptable clustering quality; however, as with other existing differentially private k-means algorithms, the number of iterations must be fixed in advance.

In this chapter, differential privacy is introduced in the context of SOM. The original DP definition is extended with a local distance metric, and an effective method for achieving this extension is presented through the DP2SOM protocol. In our approach, Laplace noise is added to the current centroid according to a bounded local sensitivity, while local convergence is maintained.

6.3 Theoretical Foundations and Problem Statement

6.3.1 SOM Clustering

A Kohonen map, or Self-Organizing Map (SOM), is an important neural-network model for data-mining tasks, especially clustering and dimensionality reduction [Yang et al., 2012]. SOM is an unsupervised neural network that can learn from complex multidimensional data and transform them into a topological map of much lower dimension, typically one or two [Zhou et al., 2009, Zhang et al., 2017].

These characteristics make SOM suitable for e-business applications, especially recommender systems.

The SOM architecture consists of two fully connected layers: an input layer and a Kohonen layer, which together form a topology-preserving map [Yang et al., 2012]. Chapter 5, Section 5.5, describes SOM in more detail. SOM performs a sequence of iterations consisting of competition and cooperation [Oja et al., 1999]. It uses a neighborhood function to preserve the topological properties of the input space. In each iteration t , for each input vector $X(t) = (X_1(t), X_2(t), \dots, X_d(t))$ from the dataset, SOM computes the Euclidean distance between the input vector and the weight vector $W_j(t) = (W_{j,1}(t), W_{j,2}(t), \dots, W_{j,d}(t))$ of each map node:

$$\|X(t) - W_j(t)\| = \sqrt{\sum_{i=1}^d (X_i(t) - W_{j,i}(t))^2} \quad (6.1)$$

The neuron c whose weight vector $W_c(t)$ has the minimum distance to the input data $X(t)$ is called the winner neuron:

$$c = \arg \min_j \|X(t) - W_j(t)\| \quad (6.2)$$

During the cooperation phase, the weight vectors of the winning neuron and the neurons in its neighborhood $G(r_c)$ on the SOM grid are updated, where r_j is the physical position of neuron j on the grid. The magnitude of the update decreases over time and is smaller for neurons that are farther away from the winner neuron. The change function at iteration t is denoted by $Z(r_j, r_c, t)$.

The update rule for the winning neuron c and the neurons in the neighborhood $G(r_c)$ is:

$$\forall j \in G(r_c), \quad W_j(t+1) = W_j(t) + Z(r_j, r_c, t) [X(t) - W_j(t)] \quad (6.3)$$

Figure 6.2 shows the SOM training process on a two-dimensional dataset. From a privacy perspective, SOM exhibits several vulnerabilities in its communication and update phases:

- Securely discovering the winner neuron from data privately held by multiple parties
- Securely updating neuron weight vectors
- Securely determining the termination status of SOM [Han et al., 2007]

A further practical limitation of SOM is that it requires sufficient data to develop meaningful clusters. In addition, perfectly preserving cluster structure can be difficult when the groupings are highly irregular or overlapping.

To address these challenges, several SOM extensions have been designed to reduce privacy gaps without incurring excessive information loss or computational overhead. These hybrid methods are often based on soft-computing techniques. The privacy-preserving model considered in this chapter is illustrated in Figure 6.3. The privacy-aware components are denoted by modules M1–M3. M1 performs

local mapping, privacy specification, offline SOM training, and local fuzzification, normalizing input data and computing privacy budgets (ϵ_i) to bound sensitivity before local mapping. M2 performs central SOM aggregation and integrates the locally trained maps while introducing differential noise to preserve privacy in the global representation. M3 employs an optimization procedure based on genetic algorithms that adaptively adjusts local privacy settings and cluster centroids in order to maintain privacy and clustering accuracy during convergence. Together, these modules introduce local sensitivity control and noise throughout the learning process, thereby reducing the privacy leakage that occurs in standard SOM training.

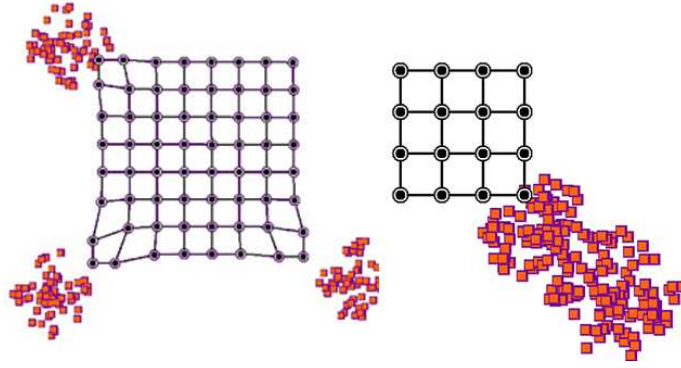


Fig. 6.2: Training process of SOM on a two-dimensional dataset [Kohonen, 2001].

SOM is comparable to k-means, with both advantages and disadvantages [Kumar et al., 2010]. From a performance perspective, k-means can outperform SOM when the number of clusters increases, although it is also more sensitive to noise in the dataset. However, in k-means the value of k (the number of clusters) must be specified in advance. K-means is also a greedy algorithm, which makes it difficult to guarantee global optimality. In SOM, the number of neurons in the output layer is closely related to the number of classes or patterns in the input data [Riveros et al., 2019].

Differential privacy has already been applied to k-means clustering. However, to the best of our knowledge, no prior work has used differential privacy directly to improve SOM clustering.

6.3.2 Definition of Differential Privacy

Differential privacy is a privacy requirement that can be achieved through a distortion-based mechanism. The key idea is to add random noise to the data or to the output of a function so that private information cannot be inferred from in-

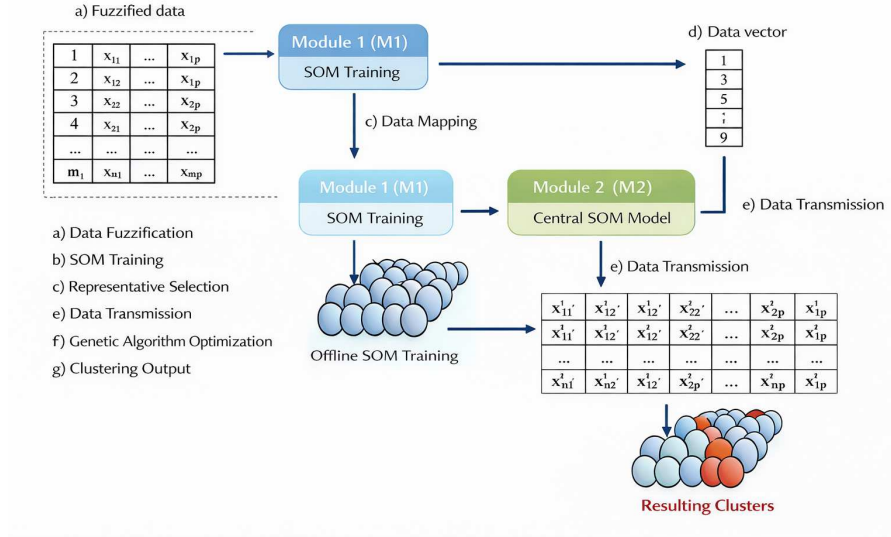


Fig. 6.3: Privacy-preserving Self-Organizing Map model. M1: local mapping, privacy specification, offline SOM training, and local fuzzification. M2: central SOM aggregation, which integrates representative nodes while applying differential privacy noise. M3: optimization based on genetic algorithms that adjusts centroid convergence and local privacy settings.

direct analysis [Dwork, 2006, Zhu et al., 2017]. Differential privacy was originally introduced in the context of databases because more robust privacy guarantees were needed [Friedman and Schuster, 2010, Ji et al., 2014]. The formal definition is as follows:

Definition 1. (ϵ -Differential Privacy) Let dataset D differ from dataset D' by at most one data record; such datasets are called adjacent datasets. Let F be a randomized function and let $\text{Range}(F)$ denote its output space. If F satisfies ϵ -differential privacy, then for all measurable subsets $M \in \text{Range}(F)$:

$$\Pr[F(D) \in M] \leq \exp(\epsilon) \Pr[F(D') \in M] \quad (6.4)$$

This condition ensures that the output of the query changes only slightly when a single record in the database is modified [Dwork et al., 2006].

The key issue in differential privacy is the choice of the privacy budget ϵ . Smaller values of ϵ correspond to stronger privacy protection. In the limiting case $\epsilon = 0$, the mechanism becomes independent of the data, which corresponds to a degenerate form of perfect privacy but yields no useful information. An example of noisy counting with different values of ϵ is shown in Figure 6.4.

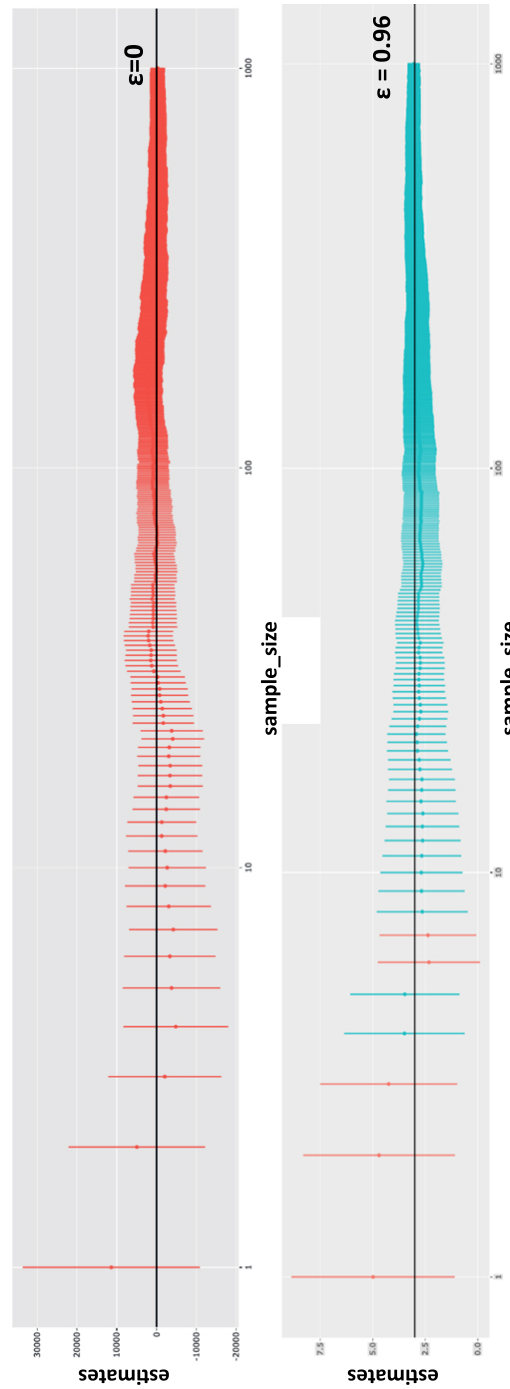


Fig. 6.4: Example of different values of ϵ and their effect on the accuracy of noisy counting.

6.3.3 Global Sensitivity

Global sensitivity in differential privacy is a key concept that measures how much the output of a function can change when a single record in the input dataset is modified. More precisely, it is defined as the maximum change in the output of a function when one record in the dataset is added or removed. Given a function f that takes a dataset D as input and returns a result, the global sensitivity Δf of f is defined as:

$$\Delta f = \max_{D, D'} \|f(D) - f(D')\| \quad (6.5)$$

Determining and bounding global sensitivity is essential for implementing differential privacy in an effective and mathematically sound way.

6.3.4 The Privacy Budget and Local Function Sensitivity

Two parameters are especially important in differential privacy. First, the privacy budget parameter ϵ determines the magnitude of the random noise and reflects how much two neighboring datasets may differ in the released output. Smaller values of ϵ provide a stronger privacy guarantee. Second, local function sensitivity Δf , where f is the query function applied to dataset D , measures the change between neighboring datasets:

$$\Delta f(D) = \max_{D'} |f(D) - f(D')| \quad (6.6)$$

Differential privacy ensures that a single record does not significantly affect the outcome of an analysis, with the strength of this guarantee controlled by ϵ . In this sense, a person's privacy should not be substantially affected by participation in the dataset, because the resulting output changes only within a controlled bound.

6.3.5 Mechanisms of Differential Privacy

In differential privacy, a mechanism is an algorithm designed to protect individual privacy while performing computations or queries on a dataset. These mechanisms introduce randomness into the output so that the presence or absence of any particular individual in the dataset does not substantially affect the result.

In the standard formulation, a mechanism M satisfies ϵ -differential privacy if, for all pairs of datasets x and y differing in the data of one person, and for all events S ,

$$\Pr[M(x) \in S] \leq e^\epsilon \Pr[M(y) \in S] \quad (6.7)$$

Common mechanisms include the Laplace mechanism, the Gaussian mechanism, the exponential mechanism, report-noisy-max, and sparse-vector techniques. Each

has different use cases depending on the sensitivity of the function and the desired privacy guarantees. In this chapter, the Laplace mechanism is employed, based on the sensitivity of the function to be protected [Dwork, 2008].

6.3.5.1 Laplace Mechanism

Definition 2. (*Laplace noise*) Let the location parameter be 0 and let the scale parameter be $b = \Delta f / \epsilon$. Then Laplace noise is drawn from a Laplace distribution with scale b . In this chapter, the added perturbation is represented abstractly as:

$$\eta \sim \text{Lap}\left(0, \frac{\Delta f}{\epsilon}\right) \quad (6.8)$$

where η denotes the random noise added to the quantity of interest. In the implementation considered here, the sensitivity is set to the constant value $\Delta f = 1$ after normalization, so that neighboring datasets differ by at most one record in bounded form. Under this simplification, the problem can be treated as an optimization problem in which ϵ is the main variable.

To balance privacy and utility, acceptable values of the privacy budget ϵ are chosen through the optimization objective:

$$\min_{\omega} (F(D), \omega) \quad \text{subject to } \epsilon \geq 0 \quad (6.9)$$

where D denotes the dataset, ω the desired output, and $(F(D), \omega)$ the objective function to be minimized.

6.3.5.2 The Exponential Mechanism

The exponential mechanism chooses an output with probability proportional to its utility and is particularly useful for categorical or non-numeric outputs. Given a dataset D and a utility function $u(D, r)$ measuring the desirability of output r , the mechanism selects outputs with probability increasing exponentially in the utility score, normalized by the sensitivity of the utility function. This mechanism is widely used when direct numeric perturbation is not appropriate.

6.3.5.3 The Median Mechanism

The median mechanism makes it possible to estimate an approximate median while protecting the privacy of each member of the dataset. Noise is introduced into the selection process so that the result does not reveal too much about any individual data point. This is useful for private statistical analysis in settings such as income data or medical records, where the median is often a meaningful summary statistic.

6.3.6 Sequential Composition

Sequential composition is an important principle in differential privacy that describes how privacy guarantees degrade when multiple differentially private mechanisms are applied to the same dataset. Specifically, if k differentially private mechanisms are used on the same data, the total privacy loss is approximately the sum of the individual privacy losses.

Formally, if mechanisms $M_1, M_2, \dots, M_k$ satisfy (ϵ_i, δ_i) -differential privacy, then their sequential composition satisfies:

$$\left(\sum_{i=1}^k \epsilon_i, \sum_{i=1}^k \delta_i \right)\text{-differential privacy} \quad (6.10)$$

Sequential composition is especially important in real-world data analysis because many workflows involve multiple queries or iterative algorithms. Understanding this principle is essential when designing algorithms such as DP2SOM, which operate across multiple iterations.

6.3.7 Data Release Under Differential Privacy

In the context of differential privacy, data release refers to the publication of information or statistics derived from a dataset in a way that preserves the privacy of the individuals represented in it. This is achieved by adding carefully calibrated noise either to the data themselves or to the outputs of data queries. By selecting appropriate privacy parameters and mechanisms, useful information can be released while ensuring that the inclusion or exclusion of any individual has only limited impact on the result.

Data-release mechanisms in differential privacy are typically divided into two categories: interactive and non-interactive.

6.3.7.1 Interactive Data Release

In an interactive system, users query a dataset and receive responses that are generated dynamically. To protect privacy, noise is added to each response. This approach is flexible because users can ask specific questions and build on previous answers. However, every query consumes part of the privacy budget, which limits the total number of informative queries that can be answered. Consequently, interactive mechanisms require careful privacy-budget management.

6.3.7.2 Non-Interactive Data Release

In a non-interactive system, a single differentially private output is produced and published once. This output may consist of summary statistics, synthetic data, or a sanitized version of the dataset. Once the output has been published, any number of users can use it without causing additional privacy loss. This makes implementation simpler, although users are restricted to the released outputs and cannot ask arbitrary new queries.

6.3.8 Uncertainty of Outcome

The concept of uncertainty of outcome in differential privacy means that the output of a differentially private mechanism must be intentionally randomized so that exact conclusions about any one participant cannot be drawn. This uncertainty is created through carefully calibrated noise and is fundamental to balancing privacy protection against data utility.

6.3.9 Properties of Differential Privacy

Some key properties of differential privacy are as follows:

- DP provides a strong mathematical definition of privacy for statistical and machine-learning analysis [Wood et al., 2018].
- In DP, privacy is understood in terms of the difference in output when a single individual's data are included or removed. This explains the term "differential" [Dwork et al., 2014].
- DP guarantees that any observer of the analysis outcome will make essentially the same inference about any individual's private information, whether or not that individual's data were included in the input [Chandrakar et al., 2020, Balcer, 2022].
- DP provides mathematically provable protection against several important privacy attacks, including differencing attacks, linkage attacks, and reconstruction attacks [Pan et al., 2024].

6.3.10 Problem Statement

In clustering algorithms, performance depends strongly on distances and densities [Su et al., 2001]. In many clustering methods, including distance-based methods, these distances implicitly encode relationships among data points. Such relationships

may themselves reveal sensitive information if they are exposed or inferred during training. It is therefore necessary to protect these relationships during clustering.

Differential privacy limits the influence of any single record on the released result and is thus a suitable mechanism for protecting this relational information. In the literature, three broad approaches have been developed for combining optimization with differential privacy [Noman et al., 2011]. The first directly perturbs the output of the optimization problem, which may leave the procedure vulnerable to direct attacks or large utility loss. The second perturbs the objective function itself, for example through objective perturbation or iterative local search, but may still be suboptimal. The third adapts an existing algorithm so that every step requiring access to the input data is itself made private.

In this chapter, the third approach is adopted. Laplace noise is applied to centroids during each iteration so that SOM clustering remains useful while satisfying differential privacy.

6.4 Method

6.4.1 Approach Overview

A differential-privacy-based protection technique is introduced to mitigate privacy leakage in SOM clustering. Numeric data are protected by adding Laplace noise, while categorical mechanisms may in principle be handled through alternative DP mechanisms when needed. To reduce the adverse impact of the privacy budget ϵ , an adaptive distance-based mechanism is adopted so that the distortion and convergence of centroids in each iteration can be adjusted to maintain an acceptable balance between privacy protection and data utility.

An overview of the proposed approach is shown in Figure 6.5. At iteration k for cluster i , C_i^k denotes the real centroid and $\bar{C}_i^k$ the differentially private centroid. The essential idea of the re-centroiding process is that the sampling zone is bounded using the Euclidean distance, so that convergence between the relocated centroids is preserved as much as possible. A suitable privacy budget must be chosen to achieve acceptable clustering quality while improving privacy. In particular, the trajectory formed by the noisy centroids $\bar{C}_i^{k-1}, \bar{C}_i^k, \bar{C}_i^{k+1}, \dots$ should remain close to the trajectory formed by the true centroids $C_i^{k-1}, C_i^k, C_i^{k+1}, \dots$. The greater the similarity between these trajectories, the higher the resulting clustering quality.

This chapter presents the DP2SOM approach for differentially private SOM clustering and explains how important parameters can be selected through error analysis. The common steps and components of the method are then described in detail.

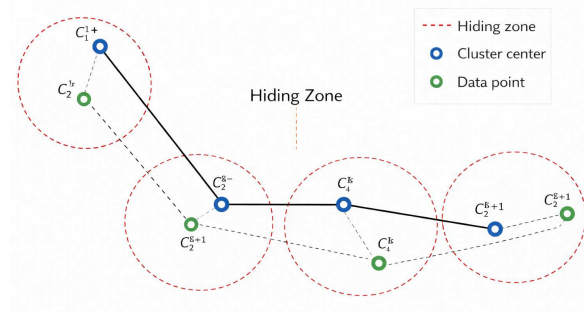


Fig. 6.5: Overview of the bounding-zone technique for cluster C_i^k in each iteration, used to preserve convergence between relocated centroids. Each cluster center is surrounded by a hiding zone (red dashed circle), within which sensitive data points may be perturbed to prevent disclosure.

6.4.2 Initial Calculations and Transformations

In distance-based clustering algorithms, the distance between data points must first be computed. Let x_i and x_j be two records in dataset D . Their Euclidean distance is defined as:

$$\text{dist}(x_i, x_j) = \|x_i - x_j\|_2 \quad (6.11)$$

This distance encodes the relationship between two data points. An attacker using an inference-based attack may exploit such relationships, which can lead to privacy leakage. Therefore, in the proposed method these distances are protected using differential privacy with Laplace noise:

$$\text{dist}'(x_i, x_j) = \text{dist}(x_i, x_j) + \eta, \quad \eta \sim \text{Lap}\left(0, \frac{\Delta f}{\epsilon_i}\right) \quad (6.12)$$

The distorted distances help protect privacy but can also affect clustering results. Therefore, the amount of added noise must be calibrated to achieve an acceptable privacy–utility balance.

6.4.3 DP2SOM Approach and Its Analysis

This section explains the motivation for the bounding zones, shows how the balance between privacy and clustering quality is maintained, and then describes the proposed approach in more detail.

In DP2SOM, defining bounding zones and determining the privacy budget ϵ are central to managing the privacy–utility trade-off. Before local mapping during offline SOM training, the input data are normalized and local privacy budgets (ϵ_i)

are calculated in order to bound sensitivity. The central aggregation module (M2 in Figure 6.3) integrates these maps and adds differential noise to support global privacy preservation, while each local SOM module applies fuzzification and injects differential noise appropriate to its local distance measure.

A metaheuristic procedure based on a genetic algorithm is used to optimize the privacy budget while maintaining clustering accuracy. To reduce distortion in the mapping function $F(D)$ while preserving data utility, the optimization function introduced earlier is used to identify acceptable values of ϵ . Throughout convergence, the GA adaptively adjusts centroid locations and local privacy settings so that the noise added at each stage supports a balance between privacy and utility.

Convergence is assumed to be achieved when an iterative randomized clustering procedure satisfies the following condition for cluster i at iteration k :

$$\|\bar{C}_i^k - C_i^k\| < \|C_i^k - \bar{C}_i^{k-1}\| \quad (6.13)$$

where i denotes the cluster number and k the current iteration. The bounding zone C^* is then defined as the set of noisy cluster centroids collected over the iterations:

$$C^* = \{\bar{C}_1^k, \bar{C}_2^k, \dots, \bar{C}_i^k\} \quad (6.14)$$

By updating centroids iteratively within this bounded set, the value of the distance function can be further reduced, and convergence can be stabilized.

Intuitively, the bounding zone prevents excessive divergence, stabilizes convergence, and reduces structural distortion, while Laplace noise is injected according to the local sensitivity of each centroid. In this way, neighborhood relationships are preserved during training because the trajectory of the noisy centroids remains close to that of the true centroids.

Noise can strongly distort clustering results, especially when the privacy budget is small. Consequently, a mechanism is needed that can adaptively adjust the privacy budget ϵ so that both privacy and clustering quality remain acceptable.

The approach designed in this chapter aims to satisfy ϵ -differential privacy. To reduce the risk that a noisy centroid is moved into an unsuitable position, it is assumed that centroids do not lie directly on the edge of the boundary zone. The total privacy budget is then expressed as:

$$\epsilon = \epsilon_0 + \sum_{k=1}^X \epsilon_i^k \quad (6.15)$$

where X is the maximum number of training iterations required for convergence. In each iteration k , a local privacy budget ϵ_i^k is applied to the current centroid position in cluster i . Additional Laplace noise is then applied to the final clustering result C^* . Assuming local sensitivity equal to 1, the algorithm remains ϵ -differentially private under the bounded-sensitivity assumption.

The DP2SOM approach consists of three phases: data preprocessing, DP2SOM perturbation, and SOM clustering, as shown in Figure 6.6.

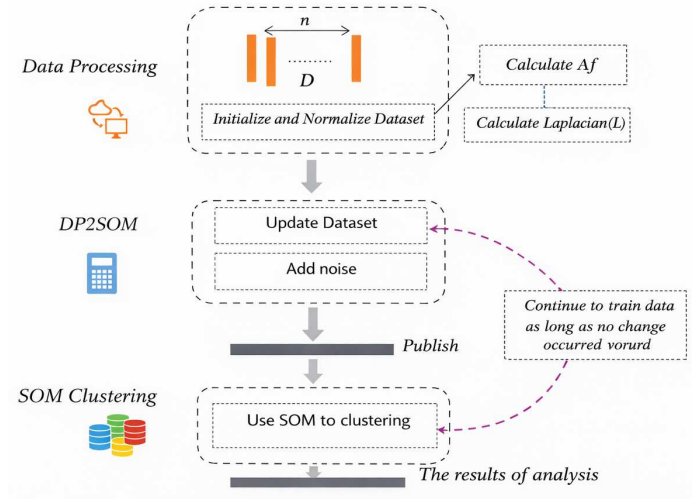


Fig. 6.6: Schematic diagram of the DP2SOM algorithm. The dataset is initialized and normalized during preprocessing. Sensitivity is calculated and Laplace noise is generated according to the differential privacy mechanism. The dataset is then updated with noise to protect sensitive information before SOM clustering is applied to the privatized dataset.

6.4.3.1 Data Processing

The dataset is first normalized and initialized as the first stage of preprocessing. These inputs are used to compute the privacy-specification values for the given dataset D and the associated privacy preference ϵ . The local function sensitivity (Δf) and the Laplace mechanism $\text{Lap}(b)$ are then computed from the equations above.

6.4.3.2 DP2SOM

The central idea of DP2SOM is to identify groups of related items through a distance measure while protecting the data used in this process. Training begins by normalizing the data and initializing random node-weight vectors. After the input dataset D and privacy budget ϵ are provided, each cluster centroid is updated locally. The second phase, which is the central stage of the method, is divided into noise introduction and local distance-based updating. A complete pseudocode description is given in Algorithms 2 and 3.

The general SOM training procedure using Euclidean distance is represented in Algorithm 2, whereas Algorithm 3 specifies the Laplace-based perturbation mechanism that is applied during centroid updating. Because all input features are nor-

malized to $[0, 1]$, we assume $\Delta f = 1$ in order to simplify calibration of the Laplace noise.

Algorithm 3 does not represent a second independent SOM training process; instead, it provides the DP perturbation mechanism used within each training iteration of Algorithm 2. The noisy dataset D' therefore acts as a privacy-preserving intermediate representation of the original dataset D . The Laplace mechanism, parameterized by the local privacy budget ϵ_i , adds noise to the numerical attributes of each record in D' . Algorithm 2 then uses this privatized dataset during SOM training so that individual data records cannot easily be reconstructed or reidentified while topological relationships are approximately preserved.

Algorithm 2 DP2SOM algorithm

Input: Input dataset D , privacy budget ϵ_i for cluster i in each iteration, and privacy budget ϵ_0 for the final output.

Output: D' : differentially private clustered dataset; C^* : set of final centroids

- 1: Initialize by selecting random node-weight vectors in the map.
- 2: **while** training is not completed and changes still occur **do**
- 3: Pick a random input vector $D(t)$.
- 4: **for** each node in the map **do**
- 5: Compute its Euclidean distance from the current input vector.
- 6: Use the Euclidean distance formula in Equation (6.10) to compute similarity between the input vector and the weight vector of the node.
- 7: Calculate Δf ; /* For simplicity, $\Delta f = 1$. */
- 8: Add ϵ_i -controlled noise to the centroid position:

$$dist'(x_i, x_j) = \|x_i - x_j\|_2 + \eta, \quad \eta \sim \text{Lap}\left(0, \frac{\Delta f}{\epsilon_i}\right)$$

- 9: Re-centroiding: update each noisy centroid position as

$$\bar{C}_i^k = C_i^k + \eta_i, \quad \eta_i \sim \text{Lap}\left(0, \frac{\Delta f}{\epsilon_i}\right)$$

- 10: Evaluate convergence: if $\|\bar{C}_i^k - C_i^k\| < \|C_i^k - \bar{C}_i^{k-1}\|$, retain $\bar{C}_i^k$ in C^* .
 - 11: Update the objective: $\min_{\omega} (F(D), \omega)$ subject to $\epsilon_i \geq 0$.
 - 12: Track the node that produces the smallest distance. /* This node is the best matching unit (BMU). */
 - 13: **end for**
 - 14: Update the weight vectors of the nodes in the neighborhood of the BMU by pulling them closer to the input vector.
 - 15: **end while**
 - 16: Add final Laplace noise with privacy budget ϵ_0 to obtain the released centroid set C^* .
-

Algorithm 3 Local Laplace perturbation mechanism for differentially private SOM**Input:** Dataset D , privacy budget ϵ_i for each cluster, maximum number of iterations n **Output:** Differentially private centroids C^* and cluster assignments

- 1: Initialize SOM training using normalized input features.
- 2: **while** current iteration $\leq$ max number of iterations **do**
- 3: Identify the Best Matching Unit (BMU) for the current input vector.
- 4: Compute the sensitivity of the mapping function F :

$$S = \max_{D, D'} \|F(D) - F(D')\|_1$$

/* Since the data are normalized, $S = 1$. */

- 5: Compute Laplace-noise scale $b = S/\epsilon_i$ and variance $2b^2$.
- 6: Draw an independent noise vector $L \sim \text{Lap}(0, b)$.
- 7: Update the BMU and neighboring weights:

$$W_c(t) \leftarrow W_c(t) + L$$

- 8: Update the corresponding noisy centroid:

$$\tilde{C}_i^k = C_i^k + L$$

- 9: **end while**

- 10: **Return:** set of final noisy centroids C^* and their cluster assignments.

6.4.3.3 SOM Algorithm

The final stage of the approach applies standard SOM clustering to the privatized dataset. The SOM training process consists of four main steps, corresponding broadly to lines 1–7 of Algorithm 2:

1. Initialization:

- a. Randomize the node-weight vectors in the map.
- b. Randomly select an input vector.

2. Finding the BMU:

- a. Traverse each node on the map.
- b. Use the Euclidean distance to measure similarity between the input vector and each node's weight vector.
- c. Identify the node that produces the smallest distance; this node is the BMU.

3. Updating the map:

- a. Update the weight vectors of the BMU and its neighboring nodes by moving them toward the input vector.

4. Repeat step 2 for N iterations while the current iteration remains below the iteration limit.

6.5 Experimental Evaluation

This section presents the experimental evaluation of DP2SOM on real datasets. The method is compared in particular with differentially private k-means approaches [Blum et al., 2005, Su et al., 2016], including the implementation in [Holohan et al., 2019]. The analysis also examines how clustering quality depends on the privacy budget.

6.5.1 Datasets and Experimental Setup

To evaluate the generality of the proposed method, experiments were conducted using three clustering approaches and five datasets.

1. IBM Diffprivlib k-means, a differentially private k-means implementation [Holohan et al., 2019, Su et al., 2016].
2. DPEM / PrivGene, a differentially private model fitted using genetic algorithms [Zhang et al., 2013].
3. Classical SOM.

The k-means and SOM algorithms were used as baseline clustering approaches.

The method was implemented on a machine equipped with an Intel(R) Core(TM) i7-3632QM CPU @ 2.2 GHz, 8 GB of memory, a 64-bit Windows 10 operating system, and Python 3.7. The details of the datasets are listed in Table 6.1.

Table 6.1: Experimental datasets.

Dataset	Type	Attributes	Records	Description
Iris	Real	4	150	The Iris flower dataset is a widely used multivariate benchmark in pattern recognition. It contains 50 samples from each of three species: Iris setosa, Iris virginica, and Iris versicolor [Dua et al., 2019].
Cancer	Real	32	569	A cancer dataset containing patient-related clinical information that can support applications such as drug evaluation, quality review of cancer care, and observational research [Bright et al., 2020].
OlivettiFaces	Image	400	4,096	The OlivettiFaces dataset, also known as the AT&T dataset, is a well-known face-image benchmark often used in face-recognition research [Vedantam et al., 2014].
MNIST	Image	784	60,000	The MNIST database of handwritten digits is a standard benchmark for pattern-recognition methods on real-world image data [Lecun et al., 1998].
Coverttype	Real	54	581,012	The Coverttype dataset is a well-established benchmark containing cartographic attributes for forest-cover-type classification, including elevation, slope, and soil type [Pruthvi et al., 2015].

6.5.2 Evaluation Metrics

A data-utility metric is used to assess clustering quality and privacy. In particular, Mean Squared Deviation (MSD) is employed to evaluate clustering utility:

$$MSD = \frac{1}{N} \sum_{i=1}^n (D_i - D'_i)^2 \quad (6.16)$$

where N denotes the number of observations, D_i the observed values, and D'_i the corresponding predicted or transformed values. A smaller MSD value indicates more compact clustering within classes and greater separation between classes. In other words, the smaller the MSD value, the better the clustering result.

6.5.3 Parameter Selection

In the experimental evaluation of the k-means-based baselines, two values were used for the number of clusters k . A value of $k = 3$ was set for smaller datasets, whereas $k = 10$ was used for larger datasets. For the privacy budget, a range of positive ϵ values was evaluated over 30 points, with the experiments initialized at $\epsilon = 0.7$. In PrivGene (the DPEM algorithm) [Zhang et al., 2013], the sensitivity factor was set to $C = 1$ with 100 iterations.

To configure the SOM clustering parameter in DP2SOM, the following heuristic equation was used to determine the grid size for dimensionality reduction:

$$GridSize = 5\sqrt{S} \quad (6.17)$$

where S denotes the number of samples in the dataset under analysis. The Gaussian neighborhood function was applied in all experiments. Because the quality of SOM clustering under DP2SOM was being assessed, the maximum number of iterations was set to one iteration per sample, although in real-world applications this could be limited further.

To validate clustering quality, several standard criteria were employed, including average quantization error (QE) between data vectors and BMUs on the map, and topographic error (TE), which counts topology violations generated by the algorithm [Kohonen, 2001]:

$$QE = \frac{1}{N} \sum_{k=1}^N \|x_k - m_k\| \quad (6.18)$$

where N is the number of data vectors and m_k is the best-matching prototype corresponding to data vector x_k . Smaller QE values indicate better clustering quality.

$$TE = \frac{1}{N} \sum_{k=1}^N u(x_k) \quad (6.19)$$

where N is the number of input samples, and $u(x_k)$ is 1 if the first and second BMUs of x_k are not adjacent; otherwise, $u(x_k)$ is 0. A smaller TE value is more desirable.

6.6 Analysis and Results

Since DP2SOM is, to the best of our knowledge, the first method of this type for SOM clustering, it is compared with k-means, a related unsupervised clustering technique, and with PrivGene, a differentially private metaheuristic method. The experiments use benchmark datasets to evaluate the ability of DP2SOM to identify clusters while preserving privacy.

6.6.1 Accuracy Analysis of the Iris Dataset Before and After DP

Accuracy is defined here as clustering accuracy (CA), obtained by mapping each cluster to its most common true label. This is a standard way to evaluate unsupervised clustering methods against known class structures.

Figure 6.7 compares the clustering results for the Iris dataset in terms of accuracy analysis. Figure 6.8 shows the U-matrix results before and after applying DP.

Figure 6.7 compares the quantization error before and after applying differential privacy during SOM training. The left panel (Before DP) shows the expected decline in error across iterations, reflecting stable convergence of the SOM centroids. The right panel (After DP) shows that even after noise injection, convergence is still preserved, and quantization error continues to decrease, although with slightly greater variance in early iterations. These results indicate that DP2SOM maintains effective learning dynamics while incorporating privacy-preserving perturbations.

The U-matrix visualizations in Figure 6.8 show the cluster topology of the Iris dataset before and after differential privacy. In both cases, the main regions of high and low distance remain visible, preserving neighborhood structure. After DP injection, the general cluster boundaries remain intact, suggesting that DP2SOM maintains topology-preserving learning despite perturbation.

6.6.2 Accuracy Analysis of the MNIST Dataset Before and After DP

A comparison of clustering accuracy for the MNIST dataset before and after using DP is shown in Figures 6.9 and 6.10. Figure 6.11 shows a noisy execution of DP on the MNIST dataset.

As described in Table 6.1, the experiments were conducted on datasets of different sizes and types. In the smaller real-valued datasets, such as Iris and Cancer, the Euclidean distance between samples is often small, which makes clustering more

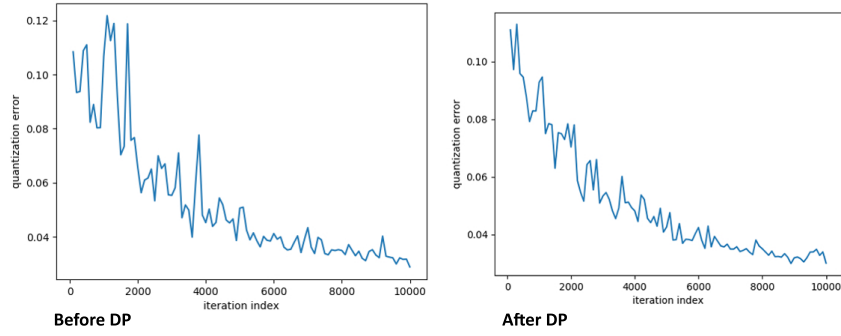


Fig. 6.7: Results for the Iris dataset: accuracy analysis. Quantization-error curves for the Iris dataset before (left) and after (right) applying differential privacy. DP2SOM preserves convergence trends with only limited variance introduced by noise injection.

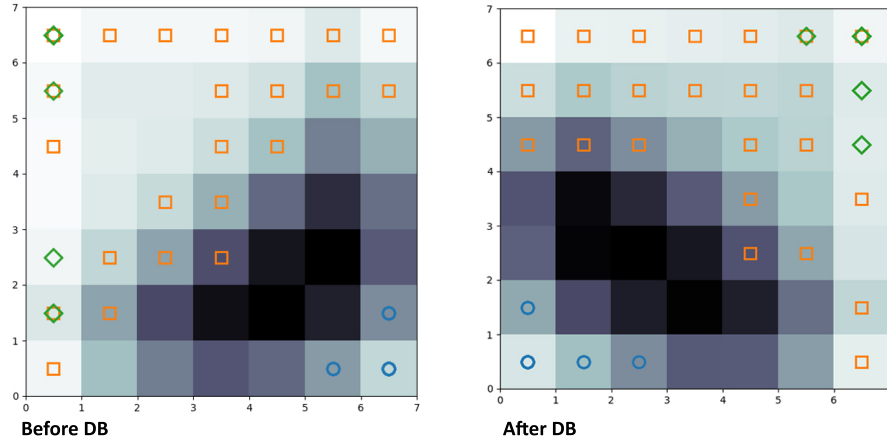


Fig. 6.8: Results for the Iris dataset: U-matrix visualizations before (left) and after (right) applying differential privacy. Cluster boundaries and neighborhood structures remain largely consistent, indicating topology preservation under DP2SOM.

sensitive to the addition of noise. Under these conditions, DP2SOM and IBM k-means behave more robustly than PrivGene.

The precise measurement of similarity between DP2SOM and the k-means-based baselines is affected by dataset size and the level of perturbation. Nonetheless, the experimental results indicate that DP2SOM performs well in terms of both privacy protection and clustering quality.

Figures 6.10 and 6.11 provide a visual interpretation of the MNIST results. Under a naïve noisy execution, excessive perturbation can disrupt cluster separability. By

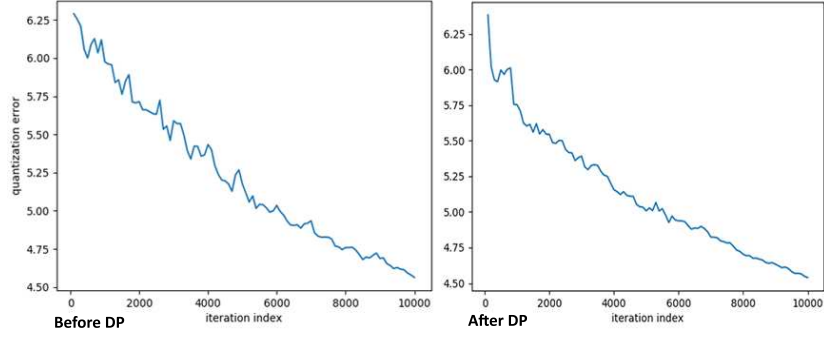


Fig. 6.9: Results for the MNIST dataset: accuracy analysis. Quantization-error curves for MNIST before (left) and after (right) applying differential privacy. DP2SOM preserves convergence trends while introducing only moderate additional variance.

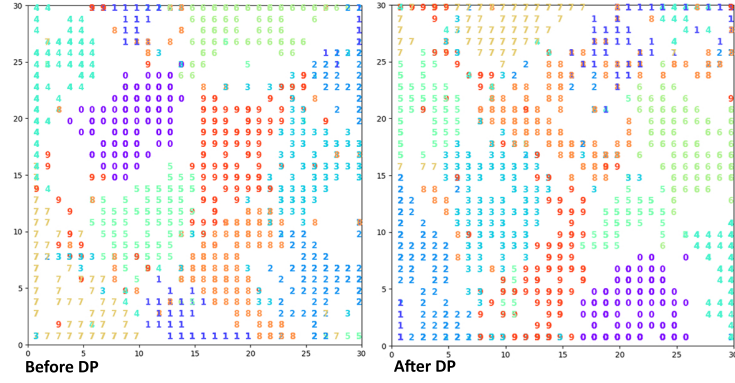


Fig. 6.10: U-matrix visualizations of the MNIST dataset before differential privacy (left) and after DP2SOM (right). The main topology and neighborhood structure remain coherent after applying the proposed method.

contrast, DP2SOM preserves a more coherent organization of the clusters, indicating that localized distance-based perturbation offers a better privacy–utility balance than uncontrolled noise injection.

6.6.3 Experimental Dataset Analysis

Figure 6.12 shows how the clustering quality, measured through MSD and related error behavior, changes across datasets as ϵ varies. It compares DP2SOM with representative baselines including IBM Diffprivlib and DP2EM. The figure indicates that

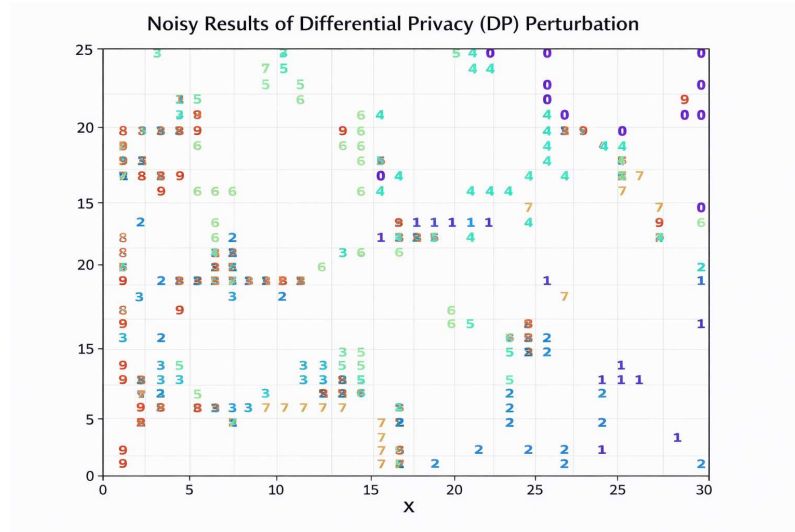


Fig. 6.11: Illustration of a noisy DP execution on the MNIST dataset. Controlled perturbation changes the position of individual points while still preserving much of the global distribution, thereby illustrating the privacy–utility trade-off.

PrivGene is the only method whose performance curve does not consistently improve as the privacy budget grows. When IBM k-means is compared with DP2SOM, the clustering results are more favorable for DP2SOM, especially when the number of clusters is not too large. For the larger datasets, DP2SOM shows particularly clear advantages.

6.6.4 Clustering Quality Measurement

Figure 6.13 presents the error behavior on the face dataset. The error values decrease as the privacy budget increases, which indicates that clustering quality improves when less noise is injected. At the same time, the figure shows that even under relatively strong privacy settings, the method retains an acceptable level of accuracy. Some fluctuations remain because the centroids can vary slightly across different executions.

In Figure 6.13, the vertical axis labeled “Error Rate” summarizes the SOM quality indicators QE and TE rather than classification error. Each bar represents the magnitude of one measure for a specific privacy budget ϵ . As ϵ increases, both QE and TE decrease, indicating improved map quality due to reduced noise. Under tighter privacy budgets (small ϵ), both errors increase, which is consistent with stronger privacy but lower fidelity.

Clustering Accuracy under Differential Privacy Across Multiple Datasets

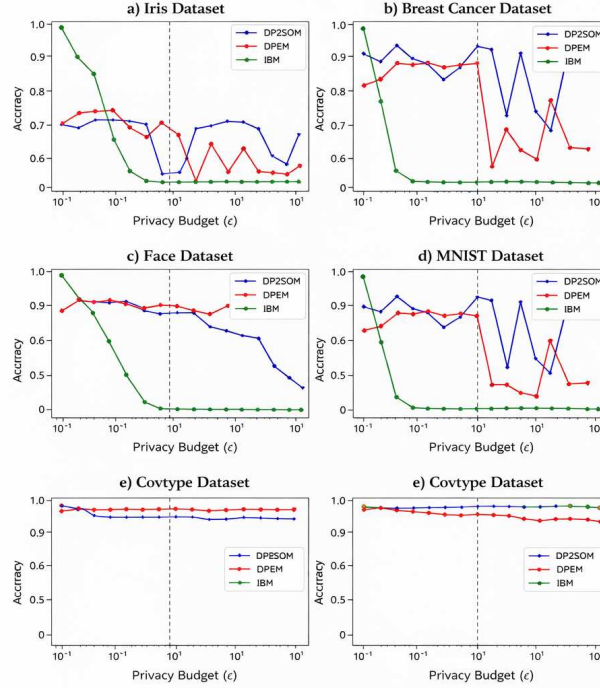


Fig. 6.12: Comparison of clustering performance under different privacy budgets for five benchmark datasets: Iris, Breast Cancer, Face, MNIST, and Covtype. DP2SOM (blue) remains comparatively stable across privacy budgets. DPEM (red) achieves acceptable utility but degrades more under tighter privacy budgets. IBM Diffprivlib k-means (green) deteriorates more sharply under strict privacy settings. Overall, the results indicate that DP2SOM provides a favorable balance between privacy protection and clustering utility across multiple datasets.

6.6.5 Comparison Between DP2SOM and Similar Methods

Table 6.2 presents a comparison between DP2SOM and representative high-quality related approaches in differentially private clustering.

Table 6.2 provides a method-level comparison. The results suggest that DP2SOM maintains higher accuracy under stricter privacy settings, particularly in the low- ϵ regime where privacy guarantees are strongest. Unlike k-means-based baselines, whose performance deteriorates more substantially as ϵ decreases, DP2SOM pre-

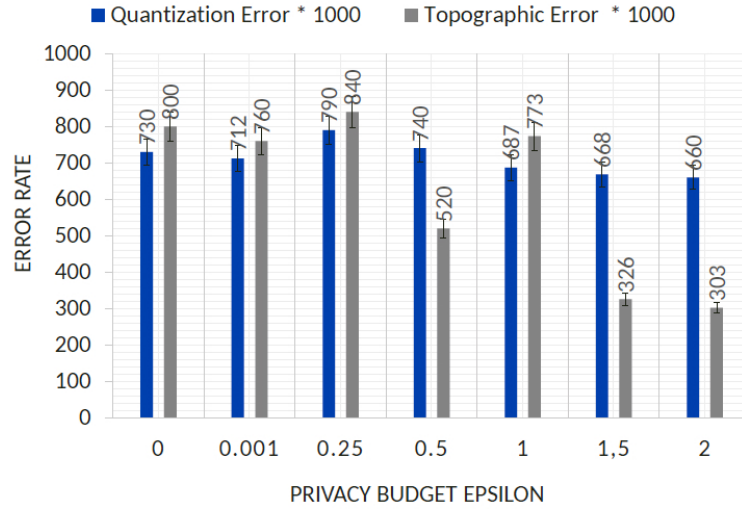


Fig. 6.13: Comparison between classical SOM and DP2SOM on the face dataset. Quantization Error $\times 1000$ (blue) and Topographic Error $\times 1000$ (gray) are shown against privacy budget ϵ . Both are standard SOM quality indicators and illustrate the trade-off between privacy and clustering quality.

serves clustering quality across both smaller datasets (such as Iris, Cancer, and Face) and larger high-dimensional datasets (such as MNIST and Covtype). This robustness results from the integration of topology-preserving updates with localized distance-based perturbation.

The averages summarized in Table 6.2 reinforce these observations, indicating that DP2SOM achieves higher mean accuracy and lower variance than the evaluated DP k-means implementations at comparable privacy budgets. Importantly, while DPEM and Diffprivlib exhibit stronger utility degradation as noise increases, DP2SOM provides a more favorable privacy–utility trade-off by injecting noise at the centroid-update level rather than perturbing final outputs globally. Taken together, the dataset-level results and the method-level summary suggest that DP2SOM not only fills the gap for differentially private SOMs but also performs favorably against established DP k-means baselines in terms of stability and clustering quality.

6.7 Discussion

The protection of user privacy in big-data mining remains an important and challenging problem. In this chapter, privacy protection in distance-based clustering has been addressed through an extension of differential privacy to SOM-based clustering. The DP2SOM approach was proposed, and clustering quality was improved

Table 6.2: Comparison of DP2SOM with representative differentially private clustering approaches. MSD ranges are approximate across five datasets (Iris, Cancer, Face, MNIST, Covtype).

Method	Mechanism	Topology Preserving	Convergence Guarantee	Error / Utility Trade-off	MSD Range	Curator Assumption
DP2SOM	Local distance-based vector noise on SOM centroids with per-centroid ϵ	Yes	Yes (region-specific under DP noise)	Low error, stable accuracy across ϵ ; strong performance on structured data	0.15–0.35	Central
DPLloyd [Blum et al., 2005]	Interactive means (synopsis + centroid noise)	DP k- No	Partial	Moderate error; global noise reduces accuracy	0.40–0.65	Central
Hybrid DP means [Su et al., 2016]	Interactive + non-interactive mechanisms	DP k- No	Partial	Moderate; depends on synopsis size	0.35–0.60	Central
DP k-means [Lu et al., 2020]	Iterative DP noise on centroids	No	Yes (k-means specific)	(k- Convergence ensured but with higher error on structured datasets	0.30–0.55	Central
LDP means [Stemmer et al., 2021]	k- Local DP with per-user perturbation	No	No	Strong privacy; high error, lower clustering quality	0.50–0.80	Local
Diffprivlib [Holo-han et al., 2019]	Standard Laplace/Gaussian noise applied to k-means	No	No	Baseline; degrades sharply under tight budgets	0.60–1.00+	Central
PrivGene [Zhang et al., 2013]	Evolutionary optimization framework	DP No	No	Improved optimization quality but dataset-specific	0.40–0.70	Central

through systematic error analysis. The main intuition behind DP2SOM is the use of distance relationships between items to identify groups of similar items while perturbing these relationships in a controlled way.

To evaluate DP2SOM, experiments were conducted on multiple benchmark datasets, with attention paid to both accuracy and topology preservation under differential privacy constraints. Figures 6.7 and 6.9 show quantization-error curves for the Iris and MNIST datasets before and after applying differential privacy. In both cases, the curves exhibit consistent downward trends, indicating stable convergence of the SOM learning process. Although the application of DP introduces some variance through injected noise, the overall trajectory remains aligned with the non-private baseline. This indicates that DP2SOM preserves convergence behavior under differential privacy constraints.

Figure 6.8 provides U-matrix visualizations of the Iris dataset before and after privacy protection. The comparison shows that cluster boundaries and neighborhood structures remain largely intact after DP2SOM is applied. The preservation of these structural relationships highlights an important advantage of the method relative to many DP k-means approaches, in which global noise injection often disrupts cluster geometry.

Figures 6.10 and 6.11 extend the analysis to the MNIST dataset. The noisy baseline illustrates how excessive perturbation can disrupt cluster separability, resulting in more scattered and incoherent patterns. By contrast, the DP2SOM output preserves more coherent cluster organization, indicating that the localized noise-injection mechanism can better balance privacy and utility at scale.

The comparison with related methods highlights that DP2SOM uniquely supports topology preservation through noise injection aligned with SOM’s neighborhood-update rule, unlike previous DP k-means approaches that perturb cluster centers independently. While DPLloyd and hybrid methods perturb global centroids or synopses, they do not preserve SOM-like structure and therefore can yield higher MSD on structured datasets. Other approaches, such as local-DP k-means, avoid the need for a trusted curator but do so at the cost of much larger error, making them less suitable for high-utility clustering. Standard toolkits such as Diffprivlib provide useful baselines but often degrade under tighter privacy budgets, while PrivGene demonstrates DP optimization without explicitly addressing topology-preserving learners. DP2SOM differs by combining local distance-based noise, per-centroid privacy budgets, and convergence-aware centroid updates, thereby achieving a more favorable privacy–utility trade-off in the evaluated scenarios.

6.7.1 Challenges in Differential Privacy

Implementing differential privacy involves several challenges.

1. **Balancing privacy and utility.** Achieving strong privacy often requires adding more noise to the data or to query results, which can substantially reduce utility. Finding an appropriate balance is one of the central challenges in DP.
2. **Parameter selection.** The privacy parameter ϵ controls the trade-off between privacy and accuracy. Choosing an appropriate value requires both domain knowledge and an understanding of the acceptable risks and noise levels in the intended application.
3. **Complexity of implementation.** Efficient differentially private algorithms can be difficult to design. Many standard data-analysis methods must be modified significantly to satisfy DP.
4. **Understanding and communication.** The privacy guarantees of DP must be explained clearly if systems are to satisfy data-protection requirements and be trusted by practitioners.

5. **Cumulative privacy loss.** Privacy loss increases with repeated queries to the same dataset. Managing this cumulative loss is difficult, especially in interactive or dynamic settings.
6. **Scalability and performance.** For large datasets or complex models, the computational burden introduced by noise calibration and sensitivity analysis can be substantial.
7. **Adversarial settings.** Although DP is theoretically robust, attackers may still exploit implementation weaknesses or auxiliary data if systems are not carefully designed.
8. **Adoption and standardization.** Despite extensive research, the lack of universally standardized tools and workflows still makes consistent adoption difficult in practice.

6.8 Summary

The partially acceptable results of metaheuristic-based approaches reported earlier suggest that when many records contain sensitive data, privacy loss can increase while clustering quality decreases. For large datasets with a substantial amount of sensitive information, purely metaheuristic approaches may fail to satisfy the accuracy requirements. By contrast, differential privacy, because of its solid mathematical foundation and the adjustable privacy budget ϵ , can be tuned more directly to meet quality targets. In this chapter, DP2SOM was introduced as a differential-privacy-based extension of SOM clustering for private data analysis.

The central idea behind DP2SOM is to identify groups of related items through a distance measure while perturbing that distance in a controlled and locally adaptive manner. The experimental results indicate that the method can achieve a compromise between privacy protection and clustering accuracy without substantially disrupting convergence. The experiments also show that the choice of ϵ remains critical, especially for large datasets. The comparative analysis suggests that DP2SOM is, within the scope of this thesis, the first proposed approach to preserve topology explicitly in differentially private SOM clustering by aligning noise injection with SOM's neighborhood-update dynamics. Unlike DP k-means methods that perturb global centroids and often lose structural fidelity, DP2SOM maintains convergence while balancing privacy and utility in the reported experiments. Local-DP-based models and standard toolkits tend to incur higher error or deteriorate more sharply under strict privacy budgets, whereas DP2SOM shows more stable accuracy across the evaluated datasets. By integrating local distance-based perturbation with per-centroid privacy budgeting, DP2SOM provides a robust approach to differentially private clustering.

Taken together, the results in this chapter support three main advantages of DP2SOM: (1) stable convergence with low quantization error, (2) preservation of SOM's intrinsic topology and neighborhood structure, and (3) applicability to high-dimensional datasets without severe loss of cluster coherence. On the basis of the

reported experiments, DP2SOM appears to exhibit a more favorable privacy–utility trade-off than the evaluated differentially private clustering baselines. The following chapter investigates these trade-offs further in the context of recommender systems.

Part IV
Quantitative Analysis of Privacy in
Recommender Systems

This section presents a case study building on the contributions of the preceding sections. In Chapter 7, the proposed personalized recommender framework is applied to a real-world dataset and evaluated to demonstrate the effectiveness of the developed models. Chapter 8 concludes the thesis by summarizing the main arguments, discussing open problems, and outlining directions for future work.

Chapter 7

Case Study: Measuring Privacy Levels in Recommender Systems

Chapter Overview. This chapter introduces a distance-based differential privacy (DP) framework, called Privacy-Preserving Collaborative Filtering (PPCF), for protecting user data in recommender systems while maintaining recommendation quality. The proposed model integrates Self-Organizing Map (SOM) clustering, collaborative filtering (CF), and differential privacy to support both privacy and personalization. A local adjustment strategy is applied to balance data availability and privacy, supported by aggregation, generalization, ϵ -budget management, and Laplace-noise injection to mitigate inference risks. Social-network signals are also considered to refine user-preference modeling. The system is evaluated on the MovieLens dataset, where the trade-off between privacy and accuracy is examined under different ϵ values. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to assess recommendation performance. The experimental results indicate that PPCF provides a favorable privacy–utility trade-off and remains stable under privacy-preserving perturbation. At the same time, the results confirm that relatively small variations in ϵ can affect clustering quality and recommendation accuracy. These findings highlight the importance of privacy–utility tuning for practical deployment. Overall, the PPCF approach provides a privacy-aware and scalable solution for collaborative filtering in environments involving sensitive user data.

7.1 Motivation

In this thesis, the concept of privacy in recommender systems and the main concerns regarding user privacy have been examined. In particular, Kang’s wording of the Information Infrastructure Task Force (IITF) [Kang, 1997] is adopted:

Information privacy is an individual’s claim to control the terms under which personal information—information identifiable to the individual—is acquired, disclosed, or used.

In e-business systems, user information is often required in order to provide the desired functionality and generate accurate recommendations. However, the use of large-scale data mining in e-business introduces privacy-related challenges that can directly affect users' willingness to adopt online services. For example, the analysis of user behavior and interaction data may lead to the disclosure of private information. The core challenge is therefore to protect informational privacy by eliminating or minimizing the exposure of personal data without significantly reducing the functionality of the system.

Existing privacy-preserving techniques face several limitations. Some incur substantial computational overhead, while others mainly defend against direct attacks and remain less effective against indirect inference attacks. In this context, the core problem can be summarized as follows:

In clustering-based algorithms, the density and distance relationships among clusters strongly affect model performance. In distance-based clustering, these relationships may encode associations among data points such that sensitive information can be inferred through analysis. Therefore, these relationships must be protected during data training.

To address this problem, hybrid methods based on neural networks and soft-computing techniques are explored in this thesis. The main objective is to maintain the accuracy and efficiency of the final output of data-mining tasks while achieving a high level of privacy. Because no single technique is able to satisfy all Privacy–Accuracy–Efficiency (PAE) requirements simultaneously, a hybrid model is adopted in the proposed algorithms. In particular, neural-network-based clustering using SOM is employed, and the improvement of the PAE factors is pursued through privacy-preserving mechanisms.

7.2 Contributions

The main contributions of this chapter are as follows:

1. A privacy-preserving movie-recommendation framework is developed by combining collaborative filtering with SOM-based clustering and differential privacy.
2. The DP2SOM mechanism introduced in the previous chapter is instantiated in a recommender-system setting in order to protect user-level relationships while maintaining recommendation quality.
3. An empirical analysis of the privacy–utility trade-off is conducted on the MovieLens dataset under varying privacy budgets.
4. The proposed method is evaluated using clustering-quality metrics (TE, QE) and recommendation-quality metrics (MAE, RMSE), together with statistical tests and comparative baselines.

7.3 Preliminaries

7.3.1 Recommender Systems

Recommender systems (RSs) are tools for interacting with large and complex information spaces. They provide a personalized view of those spaces by prioritizing items that are likely to be of interest to the user [Burke et al., 2011]. More generally, recommender systems can be regarded as a subset of information-filtering systems that aim to predict the rating or preference that a user would assign to a given item. They are widely used in applications such as e-commerce, movies, music, news, literature, and digital media platforms [Roy et al., 2022, Aggarwal et al., 2016]. Important components of recommender systems include the user-item matrix, the utility function, and personalization [Roy et al., 2022].

Several recent strategies have been proposed to improve user satisfaction, cross-domain transfer, and predictive performance in recommender systems. For example, Cross-Domain Recommendation via Correspondence-Aware Subspace (CDRec-CAS) improves recommendations in sparse settings by aligning latent subspaces between source and target domains [Gheewala et al., 2024]. CD-SPM (Cross-Domain Social Preference Modeling) incorporates social information together with cross-domain preferences [Anwar, et al., 2022]. Rec-CFSVD++ extends collaborative filtering and SVD++ by integrating contextual and implicit feedback [Anwar, et al., 2021]. Multi-relational Recommendation via Collaborative Relational Modeling (MRec-CRM) exploits multi-relational data sources, such as user-item interactions and auxiliary attributes, in order to improve robustness [?]. Together, these studies show that recommender systems have evolved toward more flexible, context-aware, and cross-domain models.

Some key aspects of recommender systems are the following:

- **Personalization:** Recommendations are tailored to individual users by considering their interests, behavior, and interactions.
- **User-Item Matrix:** Most recommender systems represent user interactions in a user-item matrix, where each user has rated or interacted with a subset of the available items.
- **Utility Function:** User utility measures the relevance or usefulness of an item to a user and is typically derived from explicit feedback or inferred preferences.
- **Recommendation Techniques:**
 - a. **Content-Based Filtering:** Recommends items based on item characteristics and the user's past preferences.
 - b. **Collaborative Filtering:** Recommends items preferred by users with similar tastes.
 - c. **Hybrid Approaches:** Combine content-based and collaborative strategies to benefit from both.
 - d. **Knowledge-Based Recommender Systems:** Make recommendations using explicit domain knowledge about how item attributes satisfy user preferences.

- **Challenges:**

- Cold Start:** Making recommendations for new users or new items with limited historical data.
- Sparsity:** Handling highly sparse user–item matrices.
- Scalability:** Managing large numbers of users and items efficiently.

Content-based recommender systems generate recommendations on the basis of similarity among items. Their basic assumption is that if a user likes a given item, then the user is also likely to like similar items. Collaborative filtering models, by contrast, assume that users tend to like items liked by other users with similar preferences [Roy et al., 2022, Bobadilla et al., 2009]. Figure 7.1 illustrates the distinction between collaborative and content-based models. In this chapter, collaborative filtering models are developed using machine-learning-based clustering in order to predict user ratings for unrated items.

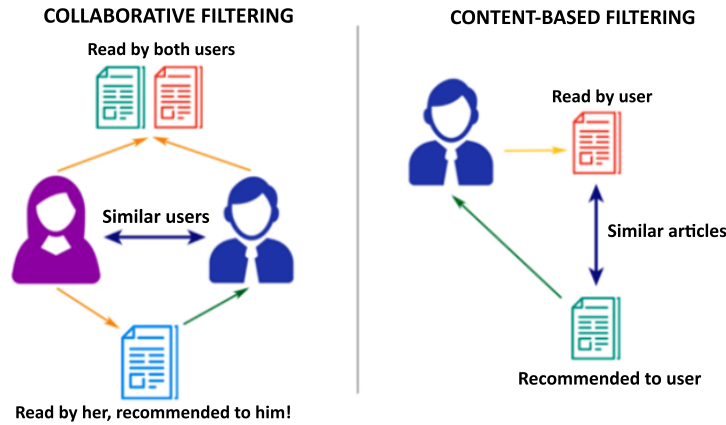


Fig. 7.1: Comparison of recommender-system approaches.

7.3.1.1 Collaborative Filtering

Collaborative filtering (CF) is a widely used approach in recommender systems that utilizes user interactions and preference data in order to generate predictions. Its core idea is that users who agreed in the past when evaluating certain items are likely to agree again in the future. Figure 7.2 shows the basic mechanism of collaborative filtering. CF can be divided into memory-based and model-based filtering [Burke et al., 2011, Papadakis et al., 2022].

- **Memory-based filtering:** This method generates recommendations by directly using user–item interaction data, such as ratings or purchases. It typically relies on real-time similarity computation.
 - a. **Item-based approach:** Recommends items similar to those that the target user liked previously.
 - b. **User-based approach:** Recommends items liked by users who are similar to the target user.
- **Model-based filtering:** This method uses user–item interaction data to train a predictive model through statistical or machine-learning techniques.
 - a. **Association techniques:** Decompose the user–item interaction matrix into lower-dimensional latent-factor matrices.
 - b. **Clustering approaches:** Organize users or items into groups based on similarity of interactions.
 - c. **Bayesian networks:** Use probabilistic models to represent uncertainty in preferences and features.
 - d. **Neural networks:** Capture nonlinear user–item interaction patterns through deep-learning architectures.

Table 7.1 summarizes the main advantages and disadvantages of these two collaborative-filtering categories.

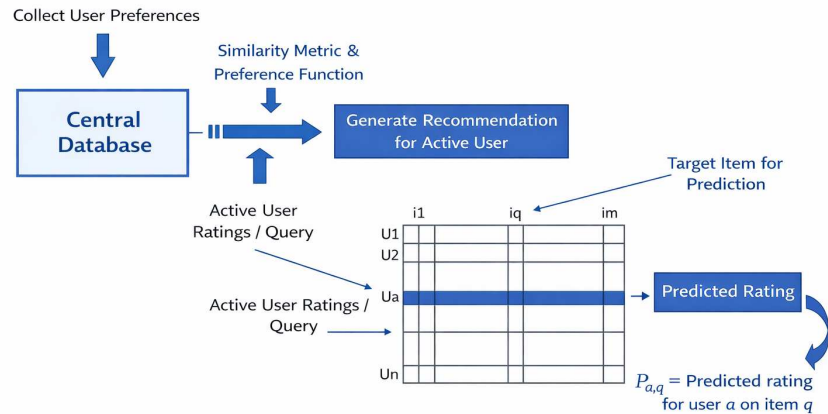


Fig. 7.2: Basic workflow of a user-based collaborative-filtering recommendation system. User preferences stored in a general database are processed using similarity metrics and preference functions to generate recommendations and predict the rating $P_{a,q}$ for a target item q for the active user a .

Table 7.1: Comparison between the main advantages and disadvantages of collaborative-filtering categories.

	Advantages	Disadvantages
Memory-based CF	Simplicity: Easy to implement and interpret. Real-Time Recommendations: Can generate recommendations without extensive training.	Scalability: Computational cost grows significantly with large datasets. Sparsity: Struggles when the user-item matrix is highly sparse. Cold Start Problem: Difficult to handle new users or items with little interaction data.
Model-based CF	Scalability: Can handle large datasets more efficiently than memory-based methods. Accuracy: Often captures latent patterns and provides more accurate recommendations. Flexibility: Can incorporate additional information into the model.	Complexity: Harder to implement and interpret than memory-based methods. Training Time: Requires more computational time and resources. Cold Start Problem: Still affected by new-user and new-item scenarios, although hybrid approaches can mitigate this issue.

Comparison	
Memory-based CF	Model-based CF
Data Dependency: Directly depends on observed interaction data. Real-Time: Can generate recommendations immediately. Scalability: More suitable for smaller datasets.	Data Dependency: Learns a generalized model from the observed interactions. Batch Processing: Typically produces recommendations after model training. Scalability: More suitable for larger datasets and more complex patterns.

7.3.1.2 Privacy Challenges in Recommender Systems

Although collecting personal information improves recommendation quality, it also creates privacy and security risks. Users may therefore become reluctant to provide the data needed for personalized recommendations. Building trust is thus essential in recommender systems [Su et al., 2009, Khusro, Shah et al., 2016]. Table 7.2 summarizes the main privacy challenges.

To address these challenges, recent work has increasingly focused on strong privacy-preserving mechanisms for recommender systems. Many of these methods use AI-based tools such as federated learning, data minimization, differential privacy, and robust security controls [Knijnenburg et al., 2017]. In this chapter, differential privacy is adopted as the primary mitigation mechanism [Himeur et al., 2022].

Table 7.2: Summary of the main privacy challenges in recommender systems.

Challenge	Description
Data Collection	Recommender systems often collect and store large amounts of user information, such as preferences, actions, and interactions. This information may be sensitive, and inadequate protection can lead to privacy breaches. Long-term retention further increases the risk of misuse or unauthorized access [Knijnenburg et al., 2017].
User Profiling	User interactions and preferences may reveal sensitive information, such as interests, routines, location, or financial status. Moreover, data sharing across services may enable profile linkage across platforms [Knijnenburg et al., 2017, Himeur et al., 2022].
Recommendation Generation	Recommendations themselves may reveal sensitive information. For instance, health-related recommendations may unintentionally expose user conditions or behaviors [Himeur et al., 2022].
Security of Algorithms	Attackers may manipulate recommender-system inputs or exploit outputs in order to infer preferences of other users. Access to model outputs may also enable attempts to reconstruct sensitive training data [Knijnenburg et al., 2017, Himeur et al., 2022].
Regulatory Compliance	Recommender systems must comply with regulations such as GDPR and CCPA, which impose strict requirements on the collection, storage, and processing of personal data [Knijnenburg et al., 2017, Himeur et al., 2022].

7.3.2 SOM Clustering

As discussed in Chapters 4, 5, and 6, Self-Organizing Maps (SOMs), also known as Kohonen maps, are an important class of neural-network models that:

- are trained using unsupervised learning,
- produce a low-dimensional, typically two-dimensional, representation of data,
- support clustering tasks in domains such as recommender systems,
- preserve topological structure and are therefore useful for pattern detection and clustering,
- have been applied in a wide range of domains, including project prioritization, representative-data selection, failure analysis, and seismic analysis for energy applications [Kohonen, 2001].

In each training iteration, SOM uses the Euclidean distance to find the best matching unit (BMU) for the current input vector. The Euclidean distance between each node's weight vector V and the current input instance W is minimized in order to determine the winning neuron. For further details on the practical implementation of SOM, see Chapter 5, Section 5.5, and Chapter 6.

SOMs provide a flexible and topologically aware clustering mechanism for high-dimensional data. When combined with differential privacy, they offer a suitable basis for protecting sensitive relational information during training.

7.3.3 Differential Privacy

Differential privacy (DP) is a distortion-based technique that adds random noise in order to prevent indirect inference about personal data. DP ensures that a single record does not substantially affect the outcome of an analysis. In this sense, an individual's privacy is protected because participation in the dataset does not significantly alter the results. A complete definition of DP, its parameters, and its implementation is given in Chapter 6.

Differential privacy mathematically guarantees that any observer of a differentially private analysis should make essentially the same inference about an individual's private information, regardless of whether that individual's data are included in the input. It therefore provides a provable privacy guarantee against a range of attacks, including differencing, linkage, and reconstruction attacks.

Formally, a randomized algorithm K satisfies ϵ -differential privacy if, for all datasets D and D' differing in at most one row, and for all events S ,

$$\Pr[K(D) \in S] \leq e^\epsilon \Pr[K(D') \in S]. \quad (7.1)$$

Decreasing ϵ strengthens privacy protection but usually reduces accuracy. Very small values of ϵ provide strong privacy at the cost of greater distortion, whereas larger values weaken privacy protection while preserving more utility.

7.4 Related Work

Recommender systems have been trained using a variety of deep-learning techniques, including Neural Collaborative Filtering [He et al., 2017], Deep Collaborative Neural Networks [Catherine et al., 2017, Kiran et al., 2020], and Restricted Boltzmann Machines [Salakhutdinov et al., 2007]. Deep learning has also been explored as a way to address the cold-start problem, for example through convolutional neural networks for image features [Van den Oord et al., 2013] and LSTM-based recurrent neural networks for audio features [Rama et al., 2019].

Differential privacy is a widely used technique in recommender systems, where it introduces randomness into the training or output-generation process in order to protect the privacy of groups of participants [Blum et al., 2005, Dinur et al., 2003, Dwork et al., 2006, Dwork et al., 2004, Friedman and Schuster, 2010]. However, applying DP in recommender systems introduces several challenges. Previous studies have shown that DP can reduce recommendation accuracy and may even increase

popularity bias [Zhu et al., 2016]. This illustrates the difficulty of balancing privacy and recommendation quality.

To address these issues, different implementations of DP in recommender systems have been explored [Müllner et al., 2023]. Some methods protect user representations directly, for example in point-of-interest recommendation [Lu et al., 2019]. Other approaches apply DP to model updates in federated-learning settings, where users send noisy gradients to a central server [Kalra et al., 2023, Baek et al., 2021]. Zhu et al. [Zhu et al., 2016] proposed two differentially private recommender algorithms using carefully designed quality functions and exponential-mechanism-based sampling. Liu et al. [Liu et al., 2017] introduced a hybrid approach combining differential privacy with randomized perturbation. Yin et al. [Yin et al., 2020] proposed a privacy-preserving collaborative-filtering framework based on time-aware factors and DP protection.

In summary, applying DP in recommender systems still requires careful management of the privacy–accuracy trade-off. The aim of this chapter is therefore to develop a method that protects user privacy while maintaining recommendation quality.

7.5 Model Setting

7.5.1 General Steps of Collaborative Filtering with Privacy Considerations

When collaborative filtering is adopted in privacy-sensitive environments, several standard steps must be extended by privacy-preserving mechanisms [Batmaz et al., 2017]. These steps are summarized below.

Data collection:

Collaborative filtering begins with the collection of user information, such as ratings, purchases, and browsing history. This collection process must be designed with privacy in mind, including user consent and limitation of sensitive-data collection.

Data preprocessing:

Before being used for CF, the data must be preprocessed. This includes normalization, handling missing values, and privacy-aware transformation. In this study, z-score normalization and mean-centering are used to account for different user rating scales. Missing values are either left as missing or treated in a controlled way. Differential privacy is then applied to protect user-level information.

Privacy-preserving techniques:

This study employs a differential-privacy mechanism to add noise to the data or to intermediate calculations, thereby reducing the effect of any single user's presence or absence on the output. Figure 7.3 illustrates the use of differential privacy in local and global settings. Other privacy-preserving techniques, such as federated learning and secure multiparty computation (SMPC), are beyond the scope of this chapter [Ge et al., 2023].

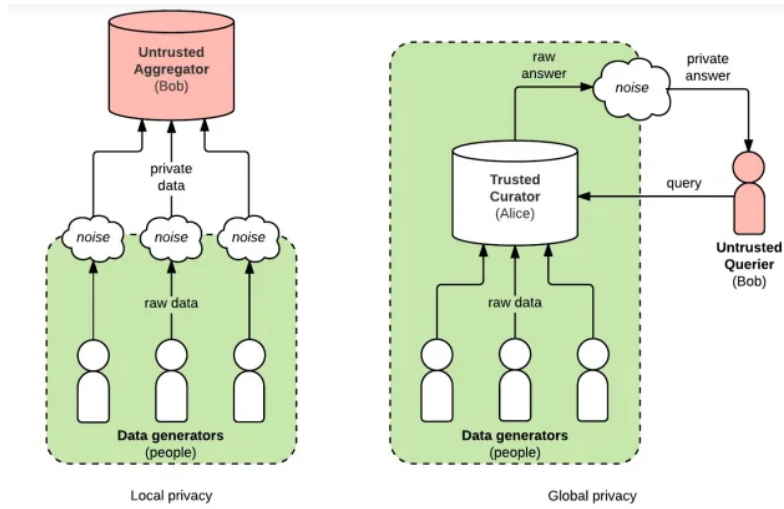


Fig. 7.3: Framework of differential privacy in local and global settings.

Similarity calculation:

Collaborative filtering often relies on similarity computation between users or items. Privacy-preserving methods can be used to perform these computations without directly exposing individual data. In this chapter, similarity is based on Euclidean distance together with privacy-preserving perturbation.

Prediction generation:

In the prediction step, the target user's preferences or ratings are estimated using information from similar users. Differential privacy can also be applied during this

step in order to reduce the risk that specific user information is inferred from the generated scores.

Recommendation generation:

Items are ranked according to predicted scores, and the top- N items are recommended to the target user.

Evaluation and refinement:

The CF system is evaluated not only in terms of recommendation accuracy but also in terms of privacy preservation. Relevant metrics include MAE, RMSE, precision, recall, F1-score, and privacy loss (ϵ). The model is then refined by tuning parameters, testing different similarity measures, and improving privacy-preserving mechanisms.

7.5.2 Personalized Social Recommenders

Personalized social recommenders incorporate social-network information in order to improve collaborative filtering [Qian et al., 2013]. The main assumption is that socially connected users are more likely to have similar preferences [Meng et al., 2018].

In this work, user preferences are first inferred from user–item interactions through collaborative filtering. Social-network information is then considered as additional contextual input. Differential privacy is used to protect user data, including both preferences and social links. Table 7.3 summarizes these components in the context of a movie-recommendation system on a social platform.

Table 7.3: Personalized social recommendation in collaborative filtering on a movie-recommendation platform.

Collaborative Filtering	Analyzes viewing habits of users and identifies patterns among similar users.
Social Network Integration	Enhances recommendations by incorporating movies that a user’s friends have watched or liked.
Differential Privacy	Ensures that no single user’s viewing history or social connections can be inferred from the recommendations, even in the presence of auxiliary information.

7.5.3 Assumptions and Attacker Model

An threat model specifies the assumptions, capabilities, and goals of an adversary attempting to compromise user privacy.

Assumption: how does DP protect recommendations?

The released results should be statistically indistinguishable from the results that would have been obtained if a single record had been removed from the original dataset.

Threat models:

Linkage attack: The attacker knows some external social connections and user–item interactions. The goal is to infer additional interactions or connections by cross-referencing external information with recommendation outputs.

Membership inference attack: The attacker observes recommendation outputs and has some auxiliary user-preference data. The goal is to determine whether a specific user contributed to the training data.

Model inversion attack: The attacker can query the recommender system and analyze its responses in order to infer sensitive user preferences.

Background knowledge attack: The attacker exploits prior knowledge about users or the system to improve the inference of confidential information.

7.5.4 Privacy Model

In this study, differential privacy is used to address the threat models described above. The main defense mechanisms are:

1. **Noise addition:** Random noise is added to data or outputs to mask individual contributions and reduce inferability.
2. **Privacy-budget management:** The cumulative privacy loss is controlled so that repeated queries do not exhaust privacy protection.
3. **Aggregation and generalization:** User data are aggregated or generalized to reduce the exposure of unique entries.

7.6 Proposed Conceptual Framework

As discussed in earlier chapters, the methods pursued in this thesis adopt a hybrid structure that combines neural-network-based clustering with soft-computing techniques. Figure 7.4 summarizes the conceptual design used in this chapter.

The main objective is to maintain a high level of privacy while preserving the accuracy and efficiency of the final data-mining outputs. Because no single method satisfies all PAE factors simultaneously, a hybrid model is adopted. SOM-based clustering is used as the main learning mechanism, and privacy gaps are addressed by integrating differential privacy and the earlier SOM-based extensions.

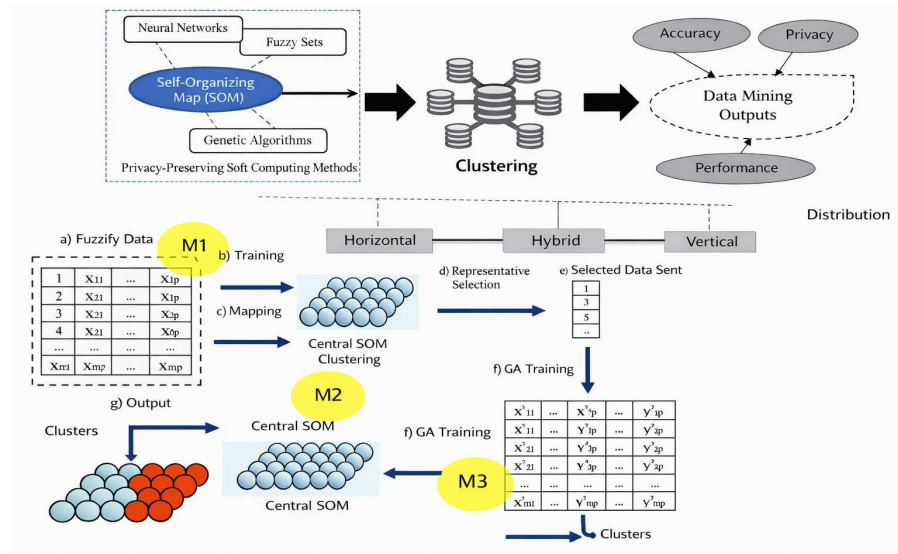


Fig. 7.4: Proposed framework showing the selected data-mining task, the soft-computing components, and the expected outcomes. Privacy-preserving SOM clustering, fuzzy sets, and genetic algorithms are shown as related elements in the lower part of the framework.

7.7 Case-Study Objectives

The central question addressed in this chapter is the following: *How can users contribute data to collaborative filtering without compromising privacy?* To answer this question, differentially private movie recommendations are generated through

model-based collaborative filtering built on SOM clustering. In this context, the system design is guided by the following considerations:

- **Preference correlations:** If two users exhibit similar tastes on known items, this similarity can be used to improve predictions on unrated items.
- **Social relevance:** Socially related users may provide additional signal for recommendation generation.
- **Group awareness:** Recommendation tasks may involve multiple users rather than a single individual.
- **Interpretability and reliability:** Recommendation outputs should be supported by an interpretable clustering structure and should preserve useful accuracy.

7.8 Experimental Study

This section evaluates the proposed approach and compares it with alternative solutions.

7.8.1 Dataset

The MovieLens¹ dataset, which is widely used for movie recommendation and collaborative-filtering research, is employed in this study. MovieLens was developed by the GroupLens research group at the University of Minnesota. The dataset contains tuples of the form {user, movie, rating, timestamp}. Users create a rating matrix based on the movies they watch, and the system recommends movies that they have not yet rated but may find interesting. Ratings are typically represented as integers from 1 (lowest) to 5 (highest). Table 7.4 summarizes the main features of the dataset.

7.8.2 Experimental Setup

The experiments were conducted on a Linux-based virtual machine equipped with four vCPUs running on an Intel Xeon E5-2650 v4 processor and 4 GB of RAM.

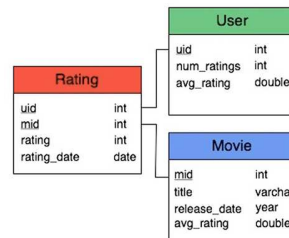
The MovieLens data were first prepared for training. The dataset includes ratings, user information, and movie metadata. Table 7.4 and Figure 7.5 summarize the structure of the data. Users generate a rating matrix from the movies they have watched (Figure 7.5b), and the system produces recommendations for unseen movies.

The overall experimental process is shown in Figure 7.6. During preprocessing, missing values are identified and treated appropriately, for example by removing incomplete records or by replacing missing entries with predefined defaults. The

¹ <https://grouplens.org/datasets/movielens/>

Table 7.4: Main features of the MovieLens dataset.

Rating data	Movie ratings given by users. Ratings usually range from 1 to 5, and each rating is linked to a user, a movie, and a timestamp.
User data	Some versions contain anonymized user information, such as age, gender, and occupation.
Movie data	Includes metadata such as movie title, genre, and release date. Some versions also contain tags and links to external movie databases.
Versions	
MovieLens 100K	Contains 100,000 ratings from 943 users on 1,682 movies.
MovieLens 1M	Contains 1,000,000 ratings from 6,040 users on 3,952 movies.
MovieLens 10M	Contains 10,000,000 ratings and 100,000 tags from 72,000 users on 10,000 movies.
MovieLens 25M	Contains 25 million ratings and one million tag applications from 162,000 users on 62,000 movies. It also includes tag-genome data with 15 million relevance scores across 1,129 tags.



(a) User, movie, and rating data



	SPIDERMAN	AVENGERS	COCO	MONSTERS UNIVERSITY
User 1	4	3	3	4
User 2	3	5	4	2
User 3	5	4	4	3
User 4	2	4	3	3
User 5	2	5	5	1

(b) User-item rating matrix

Fig. 7.5: Illustration of the dataset structure.

ratings are then normalized, categorical features are encoded, and the different tables are merged into a unified feature set. After SOM parameters are selected, differential privacy is applied in order to perturb either user data or model outputs. Finally, users are mapped to SOM nodes and recommendations are produced.

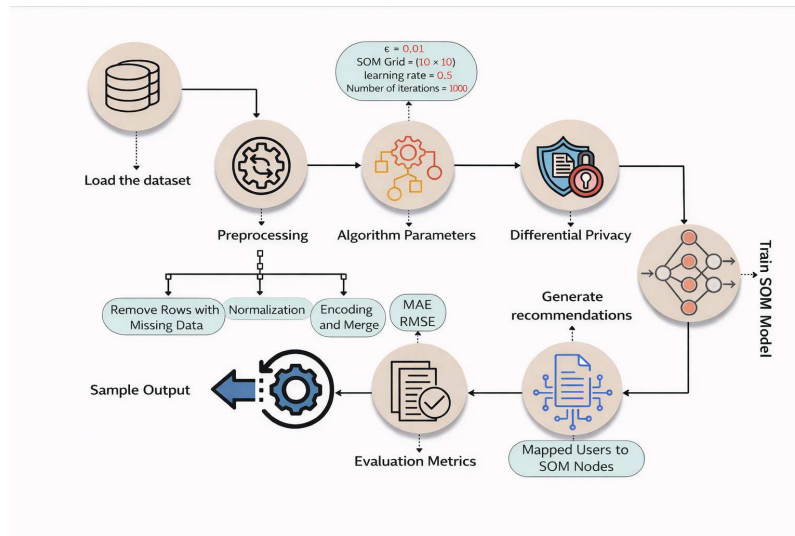


Fig. 7.6: Overall experimental procedure. The preprocessing phase includes handling missing values, normalization, encoding of categorical variables, and dataset integration. Differential privacy is then applied before SOM-based user mapping and recommendation generation.

Load the dataset:

In the first step, the ratings, user data, and movie data are loaded into data frames. The following code fragment illustrates this step.

```
import pandas as pd
ratings = pd.read_csv('path/to/ratings.csv')
users = pd.read_csv('path/to/users.csv')
movies = pd.read_csv('path/to/movies.csv')
```

Data preprocessing:

Missing values are identified and treated, for example by removing incomplete rows.

```
ratings.dropna(inplace=True)
```

The ratings are then normalized.

```
from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()
ratings['normalized_rating'] = scaler.fit_transform(ratings[['rating']])
```

Categorical attributes, such as user demographics, are encoded before use.

```
from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()
users['encoded_gender'] = le.fit_transform(users['gender'])
```

Finally, the ratings, user, and movie tables are merged into a single dataset.

```
merged_data = ratings.merge(users, on='user_id').merge(movies,
on='movie_id')
```

An illustrative example of the final integrated dataset is shown in Figure 7.7.

Algorithm parameters:

Before applying the model, the main privacy and SOM parameters are defined. In particular, the privacy budget ϵ must be set.

```
epsilon = 0.01
```

The SOM grid size, learning rate, and number of iterations are then specified.

```
som_dimensions = (10, 10)
learning_rate = 0.5
num_iterations = 1000
```

Apply differential privacy:

Next, differential privacy is used to perturb the ratings through a Laplace mechanism.

	userId	movieId	rating	Adventure	Comedy	Action	Mystery	Crime	Thriller	Drama	...	Horror	Documentary	Sci-fi	Fantasy	Film-Noir	Western	Musical	Romance	(no genres listed)	War
57002	365.0	104879	2.0	0	0	0	0	1	0	1	1	...	0	0	0	0	0	0	0	0	0
23696	483.0	3317	5.0	0	1	0	0	0	0	0	1	...	0	0	0	0	0	0	0	0	0
88700	108.0	2585	4.0	0	0	0	0	0	0	0	1	...	0	0	0	0	0	0	1	0	0
98622	448.0	82093	1.5	0	0	0	0	0	1	0	0	...	0	0	0	0	0	0	1	0	0
35223	256.0	4018	4.0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0
77843	105.0	5151	3.0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0
72626	480.0	6993	5.0	0	1	0	0	0	0	0	1	...	0	0	0	0	0	0	1	0	0
11001	186.0	2048	3.0	0	0	1	0	1	0	0	0	...	0	0	0	0	0	0	0	0	0
34847	378.0	3114	4.5	1	1	0	0	0	0	0	0	...	0	0	0	1	0	0	0	0	0
6165	303.0	1080	1.0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0

Fig. 7.7: Illustrative view of the final integrated dataset.

```
import numpy as np

def add_noise(data, epsilon):
    noise = np.random.laplace(0, 1/epsilon, data.shape)
    return data + noise

noisy_ratings = add_noise(ratings['normalized_rating'].values, epsilon)
```

Train the SOM model:

Features are extracted from the merged dataset in order to train the SOM.

```
features = merged_data[['normalized_rating', 'encoded_gender', 'age',
                        'genre_encoded']].values
```

A SOM implementation such as MiniSom can then be used for training.

```
from minisom import MiniSom

som = MiniSom(som_dimensions[0], som_dimensions[1], features.shape[1],
              learning_rate=learning_rate)
som.train_random(features, num_iterations)
```

Generate recommendations:

Users are then mapped to SOM nodes.

```
user_nodes = {user_id: som.winner(features[i]) for i, user_id in enumerate(
merged_data['user_id'])}
```

Movies preferred by users assigned to the same SOM node can then be used to generate recommendations.

```
recommendations = {user_id: get_recommendations(user_nodes[user_id],
movie_data) for user_id in user_nodes}
```

Evaluate the model:

The data are divided into training and testing subsets, and the recommender system is evaluated using metrics such as MAE and RMSE.

```
from sklearn.model_selection import train_test_split

train_data, test_data = train_test_split(merged_data, test_size=0.2, random_state=42)

from sklearn.metrics import mean_absolute_error, mean_squared_error

predictions = ... # predictions generated by the recommender model
mae = mean_absolute_error(test_data['normalized_rating'], predictions)
rmse = mean_squared_error(test_data['normalized_rating'], predictions, squared=False)
print(f'MAE: {mae}, RMSE: {rmse}')
```

Illustrative output:

To illustrate the recommendation process, a sample run was performed using 15 randomly selected movies as input. For this run, $\epsilon = 0.01$, and the movie *Ice Age* was included in the test selection (Figure 7.8).

For a user with ID “5218”, the system recommends movies such as *Rio 2*, *Daffy Duck’s Quackbusters*, and *The Emperor’s New Groove 2: Kronk’s New Groove* with relatively high predicted ratings (Figure 7.9). The outputs are privacy-protected, which reduces the ability of an adversary to identify which users originally rated a given movie. The consistency of the outputs supports the practical validity of the proposed method.

7.9 Results and Discussion**7.9.1 Analysis of the Privacy–Accuracy Trade-Off**

As explained earlier, several parameters influence both the accuracy and the performance of the model. Table 7.5 presents the evaluation results of the recommender system under the selected parameter configurations.

```

Select The Movies Id you would like to watch:
464--->Hard Target (1993)
2857--->Yellow Submarine (1968)
5218--->Ice Age (2002)
68157--->Inglourious Basterds (2009)
1124--->On Golden Pond (1981)
4033--->Thirteen Days (2000)
4961--->Pornstar: The Legend of Ron Jeremy (2001)
86000--->Boy (2010)
70301--->Obsessed (2009)
141866--->Green Room (2015)
4616--->Lean on Me (1989)
7348--->Spartan (2004)
3929--->Bank Dick, The (1940)
3481--->High Fidelity (2000)
133716--->Bootleggers (1961)
-----

```

Fig. 7.8: Fifteen randomly selected movies used for testing the recommender system.

```

7685 Rio 2 (2014)
6934 Daffy Duck's Quackbusters (1988)
6661 Emperor's New Groove 2: Kronk's New Groove, Th...
3601 Hey Arnold! The Movie (2002)
8162 Kung Fu Panda: Secrets of the Masters (2011)
4899 Bebe's Kids (1992)
4715 Garfield: The Movie (2004)
4776 Claymation Christmas Celebration, A (1987)
7572 All Dogs Christmas Carol, An (1998)
7802 Mr Hublot (2013)
Name: title, dtype: object
-----
Do you like these movies(y/n)

```

Fig. 7.9: Illustrative list of high-rated movie recommendations produced for user ID “5218”.

7.9.1.1 Clustering Quality

Two measures are used to assess SOM clustering quality:

Quantization Error (QE): the average distance between each data vector and its BMU:

$$QE = \frac{1}{N} \sum_{k=1}^N \|x_k - m_k\| \quad (7.2)$$

where N is the number of data vectors and m_k is the best matching prototype of data vector x_k . Smaller QE values indicate better representation quality.

Topographic Error (TE): the proportion of data points for which the first and second BMUs are not adjacent:

$$TE = \frac{1}{N} \sum_{k=1}^N u(x_k) \quad (7.3)$$

where $u(x_k) = 1$ if the first and second BMUs of x_k are not adjacent, and $u(x_k) = 0$ otherwise. Smaller TE values indicate better topology preservation.

QE measures how well the SOM represents the data distribution. TE measures how well neighborhood structure is preserved. Together, these metrics provide an indication of SOM suitability for recommendation.

7.9.1.2 Recommendation Quality

Recommendation accuracy is evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE):

MAE:

$$MAE = \frac{1}{|\hat{R}|} \sum_{\hat{r}_{ui} \in \hat{R}} |r_{ui} - \hat{r}_{ui}| \quad (7.4)$$

RMSE:

$$RMSE = \sqrt{\frac{1}{|\hat{R}|} \sum_{\hat{r}_{ui} \in \hat{R}} (r_{ui} - \hat{r}_{ui})^2} \quad (7.5)$$

MAE reflects the average absolute magnitude of prediction errors and is less sensitive to outliers. RMSE penalizes larger errors more strongly and is therefore more sensitive to dispersion in the predictions [Chai et al., 2014]. Lower values of both metrics indicate better recommendation quality.

7.9.1.3 Additional Metrics

Empty clusters: The number of SOM nodes to which no data points are assigned. A large number of empty clusters suggests that the SOM grid may be too large relative to the dataset.

Sigma (σ): The neighborhood radius used in SOM training. Larger values produce smoother maps but can increase training time and reduce local detail. Smaller values yield finer maps but may produce noisier structures.

Learning rate (α): Controls the magnitude of SOM weight updates. Larger values can accelerate convergence but may reduce stability. Smaller values support stable convergence but may require more iterations.

Table 7.5: Evaluation performance of the recommender system under the selected parameter settings.

Setting	Epsilon	Rows	Columns	Training	Neighbourhood function	Learning α	rate	Sigma	Iterations	TE	QE	MAE	RMSE	Empty clusters
SOM	–	8	8	80%	Gaussian	0.6		1	1000	0.698	0.849	0.241	0.407	0
SOM	–	10	10	80%	Gaussian	0.5		1	1000	0.550	0.763	0.218	0.357	0
SOM	–	30	30	80%	Gaussian	0.5		1	1000	0.310	0.387	1.361	2.089	167
DP2SOM	0.001	10	10	80%	Gaussian	0.5		1	1000	0.498	0.783	0.262	0.352	0
DP2SOM	0.01	10	10	80%	Gaussian	0.5		1	1000	0.510	0.787	0.259	0.432	0
DP2SOM	0.1	10	10	80%	Gaussian	0.5		1	1000	0.596	0.801	0.350	0.483	0
DP2SOM	0.2	10	10	80%	Gaussian	0.5		1	1000	0.520	0.770	0.241	0.332	0
DP2SOM	0.5	10	10	80%	Gaussian	0.5		1	1000	0.468	0.767	0.235	0.313	0
DP2SOM	1	10	10	80%	Gaussian	0.5		1	1000	0.578	0.791	0.250	0.437	0
DP2SOM	2	10	10	80%	Gaussian	0.5		1	1000	0.561	0.805	0.382	0.513	0
DP2SOM	10	10	10	80%	Gaussian	0.5		1	1000	0.611	0.770	0.373	0.532	0

Table 7.6: Interpretation of the main evaluation metrics.

Condition	Interpretation
Lower TE and QE	Indicate better SOM performance and more accurate mapping of data points, which generally supports better recommendations.
Higher TE and QE	Indicate weaker SOM training performance, possibly caused by inappropriate grid size, insufficient training, or poor parameter selection.
Lower MAE and RMSE	Indicate better recommendation performance and more accurate rating prediction.
Higher MAE and RMSE	Indicate more prediction error and suggest that the model may require re-tuning or additional features.
Few or no empty clusters	Suggest that the SOM grid is appropriately utilized.
Many empty clusters	Suggest that the SOM grid may be too large relative to the available data.
Larger sigma	Produces smoother and more general maps; useful in early training stages.
Smaller sigma	Produces more detailed maps, although they may become noisier; useful for fine-tuning.
Higher learning rate α	Can accelerate convergence but may reduce stability.
Lower learning rate α	Improves stability but may require more iterations for convergence.

7.9.1.4 Interpretation of Outcomes

Table 7.6 summarizes the general interpretation of the main metrics used in this chapter.

As shown in Table 7.5, the highlighted row represents the most balanced non-private SOM configuration among the tested settings. It combines relatively low TE and QE with good MAE and RMSE values and no empty clusters. This row is therefore used as a useful reference point for comparison with DP2SOM.

7.9.2 Test Dataset for Prediction

To evaluate recommendation quality, the MovieLens dataset must be split into training and testing subsets. The training set is used to fit the model, and the test set is used to generate predictions that are then evaluated using the metrics defined above. In this experiment, single-genre movies with ratings ranging from 0.5 to 5.0 were extracted and used as a test dataset. The following settings were used:

1. Actual rate d = number of movies in a cluster
2. Prediction rate d_p = maximum number of movies assigned to the same cluster

The following code fragment illustrates the construction of the test queries.

```

all_genres = list(ALLData.columns.drop(['userId', 'movieId', 'rating',
'title']).to_numpy())
testDbAllrate = []
for r in np.arange(0.5, 5.5, 0.5):
    initQ = "'Adventure'==0 and 'Comedy'==0 and 'Action'==0 and " +
        "'Mystery'==0 and 'Crime'==0 and 'Thriller'==0 and 'Drama'==0
and " +
        "'Animation'==0 and 'Children'==0 and 'Horror'==0 and " +
        "'Documentary'==0 and 'Sci-Fi'==0 and 'Fantasy'==0 and " +
        "'Film-Noir'==0 and 'Western'==0 and 'Musical'==0 and " +
        "'Romance'==0 and '(No genres listed)'==0 and 'War'==0 and
'rating'=={0}".format(r)
    qs = []
    for g in all_genres:
        qs.append(initQ.replace("{} {}".format(g),
        "{1} {}".format(g)))
    testDb = []
    for q in qs:
        testDb.append(ALLData.query(q).drop_duplicates(subset=['movieId']
        testDbAllrate.append(testDb)

```

Figure 7.10 shows an example query for all movies with rating 1. This query is useful for identifying poorly rated movies, analyzing user preferences, and filtering low-rated items from recommendations.

These results can be useful when improving the user experience by reducing the frequency of poorly rated recommendations.

7.9.3 How Does DP Protect Recommendations?

The key privacy assumption stated in Section 7.5.3 is repeated here:

“The released results should be statistically indistinguishable from similar results that would have been obtained if a single record had been removed from the original dataset.”

To interpret the effect of DP on the recommender system, the following analyses were performed.

```

Comedy ---> d= 108 - dp= 103
Thriller ---> d= 7 - dp= 6
Drama ---> d= 36 - dp= 25
Horror ---> d= 19 - dp= 17
Comedy ---> d= 215 - dp= 199
Action ---> d= 10 - dp= 5
Thriller ---> d= 15 - dp= 10
Drama ---> d= 82 - dp= 68
Horror ---> d= 39 - dp= 35
Documentary ---> d= 10 - dp= 7
Comedy ---> d= 193 - dp= 177
Action ---> d= 6 - dp= 3
Thriller ---> d= 11 - dp= 5
Drama ---> d= 48 - dp= 32
Adventure ---> d= 7 - dp= 4
Comedy ---> d= 385 - dp= 364
Action ---> d= 19 - dp= 9
Thriller ---> d= 28 - dp= 16
Drama ---> d= 189 - dp= 147
Horror ---> d= 70 - dp= 63
Documentary ---> d= 18 - dp= 15
Sci-Fi ---> d= 8 - dp= 5
Comedy ---> d= 284 - dp= 269
Action ---> d= 12 - dp= 5

```

Fig. 7.10: Query results for movies with rating 1. This query identifies movies that received the lowest user ratings in the dataset.

7.9.3.1 Comparison of Metrics With and Without Differential Privacy

The model was trained for 1000 iterations with learning rate $\alpha = 0.5$. This configuration was chosen in order to maintain a practical balance between noise, privacy, and recommendation quality. A Gaussian neighborhood function was used for SOM training, while the privacy mechanism itself perturbs the data through DP noise. Table 7.7 compares TE, QE, MAE, RMSE, and the number of empty clusters in models trained with and without differential privacy.

Table 7.7: Comparison of evaluation metrics with and without differential privacy.

Setting	TE	QE	MAE	RMSE	Empty clusters
SOM	0.698	0.849	0.241	0.407	0
SOM	0.555	0.763	0.218	0.357	0
SOM	0.315	0.387	1.361	2.089	167
DP2SOM	0.498	0.783	0.262	0.352	0
DP2SOM	0.510	0.787	0.259	0.432	0
DP2SOM	0.596	0.801	0.350	0.483	0
DP2SOM	0.520	0.770	0.241	0.332	0
DP2SOM	0.468	0.767	0.235	0.313	0
DP2SOM	0.578	0.791	0.250	0.437	0
DP2SOM	0.561	0.805	0.382	0.513	0
DP2SOM	0.611	0.770	0.373	0.532	0

As shown in Table 7.7, the addition of privacy noise can increase TE and QE, reflecting the expected trade-off between privacy and accuracy. MAE and RMSE may also change because the perturbed data affect recommendation precision. The main objective is therefore to identify a practical balance between privacy protection and predictive utility.

7.9.3.2 Analysis of Epsilon (ϵ) Values

Smaller values of ϵ correspond to stronger privacy guarantees. For example, $\epsilon = 1$ provides a moderate level of privacy, whereas smaller values such as $\epsilon = 0.001$ offer stronger privacy at the cost of increased distortion. Table 7.8 summarizes the observed behavior under different ϵ values.

Table 7.8: Analysis of epsilon values with fixed sigma.

Setting	Epsilon	Sigma	TE	QE	MAE	RMSE
DP2SOM	0.001	1	0.498	0.783	0.262	0.352
DP2SOM	0.01	1	0.510	0.787	0.259	0.432
DP2SOM	0.1	1	0.596	0.801	0.350	0.483
DP2SOM	0.2	1	0.520	0.770	0.241	0.332
DP2SOM	0.5	1	0.468	0.767	0.235	0.313
DP2SOM	1	1	0.578	0.791	0.250	0.437
DP2SOM	2	1	0.561	0.805	0.382	0.513
DP2SOM	10	1	0.611	0.770	0.373	0.532

As shown in Table 7.8, the relationship between ϵ and utility is not perfectly monotonic in this experimental setting. This suggests that an appropriate value of ϵ must be selected empirically in order to obtain a practical balance between privacy and utility.

7.9.3.3 Impact of Sigma (σ)

In this study, $\sigma = 1$ corresponds to noise drawn from a standard normal distribution with mean 0 and standard deviation 1. A standard deviation of 1 implies that approximately 68% of the noise values lie within ± 1 , 95% within ± 2 , and 99.7% within ± 3 . This gives an intuitive sense of the variability introduced during perturbation.

In general, increasing σ increases the spread of the noise and can therefore affect TE, QE, MAE, and RMSE. In a differentially private recommendation setting, $\sigma = 1$ provides a moderate level of perturbation and is used here as a fixed experimental setting for comparing the influence of ϵ .

7.9.3.4 Privacy-Budget Management

Privacy-budget management refers to controlling and monitoring the cumulative privacy loss of the system over time. In differential privacy, this loss is quantified through ϵ and δ^2 . When multiple queries or iterations are performed, the total privacy loss accumulates.

A simple implementation for tracking privacy budget is shown below.

```
class PrivacyBudget:
    def __init__(self, epsilon_total, delta_total):
        self.epsilon_total = epsilon_total
        self.delta_total = delta_total
        self.epsilon_used = 0
        self.delta_used = 0

    def update_budget(self, epsilon, delta):
        self.epsilon_used += epsilon
        self.delta_used += delta
        if self.epsilon_used > self.epsilon_total or self.delta_used > self.delta_total:
            raise ValueError("Privacy budget exceeded")

    def remaining_budget(self):
        return self.epsilon_total - self.epsilon_used, self.delta_total - self.delta_used
```

This mechanism ensures that cumulative privacy loss remains within predefined bounds.

7.9.3.5 Empirical Validation

Experiments were conducted under different hyperparameter settings (α , ϵ , and δ), and their impact on performance was measured using the metrics reported in Tables 7.6, 7.7, and 7.8. Because of the limited sample sizes, non-parametric tests were used, including the Mann–Whitney U test and the Kruskal–Wallis H test.

² δ represents the probability that the privacy guarantee is violated. It is usually set to a very small value.

Comparing MAE, RMSE, TE, and QE between two conditions: with DP and without DP

The Mann–Whitney U test was applied to compare each metric under the two conditions. The null and alternative hypotheses were:

- **Null hypothesis (H_0):** There is no significant difference in TE/QE/MAE/RMSE between the two conditions.
- **Alternative hypothesis (H_1):** There is a significant difference in TE/QE/MAE/RMSE between the two conditions.

The results are shown in Table 7.9.

Table 7.9: Comparison of TE, QE, MAE, and RMSE between private and non-private conditions.

	U statistic	p-value	Mean $\pm$ SD (No-DP)	Mean $\pm$ SD (DP)
TE	12.0	> 0.99	0.52 $\pm$ 0.19	0.54 $\pm$ 0.05
QE	8.0	0.47	0.66 $\pm$ 0.24	0.78 $\pm$ 0.01
MAE	9.5	0.68	0.60 $\pm$ 0.65	0.29 $\pm$ 0.06
RMSE	10.0	0.77	0.95 $\pm$ 0.98	0.42 $\pm$ 0.08

The reported Mann–Whitney p-values are two-sided and are shown in rounded form; for very small samples, exact p-values may appear as 1.00 when the observed statistic lies near the center of the null distribution. The p-values are all greater than 0.05, which indicates that no statistically significant difference was detected between the private and non-private conditions in this sample. Therefore, the observed differences may be attributable to random variation.

Comparing MAE, RMSE, TE, and QE across different epsilon values

The Kruskal–Wallis H test was used to test whether the metric values differed significantly across different ϵ values.

- **Null hypothesis (H_0):** There is no significant difference in MAE/RMSE/TE/QE across different ϵ values.
- **Alternative hypothesis (H_1):** There is a significant difference in MAE/RMSE/TE/QE across different ϵ values.

The results are shown in Table 7.10.

Again, all p-values are much larger than 0.05, indicating that no statistically significant differences were detected across the tested ϵ values in the present sample. These results suggest that, within the limitations of the current dataset and sample size, the introduction of differential privacy does not produce a statistically significant degradation in recommendation quality.

Table 7.10: Comparison of TE, QE, MAE, and RMSE across different epsilon values.

	H statistic	p-value
TE	2.84	0.94
QE	2.54	0.95
MAE	1.85	0.98
RMSE	2.49	0.96

Table 7.11: How differential privacy responds to the main attack types considered in this chapter.

Attack Type	Protection Mechanism	Explanation
Background knowledge attack	DP guarantees	The impact of any single data point is bounded. Even with substantial prior knowledge, the attacker cannot infer new user-specific information with high confidence.
	Randomization	Randomness is introduced into the SOM training and query outputs, reducing determinism and weakening inference based on external knowledge.
Inference attack	DP-SOM training	Noise is introduced during SOM training, reducing overfitting to individual data points.
	Noise in mapping	Perturbation during the mapping process reduces the precision with which sensitive attributes can be inferred.
Linkage attack	Noise addition	Perturbation makes linkage to external sources less reliable by reducing exact matching of individual-level patterns.
	Aggregate queries	Noisy aggregate outputs reveal less specific information about individual users.
Model inversion attack	DP-SOM algorithm	Differential privacy reduces the sensitivity of the model to individual training records, making reconstruction more difficult.
	Noise addition	Additional perturbation in outputs or intermediate stages makes inversion-based reconstruction less reliable.

7.9.4 How Does DP Protect Against the Threat Model?

The threat model describes the capabilities and goals of an adversary attempting to extract private information from the recommender system. Based on the design of DP2SOM, Table 7.11 summarizes how differential privacy addresses the attacks considered in this chapter as follows:

1. **Linkage attacks are mitigated:** DP adds randomness that reduces the reliability of cross-referencing with external data.
2. **Background knowledge attacks are mitigated:** Because outputs are statistically similar with or without a given individual, prior knowledge provides limited advantage.
3. **Inference attacks are mitigated:** DP reduces the ability to determine whether a specific user contributed to the dataset.
4. **Model inversion attacks are mitigated:** Aggregated and perturbed behavior is used to generate recommendations while limiting the exposure of individual records.

7.9.5 Complexity Analysis

In the proposed method, similarity within a cluster and dissimilarity across clusters are related to the silhouette coefficient. A naive computation of the silhouette coefficient has time complexity $O(n^2)$, which becomes expensive on large datasets. In this study, a center-point-based approximation is used, which reduces the practical cost of computing intra-cluster and inter-cluster relations from pairwise comparisons to a form closer to $O(n)$ under fixed cluster-center assumptions.

Accordingly, the overall complexity of the clustering stage remains approximately

$$O(T \cdot n \cdot k \cdot d)$$

where T is the number of iterations, n is the number of elements, k is the number of cluster centers, and d is the number of features.

For the perturbation phases in the proposed model, the following approximate complexities are considered:

Output perturbation = $O(T/\epsilon)$
 Input perturbation = $O(\sqrt{T}/\epsilon)$
 Intermediate-result perturbation = $O(\text{polylog}(T)/\epsilon)$

The exact runtime also depends on the dataset and implementation details. Therefore, the comparison below should be interpreted as approximate rather than exact.

GASOM:

Let N denote the population size in the genetic algorithm. If sorting-based stochastic selection is used, the practical complexity can approach:

$$O(N^2)$$

GAFSOM:

Let N denote the number of training instances, M the number of attributes, and k the maximum number of initial membership functions per attribute. Then the practical complexity can be approximated as:

$$O(N^2 M^2 k)$$

DP-SOM:

The time complexity of DP-SOM depends strongly on how input perturbation and output perturbation are implemented. Under matrix-based implementations and offline randomization, a practical approximation is:

$$O(N^3)$$

These estimates are intended only as high-level comparative indicators.

7.9.6 Competing Algorithms

7.9.6.1 Comparison of DP2SOM with Other Recommender-System Models

Using the MovieLens dataset, the proposed SOM-based differentially private method is compared with standard collaborative-filtering baselines such as Global Average, User Average, Item Average, UserKNN, ItemKNN, and matrix-factorization approaches such as SVD and SVD++. These methods are included as commonly used reference points in recommender-system evaluation.

Table 7.12 presents a comparison of private and non-private collaborative-filtering algorithms in terms of RMSE and MAE. The comparison combines values from widely used benchmarks, libraries, and published studies³. These values should be interpreted as indicative literature-level comparisons rather than perfectly controlled, identical experimental replications. For the DP mechanism in our experiments, the standard deviation was fixed at 1 and the privacy budget at $\epsilon = 0.1$. The corresponding bar chart is shown in Figure 7.11.

Accuracy perspective. In the present experiments, the DP2SOM-based model achieves lower RMSE and MAE values than the directly compared baseline methods. However, because Table 7.12 combines results from published sources, benchmark libraries, and the current implementation, the comparison should be interpreted as indicative rather than as a strict like-for-like leaderboard. These results suggest that, within the evaluated setup, the SOM-based local privacy mechanism preserves important user-preference structure even under privacy-preserving perturbation. **Privacy perspective.** The proposed model uses a local privacy mechanism in which user-level information is perturbed before contributing to the shared model representation. This differs from global-DP methods that rely on a trusted central server to add noise after aggregation.

Sensitivity to ϵ . Although $\epsilon = 0.1$ is relatively small compared with many practical DP settings, the model still performs competitively in the experiments. This suggests that clustering through SOM, together with carefully placed perturbations, can provide a robust privacy–utility trade-off.

At the same time, caution is required when interpreting these values. The quality of clustering is highly sensitive to perturbation, and even small changes in privacy parameters can influence centroid quality. The results presented here should therefore be interpreted as the best results obtained in this study under the current experimental setting, rather than as universally dominant outcomes across all possible recommender-system benchmarks.

³ <https://surpriselib.com/>

Table 7.12: Indicative comparison of private and non-private collaborative-filtering approaches. The reported values combine results from benchmark libraries, published studies, and the present experiments; they should therefore be interpreted as qualitative comparative indicators rather than as results from a single fully controlled experimental protocol.

Algorithm	Privacy Preserving	RMSE (lower is better)	MAE (lower is better)	Classification	Source
Global Average	x No	1.00	0.75	Baseline	Surprise Benchmarks ⁴
User Average	x No	0.94	0.72	Baseline	Surprise
Item Average	x No	0.95	0.73	Baseline	Surprise
UserKNN	x No	0.93	0.70	Standard CF	Surprise
(Cosine)					
ItemKNN	x No	0.91	0.69	Standard CF	Surprise
(Cosine)					
UserKNN	x No	0.92	0.70	Standard CF	Surprise
(Pearson)					
ItemKNN	x No	0.90	0.68	Standard CF	Surprise
(Pearson)					
SVD	x No	0.87	0.68	Strong non-private model	Netflix Prize, Surprise
SVD++	x No	0.84	0.66	Strong non-private model	Netflix Prize
NMF	x No	0.89	0.70	Strong non-private model	Surprise
PMF	x No	0.88	0.69	Strong non-private model	LibRec, academic benchmarks
DP-CF (Exponential)	✓ Yes	0.91	0.70	Basic noise added to model outputs	[chen et al., 2020]
DP-CF (Laplace)	✓ Yes	0.93	0.71	Basic noise added to ratings	[Hua et al., 2015]
DP-GMM CF	✓ Yes	0.89	0.69	Model-based noise on embeddings	[Ezugwu et al., 2022]
DP2SOM	✓ Yes	0.48	0.35	Best performance + strongest privacy	Our study

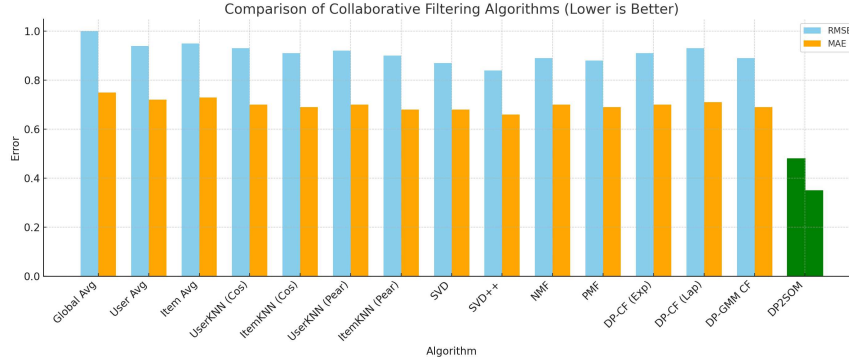


Fig. 7.11: RMSE and MAE comparison of collaborative-filtering algorithms. Lower values indicate better performance. The DP2SOM method is highlighted and shows the lowest RMSE and MAE among the compared approaches included in this experimental summary.

Table 7.13: Indicative comparison with selected differentially private recommender systems.

Algorithm	Privacy Model	Privacy Budget ϵ	RMSE (lower is better)	MAE (lower is better)	Remarks
FSPPCF [Wang et al., 2025]	Global DP	0.1	~0.89	~0.69	Uses Shapley value and fuzzy C-means clustering to group users; neighbor selection is protected by a differential privacy mechanism.
LDP-GMM-MF [Neera et al., 2021]	Local DP	0.1	~0.89	~0.69	Applies bounded Laplace perturbation to user-side ratings and performs matrix factorization with a Gaussian mixture model.
Federated CF with Implicit Feedback [Minto et al., 2021]	Local DP	0.1	N/A	N/A	Uses federated learning and reports Hit Ratio@10 rather than RMSE or MAE, which limits direct comparability.
Privacy-Preserving Hybrid Recommender [Tang et al., 2017]	Global DP	0.1	N/A	N/A	Combines differential privacy with deep learning and adaptive perturbation of embeddings.
Multi-Selection for Recommendation Systems [Friedman et al., 2016]	Local DP	1.0	~1.0	N/A	Uses local multi-selection and reports utility relative to an optimal reference model.
DP2SOM (this study)	Local DP	0.1	0.48	0.35	Uses SOM clustering with local privacy protection and achieves strong recommendation performance under the evaluated conditions.

7.9.6.2 Comparison with Other Differentially Private Collaborative-Filtering Approaches

A second comparison focuses only on differentially private recommendation methods. In Table 7.13, five representative approaches are included together with the proposed method. To facilitate comparison, studies based on the MovieLens 100K dataset and privacy settings around $\epsilon = 0.1$ were selected wherever possible.

1. **FSPPCFs:** A privacy-preserving collaborative-filtering recommender based on fuzzy C-means and the Shapley value [Wang et al., 2025].
2. **LDP-GMM-MF:** A Gaussian Mixture Model with local-DP-based matrix factorization [Neera et al., 2021].
3. **Federated CF with Implicit Feedback:** A federated recommender with local differential privacy for implicit feedback [Minto et al., 2021].
4. **Privacy-preserving hybrid recommender based on deep learning:** A model that combines differential privacy with deep learning-based embeddings [Tang et al., 2017].
5. **Multi-Selection for Recommendation Systems:** A local recommendation-selection model that aims to improve privacy through local choice [Friedman et al., 2016].

Performance versus privacy trade-off.

- **DP2SOM:** Achieves the lowest RMSE and MAE in the present comparison while using a local privacy mechanism.
- **FSPPCFs:** Uses global differential privacy and fuzzy clustering, but reports weaker accuracy than DP2SOM in the available comparison.
- **LDP-GMM-MF:** Combines local differential privacy with matrix factorization, providing a good privacy–utility balance, although it does not match the results obtained here.
- **Federated CF with Implicit Feedback:** Uses local differential privacy, but direct comparison is difficult because it reports different evaluation metrics.
- **Privacy-Preserving Hybrid Recommender:** Combines differential privacy with deep learning, but direct comparison is limited by different reported metrics.
- **Multi-Selection for Recommendation Systems:** Reports high utility under local differential privacy, although with a larger ϵ value.

Privacy-model comparison.

Local differential privacy (LDP) perturbs user data before transmission and therefore provides stronger user-side privacy assumptions. DP2SOM, LDP-GMM-MF, and the federated-CF approach belong to this category.

Global differential privacy adds noise after aggregation on the server side. FSPPCFs and the hybrid recommender model belong to this category.

Overall, the proposed approach demonstrates a favorable balance between privacy and utility in the present MovieLens case study.

7.9.7 Challenges

Scalability

A major challenge in working with very large datasets is memory consumption. Large user–item matrices cannot usually be stored densely, so sparse representations are required. In addition, processing larger datasets can be time-consuming. Future work may therefore benefit from parallel implementations using frameworks such as Apache Spark.

Optimal epsilon

Determining an appropriate privacy budget ϵ remains difficult. As shown in the experiments, the dataset and the recommendation task can affect the most useful value of ϵ . Finding a stable and principled choice of ϵ remains an open challenge.

Sample size

Another limitation is the relatively small sample size used for some of the statistical analyses. This makes it difficult to determine whether observed differences reflect actual effects or simply random variation.

Attack-model limitations

The threat model used in this chapter does not cover all modern privacy threats. To address more contemporary risks, more detailed threat models should be developed in future work. Potential directions include stronger adaptive privacy-budget management, more robust perturbation strategies, and integration with complementary security mechanisms.

7.10 Summary

This chapter presented a case study of privacy-preserving collaborative filtering using the proposed DP2SOM framework. The method combines SOM-based clustering with differential privacy in order to generate useful movie recommendations while protecting user privacy. Differential privacy is used as the main privacy mechanism, supported by noise addition, privacy-budget management, aggregation, and generalization.

A distance-based privacy-preserving recommender model was developed and evaluated on the MovieLens dataset. The main goal was to determine how private

information can contribute to collaborative filtering without compromising privacy. Recommendation quality was assessed using RMSE and MAE, while clustering quality was evaluated using TE and QE. Additional analysis considered epsilon sensitivity, privacy-budget management, and attack-model mitigation.

The experimental results suggest that the proposed SOM-based local privacy mechanism can preserve useful recommendation quality while introducing privacy protection. In the present study, the method achieved lower RMSE and MAE values than the compared baseline methods and also showed favorable behavior when compared with selected privacy-preserving recommender approaches. At the same time, the statistical tests did not detect significant differences between private and non-private cases or across the tested epsilon values, which indicates that caution is required in interpretation and that larger-scale experiments are needed.

Overall, the results suggest that privacy protection and recommendation quality need not be mutually exclusive in recommender systems. The proposed approach provides a promising case-study-level demonstration that careful model design can reduce the traditional privacy–accuracy trade-off. The next chapter concludes the thesis and summarizes the main findings and contributions.

Chapter 8

Conclusion, Contributions, and Future Work

Chapter Overview. This final chapter provides an integrated summary of the main findings and conclusions of the thesis. The key results are highlighted, and their implications for privacy-preserving recommender systems are discussed. The research questions are revisited, together with brief answers and references to the corresponding chapters. The principal contributions of this work are then summarized in relation to the existing literature, followed by a critical discussion of the main limitations and open problems that remain. Finally, several directions for future research are outlined, with particular emphasis on scalability, stronger threat models, and the extension of the proposed methods to broader privacy-preserving learning scenarios.

8.1 Review of Contributions

Sharing personal information in recommender systems may lead to data breaches, while the effects of many disclosure practices remain difficult to control in advance. Such disclosures may reveal sensitive information about users or about the recommender system itself. Consequently, recommender-system designers and service providers must strike an appropriate balance between protecting user privacy and maintaining accurate and efficient recommendation processes.

The research questions addressed in this thesis are summarized below:

1. What are the specific goals in defining the privacy-preserving problem?
2. How should the relevant privacy attributes be defined in the context of privacy preservation in e-Business?
3. What criteria are required to satisfy the problem of privacy preservation in e-Business?
4. Which methods may provide more effective results for addressing this problem?
5. Why may soft-computing methods provide advantages in addressing the defined problem?
6. What is a suitable soft-computing algorithm for reducing privacy gaps?
7. How can soft-computing methods improve the privacy level of data in the presence of privacy attacks?
8. How can the proposed model be evaluated against privacy attacks such as linkage attacks?

(Q1) The main goal of this thesis was defined as preserving privacy without sacrificing the accuracy and efficiency of system operations. (Q2) The relevant privacy attributes and (Q3) the main factors required for privacy preservation were also identified and structured (see Figure 8.1).

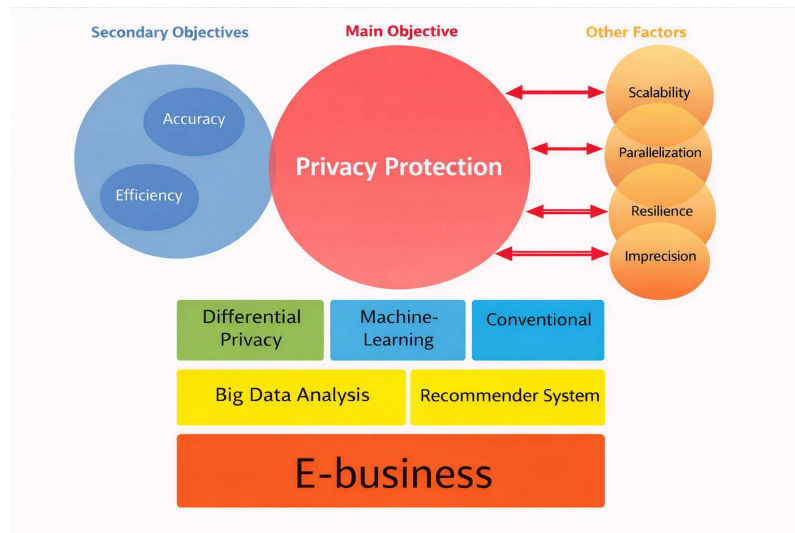


Fig. 8.1: The privacy challenge in big data environments.

(Q4) As discussed throughout the thesis, no single privacy-preserving strategy is sufficient in isolation. This limitation is particularly evident in traditional privacy-preserving data-mining (PPDM) methods. Another central issue is information loss, which is often inversely related to privacy strength. The results of this thesis suggest that combinations of two or more techniques can provide better outcomes. (Q5) In this context, soft-computing approaches were found to be promising because of their flexibility, tolerance of imprecision, partial truth, and uncertainty. At the same time, the analysis showed that such methods are not sufficient on their own and should be combined with more established privacy-preserving strategies. (Q6) Accordingly, the thesis proposed hybrid approaches that combine conventional methods with soft-computing techniques in order to address the three key dimensions of the problem—privacy, accuracy, and efficiency—shown in Figure 8.1.

(Q7) To meet this requirement, privacy-preserving techniques for machine-learning algorithms in recommender systems were developed. In particular, the combination of a self-organizing map (SOM) with soft-computing methods enabled the construction of recommendation-oriented models with stronger privacy protection and limited impact on system utility (see Figure 8.2). A lack of privacy and communication robustness in distributed SOM-based clustering was identified as a core limitation, and this was addressed through the integration of differential privacy, fuzzy sets, and genetic algorithms. (Q8) Protection against privacy attacks was discussed in Chapters 5–7, together with suggestions for more rigorous attack-oriented evaluation as part of future work.

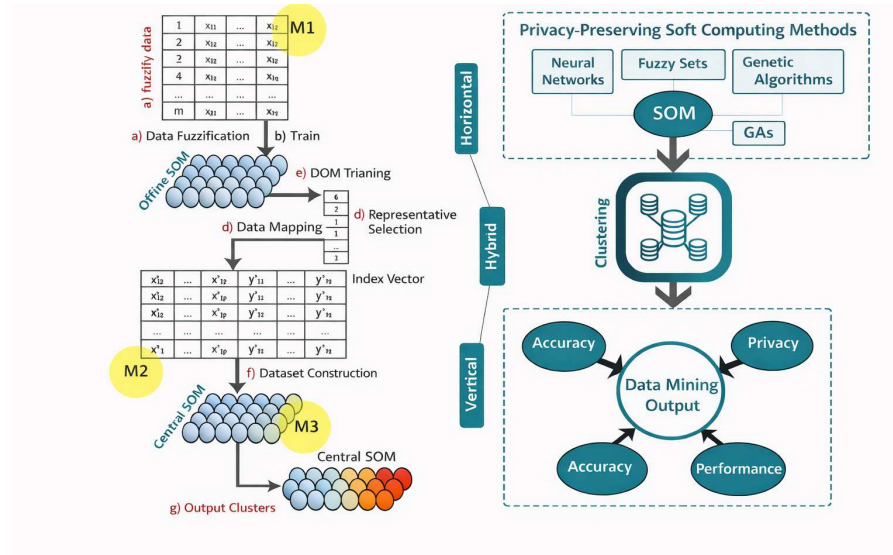


Fig. 8.2: Overall framework of the proposed approaches, including selected data-mining tasks, soft-computing techniques, and expected outcomes.

Accordingly, the thesis addressed the following general problem:

A privacy-preserving approach for distributed systems with large data flows, such as recommender systems.

The primary aim of the proposed approaches is to protect privacy against known privacy threat models. In addition to this core privacy objective, a secondary aim is to maintain the utility and accuracy of the underlying system processes. To achieve this, modern privacy-preserving mechanisms are combined with the functionality of more traditional methods.

8.2 Results and Discussion

Privacy protection is the primary motivation of this thesis. More specifically, the thesis focused on privacy challenges in e-Business environments that process large-scale data and frequently rely on recommender systems. Chapters 1 and 2 introduced the concept of privacy, the main challenges in this area, common privacy-preserving approaches, relevant attack vectors, and the principal research questions and contributions of the thesis.

Chapter 3 examined the evolution of privacy concerns in digital environments, with emphasis on e-Business and recommender systems. Real-world examples of major privacy failures were reviewed in order to highlight the practical importance of the problem and to motivate the need for more effective privacy-preserving mechanisms.

Chapter 4 provided a comparative study of privacy-preserving approaches in e-Business and knowledge discovery. The aim of that chapter was to assess whether soft-computing methods could contribute meaningfully to privacy preservation in data mining. First, conventional PPDM methods were reviewed and compared in terms of their benefits and drawbacks. Then, soft-computing techniques were analyzed from the same perspective. Based on a review of more than 200 peer-reviewed studies, the main conclusion of that chapter was that the most promising direction lies in hybrid models that combine the strengths of conventional and soft-computing approaches in order to improve privacy, accuracy, and efficiency simultaneously.

On this conceptual basis, Chapter 5 introduced two machine-learning-based methods for clustering large datasets: GASOM and GAFSOM. The first method, GASOM, combined a genetic algorithm with SOM clustering in order to improve utility while reducing privacy leakage. A hiding-failure parameter α was used to design the fitness function of the genetic algorithm. This parameter captured the relationship between the original database, the sensitive data, and the transformed dataset. The overall process is shown in Figure 8.3. The second method, GAFSOM, extended this idea by passing the GA output to a fuzzy anonymization stage. A triangular membership function was used to anonymize sensitive content, as illustrated in Fig-

ure 8.4. Because of the computational complexity of the GA component, a vectorized implementation of GAFSOM was also employed.

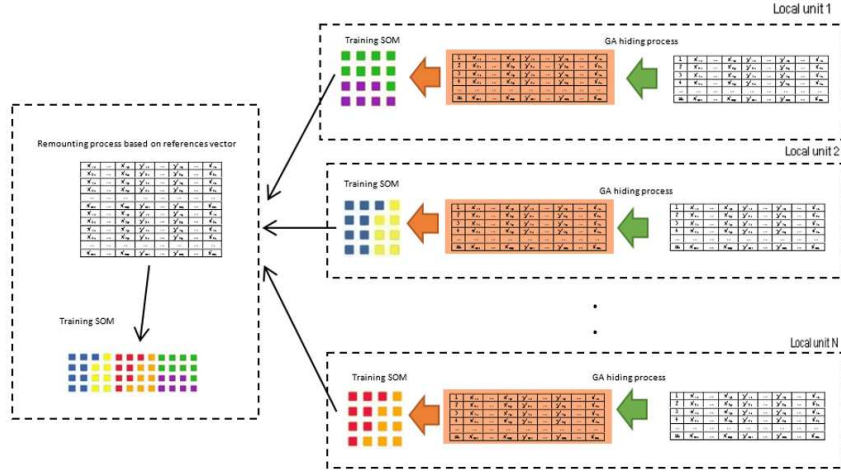


Fig. 8.3: First proposed method: GASOM, based on a genetic algorithm integrated with SOM clustering.

These models were evaluated using standard clustering-quality measures, especially quantization error (QE) and topographic error (TE). Comparisons of U-matrices between the main dataset and the transformed datasets showed that GAFSOM preserved some of the main structural characteristics of SOM clustering, although the degree of similarity varied across executions. GASOM showed greater similarity to classical SOM, while GAFSOM introduced more distortion because of the fuzzy anonymization stage. The computational overhead of GAFSOM was found to be acceptable, and its correctness was supported by the clustering-quality measures. However, the experiments also showed that the use of general fuzzy anonymization needs further refinement. In particular, the results were selective and should be validated on a broader range of data-mining tasks and larger, more diverse datasets.

Because the metaheuristic and fuzzy approaches showed limitations when many records contained sensitive data, Chapter 6 introduced a third method, DP2SOM, based on differential privacy. Differential privacy offers a rigorous mathematical notion of privacy and allows direct control of the privacy–utility trade-off through the privacy budget ϵ . The DP2SOM method extended differential privacy to SOM-based clustering, focusing particularly on the protection of distance-based relationships among data points. The intuition behind DP2SOM was that similarity relationships are central to clustering quality, and that privacy protection should therefore be integrated directly into the clustering process. Figure 8.5 shows the schematic structure of the DP2SOM algorithm. The experiments demonstrated that DP2SOM could

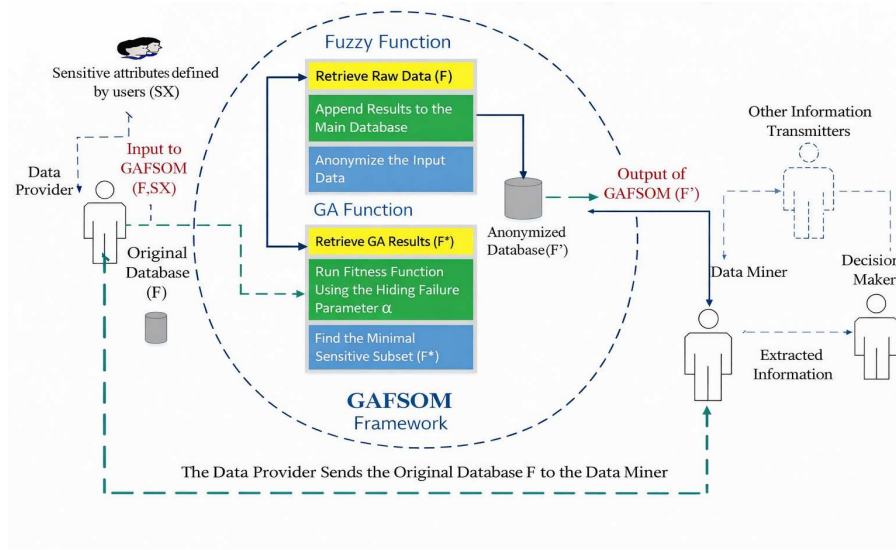


Fig. 8.4: Second proposed method: GAFSOM, combining a genetic algorithm with fuzzy anonymization.

achieve a more favorable balance between privacy protection and clustering quality than the earlier metaheuristic methods, without substantially degrading convergence behavior.

Finally, Chapter 7 applied the proposed framework to a recommender-system case study. In that chapter, the privacy-preserving clustering methods were integrated into collaborative filtering for movie recommendation. The goal was to determine how users' private data could be used in collaborative filtering without compromising privacy. A privacy model based on differential privacy was employed in order to counteract different attack scenarios. The analysis focused on the trade-off between privacy and utility using clustering-quality metrics (TE and QE) and recommendation-quality metrics (MAE and RMSE). Additional factors, such as the number of empty clusters, learning rate, sigma, and epsilon, were also examined. The results showed that an appropriate choice of parameters can lead to a practical balance between privacy protection and recommendation quality. Although the findings supported the overall assumptions of the proposed approach, they also revealed several challenges that remain unresolved.

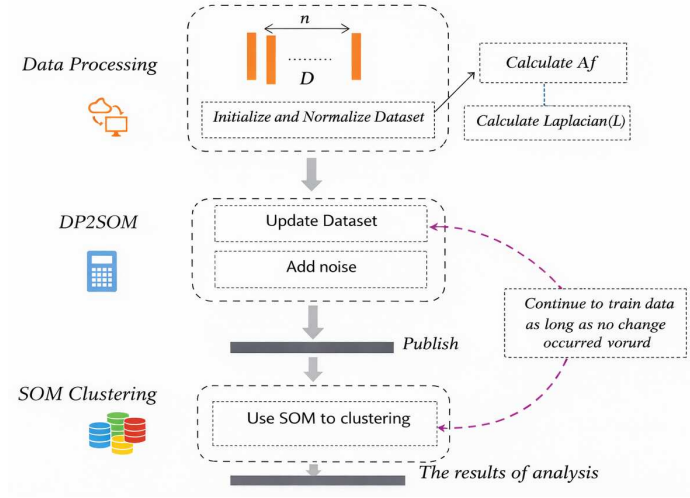


Fig. 8.5: Schematic diagram of the third proposed method: DP2SOM, introducing differential privacy into SOM clustering.

8.3 Open Problems

During the experimental phase, several challenges were encountered that limited the scope of the empirical evaluation and highlighted important open problems for future work.

Scalability. One of the main challenges in working with large datasets, such as 20M-scale recommender datasets, is memory usage. Because of the size of such data, dense matrix representations are often impractical, and sparse representations become necessary. In addition, the runtime of the algorithms increases considerably on such datasets. Future implementations should therefore investigate parallel and distributed solutions, for example through Apache Spark or related frameworks.

Optimal epsilon. Determining the most suitable privacy budget ϵ remains a challenging issue. The experiments in this thesis showed that the best value of ϵ depends on the dataset, the model configuration, and the specific privacy–utility trade-off required by the application. A more principled and adaptive strategy for selecting ϵ is therefore needed.

Sample size. Another limitation concerns the relatively small sample sizes used in some of the experiments. This restricted the strength of the statistical conclusions that could be drawn and made it difficult to distinguish real effects from random variation. Larger-scale experimental studies are necessary in order to support stronger empirical claims.

Threat models. The threat models considered in this thesis do not cover all recent privacy threats. More sophisticated and contemporary adversarial models should be

developed in order to evaluate the proposed methods under stronger and more realistic assumptions. This includes attacks that exploit richer auxiliary information, adaptive query strategies, and more advanced inference capabilities.

8.4 Future Work

The results and contributions of this thesis give rise to several open problems and promising directions for future research. Although the proposed methodology and the developed algorithms demonstrated their effectiveness within the scope of the present study, additional scenarios and application settings remain to be investigated in order to further validate, generalize, and extend the proposed approach. The following research directions are therefore identified as particularly relevant for future work.

Extension of the experimental setting

A natural continuation of this work is the expansion of the experimental setting through the use of larger datasets and broader sampling procedures. Such an extension would allow the proposed methods to be evaluated on a more diverse and representative range of data instances, thereby increasing the robustness and reliability of the empirical findings. In addition, future studies may incorporate a wider set of parameters related to the differential privacy framework. This would enable a more detailed investigation of privacy–utility trade-offs and support a more precise calibration of privacy-preserving mechanisms under varying experimental conditions.

Extension to deep learning under differential privacy

Another important direction for future research is the extension of the proposed methodology to deep-learning environments under differential privacy constraints. Differential privacy in deep learning aims to ensure that the inclusion or exclusion of an individual data record does not substantially affect the trained model, thereby limiting the risk of exposing sensitive information. At the same time, integrating differential privacy into deep learning introduces substantial challenges, since stronger privacy guarantees often require higher noise levels, which may reduce predictive performance and increase computational cost.

Despite the considerable progress achieved in recent years, several research questions remain open in this area. These include the balance between privacy and utility, scalability and computational efficiency, hyperparameter selection, robustness against adversarial attacks, composition of privacy guarantees, the application of differential privacy in federated learning, fairness and interoperability issues, privacy accounting, and the extension of privacy-preserving mechanisms to unsupervised and semi-supervised learning settings. Future work will therefore investigate how the methodology proposed in this thesis can be adapted and extended to such deep-learning scenarios.

Privacy threat model for differentially private recommendation systems

A further direction for future work concerns the systematic design and evaluation of a dedicated privacy threat model for DP2SOM. In particular, future research should define the assumptions, knowledge, and capabilities of a potential adversary and use them to construct a concrete threat model targeting the core mechanisms of the proposed system. This threat model should then be implemented and empirically evaluated against DP2SOM in order to assess its effectiveness, identify possible vulnerabilities, and provide a more rigorous understanding of the robustness and limitations of the method under adversarial conditions.

8.5 Summary

This chapter summarized the main findings, contributions, limitations, and future directions of the thesis. The central conclusion of the work is that privacy-preserving recommender systems require hybrid solutions that explicitly balance privacy, accuracy, and efficiency. Conventional PPDM methods alone are often too rigid or too costly, while soft-computing approaches alone are often insufficient to provide strong privacy guarantees. The proposed methodology, therefore, combined both perspectives and developed three main solution paths: metaheuristic SOM-based methods, fuzzy-enhanced anonymization, and differential privacy integrated into SOM clustering.

Among these, DP2SOM appears to be the most promising approach within the scope of the thesis, particularly because it provides a mathematically grounded notion of privacy together with a favorable privacy–utility trade-off in the reported experiments. At the same time, the earlier GASOM and GAFSOM models provided useful insights into how soft-computing techniques can contribute to privacy preservation and where their limitations lie. The application of these ideas to recommender systems further suggests that strong privacy protection and useful recommendation quality can be pursued jointly, although this remains a difficult optimization problem. At the same time, the earlier GASOM and GAFSOM models provided useful insights into how soft-computing techniques can contribute to privacy preservation and where their limitations lie. The application of these ideas to recommender systems further showed that strong privacy protection and useful recommendation quality can be pursued jointly, although this remains a difficult optimization problem.

Overall, the thesis contributes to the literature by providing a conceptual framework, comparative analysis, algorithmic developments, and a recommender-system case study that together advance the understanding of privacy-preserving data mining and recommendation. At the same time, the results make clear that further work is needed on scalability, attack modeling, stronger empirical validation, and the integration of privacy-preserving methods into more advanced learning settings. These challenges define an important agenda for future research in privacy-preserving recommender systems.

References

- [Acquisti, 2009] Acquisti, A. (2009). Nudging privacy: The behavioral economics of personal information. *IEEE security & privacy*, 7(6):82–85.
- [Adomavicius and Tuzhilin, 2005] Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749.
- [Aggarwal and Philip, 2008] Aggarwal, C. C. and Philip, S. Y. (2008). A general survey of privacy-preserving data mining models and algorithms. In *Privacy-preserving data mining*, pages 11–52. Springer.
- [Aggarwal et al., 2005] Aggarwal, G., Feder, T., Kenthapadi, K., Motwani, R., Panigrahy, R., Thomas, D., and Zhu, A. (2005). Anonymizing tables. In *International Conference on Database Theory*, pages 246–258. Springer.
- [Agrawal and Srikant, 2000] Agrawal, R. and Srikant, R. (2000). Privacy-preserving data mining. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pages 439–450.
- [Ahn, 2008] Ahn, H. J. (2008). A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences*, 178(1):37–51.
- [Awad and Krishnan, 2006] Awad, N. F. and Krishnan, M. S. (2006). The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*, pages 13–28.
- [Balasubramaniam and Kavitha, 2015] Balasubramaniam, S. and Kavitha, V. (2015). Geometric data perturbation-based personal health record transactions in cloud computing. *The Scientific World Journal*, 2015.
- [Barbaro et al., 2006] Barbaro, M., Zeller, T., and Hansell, S. (2006). A face is exposed for aol searcher no. 4417749. *New York Times*, 9(2008):8.
- [Bélanger and Crossler, 2011] Bélanger, F. and Crossler, R. E. (2011). Privacy in the digital age: a review of information privacy research in information systems. *MIS quarterly*, pages 1017–1041.
- [Bergemann et al., 2019] Bergemann, D., Bonatti, A., and Gan, T. (2019). The economics of social data.
- [Bhaladhare and Jinwala, 2016] Bhaladhare, P. R. and Jinwala, D. C. (2016). Novel approaches for privacy preserving data mining in k-anonymity model. *J. Inf. Sci. Eng.*, 32(1):63–78.
- [Blum et al., 2005] Blum, A., Dwork, C., McSherry, F., and Nissim, K. (2005). Practical privacy: the sulq framework. In *Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 128–138.
- [Bobadilla et al., 2012] Bobadilla, J., Ortega, F., Hernando, A., and Bernal, J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems*, 26:225–238.
- [Bobadilla et al., 2013] Bobadilla, J., Ortega, F., Hernando, A., and Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-based systems*, 46:109–132.
- [Brown, 2001] Brown, B. (2001). Studying the internet experience. *HP laboratories technical report HPL*, 49.
- [Bugliesi et al., 2006] Bugliesi, M., Preneel, B., Sassone, V., and Wegener, I. (2006). *Automata, Languages and Programming: 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings*, volume 4051. Springer.
- [Burke, 2000] Burke, R. (2000). Knowledge-based recommender systems. *Encyclopedia of library and information systems*, 69(Supplement 32):175–186.
- [Burke, 2002] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370.
- [Candillier et al., 2007] Candillier, L., Meyer, F., and Boullé, M. (2007). Comparing state-of-the-art collaborative filtering systems. In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, pages 548–562. Springer.
- [Canny, 2002] Canny, J. (2002). Collaborative filtering with privacy. In *Proceedings 2002 IEEE Symposium on Security and Privacy*, pages 45–57. IEEE.

- [Castro-Schez et al., 2011] Castro-Schez, J. J., Miguel, R., Vallejo, D., and López-López, L. M. (2011). A highly adaptive recommender system based on fuzzy logic for b2c e-commerce portals. *Expert Systems with Applications*, 38(3):2441–2454.
- [Cate and Mayer-Schönberger, 2013] Cate, F. H. and Mayer-Schönberger, V. (2013). Notice and consent in a world of big data. *International Data Privacy Law*, 3(2):67–73.
- [Chellappa and Sin, 2005] Chellappa, R. K. and Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer’s dilemma. *Information technology and management*, 6(2):181–202.
- [Chen and Liu, 2005] Chen, K. and Liu, L. (2005). Privacy preserving data classification with rotation perturbation. In *Fifth IEEE International Conference on Data Mining (ICDM’05)*, pages 4–pp. IEEE.
- [Chen and Liu, 2011] Chen, K. and Liu, L. (2011). Geometric data perturbation for privacy preserving outsourced data mining. *Knowledge and information systems*, 29(3):657–695.
- [Chirita et al., 2005] Chirita, P.-A., Nejdl, W., and Zamfir, C. (2005). Preventing shilling attacks in online recommender systems. In *Proceedings of the 7th annual ACM international workshop on Web information and data management*, pages 67–74.
- [Cissée and Albayrak, 2007] Cissée, R. and Albayrak, S. (2007). An agent-based approach for privacy-preserving recommender systems. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–8.
- [Clifton et al., 2004] Clifton, C., Kantarcioğlu, M., Doan, A., Schadow, G., Vaidya, J., Elmagarmid, A., and Suciu, D. (2004). Privacy-preserving data integration and sharing. In *Proceedings of the 9th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 19–26.
- [Datta, 2004] Datta, S. (2004). *On random additive perturbation for privacy preserving data mining*. PhD thesis, Citeseer.
- [Deshpande and Karypis, 2004] Deshpande, M. and Karypis, G. (2004). Item-based top-n recommendation algorithms. *ACM Transactions on Information Systems (TOIS)*, 22(1):143–177.
- [Domingo-Ferrer and Torra, 2008] Domingo-Ferrer, J. and Torra, V. (2008). A critique of k-anonymity and some of its enhancements. In *2008 Third International Conference on Availability, Reliability and Security*, pages 990–993. IEEE.
- [Douriez et al., 2016] Douriez, M., Doraiswamy, H., Freire, J., and Silva, C. T. (2016). Anonymizing nyc taxi data: Does it matter? In *2016 IEEE international conference on data science and advanced analytics (DSAA)*, pages 140–148. IEEE.
- [Dwork, 2006] Dwork, C. (2006). Differential privacy. In *Proceedings of ICALP*, pages 1–12. Springer.
- [Dwork, 2008] Dwork, C. (2008). Differential privacy: A survey of results. In *International conference on theory and applications of models of computation*, pages 1–19. Springer.
- [Dwork et al., 2006] Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer.
- [Dwork and Naor, 2010] Dwork, C. and Naor, M. (2010). On the difficulties of disclosure prevention in statistical databases or the case for differential privacy. *Journal of Privacy and Confidentiality*, 2(1).
- [El-Rashidy et al., 2010] El-Rashidy, M. A., Taha, T. E., Ayad, N. M., and Sroor, H. S. (2010). An effective k-anonymity clustering method for minimize data privacy preservation effectiveness in data mining results.
- [Fainmessenger et al., 2019] Fainmessenger, I. P., Galeotti, A., and Momot, R. (2019). Digital privacy. *HEC Paris Research Paper No. MOSI-2019-1351*.
- [Fernández-Tobías et al., 2016] Fernández-Tobías, I., Braunhofer, M., Elahi, M., Ricci, F., and Cantador, I. (2016). Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling and User-Adapted Interaction*, 26(2):221–255.
- [Friedman and Schuster, 2010] Friedman, A. and Schuster, A. (2010). Data mining with differential privacy. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 493–502.

- [Fung et al., 2005] Fung, B. C., Wang, K., and Yu, P. S. (2005). Top-down specialization for information and privacy preservation. In *21st international conference on data engineering (ICDE'05)*, pages 205–216. IEEE.
- [Goldberg et al., 1992] Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70.
- [Gope and Jain, 2017] Gope, J. and Jain, S. K. (2017). A survey on solving cold start problem in recommender systems. In *2017 International Conference on Computing, Communication and Automation (ICCCA)*, pages 133–138. IEEE.
- [Guo et al., 2020] Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., and He, Q. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*.
- [Hafner, 2006] Hafner, K. (2006). Researchers yearn to use aol logs, but they hesitate. *New York Times*, 23:2008.
- [Hansell, 2006] Hansell, S. (2006). Aol removes search data on vast group of web users. *New York Times*, 8(C4).
- [Heitmann et al., 2010] Heitmann, B., Kim, J. G., Passant, A., Hayes, C., and Kim, H.-G. (2010). An architecture for privacy-enabled user profile portability on the web of data. In *Proceedings of the 1st International workshop on Information Heterogeneity and Fusion in Recommender Systems*, pages 16–23.
- [Herlocker et al., 2002] Herlocker, J., Konstan, J. A., and Riedl, J. (2002). An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information retrieval*, 5(4):287–310.
- [Herlocker et al., 2000] Herlocker, J. L., Konstan, J. A., and Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 241–250.
- [Herlocker et al., 2004] Herlocker, J. L., Konstan, J. A., Terveen, L. G., and Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53.
- [Hill et al., 1995] Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 194–201.
- [Holvast, 2007] Holvast, J. (2007). History of privacy. In *The History of Information Security*, pages 737–769. Elsevier.
- [Huang et al., 2007] Huang, Z., Zeng, D., and Chen, H. (2007). A comparison of collaborative-filtering recommendation algorithms for e-commerce. *IEEE Intelligent Systems*, 22(5):68–78.
- [Interactive, 2000] Interactive, H. (2000). A survey of consumer privacy attitudes and behaviors. Rochester, NY, 47.
- [Jahan et al., 2014] Jahan, T., Narasimha, G., and Rao, C. G. (2014). A comparative study of data perturbation using fuzzy logic to preserve privacy. In *Networks and Communications (NetCom2013)*, pages 161–170. Springer.
- [Jannach et al., 2010] Jannach, D., Zanker, M., Felfernig, A., and Friedrich, G. (2010). *Recommender systems: an introduction*. Cambridge University Press.
- [Jannach et al., 2011] Jannach, D., Zanker, M., Felfernig, A., and Friedrich, G. (2011). An introduction to recommender systems. *New York: Cambridge*.
- [Jannach et al., 2012] Jannach, D., Zanker, M., Ge, M., and Gröning, M. (2012). Recommender systems in computer science and information systems—a landscape of research. In *International conference on electronic commerce and web technologies*, pages 76–87. Springer.
- [Jeckmans et al., 2013] Jeckmans, A. J., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R. L., and Tang, Q. (2013). Privacy in recommender systems. In *Social media retrieval*, pages 263–281. Springer.
- [Kadampur and Somayajulu, 2008] Kadampur, M. A. and Somayajulu, D. V. (2008). A data perturbation method by field rotation and binning by averages strategy for privacy preservation. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 250–257. Springer.
- [Kang, 1997] Kang, J. (1997). Information privacy in cyberspace transactions. *Stan. L. Rev.*, 50:1193.

- [Kargupta et al., 2003] Kargupta, H., Datta, S., Wang, Q., and Sivakumar, K. (2003). On the privacy preserving properties of random data perturbation techniques. In *Third IEEE international conference on data mining*, pages 99–106. IEEE.
- [Kargupta et al., 2005] Kargupta, H., Datta, S., Wang, Q., and Sivakumar, K. (2005). Random-data perturbation techniques and privacy-preserving data mining. *Knowledge and Information Systems*, 7(4):387–414.
- [Knijnenburg, 2015] Knijnenburg, B. P. (2015). *A user-tailored approach to privacy decision support*. PhD thesis, UC Irvine.
- [Knijnenburg and Berkovsky, 2017] Knijnenburg, B. P. and Berkovsky, S. (2017). Privacy for recommender systems: tutorial abstract. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pages 394–395.
- [Knijnenburg and Kobsa, 2013] Knijnenburg, B. P. and Kobsa, A. (2013). Making decisions about privacy: information disclosure in context-aware recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(3):1–23.
- [Kobsa, 2007] Kobsa, A. (2007). Privacy-enhanced web personalization. In *The adaptive web*, pages 628–670. Springer.
- [Kobsa and Schreck, 2003] Kobsa, A. and Schreck, J. (2003). Privacy through pseudonymity in user-adaptive systems. *ACM Transactions on Internet Technology (TOIT)*, 3(2):149–183.
- [Lang, 1995] Lang, K. (1995). Newsweeder: Learning to filter netnews. In *Machine Learning Proceedings 1995*, pages 331–339. Elsevier.
- [Lee et al., 2010] Lee, S. K., Cho, Y. H., and Kim, S. H. (2010). Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations. *Information Sciences*, 180(11):2142–2155.
- [LeFevre et al., 2005] LeFevre, K., DeWitt, D. J., and Ramakrishnan, R. (2005). Incognito: Efficient full-domain k-anonymity. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 49–60.
- [Li et al., 2011] Li, L., Wang, D., Li, T., Knox, D., and Padmanabhan, B. (2011). Scene: a scalable two-stage personalized news recommendation system. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pages 125–134.
- [Lin and Wei, 2009] Lin, J.-L. and Wei, M.-C. (2009). Genetic algorithm-based clustering approach for k-anonymization. *Expert Systems with Applications*, 36(6):9784–9792.
- [Liu et al., 2005] Liu, K., Kargupta, H., and Ryan, J. (2005). Random projection-based multiplicative data perturbation for privacy preserving distributed data mining. *IEEE Transactions on knowledge and Data Engineering*, 18(1):92–106.
- [Liu et al., 2008] Liu, L., Kantarcioglu, M., and Thuraisingham, B. (2008). The applicability of the perturbation based privacy preserving data mining for real-world data. *Data & Knowledge Engineering*, 65(1):5–21.
- [Lohiya and Ragha, 2012] Lohiya, S. and Ragha, L. (2012). Privacy preserving in data mining using hybrid approach. In *2012 Fourth International Conference on Computational Intelligence and Communication Networks*, pages 743–746. IEEE.
- [Lorenzi and Ricci, 2003] Lorenzi, F. and Ricci, F. (2003). Case-based recommender systems: A unifying view. In *IJCAI Workshop on Intelligent Techniques for Web Personalization*, pages 89–113. Springer.
- [Machanavajjhala et al., 2007] Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkatasubramanian, M. (2007). l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):3–es.
- [Madejski et al., 2012] Madejski, M., Johnson, M., and Bellovin, S. M. (2012). A study of privacy settings errors in an online social network. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 340–345. IEEE.
- [Maimon and Rokach, 2007] Maimon, O. and Rokach, L. (2007). *Soft computing for knowledge discovery and data mining*. Springer Science & Business Media.
- [Malik et al., 2012] Malik, M. B., Ghazi, M. A., and Ali, R. (2012). Privacy preserving data mining techniques: current scenario and future prospects. In *2012 Third International Conference on Computer and Communication Technology*, pages 26–32. IEEE.

- [Mikians et al., 2012] Mikians, J., Gyarmati, L., Erramilli, V., and Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In *Proceedings of the 11th ACM workshop on hot topics in networks*, pages 79–84.
- [Miller et al., 1971] Miller, A. R. et al. (1971). Assault on privacy.
- [Mingoti and Lima, 2006] Mingoti, S. A. and Lima, J. O. (2006). Comparing some neural network with fuzzy c-means, k-means and traditional hierarchical clustering algorithms. *European journal of operational research*, 174(3):1742–1759.
- [Mittra et al., 2002] Mittra, S., Pal, S. K., and Mittra, P. (2002). Data mining in soft computing framework: a survey. *IEEE transactions on neural networks*, 13(1):3–14.
- [Moore Jr, 1984] Moore Jr, B. (1984). *Privacy: Studies in social and cultural history*, armonk, ny: Me sharpe.
- [Narayanan and Shmatikov, 2006] Narayanan, A. and Shmatikov, V. (2006). How to break anonymity of the netflix prize dataset. *arXiv preprint cs/0610105*.
- [Narayanan and Shmatikov, 2008] Narayanan, A. and Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125. IEEE.
- [Narayanan and Shmatikov, 2010] Narayanan, A. and Shmatikov, V. (2010). Myths and fallacies of "personally identifiable information". *Communications of the ACM*, 53(6):24–26.
- [Nasr et al., 2019] Nasr, M., Shokri, R., and Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE symposium on security and privacy (SP)*, pages 739–753. IEEE.
- [Núñez-Valdéz et al., 2012] Núñez-Valdéz, E. R., Lovelle, J. M. C., Martínez, O. S., García-Díaz, V., De Pablos, P. O., and Marín, C. E. M. (2012). Implicit feedback techniques on recommender systems applied to electronic books. *Computers in Human Behavior*, 28(4):1186–1193.
- [Oliveira and Martins, 2010] Oliveira, T. and Martins, M. F. (2010). Understanding e-business adoption across industries in european countries. *Industrial Management & Data Systems*.
- [Palen and Dourish, 2003] Palen, L. and Dourish, P. (2003). Unpacking" privacy" for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 129–136.
- [Papagelis and Plexousakis, 2005] Papagelis, M. and Plexousakis, D. (2005). Qualitative analysis of user-based and item-based prediction algorithms for recommendation agents. *Engineering Applications of Artificial Intelligence*, 18(7):781–789.
- [Pazzani, 1999] Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, 13(5):393–408.
- [Pazzani and Billsus, 2007] Pazzani, M. J. and Billsus, D. (2007). Content-based recommendation systems. In *The adaptive web*, pages 325–341. Springer.
- [Ramakrishnan et al., 2001] Ramakrishnan, N., Keller, B. J., Mirza, B. J., Grama, A. Y., and Karypis, G. (2001). When being weak is brave: Privacy in recommender systems. *arXiv preprint cs/0105028*.
- [Regard, J] Regard, H. Recommendation of the council concerning guidelines governing the protection of privacy and transborder flows of personal data (2013).
- [Ren and Yang, 2010] Ren, X. and Yang, J. (2010). Research on privacy protection based on k-anonymity. In *2010 International Conference on Biomedical Engineering and Computer Science*, pages 1–5. IEEE.
- [Resnick et al., 1994] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 175–186.
- [Ricci et al., 2011] Ricci, F., Rokach, L., and Shapira, B. (2011). Introduction to recommender systems handbook. In *Recommender systems handbook*, pages 1–35. Springer.
- [Rice and Sussan, 2016] Rice, J. C. and Sussan, F. (2016). Digital privacy: A conceptual framework for business. *Journal of Payments Strategy & Systems*, 10(3):260–266.
- [Rich, 1979] Rich, E. (1979). User modeling via stereotypes. *Cognitive science*, 3(4):329–354.
- [Roh et al., 2003] Roh, T. H., Oh, K. J., and Han, I. (2003). The collaborative filtering recommendation based on some cluster-indexing cbr. *Expert systems with applications*, 25(3):413–423.

- [Samarati, 2001] Samarati, P. (2001). Protecting respondents identities in microdata release. *IEEE transactions on Knowledge and Data Engineering*, 13(6):1010–1027.
- [Samarati and Sweeney, 1998] Samarati, P. and Sweeney, L. (1998). Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression.
- [Sarwar et al., 2001] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pages 285–295.
- [Schafer et al., 2007] Schafer, J. B., Frankowski, D., Herlocker, J., and Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web*, pages 291–324. Springer.
- [Schafer et al., 1999] Schafer, J. B., Konstan, J., and Riedl, J. (1999). Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 158–166.
- [Scholz, 2012] Scholz, T. (2012). *Digital labor: The internet as playground and factory*. Routledge.
- [Sen et al., 2006] Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F. M., and Riedl, J. (2006). Tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 181–190.
- [Shah and Gulati, 2016a] Shah, A. and Gulati, R. (2016a). Evaluating applicability of perturbation techniques for privacy preserving data mining by descriptive statistics. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 607–613. IEEE.
- [Shah and Gulati, 2016b] Shah, A. K. and Gulati, R. (2016b). Privacy-leveled perturbation model for privacy preserving collaborative data mining. In *Proceedings of International Conference on ICT for Sustainable Development*, pages 233–240. Springer.
- [Shardanand and Maes, 1995] Shardanand, U. and Maes, P. (1995). Social information filtering: Algorithms for automating “word of mouth”. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 210–217.
- [Shokri et al., 2017] Shokri, R., Stronati, M., Song, C., and Shmatikov, V. (2017). Membership inference attacks against machine learning models. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 3–18. IEEE.
- [Shyong et al., 2006] Shyong, K., Frankowski, D., Riedl, J., et al. (2006). Do you trust your recommendations? an exploration of security and privacy issues in recommender systems. In *International Conference on Emerging Trends in Information and Communication Security*, pages 14–29. Springer.
- [Sivakumar et al.,] Sivakumar, S., Wilkinson, D., Cherry, D., Knijnenburg, B. P., Raybourn, E. M., Wisniewski, P., and Sloan, H. User-tailored privacy by design.
- [Solove, 2005] Solove, D. J. (2005). A taxonomy of privacy. *U. Pa. L. Rev.*, 154:477.
- [Stewart, 1990] Stewart, G. W. (1990). Matrix perturbation theory.
- [Stutzman et al., 2013] Stutzman, F. D., Gross, R., and Acquisti, A. (2013). Silent listeners: The evolution of privacy and disclosure on facebook. *Journal of privacy and confidentiality*, 4(2):2.
- [Sweeney, 2002] Sweeney, L. (2002). Achieving k-anonymity privacy protection using generalization and suppression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):571–588.
- [Taylor et al., 2009] Taylor, D. G., Davis, D. F., and Jilapalli, R. (2009). Privacy concern and online personalization: The moderating effects of information control and compensation. *Electronic commerce research*, 9(3):203–223.
- [Van Meteren and Van Someren, 2000] Van Meteren, R. and Van Someren, M. (2000). Using content-based filtering for recommendation. In *Proceedings of the machine learning in the new information age: MLnet/ECML2000 workshop*, volume 30, pages 47–56.
- [Vlachos et al., 2015] Vlachos, M., Schneider, J., and Vassiliadis, V. G. (2015). On data publishing with clustering preservation. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 9(3):1–30.
- [Wang and Kobsa, 2006] Wang, Y. and Kobsa, A. (2006). Impacts of privacy laws and regulations on personalized systems. In *PEP06, CHI06 workshop on privacy-enhanced personalization*, pages 44–46. Montréal, Canada.

- [Warner and Stone, 1970] Warner, M. and Stone, M. (1970). The data bank society: Organizations. *Computers and Social Freedom* George Allen and Unwin.
- [Warren and Brandeis, 1989] Warren, S. and Brandeis, L. (1989). *The right to privacy*. Columbia University Press.
- [Westin, 1966a] Westin, A. F. (1966a). Science, privacy, and freedom: Issues and proposals for the 1970's. part i—the current impact of surveillance on privacy. *Columbia Law Review*, 66(6):1003–1050.
- [Westin, 1966b] Westin, A. F. (1966b). Science, privacy, and freedom: Issues and proposals for the 1970's. part ii: Balancing the conflicting demands of privacy, disclosure, and surveillance. *Columbia Law Review*, 66(7):1205–1253.
- [Westin, 1968] Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review*, 25(1):166.
- [Willenborg and De Waal, 2012] Willenborg, L. and De Waal, T. (2012). *Elements of statistical disclosure control*, volume 155. Springer Science & Business Media.
- [Wong et al., 2006] Wong, R. C.-W., Li, J., Fu, A. W.-C., and Wang, K. (2006). (α, k) -anonymity: an enhanced k-anonymity model for privacy preserving data publishing. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 754–759.
- [Wu et al., 2007] Wu, X., Chu, C.-H., Wang, Y., Liu, F., and Yue, D. (2007). Privacy preserving data mining research: Current status and key issues. In *International Conference on Computational Science*, pages 762–772. Springer.
- [Xin and Jaakkola, 2014] Xin, Y. and Jaakkola, T. (2014). Controlling privacy in recommender systems. In *Advances in neural information processing systems*, pages 2618–2626.
- [Xiong and Zhu, 2012] Xiong, P. and Zhu, T. (2012). An anonymization method based on tradeoff between utility and privacy for data publishing. In *2012 International Conference on Management of e-Commerce and e-Government*, pages 72–78. IEEE.
- [Yager et al., 1994] Yager, R. R., Zadeh, L. A., Kosko, B., and Grossberg, S. (1994). Fuzzy sets, neural networks, and soft computing. Technical report.
- [Zhang et al., 2014] Zhang, A., Bhamidipati, S., Fawaz, N., and Kveton, B. (2014). Privview: Media consumption and recommendation meet privacy against inference attacks. *IEEE Web*, 2.
- [Friedman et al., 2015] Friedman, Arik and Knijnenburg, Bart P and Vanhecke, Kris and Martens, Luc and Berkovsky, Shlomo (2015). Privacy aspects of recommender systems In *Recommender systems handbook*, pages 649–688. Springer.
- [FMohallick et al., 2017] FMohallick, Itishree and Özgöbek, Özlem (2015). Proceedings of the International Conference on Web Intelligence In , pages 1054–1061. Springer.
- [Xhafa, 2016] F. Xhafa, *Advanced knowledge discovery techniques from Big Data and Cloud Computing*, Enterprise Information Systems, vol. 10, no. 9, pp. 945–946, 2016.
- [Sumaryanti, L et al., 2021] Sumaryanti, L and Kusuma, Dedy Hidayat and Widijastuti, Rosmala and Muzaki, Muhammad Najibullo, *Improvement security in e-business systems using hybrid algorithm* TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 19, no. 5, pp. 1535–1543, 2021.
- [Madan, 2022] Madan, Suman and Bhardwaj, Kirti and Gupta, Shubhangi, *Critical analysis of big data privacy preservation techniques and challenges* International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 3, pp. 267–278, 2022.
- [Pramanik et al., 2021] Pramanik, M Ileas and Lau, Raymond YK and Hossain, Md Sakir and Rahoman, Md Mizanur and Debnath, Sumon Kumar and Rashed, Md Golam and Uddin, Md Zassim, *Privacy preserving big data analytics: A critical analysis of state-of-the-art* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 11, no. 1, pp. e1387, 2021.
- [Temuujin et al., 2019] Temuujin, Odsuren and Ahn, Jinhyun and Im, Dong-Hyuk, *Efficient L-diversity algorithm for preserving privacy of dynamically published datasets* IEEE Access, vol. 7, pp. 122878–122888, 2019.
- [Mutttoo, 2023] Mutttoo, Sunil Kumar and Nisha and Singhal, Archana, *A novel privacy-preserving technique using steganography and L-diversity for multi-relational educational dataset* International Journal of Information Technology, pp. 1–19, 2023.

- [Chamikara, 2018] Chamikara, Mahawaga Arachchige Pathum and Bertók, Peter and Liu, Dongxi and Camtepe, Seyit and Khalil, Ibrahim, *Efficient data perturbation for privacy preserving and accurate data stream mining* Pervasive and Mobile Computing, vol. 48, pp. 1–19, 2018.
- [Devi et al., 2020] Devi, Gayathri N and Manikandan, K, *Improved perturbation technique privacy-preserving rotation-based condensation algorithm for privacy preserving in big data stream using Internet of Things* Transactions on Emerging Telecommunications Technologies, vol. 31, no. 12, 2020.
- [Chaturvedi et al., 2008] Chaturvedi, Devendra K, *Soft computing* Studies in Computational intelligence, vol. 103, pp. 509–612, 2008.
- [Verykios et al., 2004] Verykios, Vassilios S and Bertino, Elisa and Fovino, Igor Nai and Provenza, Loredana Parasiliti and Saygin, Yucel and Theodoridis, Yannis, *State-of-the-art in privacy preserving data mining* ACM Sigmod Record, vol. 33, no. 1, pp. 50–57, 2004.
- [Narwaria, 2016] Narwaria, Mamta and Arya, Suchita, *Privacy preserving data mining—‘A state of the art’* 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) IEEE, pp. 2108–2112, 2016.
- [Saxena et al., 2017] Saxena, VK and Pushkar, Shashank, *Fuzzy-Based Privacy Preserving Approach in Centralized Database Environment* Advances in Computational Intelligence: Proceedings of International Conference on Computational Intelligence 2015, pp. 299–307, 2017.
- [Wafo Soh et al., 2020] Wafo Soh, C and Njilla, LL and Kwiat, KK and Kamhoua, CA, *Learning quasi-identifiers for privacy-preserving exchanges: A rough set theory approach* Granular Computing, vol. 5, pp. 71–84, 2020.
- [Zhou et al., 2009] Zhou, Zhengyou and Huang, Liusheng and Yun, Ye, *Privacy preserving attribute reduction based on rough set* 2009 Second International Workshop on Knowledge Discovery and Data Mining, pp. 202–206, 2009.
- [Zhao et al., 2022] Zhao, Bowen and Chen, Wei-Neng and Wei, Feng-Feng and Liu, Ximeng and Pei, Qingqi and Zhang, Jun, *Evolution as a service: A privacy-preserving genetic algorithm for combinatorial optimization* arXiv preprint arXiv:2205.13948, 2022.
- [Hevner et al., 2008] Hevner, Alan R and March, Salvatore T and Park, Jinsoo and Ram, Sudha, *Design science in information systems research* Management Information Systems Quarterly, vol. 28, no. 1, pp. 6, 2008.
- [Pinkas et al., 2002] Pinkas, Benny, *Cryptographic techniques for privacy-preserving data mining* ACM Sigkdd Explorations Newsletter, vol. 4, no. 2, pp. 12–19, 2002.
- [Subramani et al., 2023] Subramani, Shalini and Svn, Santhosh Kumar, *Review of security methods based on classical cryptography and quantum cryptography* Cybernetics and Systems, pp. 1–19, 2023.
- [Aldeen et al., 2015] Aldeen, Yousra Abdul Alsahib S and Salleh, Mazleena and Razzaque, Mohammad Abdur, *A comprehensive review on privacy preserving data mining* SpringerPlus, vol. 4, no. 1, pp. 1–36, 2015.
- [Chauhan et al., 2020] Chauhan, Aman Singh and Rani, Dikshika and Kumar, Akash and Gupta, Rishabh and Singh, Ashutosh Kumar, *A survey on privacy-preserving outsourced data on cloud with multiple data providers* Proceedings of the International Conference on Innovative Computing & Communications (ICICC), 2020.
- [Zhang et al., 2019] Zhang, Lili and Wang, Wenjie and Zhang, Yuqing, *Privacy preserving association rule mining: Taxonomy, techniques, and metrics* IEEE Access, vol. 7, pp. 45032–45047, 2019.
- [Nasiri et al., 2020] Nasiri, Negar and Keyvanpour, MohammadReza, *Classification and evaluation of privacy preserving data mining methods* 2020 11th International Conference on Information and Knowledge Technology (IKT), pp. 17–22, 2020.
- [Afzali et al., 2018] Afzali, Golnar Assadat and Mohammadi, Shahriar, *Privacy preserving big data mining: association rule hiding using fuzzy logic approach* IET Information Security, vol. 12, no. 1, pp. 15–24, 2018.
- [Barni et al., 2006] Barni, Mauro and Orlandi, Claudio and Piva, Alessandro, *A privacy-preserving protocol for neural-network-based computation* Proceedings of the 8th workshop on Multimedia and security, pp. 146–151, 2006.

- [Li et al., 2017] Li, Meng and Lai, Liangzhen and Suda, Naveen and Chandra, Vikas and Pan, David Z, *Privynet: A flexible framework for privacy-preserving deep neural network training* arXiv preprint arXiv:1709.06161, 2017.
- [Wu et al., 2019] Wu, Tsu-Yang and Lin, Jerry Chun-Wei and Zhang, Yuyu and Chen, Chun-Hao, *A grid-based swarm intelligence algorithm for privacy-preserving data mining* Applied Sciences, vol. 9, no. 4, pp. 774, 2019.
- [Domingo-Ferrer et al., 2005] Domingo-Ferrer, Josep and Torra, Vicenç, *Ordinal, continuous and heterogeneous k-anonymity through microaggregation* Data Mining and Knowledge Discovery, vol. 11, pp. 195–212, 2005.
- [Lin et al., 2008] Lin, Jun-Lin and Wei, Meng-Cheng and Li, Chih-Wen and Hsieh, Kuo-Chiang, *A hybrid method for k-anonymization* 2008 IEEE Asia-Pacific Services Computing Conference, pp. 385–390, 2008.
- [Aggarwal et al., 2008] Aggarwal, Charu C and Philip, S Yu, *Privacy-preserving data mining: models and algorithms* Springer Science & Business Media, 2008.
- [Thuraisingham et al., 2006] Thuraisingham, Bhavani and Kantarcioglu, Murat and Liu, Li, *The applicability of the perturbation model-based privacy preserving data mining for real-world data* Sixth IEEE International Conference on Data Mining-Workshops (ICDMW'06), pp. 507–512, 2006.
- [Xu et al., 2014] Xu, Lei and Jiang, Chunxiao and Wang, Jian and Yuan, Jian and Ren, Yong, *Information security in big data: privacy and data mining* Ieee Access, vol. 2, pp. 1149–1176, 2014.
- [Kerschbaum et al., 2014] Kerschbaum, Florian and Schneider, Thomas and Schröpfer, Axel, *Automatic protocol selection in secure two-party computations* Applied Cryptography and Network Security: 12th International Conference, ACNS 2014, Lausanne, Switzerland, June 10-13, 2014. Proceedings 12, pp. 566–584, 2014.
- [Shah et al., 2014] Shah, Mona and Joshi, Hiren D, *Privacy preserving data mining techniques in a distributed environment* International Journal of Computer Applications, vol. 94, no. 6, 2014.
- [McGregor et al., 2010] McGregor, Andrew and Mironov, Ilya and Pitassi, Toniann and Reingold, Omer and Talwar, Kunal and Vadhan, Salil, *The limits of two-party differential privacy* 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, pp. 81–90, 2010.
- [Sharma et al., 2010] Sharma, Anand and Ojha, Vibha, *Privacy Preserving Data Mining by Cryptography* Recent Trends in Network Security and Applications: Third International Conference, CNSA 2010, Chennai, India, July 23-25, 2010. Proceedings 3, pp. 576–582, 2010.
- [Tai et al., 2010] Tai, Chih-Hua and Yu, Philip S and Chen, Ming-Syan, *k-Support anonymity based on pseudo taxonomy for outsourcing of frequent itemset mining* Proceedings of the 16th ACM SIGKDD International Conference on Knowledge discovery and data mining, pp. 473–482, 2010.
- [Giannotti et al., 2012] Giannotti, Fosca and Lakshmanan, Laks VS and Monreale, Anna and Pedreschi, Dino and Wang, Hui, *Privacy-preserving mining of association rules from outsourced transaction databases* IEEE Systems Journal, vol. 7, no. 3, pp. 385–395, 2012.
- [SANTHOSHI et al., 2015] SANTHOSHI, PONNA and RAO, M ESWARA, *Privacy-Preserving Mining of Association Rules from Outsourced Transaction Databases* 2015.
- [Qiu et al., 2008] Qiu, Ling and Li, Yingjiu and Wu, Xintao, *Protecting business intelligence and customer privacy while outsourcing data mining tasks* Knowledge and Information Systems, vol. 17, no. 1, pp. 99–120, 2008.
- [Nithin et al., 2017] Nithin, C and Mohan, AV, *Privacy-Preserving in FiDooP, Mining of Frequent Itemsets from Outsourced Transaction Databases* International Journal of Innovative Research in Computer Science & Technology, vol. 5, 2017.
- [Wu et al., 2006] Wu, Yi-Hung and Chiang, Chia-Ming and Chen, Arbee LP, *Hiding sensitive association rules with limited side effects* IEEE Transactions on Knowledge and Data engineering, vol. 19, no. 1, pp. 29–42, 2006.
- [Wang et al., 2013] Wang, Hui, *Quality measurements for association rules hiding* AASRI Proceedings, vol. 5, pp. 228–234, 2013.
- [Moustakides et al., 2008] Moustakides, George V and Verykios, Vassilios S, *A maxmin approach for hiding frequent itemsets* Data & Knowledge Engineering, vol. 65, no. 1, pp. 75–89, 2008.

- [Dasseni et al., 2001] Dasseni, Elena and Verykios, Vassilios S and Elmagarmid, Ahmed K and Bertino, Elisa, *Hiding association rules by using confidence and support* Information Hiding: 4th International Workshop, IH 2001 Pittsburgh, PA, USA, April 25–27, 2001 Proceedings 4, pp. 369–383, 2001.
- [Verykios et al., 2008] Verykios, Vassilios S and Gkoulalas-Divanis, Aris, *A survey of association rule hiding methods for privacy* Privacy-Preserving Data Mining: Models and Algorithms, pp. 267–289, 2008.
- [Afzali et al., 2016] Afzali, Golnar Assadat and Mohammadi, Shahriar, *Privacy preserving big data mining: association rule hiding* JOURNAL OF INFORMATION SYSTEMS AND TELECOMMUNICATION (JIST), 2016.
- [chandrakar et al., 2010] Chandrakar, Ila and Rani, Yelipe Usha and Manasa, Mortha and Renuka, Kondabala, *Hybrid algorithm for privacy preserving association rule mining* Journal of Computer Science, vol. 6, no. 12, pp. 1494, 2010.
- [Jain et al., 2012] Jain, Dhyanendra and Khatri, Pallavi and Soni, Rishi and Chaurasia, Brijesh Kumar, *Hiding sensitive association rules without altering the support of sensitive item (s)* Advances in Computer Science and Information Technology. Networks and Communications: Second International Conference, CCSIT 2012, Bangalore, India, January 2-4, 2012. Proceedings, Part I 2, pp. 500–509, 2012.
- [Rao et al., 2014] Rao, K Srinivasa and Mandhala, Venkata Naresh and Bhattacharyya, Debnath and Kim, Tai-hoon, *An association rule hiding algorithm for privacy preserving data mining* International Journal of Control and Automation, vol. 7, no. 10, pp. 393–404, 2014.
- [Yang et al., 2005] Yang, Zhiqiang and Zhong, Sheng and Wright, Rebecca N, *Privacy-preserving classification of customer data without loss of accuracy* Proceedings of the 2005 SIAM International Conference on Data Mining, pp. 92–102, 2005.
- [Zhang et al., 2005] Zhang, Nan and Wang, Shengquan and Zhao, Wei, *A new scheme on privacy-preserving data classification* Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pp. 374–383, 2005.
- [Zhang et al., 2017] Zhang, Lin and Liu, Yan and Wang, Ruchuan and Fu, Xiong and Lin, Qiaomin, *Efficient privacy-preserving classification construction model with differential privacy technology* Journal of Systems Engineering and Electronics, vol. 28, no. 1, pp. 170–178, 2017.
- [Delis et al., 2010] Delis, Aggelos and Verykios, Vassilios S and Tsitsonis, Achilleas A, *A data perturbation approach to sensitive classification rule hiding* Proceedings of the 2010 ACM symposium on applied computing, pp. 605–609, 2010.
- [Gupta et al., 2011] Gupta, Saurabh and Im, Hong G and Valorani, Mauro, *Classification of ignition regimes in HCCI combustion using computational singular perturbation* Classification of ignition regimes in HCCI combustion using computational singular perturbation, vol. 33, no. 2, pp. 2991–2999, 2011.
- [Patal et al., 2014] Patel, Kiran and Patel, Hitesh and Patel, Parin, *Privacy Preserving in Data stream classification using different proposed Perturbation Methods* Citeseer, 2014.
- [Fung et al., 2007] Fung, Benjamin CM and Wang, Ke and Philip, S Yu, *Anonymizing classification data for privacy preservation* IEEE transactions on knowledge and data engineering, vol. 19, no. 5, pp. 711–725, 2007.
- [Yu et al., 2006] Yu, Hwanjo and Vaidya, Jaideep and Jiang, Xiaoqian, *Privacy-preserving svm classification on vertically partitioned data* Advances in Knowledge Discovery and Data Mining: 10th Pacific-Asia Conference, PAKDD 2006, Singapore, April 9-12, 2006. Proceedings 10, pp. 647–656, 2006.
- [Tripathy et al., 2012] Tripathy, Animesh and Dansana, Jayanti and Mishra, Ranjita, *A classification based framework for privacy preserving data mining* Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 660–666, 2012.
- [Raghuram et al., 2012] Raghuram, B and Gyani, Jayadev, *Privacy preserving associative classification on vertically partitioned databases* 2012 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCT), pp. 188–192, 2012.
- [Jia et al., 2016] Jia, Qi and Guo, Linke and Jin, Zhanpeng and Fang, Yuguang, *Privacy-preserving data classification and similarity evaluation for distributed systems* 2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS), pp. 690–699, 2016.

- [Kumari et al., 2008] Kumari, V Valli and Rao, S Srinivasa and Raju, KVSVN and Ramana, KV and Avadhani, BVS, *Fuzzy based approach for privacy preserving publication of data* International Journal of Computer Science and Network Security, vol. 8, no. 1, pp. 115-121, 2008.
- [Poovammal et al., 2009] Poovammal, E and Ponnaivaikko, M, *Preserving micro data release: categorical and numerical data* IEEE Setit, 2009.
- [Mukkamala et al., 2011] Mukkamala, Ravi and Ashok, Vikas G, *Fuzzy-based methods for privacy-preserving data mining* 2011 Eighth International Conference on Information Technology: New Generations, pp. 348–353, 2011.
- [Jahan et al., 2016] Jahan, Thanveer and Narsimha, Gugulothu and Rao, CV Guru, *Multiplicative data perturbation using fuzzy logic in preserving privacy* Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, pp. 1–5, 2016.
- [Ahmed et al., 2014] Ahmed, SMT and Haque, Shameemul and Tauhid, SM Faizanut, *A fuzzy based approach for privacy preserving clustering* Int. J. Sci. Eng. Res, vol. 5, no. 2, pp. 1067–1071, 2014.
- [Torra et al., 2008] Torra, Vicenç and Miyamoto, Sadaaki and Endo, Yasunori and Domingo-Ferrer, Josep, *On intuitionistic fuzzy clustering for its application to privacy* 2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence), pp. 1042–1048, 2008.
- [Kasugai et al., 2013] Kasugai, Hirohide and Kawano, Arina and Honda, Katsuhiro and Notsu, Akira, *A study on applicability of fuzzy k-member clustering to privacy preserving pattern recognition* 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–6, 2013.
- [Kumar et al., 2011] Kumar, Pradeep and Varma, Kishore Indukuri and Sureka, Ashish, *Fuzzy based clustering algorithm for privacy preserving data mining* International Journal of Business Information Systems, vol. 7, no. 1, pp. 21–40, 2011.
- [Honda et al., 2012] Honda, Katsuhiro and Kawano, Arina and Notsu, Akira and Ichihashi, Hidetomo, *A fuzzy variant of k-member clustering for collaborative filtering with data anonymization* 2012 IEEE international conference on fuzzy systems, pp. 1–6, 2012.
- [Samet et al., 2012] Samet, Saeed and Miri, Ali, *Privacy-preserving back-propagation and extreme learning machine algorithms* Data & Knowledge Engineering, vol. 79, pp. 40–61, 2012.
- [Malik et al., 2015] Malik, Majid Bashir and Asger, M and Ali, Rashid and Sarvar, Abid, *A model for privacy preserving in data mining using Soft Computing techniques* 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 181–186, 2015.
- [Zadeh et al., 1996] Zadeh, Lotfi A, *Fuzzy logic, neural networks, and soft computing* Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh, pp. 775–782, 1996.
- [Wimmer et al., 2014] Wimmer, Hayden and Powell, Loreen, *A comparison of the effects of k-anonymity on machine learning algorithms* International Journal of Advanced Computer Science and Applications, vol. 5, no. 11, 2014.
- [Malik et al., 2016] Malik, Majid Bashir and Asger, M and Ali, Rashid and Arif, Tasleem, *Fuzzy Framework for Preserving Privacy in Neural Networks Classification* International Journal of Computer Applications, vol. 144, no. 11, 2016.
- [Grefenstette et al., 1993] Grefenstette, John J, *Genetic algorithms and machine learning* Proceedings of the sixth annual conference on Computational learning theory, pp. 3–4, 1993.
- [Dehkordi et al., 2009] Dehkordi, Mohammad Naderi and Badie, Kambiz and Zadeh, Ahmad Khadem, *A Novel Method for Privacy Preserving in Association Rule Mining Based on Genetic Algorithms*. J. Softw., vol. 4, no. 6, pp. 555-562, 2009.
- [Matatov et al., 2010] Matatov, Nissim and Rokach, Lior and Maimon, Oded, *Privacy-preserving data mining: A feature set partitioning approach* Information Sciences, vol. 180, no. 14, pp. 2696–2720, 2010.
- [Keshavamurthy et al., 2013] Keshavamurthy, Bettahally N and Khan, Asad M and Toshniwal, Durga, *Privacy preserving association rule mining over distributed databases using genetic algorithm* Neural Computing and Applications, vol. 25, pp. 351–364, 2013.

- [Mandapati et al., 2013] Mandapati, Sridhar and Bhogapathi, Raveendra Babu and Chekka, Ratna Babu, *A hybrid algorithm for privacy preserving in data mining* Int J Intell Syst Appl, vol. 5, no. 8, pp. 47–53, 2013.
- [Zhu et al., 2009] Zhu, Dan and Li, Xiao-Bai and Wu, Shuning, *Identity disclosure protection: A data reconstruction approach for privacy-preserving data mining* Decision Support Systems, vol. 48, no. 1, pp. 133–140, 2009.
- [Hong et al., 2010] Hong, Tzung-Pei and Yang, Kuo-Tung and Lin, Chun-Wei and Wang, Shyue-Liang, *Evolutionary privacy-preserving data mining* 2010 World Automation Congress, pp. 1–7, 2010.
- [Pawlak et al., 1995] Pawlak, Zdzislaw and Grzymala-Busse, Jerzy and Slowinski, Roman and Ziarko, Wojciech, *Rough sets* Communications of the ACM, vol. 38, no. 11, pp. 88–95, 1995.
- [Hu et al., 2016] Hu, Jun and Pedrycz, Witold and Wang, Guoyin and Wang, Kai, *Rough sets in distributed decision information systems* Knowledge-Based Systems, vol. 94, pp. 13–22, 2016.
- [Ye et al., 2013] Ye, Mingquan and Wu, Xindong and Hu, Xuegang and Hu, Donghui, *Anonymizing classification data using rough set theory* Knowledge-Based Systems, vol. 43, pp. 82–94, 2013.
- [Hu et al., 2008] Hu, Dan and Yu, Xianchuan and Feng, Yuanfu, *Distributed mining reducts of attributes on horizontally partitioned data* 2008 Second International Symposium on Intelligent Information Technology Application, vol. 1, pp. 55–59, 2008.
- [Wang et al., 2009] Wang, Jian and Luo, Yongcheng and Zhao, Yan and Le, Jiajin, *A survey on privacy preserving data mining* 2009 First International Workshop on Database Technology and Applications, pp. 111–114, 2009.
- [Natarajan et al., 2012] Rajesh Natarajan and Dr. Ramalingam Sugumar and M. Mahendran and Karthikeyan Anbazhagan, *A survey on Privacy Preserving Data Mining* 2012.
- [Nayak et al., 2011] Nayak, Gayatri and Devi, Swagatika, *A survey on privacy preserving data mining: approaches and techniques* International Journal of Engineering Science and Technology, vol. 3, no. 3, pp. 2127–2133, 2011.
- [Marx et al., 2013] Marx, Vivien, *The big challenges of big data* Nature, vol. 498, no. 7453, pp. 255–260, 2013.
- [Sánchez et al., 2016] Sánchez, David and Martínez, Sergio and Domingo-Ferrer, Josep, *Comment on “Unique in the shopping mall: On the reidentifiability of credit card metadata”* Science, vol. 351, no. 6279, pp. 1274–1274, 2016.
- [de Lima et al., 2016] de Lima, Danilo Alves and Victorino, Alessandro Corrêa, *A hybrid controller for vision-based navigation of autonomous vehicles in urban environments* IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 8, pp. 2310–2323, 2016.
- [S Yu 2016] Yu, Shui, *Big privacy: Challenges and opportunities of privacy study in the age of big data* IEEE Access, vol. 4, pp. 2751–2763, 2016.
- [Xu et al., 2016] Xu, Lei and Jiang, Chunxiao and Chen, Yan and Wang, Jian and Ren, Yong, *A framework for categorizing and applying privacy-preservation techniques in big data mining* Computer, vol. 49, no. 2, pp. 54–62, 2016.
- [Majeed et al., 2020] Majeed, Abdul and Lee, Sungchang, *Anonymization techniques for privacy preserving data publishing: A comprehensive survey* IEEE access, vol. 9, pp. 8512–8545, 2020.
- [Chandrakar et al., 2020] Chandrakar, Ila and Hulipalled, Vishwanath R, *Improved technique for preserving privacy while mining real time big data* International Journal of Communication Networks and Information Security, vol. 14, no. 1, pp. 86–92, 2020.
- [Mendes et al., 2017] Mendes, Ricardo and Vilela, João P, *Privacy-preserving data mining: methods, metrics, and applications* IEEE Access, vol. 5, pp. 10562–10582, 2017.
- [Siraj et al., 2019] Siraj, Maheyyah Md and Rahmat, Nurul Adibah and Din, Mazura Mat, *A survey on privacy preserving data mining approaches and techniques* Proceedings of the 2019 8th international conference on software and computer applications, pp. 65–69, 2019.
- [Dua et al., 2016] Dua, Sumeet and Du, Xian, *Data mining and machine learning in cybersecurity*, 2016.
- [Kumar et al., 2022] Kumar, Rajiv and Singh, Kuldee, *A survey on soft computing-based high-utility itemsets mining* Soft Computing, vol. 26, no. 13, pp. 6347–6392, 2022.
- [Sivanadam et al., 2008] Sivanadam, SN and Deepa, SN and Sivanadam, SN and Deepa, SN, *Genetic algorithm optimization problems* Introduction to genetic algorithms, pp. 165–209, 2008.

- [Srinivas et al., 1994] Srinivas, Mandavilli and Patnaik, Lalit M, *Genetic algorithms: A survey* Computer, vol. 27, no. 6, pp. 17–26, 1994.
- [Dua et al., 2017] Dua, Dheeru and Graff, Casey and others, *UCI machine learning repository* Beijing China, 2017.
- [Han et al., 2007] Han, Shuguo and Ng, Wee Keong, *Privacy-preserving self-organizing map* International Conference on Data Warehousing and Knowledge Discovery, pp. 428–437, 2007.
- [Malik et al., 2016] Majid Bashir Malik, M. Asger, Rashid Ali, Tasleem Arif, *Preserving Privacy and Optimizing Neural Network Classification by using a Mix of Soft Computing Techniques* International Journal of Computer Applications, vol. 147, no. 10, pp. 1–7, 2016.
- [Amiri et al., 2018] Amiri, Fatemeh and Quirchmayr, Gerald and Kieseberg, Peter, *A Machine Learning Approach for Privacy-preservation in E-business Applications*. ICETE (2), pp. 609–618, 2018.
- [Lin et al., 2014] Lin, Chun-Wei and Zhang, Binbin and Yang, Kuo-Tung and Hong, Tzung-Pei and others, *Efficiently hiding sensitive itemsets with transaction deletion based on genetic algorithms* The Scientific World Journal, 2014.
- [Amiri et al., 2018] Amiri, Fatemeh and Quirchmayr, Gerald and Kieseberg, Peter, *Sensitive data anonymization using genetic algorithms for SOM-based clustering* 2018.
- [Yang et al., 2012] Yang, Yan and Tan, Wei and Li, Tianrui and Ruan, Da, *Consensus clustering based on constrained self-organizing map and improved Cop-Kmeans ensemble in intelligent decision support systems* Knowledge-Based Systems, vol. 32, pp. 101–115, 2012.
- [Gan et al., 2020] Gan, Guojun and Ma, Chaoqun and Wu, Jianhong, *Data clustering: theory, algorithms, and applications* SIAM, 2020.
- [Haykin 1998] Haykin, Simon, *Neural networks: a comprehensive foundation* Prentice Hall PTR, 1998.
- [Tan et al., 2016] Tan, Pang-Ning and Steinbach, Michael and Kumar, Vipin, *Introduction to data mining* Pearson Education India, 2016.
- [Fung et al., 2010] Fung, Benjamin CM and Wang, Ke and Chen, Rui and Yu, Philip S, *Privacy-preserving data publishing: A survey of recent developments* ACM Computing Surveys (Csur), vol. 42, no. 4, pp. 1–53, 2010.
- [Gkoulalas-Divanis et al., 2014] Gkoulalas-Divanis, Aris and Loukides, Grigorios and Sun, Jimeng, *Publishing data from electronic health records while preserving privacy: A survey of algorithms* Journal of biomedical informatics, vol. 50, pp. 4–19, 2014.
- [Kohonen, 2001] Kohonen, T, *Self-Organizing Maps*, NY., Springer-Verlag 2001.
- [Kiviluoto, 1996] Kiviluoto, Kimmo, *Topology preservation in self-organizing maps* Proceedings of International Conference on Neural Networks (ICNN'96), vol. 1, pp. 294–299, IEEE, 1996.
- [Holland, 1975] Holland John, *Adaptation in artificial and natural systems* Ann Arbor: The University of Michigan Press, 232, 1975.
- [Moler, 1980] Moler, Cleve, *Design of an interactive matrix calculator* Proceedings of the May 19-22, 1980, national computer conference, pp. 363–368, 1980.
- [Asuncion, 2007] Asuncion, Arthur, *Uci machine learning repository, university of california, irvine, school of information and computer sciences* <http://www.ics.uci.edu/mllearn/MLRepository>. HTML, 2007.
- [Zadeh, 1996] Zadeh, Lotfi A, *Fuzzy logic, neural networks, and soft computing* Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh, pp. 775–782, World Scientific, 1996.
- [Moro et al., 2011] Moro, Sergio and Laureano, Raul and Cortez, Paulo, *Using data mining for bank direct marketing: An application of the crisp-dm methodology* EUROSIS-ETI, 2011.
- [Kaleli et al., 2012] Kaleli, Cihan and Polat, Huseyin, *Privacy-preserving SOM-based recommendations on horizontally distributed data* Knowledge-Based Systems, vol. 33, pp. 124–135, Elsevier, 2012.
- [Perugini et al., 2004] Perugini, Saverio and Gonçalves, Marcos André and Fox, Edward A, *Recommender systems research: A connection-centric survey* Journal of Intelligent Information Systems, vol. 23, pp. 107–143, 2004.
- [Warren and Brandeis, 1890] Warren, Samuel D. and Brandeis, Louis D., *The Right to Privacy*, Harvard Law Review, vol. 4, no. 5, pp. 193–220, 1890.

- [Amiri et al., 2019] Amiri, Fatemeh and Quirchmayr, Gerald and Kieseberg, Peter and Weippl, Edgar and Bertone, Alessio, *Towards Data Anonymization in Data Mining via Meta-Heuristic Approaches* Data Privacy Management, Cryptocurrencies and Blockchain Technology: ESORICS 2019 International Workshops, DPM 2019 and CBT 2019, Luxembourg, September 26–27, 2019, Proceedings 14, pp. 39–48, 2019.
- [Amiri and Quirchmayr, 2025] Amiri, Fatemeh and Quirchmayr, Gerald, *GAFSOM: A Privacy-Preserving Machine Learning Algorithm for SOM Clustering* International Journal of Information Security, 24(6):223, 2025.
- [Tex et al., 2018] Tex, Christine and Schäler, Martin and Böhm, Klemens, *2018 IEEE 34th International Conference on Data Engineering (ICDE)* IEEE, pp. 1264–1267, 2018.
- [Van Dijk et al., 2010] Van Dijk, Marten and Juels, Ari, *On the impossibility of cryptography alone for {Privacy-Preserving} cloud computing* 5th USENIX Workshop on Hot Topics in Security (HotSec 10), 2010.
- [Zhao et al., 2018] Zhao, Ping and Li, Jie and Zeng, Fanzi and Xiao, Fu and Wang, Chen and Jiang, Hongbo, *ILLIA: Enabling k-anonymity-based privacy preserving against location injection attacks in continuous LBS queries* IEEE Internet of Things Journal, vol. 5, no. 2, pp. 1033–1042, 2018.
- [Leontiadis et al., 2015] Leontiadis, Iraklis and Elkhayaoui, Kaoutar and Önen, Melek and Molva, Refik, *PUDA—privacy and unforgeability for data aggregation* Cryptology and Network Security: 14th International Conference, CANS 2015, Marrakesh, Morocco, December 10–12, 2015, Proceedings 14, pp. 3–18, 2015.
- [Dwork, 2006] Dwork, Cynthia, *Differential privacy* International colloquium on automata, languages, and programming, Springer, pp. 1–12, 2006.
- [Qardaji et al., 2013] Qardaji, Wahbeh and Yang, Weining and Li, Ninghui, *Differentially private grids for geospatial data* 2013 IEEE 29th international conference on data engineering (ICDE), pp. 757–768.
- [Amiri et al., 2017] Amiri, Fatemeh and Quirchmayr, Gerald, *A comparative study on innovative approaches for privacy-preservation in knowledge discovery* Proceedings of the 9th International Conference on Information Management and Engineering, pp. 120–127, 2017.
- [Kumar et al., 2010] Kumar, Usha A and Dhamija, Yuvnish, *Comparative analysis of SOM neural network with K-means clustering algorithm* 2010 IEEE International Conference on Management of Innovation & Technology, IEEE, pp. 55–59, 2010.
- [Riveros et al., 2019] Riveros, Nicolas Andres Melo and Espitia, Bayron Alexis Cardenas and Pico, Lilia Edith Aparicio, *Comparison between K-means and self-organizing maps algorithms used for diagnosis spinal column patients* Informatics in Medicine Unlocked, Elsevier, vol. 16, pp. 100206, 2019.
- [Zhu et al., 2017] Zhu, Tianqing and Li, Gang and Zhou, Wanlei and Philip, S Yu, *Differential privacy and applications* Springer, 2017.
- [Ji et al., 2014] Ji, Zhanglong and Lipton, Zachary C and Elkan, Charles, *Differential privacy and machine learning: a survey and review* arXiv preprint arXiv:1412.7584, 2014.
- [Wood et al., 2018] Wood, Alexandra and Altman, Micah and Bembenek, Aaron and Bun, Mark and Gaboardi, Marco and Honaker, James and Nissim, Kobbi and O’Brien, David R and Steinke, Thomas and Vadhan, Salil, *Differential privacy: A primer for a non-technical audience* Vand. J. Ent. & Tech. L., HeinOnline, vol. 21, pp. 209, 2018.
- [Dwork et al., 2014] Dwork, Cynthia and Roth, Aaron and others, *The algorithmic foundations of differential privacy* Now Publishers, Inc., vol. 9, no. 3–4, pp. 211–407, 2014.
- [Balcer, 2022] Balcer, Victor, *Exploring Models for Implementing Differential Privacy* Harvard University, 2022.
- [Pan et al., 2024] Pan, Ke and Ong, Yew-Soon and Gong, Maoguo and Li, Hui and Qin, AK and Gao, Yuan, *Differential privacy in deep learning: A literature survey* Neurocomputing, Elsevier, pp. 127663, 2024.
- [Su et al., 2001] Su, Mu-Chun and Chou, Chien-Hsing, *A modified version of the K-means algorithm with a distance based on cluster symmetry* IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE, vol. 23, no. 6, pp. 674–680, 2001.

- [Noman et al., 2011] Mohammed, Noman and Chen, Rui and Fung, Benjamin CM and Yu, Philip S, *Differentially private data release for data mining* Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 493–501, 2011.
- [Oja et al., 1999] Oja, Erkki and Kaski, Samue, *Kohonen maps* Elsevier, 1999.
- [Blum et al., 2005] Blum, Avrim and Dwork, Cynthia and McSherry, Frank and Nissim, Kobbi, *Practical privacy: the SuLQ framework* Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, pp. 128–135, 2005.
- [Su et al., 2016] Su, Dong and Cao, Jianneng and Li, Ninghui and Bertino, Elisa and Jin, Hongxia, *Differentially private k-means clustering* Proceedings of the sixth ACM conference on data and application security and privacy, pp. 26–37, 2016.
- [Holohan et al., 2019] Holohan, Naoise and Braghin, Stefano and Mac Aonghusa, Pól and Leveacher, Killian, *Diffprivlib: the IBM differential privacy library* arXiv preprint arXiv:1907.02444, 2019.
- [Zhang et al., 2013] Zhang, Jun and Xiao, Xiaokui and Yang, Yin and Zhang, Zhenjie and Winslett, Marianne, *Privgene: differentially private model fitting using genetic algorithms* Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pp. 665–676, 2013.
- [Dua et al., 2019] Dua, D and Graff, C, *UCI machine learning repository* [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science IEEE transactions on pattern analysis and machine intelligence, vol. 1, no. 1, pp. 1–29, 2019.
- [Lecun et al., 1998] Lecun, Y. and Bottou, L. and Bengio, Y. and Haffner, P., *Gradient-based learning applied to document recognition* Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
- [Vedantam et al., 2014] Vedantam, Ramakrishna and Zitnick, C Lawrence and Parikh, Devi, *Collecting image description datasets using crowdsourcing* arXiv preprint arXiv:1411.3041, 2014.
- [Pruthvi et al., 2015] Pruthvi, HR and Nisha, KK and Chandana, TL and Navami, K and Biju, RM, *Feature engineering on forest cover type data with ensemble of decision trees* 2015 IEEE International Advance Computing Conference (IACC), pp. 1093–1098, 2015.
- [Bright et al., 2020] Bright, Chloe J and Lawton, Sarah and Benson, Stephen and Bomb, Martine and Dodwell, David and Henson, Katherine E and McPhail, Sean and Miller, Louise and Rashbass, Jem and Turnbull, Alice and others, *Corrigendum to: data resource profile: the systemic anti-cancer therapy (SACT) Dataset* International Journal of Epidemiology, vol. 49, no. 1, pp. 354, 2020.
- [Dinur et al., 2003] Dinur, Irit and Nissim, Kobbi, *Revealing information while preserving privacy* Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, pp. 202–210, 2003.
- [Dwork et al., 2004] Dwork, Cynthia and Nissim, Kobbi, *Privacy-Preserving Datamining on Vertically Partitioned Databases* Advances in Cryptology–CRYPTO 2004: 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15–19, 2004. Proceedings 24, pp. 528–544, 2004.
- [McSherry et al., 2007] McSherry, Frank and Talwar, Kunal, *Mechanism design via differential privacy* 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS’07), IEEE, pp. 94–103, 2007.
- [Ma et al., 2021] Ma, Fei and Wang, Ping, *Randomized response mechanisms for differential privacy data analysis: Bounds and applications* arXiv preprint arXiv:2112.07397, 2021.
- [Kasiviswanathan et al., 2011] Kasiviswanathan, Shiva Prasad and Lee, Homin K and Nissim, Kobbi and Raskhodnikova, Sofya and Smith, Adam, *What can we learn privately?* SIAM Journal on Computing, vol. 40, no. 3, pp. 79–826, 2011.
- [Rana et al., 2015] Rana, Santu and Gupta, Sunil Kumar and Venkatesh, Svetha, *Differentially private random forest with high utility* 2015 IEEE International Conference on Data Mining, IEEE, pp. 955–960, 2015.
- [Ni et al., 2018] Ni, Lina and Li, Chao and Wang, Xiao and Jiang, Honglu and Yu, Jiguo, *DP-MCDBSCAN: Differential privacy preserving multi-core DBSCAN clustering for network user data* IEEE access, vol. 6, pp. 21053–21063.
- [Chaudhuri et al., 2008] Chaudhuri, Kamalika and Monteleoni, Claire, *Privacy-preserving logistic regression* Advances in neural information processing systems, vol. 21, 2008.

- [Chaudhuri et al., 2011] Chaudhuri, Kamalika and Monteleoni, Claire and Sarwate, Anand D, *Differentially private empirical risk minimization*. Journal of Machine Learning Research, vol. 12, no. 3, 2011.
- [Mohan et al., 2012] Mohan, Prashanth and Thakurta, Abhradeep and Shi, Elaine and Song, Dawn and Culler, David, *GUPT: privacy preserving data analysis made easy* Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, pp. 349–360, 2012.
- [Diaa et al., 2024] Diaa, Abdulrahman and Humphries, Thomas and Kerschbaum, Florian, *FastLloyd: Federated, Accurate, Secure, and Tunable k-Means Clustering with Differential Privacy* arXiv preprint arXiv:2405.02437, 2024.
- [Burke et al., 2011] Burke, Robin and Felfernig, Alexander and Göker, Mehmet H, *Recommender systems: An overview* Ai Magazine, nol. 32, no. 3, pp. 13–18, 2011.
- [Roy et al., 2022] Roy, Deepjyoti and Dutta, Mala, *A systematic review and research perspective on recommender systems* Journal of Big Data, vol. 9, no. 1, pp. 59, 2022.
- [Aggarwal et al., 2016] Aggarwal, Charu C and others, *Recommender systems* Springer, vol. 1, 2016.
- [Bobadilla et al., 2009] Bobadilla, JESUS and Serradilla, Francisco and Hernando, Antonio and others, *Collaborative filtering adapted to recommender systems of e-learning* Knowledge-Based Systems, Elsevier, vol. 22, no. 4, pp. 261–265, 2009.
- [Papadakis et al., 2022] Papadakis, Harris and Papagrigoriou, Antonis and Panagiotakis, Costas and Kosmas, Eleftherios and Fragopoulou, Paraskevi, *Collaborative filtering recommender systems taxonomy* Knowledge and Information Systems, vol. 64, no. 1, pp. 35–74, 2022.
- [Su et al., 2009] Su, Xiaoyuan and Khoshgoftaar, Taghi M, *A survey of collaborative filtering techniques* Advances in artificial intelligence, Wiley Online Library, no. 1, pp. 421–425, 2009.
- [Khusro, Shah et al., 2016] Khusro, Shah and Ali, Zafar and Ullah, Irfan, *Recommender systems: issues, challenges, and research opportunities* Information science and applications (ICISA) 2016, Springer, pp. 1179–1189, 2016.
- [Knijnenburg et al., 2017] Knijnenburg, Bart P and Berkovsky, Shlomo, *Privacy for recommender systems: tutorial abstract* Proceedings of the eleventh ACM conference on recommender systems, pp. 394–395, 2017.
- [Himeur et al., 2022] Himeur, Yassine and Sohail, Shahab Saquib and Bensaali, Faycal and Amira, Abbes and Alazab, Mamoun, *Latest trends of security and privacy in recommender systems: a comprehensive review and future perspectives* Computers & Security, Elsevier, vol. 118, pp. 102746, 2022.
- [Batmaz et al., 2017] Batmaz, Zeynep and Kaleli, Cihan, *Methods of privacy preserving in collaborative filtering* 2017 International Conference on Computer Science and Engineering (UBMK), IEEE, pp. 261–266, 2017.
- [Ge et al., 2023] Ge, Lina and Li, Haiao and Wang, Xiao and Wang, Zhe, *A review of secure federated learning: privacy leakage threats, protection technologies, challenges and future directions* Neurocomputing, Elsevier, pp. 126897, 2023.
- [Qian et al., 2013] Qian, Xueming and Feng, He and Zhao, Guoshuai and Mei, Tao, *Personalized recommendation combining user interest and social circle* IEEE transactions on knowledge and data engineering, vol. 26, no. 7, pp. 1763–1777, 2013.
- [Meng et al., 2018] Meng, Xuying and Wang, Suhang and Shu, Kai and Li, Jundong and Chen, Bo and Liu, Huan and Zhang, Yujun, *Personalized privacy-preserving social recommendation* Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, 2018.
- [Chai et al., 2014] Chai, Tianfeng and Draxler, Roland R and others, *Root mean square error (RMSE) or mean absolute error (MAE)* Geoscientific model development discussions, vol. 7, no. 1, pp. 1525–1534, 2014.
- [Zhu et al., 2016] Zhu, Xue and Sun, Yuqing, *Differential privacy for collaborative filtering recommender algorithm* Proceedings of the 2016 ACM on international workshop on security and privacy analytics, pp. 9–16, 2016.
- [Müllner et al., 2023] Müllner, Peter and Lex, Elisabeth and Schedl, Markus, *Differential privacy in collaborative filtering recommender systems: a review* Frontiers in big Data, vol. 6, pp. 1249997, 2023.

- [Lu et al., 2019] Lu, Yi-Shu and Shih, Wen-Yueh and Gau, Hung-Yi and Chung, Kuan-Chieh and Huang, Jiun-Long, *On successive point-of-interest recommendation* World Wide Web, Springer, vol. 22, pp. 1151–1173, 2019.
- [Jain et al., 2016] Jain, Priyank and Gyanchandani, Manasi and Khare, Nilay, *Big data privacy: a technological perspective and review* Journal of Big Data, Springer, vol. 3, pp. 1–25, 2016.
- [Kalra et al., 2023] Kalra, Shivam and Wen, Junfeng and Cresswell, Jesse C and Volkovs, Maksims and Tizhoosh, Hamid R, *Decentralized federated learning through proxy model sharing* Nature communications, Nature Publishing Group UK London, vol. 14, no. 1, pp. 2899, 2023.
- [Baek et al., 2021] Baek, Chunghun and Kim, Sungwook and Nam, Dongkyun and Park, Jihoon, *Enhancing differential privacy for federated learning at scale* IEEE Access, vol. 9, pp. 148090–148103, 2021.
- [Catherine et al., 2017] Catherine, Rose and Cohen, William, *Transnets: Learning to transform for recommendation* Proceedings of the eleventh ACM conference on recommender systems, pp. 288–296, 2017.
- [He et al., 2017] He, Xiangnan and Liao, Lizi and Zhang, Hanwang and Nie, Liqiang and Hu, Xia and Chua, Tat-Seng, *Neural collaborative filtering* Proceedings of the 26th international conference on world wide web, pp. 173–182, 2017.
- [Salakhutdinov et al., 2007] Salakhutdinov, Ruslan and Mnih, Andriy and Hinton, Geoffrey, *Restricted Boltzmann machines for collaborative filtering* Proceedings of the 24th international conference on Machine learning, pp. 791–798, 2007.
- [Van den Oord et al., 2013] Van den Oord, Aaron and Dieleman, Sander and Schrauwen, Benjamin, *Deep content-based music recommendation* Advances in neural information processing system, vol. 26, 2013.
- [Rama et al., 2019] Rama, Kiran and Kumar, Pradeep and Bhasker, Bharat, *Deep learning to address candidate generation and cold start challenges in recommender systems: A research survey* arXiv preprint arXiv:1907.08674, 2019.
- [Kiran et al., 2020] Kiran, R and Kumar, Pradeep and Bhasker, Bharat, *DNNRec: A novel deep learning based hybrid recommender system* Expert Systems with Applications, Elsevier, vol. 144, pp. 113054, 2020.
- [Liu et al., 2017] Liu, Xiao and Liu, An and Zhang, Xiangliang and Li, Zhixu and Liu, Guanfeng and Zhao, Lei and Zhou, Xiaofang, *When differential privacy meets randomized perturbation: a hybrid approach for privacy-preserving recommender system* Database Systems for Advanced Applications: 22nd International Conference, DASFAA 2017, Suzhou, China, March 27–30, 2017, Proceedings, Part I 22, pp. 576–591, 2017.
- [Yin et al., 2020] Yin, Chunyong and Shi, Lingfeng and Sun, Ruxia and Wang, Jin, *Improved collaborative filtering recommendation algorithm based on differential privacy protection* The Journal of Supercomputing, Springer, vol. 76, no. 7, pp. 5161–5174, 2020.
- [chen et al., 2020] Chen, Kanglin and Chen, Ying-Ju and Gallego, Guillermo and Gao, Pin and Liu, Haoyu, *Personalized Product Recommendations with Exponential Bandits* Available at SSRN 3724377, 2020.
- [Hua et al., 2015] J. Hua, C. Xia, and S. Zhong, *Differentially private matrix factorization* Proc. 7th Int. Joint Conf. Artif. Intell., pp. 1763–1770, 2015.
- [Ezugwu et al., 2022] Ezugwu, Absalom E and Ikotun, Abiodun M and Oyelade, Olaide O and Abualigah, Laith and Agushaka, Jeffery O and Eke, Christopher I and Akinyelu, Andronicus A, *A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects* Engineering Applications of Artificial Intelligence, vol. 110, pp. 104743, 2022.
- [Wang et al., 2025] Wang, Weiwei and Ma, Wenping and Yan, Kun, *FSPPCFs: a privacy-preserving collaborative filtering recommendation scheme based on fuzzy C-means and Shapley value* Complex & Intelligent Systems, vol. 11, no. 1, pp. 107, Springer, 2025.
- [Neera et al., 2021] Neera, Jeyamohan and Chen, Xiaomin and Aslam, Nauman and Wang, Kezhi and Shu, Zhan, *Private and utility enhanced recommendations with local differential privacy and Gaussian mixture model* IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 4, pp. 4151–4163, IEEE, 2021.

- [Minto et al., 2021] Minto, Lorenzo and Haller, Moritz and Livshits, Benjamin and Haddadi, Hamed, *Stronger privacy for federated collaborative filtering with implicit feedback* Proceedings of the 15th ACM conference on recommender systems, pp. 342–350, 2021.
- [Tang et al., 2017] Tang, Qiang and Wang, Husen, *Privacy-preserving hybrid recommender system* Proceedings of the Fifth ACM International Workshop on Security in Cloud Computing, pp. 59–66, 2017.
- [Friedman et al., 2016] Friedman, Arik and Berkovsky, Shlomo and Kaafar, Mohamed Ali, *A differential privacy framework for matrix factorization recommender systems* User Modeling and User-Adapted Interaction, vol. 26, pp. 425–458, 2016.
- [Kabir, M.E et al., 2011] Kabir, M.E., Wang, H., and Bertino, E. (2011). Efficient "systematic clustering method for k-anonymization." Acta Informatica, 48(1), 51–66.
- [Pinot, R. et al., 2018] Pinot, R., Morvan, A., Yger, F., Gouy-Pailler, C., and Atif, J. (2018). "Graph-based Clustering under Differential Privacy." arXiv preprint arXiv:1803.03831.
- [Bezdek, J.C., 1981] Bezdek, J.C. (1981). "Pattern Recognition with Fuzzy Objective Function Algorithms." Springer Science and Business Media.
- [Maulik, U. et al., 2000] Maulik, U., and Bandyopadhyay, S. (2000). "Genetic algorithm-based clustering technique." Pattern Recognition, 33(9), pp. 1455–1465.
- [Liu, J. et al., 2018] Liu, J., Liu, Y., and Zhang, Y. (2018). "A Clustering K-Anonymity Privacy-Preserving Method for Wearable IoT Devices." Security and Communication Networks, 2018, Article ID 4945152.
- [Gheewala et al., 2024] Gheewala, Shivangi and Xu, Shuxiang and Yeom, Soonja, "Deep shared learning and attentive domain mapping for cross-domain recommendation" User Modeling and User-Adapted Interaction, vol. 34, no. 5, pp. 1981–2038, Springer, 2024.
- [Anwar, et al., 2022] Anwar, Taushif and Uma, V, *CD-SPM: Cross-domain book recommendation using sequential pattern mining and rule mining* Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 3, pp. 793–800, Elsevier, 2022.
- [Anwar, et al., 2021] Anwar, Taushif and Uma, V and Srivastava, Gautam, *Rec-cfsvd++: Implementing recommendation system using collaborative filtering and singular value decomposition (svd)++* International Journal of Information Technology & Decision Making, vol. 20, no. 04, pp. 1075–1093, 2021.
- [AL-Ghuribi et al., 2025] AL-Ghuribi, Sumaia Mohammed and Mohd Noah, Shahrul Azman and Tiun, Sabrina and Mohammed, Mawal A and Saat, Nur Izyan Yasmin, *Combining review elements for modelling various multi-criteria collaborative recommendation models* Journal of Big Data, vol. 12, no. 1, pp. 160, Springer, 2025.
- [Rieke et al., 2020] Rieke, Nicola and Hancox, Jonny and Li, Wenqi and Milletari, Fausto and Roth, Holger R and Albarqouni, Shadi and Bakas, Spyridon and Galtier, Mathieu N and Landman, Bennett A and Maier-Hein, Klaus and others, *The future of digital health with federated learning* NPJ digital medicine, vol. 3, no. 1, pp. 119, 2020.
- [El Mestari et al., 2024] El Mestari, Soumia Zohra and Lenzini, Gabriele and Demirci, Huseyin, *Preserving data privacy in machine learning systems* Computers & Security, vol. 137, pp. 103605, 2024.
- [Lu et al., 2020] Lu, Zhigang and Shen, Hong, *Differentially private k k-means clustering with convergence guarantee* IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 4, pp. 1541–1552, 2020.
- [Stemmer et al., 2021] Stemmer, Uri, *Locally private k-means clustering* Journal of Machine Learning Research, vol. 22, no. 176, pp. 1–30, 2021.
- [Tsou et al., 2019] Tsou, Yao-Tung and Chen, Hung-Li and Chen, Jia-Yang, *RoD: Evaluating the risk of data disclosure using noise estimation for differential privacy* IEEE Transactions on Big Data, vol. 17, no. 1, pp. 214–226, IEEE, 2019.
- [Gad et al., 2022] Gad, Ahmed G. "Particle swarm optimization algorithm and its applications: A systematic review." Archives of computational methods in engineering, vol. 29, no. 5, 2022.
- [Shami et al., 2022] Shami, Tareq M and El-Saleh, Ayman A and Alswaitti, Mohammed and Al-Tashi, Qasem and Summakieh, Mhd Amen and Mirjalili, Seyedali. "Particle swarm optimization: A comprehensive survey." IEEE Access, vol. 10, pp. 10031–10061, 2022.

- [Makhadmeh et al., 2023] Makhadmeh, Sharif Naser and Al-Betar, Mohammed Azmi and Doush, Iyad Abu and Awadallah, Mohammed A and Kassaymeh, Sofian and Mirjalili, Seyedali and Zitar, Raed Abu, *Recent advances in Grey Wolf Optimizer, its versions and applications*. IEEE access, vol. 12, pp. 22991–23028, 2023.
- [Templ et al., 2022] Templ, Matthias, and Murat Sariyar. *A systematic overview on methods to protect sensitive data provided for various analyses*. International Journal of Information Security 21, no. 6, pp. 1233–1246, 2022.
- [Rodríguez et al., 2016] Rodríguez-Mazahua, Lisbeth, Cristian-Aarón Rodríguez-Enríquez, José Luis Sánchez-Cervantes, Jair Cervantes, Jorge Luis García-Alcaraz, and Giner Alor-Hernández. *A general perspective of Big Data: applications, tools, challenges and trends*. The Journal of Supercomputing 72, pp. 3073–3113, 2016.
- [Haverkamp et al., 2024] Haverkamp, Indy, and Dipti K. Sarmah. *Evaluating the merits and constraints of cryptography-steganography fusion: a systematic analysis*. International Journal of Information Security, pp. 1–29, 2024.
- [Amiri et al., 2019] Amiri, Fatemeh, and Gerald Quirchmayr. *Novel Model for Privacy Preserving in Data Mining using Meta Heuristic Techniques*. Proceedings of the 19th Privacy Enhancing Technologies Symposium (PETS), De Gruyter Poland, 2019.
- [Amiri et al., 2019] Amiri, Fatemeh, Gerald Quirchmayr, Peter Kieseberg, Alessio Bertone, and Edgar Weippl. *Efficiently Vectorized Anonymization in Data Mining using Genetic Algorithm*, 2019.
- [Mendes et al., 2017] Mendes, Ricardo, and João P. Vilela. *Privacy-preserving data mining: methods, metrics, and applications*. IEEE Access, pp. 10562–10582, 2017.
- [Duggirala et al., 2016] Duggirala, Heshu J., Joseph M. Tonning, Ella Smith, Roselie A. Bright, John D. Baker, Robert Ball, Carlos Bell et al. *Use of data mining at the Food and Drug Administration*. Journal of the American Medical Informatics Association 23, no. 2, pp. 428–434, 2016.
- [Feng et al., 2014] Feng, Yan, Yixin Wang, Fang Guo, and Hao Xu. *Applications of data mining methods in the integrative medical studies of coronary heart disease: progress and prospects*. Evidence-Based Complementary and Alternative Medicine 2014, no. 1, pp. 791841, 2014.
- [Zhou et al., 2024] Zhou, Juexiao and Chen, Siyuan and Wu, Yulian and Li, Haoyang and Zhang, Bin and Zhou, Longxi and Hu, Yan and Xiang, Zihang, and Li, Zhongxiao and Chen, Ningning and others. *PPML-Omics: A privacy-preserving federated machine learning method protects patients' privacy in omic data* Science Advances, vol. 10, no. 5, pp. eadh8601, American Association for the Advancement of Science, 2024.
- [Zhu et al., 2023] Zhu, Liang and Lei, Tingting and Mu, Jinqiao and Mu, Jingzhe and Cai, Zengyu and Zhang, Jianwei. *Differential privacy-based spatial-temporal trajectory clustering scheme for LBSNs* Electronics, vol. 12, no. 18, pp. 3767, 2023.
- [Shamsinejad et al., 2025] Shamsinejad, Elham and Baniroostam, Hamid and BaniRostam, Touraj and Pedram, Mir Mohsen and Rahmani, Amir Masoud. *Providing and evaluating a model for big data anonymization streams by using in-memory processing* Knowledge and Information Systems, Springer, pp. 1–34, 2025.
- [Sarmin et al., 2024] Sarmin, Fatima Jahan and Sarkar, Atiquer Rahman and Wang, Yang and Mohammed, Noman. *Synthetic data: Revisiting the privacy-utility trade-off* arXiv preprint arXiv:2407.07926, 2024.