Learning from Leakage: Database Reconstruction from Just a Few Multidimensional Range Queries

Thesis MSc Computer Science

Peijie Li

Delft University of Technology

TU Delft

Learning from Leakage: Database Reconstruction from Just a Few Multidimensional Range Queries

by

Peijie Li

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Thursday August 28, 2025 at 9:30 AM.

Student number:	5952433	
Project duration:	October 8, 2024 – August 28, 2025	
Thesis committee:	Prof. Dr. G. Smaragdakis,	TU Delft, Supervisor
	Dr. J. Decouchant,	TU Delft
	Dr. K. Liang,	TU Delft
	H. Chen, MSc,	TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

I would like to express my heartfelt gratitude to all those who have supported and guided me throughout my research journey.

First, I sincerely thank my advisor, Prof. Dr. G. Smaragdakis, for being a member of my committee and for his guidance and support, which has been valuable throughout this process. I am deeply grateful to my supervisor, Prof. Kaitai Liang, who has met with me multiple times, provided extensive guidance on my thesis, and offered insightful advice on research methods. His mentorship has been instrumental in shaping my work. I would like to thank Prof. J. Decouchant for agreeing to be part of my defense committee. I greatly appreciate his time and willingness to review my work. Special thanks go to my daily supervisor, Huanhuan Chen, for meeting with me weekly, patiently guiding me through challenges, and providing continuous support and encouragement that made this research possible. I am especially indebted to Prof. Lilika Markatou, whose foundational work formed the basis of my research. She provided invaluable guidance, offered many insightful ideas, and helped me clarify and organize the structure of my thesis. Her support has been crucial in helping me complete this work.

Finally, I want to thank my family and friends for their unwavering love, encouragement, and understanding. Your support has been my source of strength throughout this journey, and I could not have completed this work without you.

Peijie Li
Delft, August 2025

Abstract

Searchable Encryption (SE) has shown a lot of promise towards enabling secure and efficient queries over encrypted data. In order to achieve this efficiency, SE inevitably leaks some information, and a big open question is how dangerous this leakage is. While prior reconstruction attacks have demonstrated effectiveness in one-dimensional settings, extending them to high-dimensional datasets remains challenging. Existing methods either demand excessive query information (e.g. an attacker that has observed all possible responses) or produce low-quality reconstructions in sparse databases.

In this work, we present REMIN, a new leakage-abuse attack against SE schemes in multi-dimensional settings, based on access and search pattern leakage from range queries. Our approach leverages unsupervised representation learning to transform query co-occurrence frequencies into geometric signals, allowing the attacker to infer relative spatial relationships between records. This enables accurate and scalable reconstruction of high-dimensional datasets under minimal leakage. Furthermore, we introduce REMIN-P, a practical variant of the attack that incorporates a poisoning strategy. By injecting a small number of auxiliary anchor points—either known or intentionally leaked—REMIN-P significantly improves reconstruction quality, particularly in sparse or boundary regions.

We evaluate our attacks extensively on both synthetic and real-world structured datasets. Compared to state-of-the-art reconstruction attacks, our reconstruction attack achieves up to 50% reduction in mean squared error (MSE), all while maintaining fast and scalable runtime. When the poisoning strategy is chosen properly, our poisoning attack further reduces MSE by an additional 50% on average. To the best of our knowledge, these are the first attacks that enables accurate multi-dimensional reconstruction under low-leakage conditions for any type of database.

Contents

Preface i

Abstract ii

1 Introduction 1

2 Related Works 4

3 Preliminaries 6
3.1 Basic Concepts . . . 6
3.2 Reconstruction Attack and Our Assumptions . . . 6
3.3 Dimension Reduction . . . 7

4 Basic Attack 8
4.1 Extracting Frequency-Based Distances . . . 8
4.2 Dimensionality Reduction for Reconstruction . . . 9
4.3 Refinement and Alignment . . . 9

5 Evaluation on REMIN attack 11
5.1 Experiment Setup . . . 11
5.2 Structural Visualization on Real-World and 3D Datasets . . . 12
5.3 Comparison with Prior Work . . . 13
5.3.1 Synthetic Dataset Comparison . . . 13
5.3.2 Quantitative Comparison on Real-World 2D Spatial Datasets . . . 16
5.4 Reconstruction in Higher Dimensions . . . 16

6 REMIN-P Attack: Leveraging Auxiliary Information for Improved Reconstruction 20
6.1 Intention . . . 20
6.2 The REMIN-P Attack Framework . . . 20
6.3 Experimental Evaluation . . . 21

7 Conclusion and Discussion 24
7.1 Conclusion . . . 24
7.2 Discussion and Future Work . . . 24

8 Ethics Considerations 25

References 26

A Parameter Sensitivity 29

B Supplementary Experiments 31

C Alignment and Refinement Algorithms 33

D REMIN-P: Radial Correction Using Poisoned Anchors 35

1

Introduction

Searchable encryption (SE) [42, 15] supports secure and efficient queries over encrypted databases. Current SE constructions incorporate various cryptographic techniques such as oblivious RAM (ORAM) [21, 43], fully homomorphic encryption (FHE) [17], and property-perserving encryption (PPE) [4] and they are widely implemented in real-world applications, for instance, MongoDB deploys Queryable Encryption (a variant of SE) [39] allowing users to perform expressive (e.g. range) queries on encrypted data.

In order to achieve its high efficiency, SE leaks some information to the server. Common types of leakage are *access pattern* (which records are returned given a query), *search pattern* (whether two queries are the identical), and *volume pattern* (the number of records returned given a query). This leakage has been shown to allow a server to infer information about the plaintext data [52, 34, 27], leading to a line of research aiming to understand how this leakage can be exploited to reveal information about the encrypted data. Among various query types in SE, range queries pose a particularly severe leakage threat. Unlike keyword queries, which typically reveal only discrete match signals, range queries leak structural information by returning clusters of spatially or semantically adjacent records. When a query is issued, a associated token will be generated to retrieve records within a specified range as the response. As a result, attackers may recover layout characteristics of the original dataset, such as density clusters, relative positioning, or even record locations. In this work, we investigate the security risks of range queries across arbitrary dimensions under limited information leakage.

Motivation. Prior work [30, 33, 23, 31, 24, 35] has demonstrated search pattern and access pattern can achieve full or approximate database reconstruction in one-dimensional (1D) datasets, where simpler relationships among records make such attacks more feasible. These attacks have achieved notable success in this regard, providing a solid foundation for understanding the potential of leakage-abuse attacks from range query leakage. However, as we move from one-dimensional to two-dimensional and even higher-dimensional databases, the complexity of the relationships between records increases significantly. To support attacks on these more complex scenarios, researchers have proposed various strategies to address the challenges posed by intricate data relationships and the exponential growth in equivalent databases (databases with the same leakage) [36, 16, 38, 37].

Across all existing methods, the core attack principle remains consistent: use the minimum amount of leaked information (e.g., partial query patterns) and maximize by the quality of the reconstructed data. As summarized in Table 1.1, attacks targeting 1D datasets have achieved this, enabling accurate reconstruction with minimal leakage. However, higher-dimensional datasets naturally introduce greater complexity. While some attacks attain high reconstruction accuracy, they often rely on strong and impractical assumptions, such as access to all possible queries [16] or access to all possible responses [38]. These methods require substantial leakage making them impractical for real-world settings.

The state-of-the-art graph-based approach by Markatou et al. [36] represents the first significant attempt to use minimal leakage in high-dimensional settings by observing very few queries and reconstructing from access pattern leakage alone. This attack has significantly weaker assumptions than earlier methods [16, 38]. But it primarily recovers local ordinal relationships using graph drawing techniques and fails to capture structural features like clustering and relative positioning, especially in generic datasets (i.e., non-dense or sparse datasets). Consequently, its performance degrades in more complex scenarios, which raises a fundamental research question:

Table 1.1: Comparison of our attack REMIN with closely related work. The "Required Leakage" column represents the number of queries needed for the attack, with fewer queries being more desirable. Both Required leakage and reconstruction quality are indicated with stars, where more stars mean fewer queries required and better query efficiency. Our approach achieves practical reconstruction quality with least required leakage. Markatou et al. [36] achieves FDR in dense databases, and achieves high quality ADR in non-sparse datasets.

Attacks	**Query Type**	**Assumption**				**Attack**			
		Query Dist.	Database	Search Pat.	Access Pat.	ADR	FDR	Required Leakage	Reconstruction Quality
Kellaris et al. [30]	1D Range	Uniform	General	✗	✓	✗	✓	☆☆☆	★★★
Lacharité et al. [33]	1D Range	Agnostic	Dense	✗	✓	✓	✓	★☆☆	★★★
Grubbs et al. [23]	1D Range	Uniform	General	✗	✓	✓	✓	★★☆	★★★
Kornaropoulos et al. [31]	1D Range	Agnostic	General	✓	✓	✓	✓	★★★	★★★
Falzon et al. [16]	2D Range	Agnostic	General	✓	✓	✗	✓	☆☆☆	★★★
Markatou et al. [38]	2D Range	Agnostic	General	✓	✓	✓	✓	☆☆☆	★★★
Markatou et al. [36]	dD Range	Agnostic	(Dense)	✗	✓	✓	(✓)	★★☆	★☆☆
REMIN	dD Range	Agnostic	General	✓	✓	✓	✗	★★★	★★☆

Under limited multi-dimensional range search leakage, what is best reconstruction for general databases that we can achieve?

New Perspective. In this work, we propose a novel and practical database reconstruction attack that operates under minimal leakage assumptions, named REMIN. Unlike prior work such as Markatou et al. [36], which explores reconstruction under low query leakage but focuses primarily on **topological** recovery (i.e., identifying neighboring records) via an unweighted graph that captures connectivity but does not account for finer-grained distance signals, our goal is more ambitious: we aim to recover both the **topological** and **geometric** structure of the original database.

By topological structure, we refer to the connectivity and neighborhood relationships among data points—i.e., who is close to whom. By geometric structure, we mean the relative distances and spatial arrangement of data points in the underlying data space.

Our method builds on the observation that records retrieved together in response to user queries are often semantically or structurally related in the original dataset. By measuring how frequently pairs of records co-occur in leaked queries, we can estimate their relative proximity in the data space. Using these co-occurrence frequencies, we construct a high-dimensional distance matrix of size $n \times n$, where n is the number of records, and each record is represented as an n-dimensional vector (described in Section 4).

We then apply unsupervised representation learning and dimensionality reduction techniques (e.g., t-SNE [46]) to embed the records into a low-dimensional space, reconstructing both the topology and geometry of the data in its original target dimensional space, while preserving the pairwise distance information as faithfully as possible.

A key challenge in this setting is edge distortion— there are fewer queries covering boundary regions than central ones, resulting in inaccurate positioning and compression of sparse areas. To address this, we propose REMIN-P, a new active attack variant that employs a poisoning strategy. We consider a more powerful attacker, that is able to inject carefully crafted anchor points—records with known or leaked positions—into the dataset. These anchors act as geometric references, correcting misalignment and improving layout fidelity, particularly in sparse or underrepresented regions. To the best of our knowledge, this is the first poisoning-based reconstruction strategy tailored to range query attacks.

We evaluate our attacks on both synthetic and real-world datasets, simulating practical scenarios under different query distributions. A brief comparison is summarized in Table 1.1, and the full experimental results are provided in Chapter 5. Our main contributions are:

- **A practical reconstruction attack under minimal leakage:** We introduce an unsupervised, co-occurrence-based attack that recovers both topological and geometric properties for general dataset using only limited information leakage. REMIN remains effective even in extremely low-leakage conditions, even on sparse datasets (See Fig. 1.1 for an example).

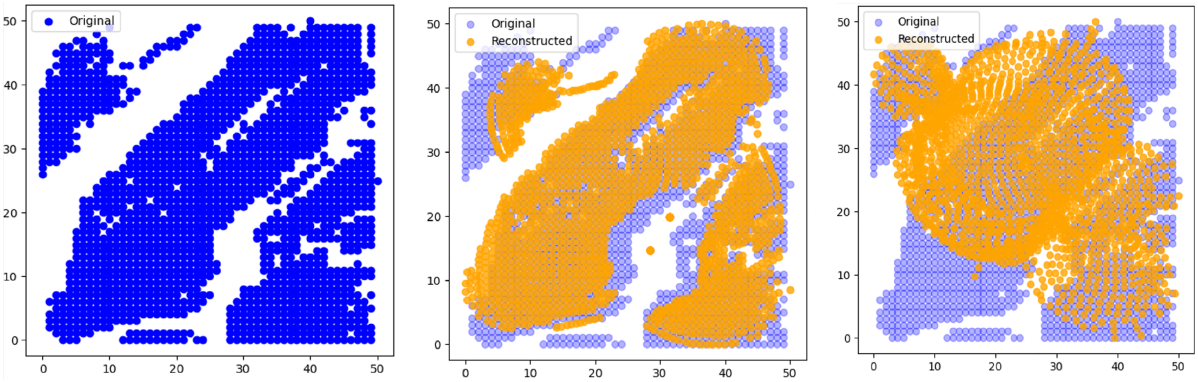


Figure 1.1: (Left) Manhattan Highway Crossings dataset. (Middle) Reconstruction by our method REMIN, achieving an MSE of 14.25. (Right) Reconstruction by Markatou et al. [36], with an MSE of 42.97. (under 1% uniformly sampled queries)

- **A novel poisoning attack:** We propose REMIN-P, the first active attack variant, where the attacker poisons the dataset, injecting anchor points to improve global alignment and correct distortions in geometric reconstruction.
- **Comprehensive experimental evaluation:** We demonstrate the effectiveness of our attacks across diverse datasets and leakage scenarios, showing significant improvements in reconstruction accuracy, neighborhood preservation, and runtime compared to prior work.

Our work shows how minimal co-occurrence leakage can be exploited to recover both how records are related (topology) and how they are spatially arranged (geometry). This highlights the need for stronger leakage mitigation techniques in searchable encryption systems.

2

Related Works

Searchable Encryption (SE) has evolved significantly from its early keyword-search origins to support increasingly complex functionality like dynamic updates and expressive queries [15, 6, 7, 9, 10, 12, 18, 19, 28, 29, 40]. Early constructions primarily supported simple keyword queries [42, 20], but were later extended to support more expressive query types, including conjunctive, boolean, and range queries [10, 22, 2, 26, 5]. In particular, Dynamic Searchable Symmetric Encryption (DSSE) [29] enables updates and deletions on encrypted databases, making SE practical for real-world applications.

Despite their theoretical security, SE schemes inevitably leak side-channel information during query execution. Several works have systematically characterized and quantified the leakage in SE protocols [34, 27, 30, 11, 3, 32], which has been shown to be particularly exploitable for reconstructing private data through range query attacks [34, 27, 30, 33, 51]. Most work on dataset reconstruction from range query leakage can be broadly categorized into three methodological classes: (1) symbolic or algebraic attacks that rely on geometric properties and deterministic constraints, (2) statistical estimation methods that infer data distributions from frequencies information, and (3) graph-theoretic approaches that exploit co-occurrence patterns to recover structural information. Each class of methods operates under different leakage assumptions and achieves varying levels of reconstruction accuracy and generality.

Early work primarily focused on determining how much information about the query inputs can be recovered from leakage [27, 11, 41]. Subsequently, several algebraic attacks have been proposed to exploit range query leakage for recovering plaintext geometry, which mainly explored the reconstruction of the encrypted dataset under strong leakage assumption-full access pattern leakage and knowledge of the query set. KKNO [30] is the first to formalize this setting by introducing a generic volume-based attack model that combines access patterns with volume leakage. By analyzing the frequency of record occurrences across query responses, KKNO is able to infer exact record positions. Lacharité et al. [33] extended this approach to Approximate Density Reconstruction (ADR) with partial inaccuracies in the inferred structure in exchange for significantly improved efficiency. Falzon et al. [16] further demonstrated leaked information in 2D databases can also be exploited to achieve exact reconstruction. These symbolic or order-based attacks offer strong recovery guarantees, but require full access pattern leakage and knowledge of the query or data distribution.

To mitigate such strong assumptions, later methods turned to statistical estimation. Instead of directly solving for positions, these approaches estimate the underlying distribution or topological layout of the data using query response patterns. Kornaropoulos et al. [31] introduced a support-size estimator that, notably, was the first to leverage search pattern leakage—a previously overlooked signal—to reconstruct the database without relying on knowledge of the underlying query or data distribution. Building on this direction, Markatou et al. [38] proposed a non-parametric estimator that aggregates access frequencies to support partial reconstruction in two-dimensional settings. While these methods reduce the required leakage compared to earlier techniques, they still assume visibility over a large fraction of the query workload and can be sensitive to dataset sparsity and distributional skew.

More recently, graph-theoretic approaches have emerged as a powerful alternative, shifting focus from precise reconstruction to structural recovery. Grubbs et al. [24] introduced a method that uses volume leakage to recon-

struct the database structure in 1D settings. While the approach targets full reconstruction, it is primarily effective on dense datasets, resulting in an ordering of records rather than detailed spatial layout. Markatou et al. [36] further proposed a framework that constructs a co-occurrence graph from access patterns and aligns it to a target topology using graph-matching techniques. These methods succeed under minimal leakage — often access pattern or volume pattern alone — and generalize well to higher dimensions. However, they model co-occurrence as binary and ignore richer signals such as relative distance or frequency, limiting their geometric fidelity.

In contrast to symbolic reasoning and graph-based approaches, machine learning offers a new paradigm for capturing geometric relationships from leakage. While ML techniques have been widely applied in privacy attacks—such as side-channel and website fingerprinting [48, 13]—they remain largely unexplored in the context of database reconstruction from searchable encryption leakage to the best of our knowledge. In particular, the effectiveness of unsupervised learning for this task remains an open question. Our work bridges this gap by leveraging co-occurrence frequency as a proxy for spatial proximity, embedding records via manifold learning, and refining the reconstruction through alignment. This enables robust and efficient recovery even under sparse leakage and agnostic query distributions.

3

Preliminaries

3.1. Basic Concepts

In a searchable encryption (SE) scheme, a user first encrypts their dataset and uploads the encrypted data to an untrusted server. Later, to perform a range query, the user generates a query token that encodes the desired range and sends it to the server. Using this token, the server searches over the encrypted dataset and returns the identifiers of the documents whose associated values fall within the specified range.

Let $\mathcal{F} = \{r_1, r_2, \ldots, r_n\}$ denote the dataset belonging to the user, where each r_i is a record in the dataset. We consider a d-dimensional dataset where each record is a unique point in a discrete space:

$$S_d = [N_1] \times [N_2] \times \cdots \times [N_d],$$

where $[N_i] = \{0, 1, \ldots, N_i - 1\}$ defines the index range along dimension i. Each record r has a domain value $x = (x_1, x_2, \ldots, x_d) \in S_d$.

Let $\mathcal{E} = \{er_1, er_2, \ldots, er_n\}$ be the encrypted dataset corresponding to $\mathcal{F}$. A range query q is defined by a set of bounds for each dimension:

$$q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d],$$

where $a_i, b_i \in [N_i]$ and $a_i \leq b_i$ for each i. The server responds with all encrypted records within the corresponding hyperrectangle:

$$\text{Response}(t_q) = \{er_{j_1}, \cdots, er_{j_q} \mid r_j \in q,\ \forall j \in [j_1, j_q]\},$$

where t_q is the token corresponding to the query q. See Table 4.1 for notation details.

We consider a setting similar to that studied in prior work on range query SE schemes [36, 30, 33, 23, 31, 38, 24, 35]. Since both the encrypted dataset and the query tokens are stored on the server, an adversarial server could observe: (1) the universe of encrypted records and query tokens, and (2) the relationship between them—specifically, whether a given encrypted record matches a given query token (i.e., whether it falls within the queried range).

Our attacks leverage both the access pattern and search pattern leakages from the encrypted dataset and associated query tokens. The access pattern reveals the relationship between each encrypted record and each query token, i.e., which records match which queries. The search pattern indicates whether two query tokens correspond to the same query, allowing the attacker to eliminate duplicate query tokens and thus avoid redundant computations in subsequent distance calculations.

3.2. Reconstruction Attack and Our Assumptions

Reconstruction attacks attempt to recover the underlying structure of the encrypted database based on the leakage observed from access and search patterns. We categorize these attacks into two types:

- **Full database reconstruction (FDR).** This attack aims to recover the **exact** positions of all records in the database, thereby reconstructing the entire dataset with complete accuracy.

- **Approximate Database Reconstruction (ADR).** This reconstruction seeks to recover the fundamental **topological** structure of the dataset, without necessarily recovering the exact positions of all record s. The focus is on preserving the overall layout, such as clustering and relative positioning, rather than achieving a perfect reconstruction.

In this work, we investigate the approximate reconstruction process under the assumption of limited leakage, assuming that only a small fraction (e.g., 1%) of the access and search patterns are available to the attacker. Specifically:

- Partial search pattern leakage: The attacker knows a subset of queries token $t_q \mid q \in \mathcal{Q}$ that were issued to the database.
- Corresponding access pattern leakage: For each observed query, the attacker also knows the set of encrypted records returned in the response, as shown in Fig. 4.1 as Observed Response.

The goal of the attacker is to recover as much information as possible about the original dataset, minimizing the reconstruction error and preserving the topological and geometric structure of the data.

3.3. Dimension Reduction

In our REMIN attack, we generate a representation matrix that captures the relative positions of records in a high-dimensional space. To convert this high-dimensional matrix into a reconstruction that reflects the original spatial distribution, we utilize dimensionality reduction techniques. Dimensionality reduction is a process used to reduce the number of dimensions in a dataset while retaining as much information as possible about the original structure. This technique is widely used in various fields, such as data visualization and pattern recognition, to simplify complex, high-dimensional data while preserving essential relationships between data points [8, 47, 14].

In the context of database reconstruction, the choice of dimensionality reduction method is critical. Traditional methods like Principal Component Analysis (PCA) [1] and Multidimensional Scaling (MDS) [45] are designed to preserve global relationships, making them less suitable for the task at hand. Specifically, PCA assumes a linear structure, while MDS relies on the assumption that global distance information is reliable. However, the distance matrix constructed from the co-occurrence frequency matrix F is likely to exhibit a nonlinear, high-dimensional manifold structure. In this case, global geometric relationships are not the most informative, and dimensionality reduction techniques should focus on recovering local manifold structures instead.

Notably, t-SNE (t-distributed Stochastic Neighbor Embedding) [46] is particularly well-suited for this scenario because it prioritizes local relationships, making it effective at uncovering the intrinsic low-dimensional geometry in high-dimensional data. By minimizing the divergence between the original and projected distance matrices, t-SNE ensures that records that are close in the high-dimensional space remain close in the lower-dimensional space. In contrast, methods like PCA and MDS, which focus on global structure, may fail to capture important local relationships, leading to less accurate reconstructions. Thus, we use t-SNE to balance local and global structures effectively.

4

Basic Attack

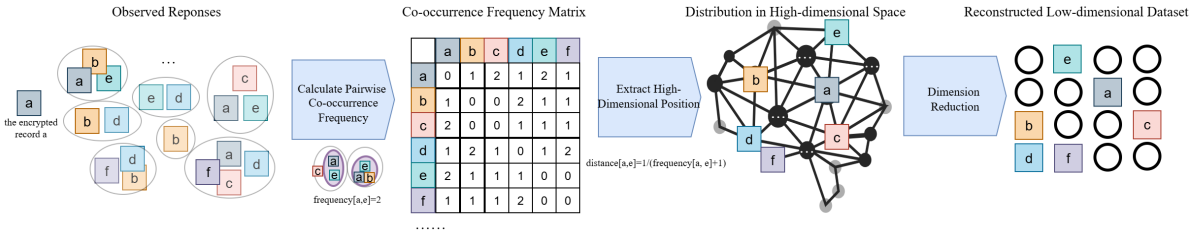


Figure 4.1: A schematic illustration of the REMIN reconstruction attack. The method first observes the model's responses to a set of query tokens and counts how frequently each pair of encrypted records (e.g., (a, b)) co-occurs in the different responses. Each cluster in the Observed Responses represents the response returned by a single query token, and each letter (e.g., a) corresponds to an individual encrypted record. These co-occurrence frequencies are transformed into pairwise distances (e.g., using the reciprocal of the frequency), resulting in a distance matrix that represents the records as points in a high-dimensional space. The high-dimensional relationships are then embedded into low-dimensional space through dimensionality reduction (e.g. t-SNE), preserving the structural relationships implied by response co-occurrence.

We present a new attack called REMIN that leverages machine learning, frequency-based co-occurrence analysis and dimensionality reduction techniques. The core insight behind this attack is that the co-occurrence frequencies of records in query responses can serve as a high-dimensional representation of their spatial relationships. By applying dimensionality reduction algorithms, we can recover the underlying geometry of the dataset in a lower-dimensional (target) space (e.g., 2D or 3D). An overview is illustrated in Fig. 4.1, outlining the attack flow.

4.1. Extracting Frequency-Based Distances

For the leakage in range query SE schemes, we consider both access pattern and search pattern leakage, as described in Section 3. In particular, the attacker leverages the leakage to infer the response set for each query token, denoted by $\{\text{Response}(t_q) \mid q \in \mathcal{Q}\}$, where each response comprises the encrypted records matching the query.

We begin by constructing a *frequency matrix* F that captures the co-occurrence relationships between encrypted records—specifically, whether two records appear together in the response to the same query. Since different queries may yield identical responses, we resolve such ambiguities using the search pattern leakage. The resulting matrix F serves as the basis for computing distances between records in subsequent steps.

To be precise, the co-occurrence frequency $F(i, j)$ is defined as the number of query tokens where encrypted records er_i and er_j are both returned in the corresponding responses:

$$F(i,j) = \sum_{t_q,\ q \in \mathcal{Q}} \mathbf{1}\left(er_i \in \text{Response}(t_q) \wedge er_j \in \text{Response}(t_q)\right).$$

An example is shown in Fig. 4.1, where $F(0, 2) = 2$, indicating that encrypted records a and c co-occur in two query responses. This matrix inherently reflects the spatial relationships among records, as records that are

Table 4.1: Notation Summary

Symbol	Description
n	Total number of records
$\mathcal{F}$	Original dataset with n records
S_d	A d-dimensional Euclidean space
N_i	Range size in dimension i ($i = 1, \ldots, d$)
$\mathcal{E}$	Encrypted dataset stored on the server
t_q	Query token corresponding to the query q
$\mathrm{Response}(t_q)$	Set of returned records of the query token t_q
$\mathcal{Q}$	Set of observed queries
F	The co-occurrence frequency matrix
$D(i, j)$	The (i, j)-th entry of the distance matrix D
Y	Representation matrix after dimension reduction

spatially closer are more likely to appear together in the same query response, resulting in higher co-occurrence frequencies. To convert this frequency-based information into a distance relationship, we apply appropriate distance metrics.

Common choices for this transformation include reciprocal distance (which we adopt for its efficiency and experimental accuracy) [44], Euclidean distance (less suitable in this case, as it assumes a real geometric relationship rather than frequency-based), and Gaussian kernel distance [25]. Reciprocal distance, defined as:

$$D(i, j) = \frac{1}{1 + F(i, j)} \tag{4.1}$$

is particularly effective, as it ensures that smaller co-occurrence frequencies (indicating weaker associations) result in larger distances, while higher frequencies (indicating stronger associations) correspond to smaller distances. The resulting distance matrix D as a proxy for the pairwise distances between records in the high-dimensional space, with its complete computation procedure detailed in Algorithm 1, where id_{er} denote the index of an encrypted record er.

4.2. Dimensionality Reduction for Reconstruction

The distance matrix D captures the high-dimensional relationships between records based on their co-occurrence. However, the true underlying structure of the dataset exists in a lower-dimensional space. To recover this structure, we apply dimensionality reduction techniques.

Dimensionality reduction is a common machine learning approach used to simplify complex, high-dimensional data while preserving its essential features. In our method, we employ t-SNE, which is particularly effective at preserving local relationships and minimizing global distortions, ensuring that the reconstructed dataset reflects the original spatial layout. The result of this step is a set of reconstructed coordinates in a low-dimensional space, approximating the original dataset's spatial relationships.

While t-SNE excels at preserving local structure, it is sensitive to parameters such as perplexity, which controls the trade-off between local and global preservation. For completeness, we refer to Appendix A for a more detailed discussion on parameter sensitivity. Our empirical analysis shows that the optimal perplexity value scales with dataset size: smaller datasets benefit from lower perplexity values, while larger datasets require higher values for optimal reconstruction. By carefully selecting perplexity, we can significantly improve the reconstruction quality, balancing local and global structure preservation.

4.3. Refinement and Alignment

The dimensionality reduction process provides an initial set of coordinates for the records in the original space. However, due to due to the lack of absolute positional information in the co-occurrence data, the reconstruction result may suffer from potential misalignment. It typically differs from the original structure by certain rotations,

Algorithm 1 GetDistanceMatrix($responses, points$)

Input: $responses$: Set of query responses to each query token
$EncRec$: Set of encrypted records in the dataset
Output: D: Distance matrix where $D(i, j)$ is the dissimilarity measure

$n \leftarrow |EncRec|$
/ Initialize the co-occurrence frequency matrix */*
An $n \times n$ symmetric matrix $F \leftarrow \{\}$
for each $response \in responses$ **do**
  for two different records $er, es \in response$ **do**
/ Update frequency of point pair co-occurrence */*
    $F[id_{er}, id_{es}] \mathrel{+}= 1$
    $F[id_{es}, id_{er}] \mathrel{+}= 1$
  end for
end for
/ Initialize distance matrix with infinity */*
An $n \times n$ symmetric matrix D initialized with all entries set to a large constant (e.g., ∞)
for each $(i, j) \in [n] \times [n]$ **do**
/ Convert frequency to distance */*
  $D(i, j) \leftarrow \frac{1}{1+F(i,j)}$
end for
for $i \leftarrow 0$ to $n - 1$ **do**
/ Set distance to itself as zero */*
  $D(i, i) \leftarrow 0$
end for
return D

scalings, and translations—affine transformations that preserve relative distances and angles between points but do not maintain their absolute positions.

To address this, we apply Procrustes analysis, a statistical method specifically designed to correct affine transformations. As shown in Algorithm 2 (see Appendix C for details), Procrustes analysis requires a reference set of points for alignment. For datasets with a relatively uniform distribution, where records are evenly spaced across the spatial domain, we construct a reference grid to serve as the target for alignment. This approach works because each record in the original database lies on a grid point, making the grid a natural representation of the true spatial structure. Mathematically, given the original coordinates X (the reference grid) and the reconstructed coordinates Y, Procrustes Analysis solves the following optimization problem:

$$\min_{R,s,t} ||X - (sYR + t)||^2 \tag{4.2}$$

where R is a rotation matrix, s is a scaling factor, and t is a translation vector. This ensures that the reconstructed embedding aligns with the original dataset's spatial layout, correcting for reasonable distortions.

For datasets requiring integer coordinate constraints, simulated annealing, a probabilistic optimization technique, can be used to further optimize point placement, guaranteeing reconstructed coordinates occupy exact integer positions. Details on this technique can be found in Algorithm 3 (see Appendix C for implementation details).

5

Evaluation on REMIN attack

5.1. Experiment Setup

We assess the performance of our reconstruction method through comprehensive experiments across various datasets, including both synthetic and real-world scenarios. These datasets are as follows:

- **Grid (Synthetic)**: A procedurally generated $n \times n$ grid dataset with uniform spacing and configurable sparsity (by randomly removing records). The grid size, denoted by n, indicates the number of cells per dimension. This dataset is used to simulate idealized structured data. *We refer to this dataset as **Grid**.*
- **California Intersections**: A real-world spatial dataset of over 21,000 road intersections in California representing urban and suburban sprawl (normalized to a 5050 grid), previously used in SE research [36]. *We refer to this dataset as **Cali**.*
- **Amsterdam Drinking Water Points**: A spatial dataset of water access locations in Amsterdam, collected from OpenStreetMap. It exhibits moderate clustering typical of urban infrastructure. *We refer to this dataset as **AMS**.* The data was normalized to a 50×50 grid, containing 289 points.
- **Manhattan Highway Crossings**: This dataset captures highway intersections in Manhattan with 1708 points normalized to 50×50 space, exhibiting prominent spatial clustering. *We refer to this dataset as **Manhattan**.*
- **Paris Shops**: Locations of shops in central Paris, characterized by non-uniform spatial clusters that reflect commercial hotspots. *We refer to this dataset as **Paris**.* The data was also normalized to a 50×50 grid, containing 1158 points.
- **Shanghai Bus Stops**: Bus stop locations in downtown Shanghai with 1046 points normalized to 50×50, forming a relatively uniform distribution shaped by the urban road network, with non-equidistant spacing between points. *We refer to this dataset as **Shanghai**.*
- **New Hampshire Elevation (3D)**: A 3D terrain dataset containing elevation samples from the White Mountains, previously used in reconstruction research [36]. Normalized to a 16×16×16 volume. *We refer to this dataset as **NH**.*

To ensure a fair comparison between reconstruction methods, we align all reconstructed coordinates to the original dataset before computing evaluation metrics. This is necessary because reconstruction from range queries is inherently non-unique: many databases can produce identical query responses, forming what is known as the *reconstruction space* [36]. Within this space, any dataset is considered indistinguishable from others based solely on query leakage. To account for this ambiguity, we apply Procrustes analysis using known point correspondences to align reconstructions to the ground truth, allowing only rotation and scaling. This alignment ensures that evaluation metrics such as MSE and neighborhood accuracy reflect meaningful geometric fidelity rather than coordinate frame discrepancies. Unlike the basic attack setting, where alignment and grid snapping are performed post-reconstruction, we avoid such post-processing as it may introduce artificial artifacts that could distort the comparison. All quantitative metrics reported in this section are computed after this alignment, ensuring that the results are directly comparable across methods and settings.

We evaluate our reconstruction method using four key metrics, each capturing a distinct aspect of reconstruction quality:

- **Mean Squared Error (MSE):** MSE measures the average squared distance between the reconstructed coordinates and the original coordinates across all records. It quantifies the overall reconstruction error, balancing the contributions of both large and small errors, thus offering a comprehensive measure of reconstruction quality.
- **Tolerant Match Rate:** This relaxed version of exact coordinate match assesses the method's ability to recover positions within a specified radius of the ground-truth location. Specifically, a predicted point is considered a match if its Euclidean distance to the true point is less than $\sqrt{2}$. This metric is particularly valuable in practical settings, as it accounts for minor distortions that remain even after alignment, providing a more meaningful measure of localization accuracy.
- **Neighbor Accuracy:** This metric evaluates the preservation of local consistency by measuring the percentage of records whose 5 nearest neighbors in the reconstructed dataset match those in the original dataset. This is crucial for assessing how well the method preserves clusters and neighborhood relationships, which are often key to the dataset's structure.
- **Chamfer Distance:** Chamfer Distance measures the average closest-point distance between two sets of points, capturing the overall structural similarity between them. Unlike MSE, which is sensitive to small shifts in coordinates, Chamfer Distance is robust to minor mismatches in point correspondence, making it an ideal metric for evaluating the structural integrity of the reconstructed dataset.

These four metrics collectively evaluate both global accuracy and local consistency, providing comprehensive assessment of reconstruction quality. To account for the inherent instability of t-SNE and ensure statistical reliability, we repeat each experimental setup ten times in our experiments, and the results are averaged across these runs.

5.2. Structural Visualization on Real-World and 3D Datasets

To qualitatively assess the reconstruction capabilities of our method, we present visual comparisons between the ground truth datasets, our reconstructed outputs, as well as those of a leading query-efficient baseline (Markatou et al. [36]). These comparisons span both real-world 2D spatial datasets obtained from OpenStreetMap and a 3D dataset, showing that our method outperforms the baseline in both geometric and topological preservation under severely limited query access.

We evaluate three real-world 2D datasets described in Section 5.1: Amsterdam drinking water access points, Manhattan highway crossings, and shops in central Paris. All reconstructions are conducted with only 1% query coverage under q uniform query distribution.

Fig. 1.1 and Fig. 5.1 depict datasets with strong clustering structures (AMS and Manhattan). Our method effectively recovers both the intra-cluster density and inter-cluster spacing, preserving neighborhood relationships with high fidelity. In the AMS dataset, we achieve 82% neighbor accuracy and a Chamfer distance of 1.46. Similar performance is observed for Manhattan. In contrast, the baseline fails to capture both the global and local structures. Its reliance on unweighted connectivity producing nearly uniform point clouds with significantly lower neighbor accuracy (below 40%) and substantially higher Chamfer distances, indicating both geometric and topological distortions.

Not all spatial datasets exhibit strong clustering. To evaluate reconstruction performance in general cases, we consider the Paris datasets (Fig. 5.2), which features more diffuse yet structurally meaningful point distributions. Despite the absence of large-scale clustering, this dataset still contains significant spatial heterogeneity and alignment patterns. Our method successfully recovers the relative positioning of points and underlying spatial organization. The reconstruction achieves consistent accuracy with 80.03% neighborhood accuracy and low MSE, demonstrating robustness beyond clustered scenarios. Conversely, the baseline once again fails to preserve spatial structure under the same query budget, producing a near-uniform random distribution of points, devoid of meaningful spatial relationships.

To demonstrate the dimensional agnosticism of our approach, we further evaluate our method on a 3D dataset (NH), embedded in $\mathbb{Z}^3$. Using the same query rate of 1% under uniform query distribution, we reconstruct the spatial layout via our co-occurrence-based dimensionality reduction pipeline. As shown in Fig. B.1, the reconstructed

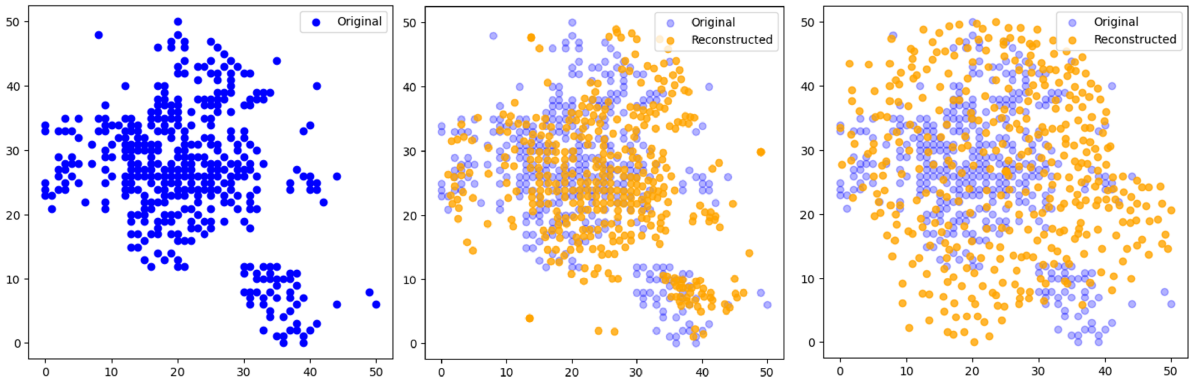


Figure 5.1: (Left) Original Amsterdam Drinking Water dataset (50×50 grid). (Middle) Reconstruction by REMIN using 1% uniformly sampled queries, achieving 82.00% neighbor accuracy and MSE of 23.52. (Right) Reconstruction by the method of Markatou et al. [36] under the same setting, achieving 50.92% neighbor accuracy and MSE of 37.86.

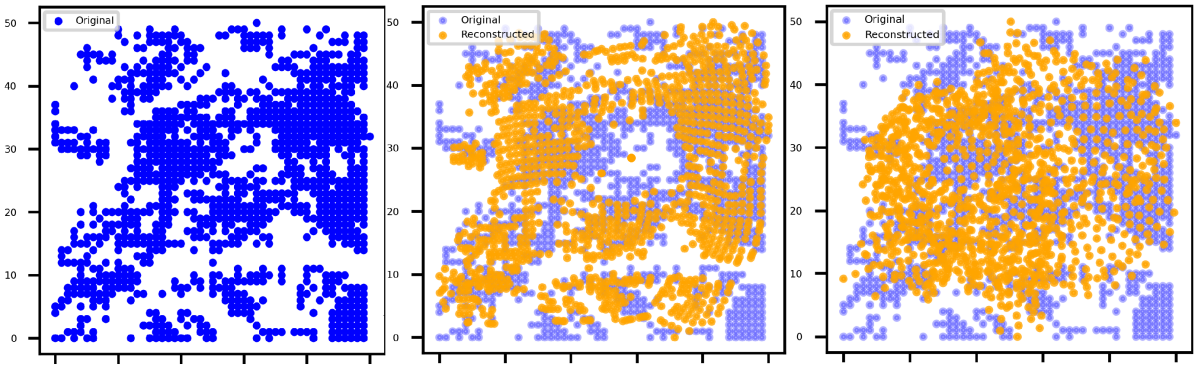


Figure 5.2: (Left) Original Paris Shops dataset (50×50 grid). (Middle) Reconstruction by REMIN using 1% uniformly sampled queries, achieving 80.03% neighbor accuracy, Chamfer distance of 0.81, and MSE of 41.1. (Right) Reconstruction by the method of Markatou et al. [36] with 30% neighbor accuracy, Chamfer distance of 1.21, and MSE of 128.4.

shape preserves most of the original **topological** structure, highlighting that our method is dimension-agnostic. Since the approach solely relies on inter-point distances rather than absolute positions, it naturally extends to any intrinsic dimensionality, as long as local neighborhood information is well represented.

5.3. Comparison with Prior Work

To quantitatively evaluate the effectiveness of our reconstruction method, we conduct a series of experiments comparing its performance against representative baselines [**1555162**, 36] across a variety of settings. Specifically, we assess reconstruction quality and scalability on both synthetic and real-world datasets.

First, we analyze performance on synthetic grid datasets, varying both dataset size and sparsity level to investigate robustness under different density and completeness conditions. Second, we perform extensive comparisons on multiple real-world 2D spatial datasets, focusing on the performance at very low (1%) and moderate (25%) query ratios under uniform query sampling. These ratios represent scenarios where an attacker has access to either very limited information or relatively more, but still insufficient, data for full reconstruction. By testing these methods on practical datasets, we aim to better demonstrate their effectiveness and applicability in real-world settings.

These experiments provide a comprehensive assessment of accuracy, robustness, and scalability, highlighting the significant advantage of our method over existing techniques under realistic leakage constraints.

5.3.1. Synthetic Dataset Comparison

To further validate the effectiveness and generality of our method, we benchmark it against Markatou et al. [36]. Although this technique also aims to recover database structure under limited query access, it tends to overlook the underlying pairwise distance in datasets, complicating alignment and structural interpretation.

We then design two evaluation scenarios to capture realistic variations in database conditions:

- **Scalability:** Datasets may vary in size or resolution, ranging from small tables to large-scale records. A robust attack should scale efficiently with dataset growth.
- **Data Density:** Real-world datasets often exhibit non-uniform density, with missing records or irregular data collection leading to sparsity and gaps.

To assess performance under these conditions, we conduct two comparative studies:

- **Varying grid size** to examine structural consistency across scales.
- **Varying missing data ratio** to evaluate robustness to sparsity and missing records.

As mentioned before, once the mapping between reconstructed results and original coordinates is established, we compute all evaluation metrics based on the matched pairs, as shown in Fig. 5.3.

Scalability. We evaluate the scalability of our approach against the baseline by testing both methods on grid-structured datasets, ranging from 20×20 to 40×40 in size. The experiments maintain a fixed query ratio of 1% and explore varying data distributions—uniform, beta, and Gaussian—to assess robustness across different query patterns. This setup isolates the impact of database scale on reconstruction performance. As dataset size increases, a robust method should maintain reconstruction accuracy while remaining computationally efficient. To fully characterize this trade-off, we augment our evaluation with runtime measurements alongside the previously introduced metrics.

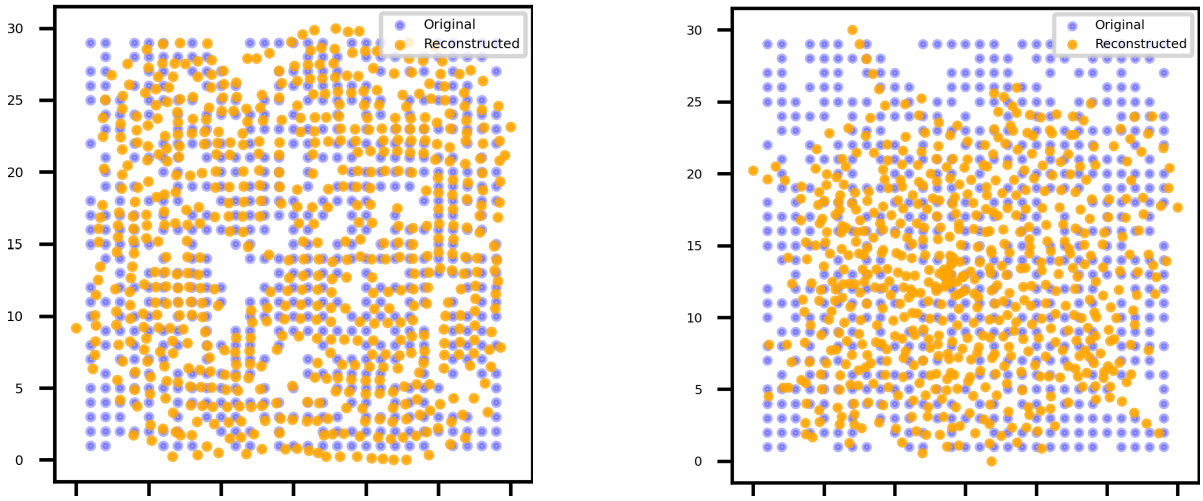


Figure 5.3: Reconstruction by REMIN (Left) and by Markatou et al. [36] (Right) using 1% uniformly sampled queries.

The experimental results in Fig. 5.4 reveal that our method consistently outperforms the baseline across all settings.

Interestingly, as the grid size increases from 20×20 to 40×40, both methods demonstrate stable performance, suggesting that both approaches exhibit scalability with respect to dataset size. However, we observe a more pronounced performance decline when varying the query distribution. Under non-uniform distributions, such as beta and Gaussian, both methods suffer in reconstruction quality due to increased sampling redundancy—fewer unique range responses result in less informative co-occurrence data. Nevertheless, our method maintains strong performance in preserving local structures. For example, neighborhood accuracy remains above 70% across all settings, demonstrating that our method is particularly resilient to distributional shifts. While other metrics, such as tolerant match rate and MSE, show greater variability, neighborhood accuracy remains relatively stable, indicating that our approach excels at capturing fine-grained local geometry, even when global information is limited.

In addition to improved accuracy, our method is significantly more efficient. While the baseline relies on iterative leakage amplification with $O(n^3)$ complexity, our pipeline reduces this to $O(n^2)$ through direct representa-

tion learning. Runtime comparisons (See Fig. B.2 in Appendix B) demonstrate an order-of-magnitude speedup, making our method particularly suitable for large-scale applications where traditional cubic-complexity methods become infeasible.

Overall, these results highlight the scalability, geometric accuracy, and topological robustness of our method. It consistently outperforms the state-of-the-art across a variety of settings, demonstrating its ability to effectively handle larger datasets and reliably preserve local structure even in the presence of sparse or biased query patterns.

Table 5.1: Comparison of reconstruction methods under increasing sparsity levels with perplexity set to 80.

Missing Data Ratio	**REMIN**				**Markatou et al. [36]**			
	MSE ↓	Tolerant Match ↑	Neighbor Accuracy ↑	Chamfer Distance ↓	MSE ↓	Tolerant Match ↑	Neighbor Accuracy ↑	Chamfer Distance ↓
0%	1.81	0.7365	0.8142	0.4375	30.10	0.3718	0.6898	0.5665
10%	2.19	0.7090	0.8103	0.4931	28.17	0.2761	0.6635	0.6035
20%	2.33	0.6715	0.8202	0.5115	28.25	0.1683	0.6096	0.6771
30%	2.44	0.6572	0.8335	0.5636	28.91	0.1146	0.5443	0.8047
40%	2.88	0.6251	0.8221	0.5951	23.59	0.1420	0.4951	0.7946
50%	10.33	0.4920	0.8132	0.6685	22.93	0.1178	0.4475	0.8575
60%	17.03	0.4109	0.8105	0.7983	22.65	0.1437	0.3944	1.0110
70%	18.79	0.2292	0.8071	0.9523	23.59	0.0859	0.3565	1.0540

Data Density. In many real-world applications, databases are not fully populated — due to missing records, data corruption, or partial collection — leading to what we refer to as gaps or missing data points. Existing methods, such as Markatou et al. [36] do not explicitly preserve local distance consistency and thus experience significant distortion when applied to incomplete datasets, struggling to infer the locations of missing points.

To evaluate the robustness of our approach under varying levels of sparsity, we simulate random missing data points on a 30×30 grid with a uniform data distribution. The query ratio is fixed at 1%, and we gradually increase the percentage of missing records from 0% to 70%. All other experimental parameters remain consistent with the previous setup.

As shown in Table 5.1, our method consistently outperforms the baseline method by Markatou et al. [36] across all sparsity levels. In the range of 0%–60% missing data, our reconstruction remains notably stable, suggesting that the method is resilient to moderate sparsity. For example, even with 60% of the records are missing, our method maintains a tolerant match rate above 49%, indicating successful structural recovery despite substantial data loss. In contrast, the baseline method exhibits a clear decline in all performance metrics as the missing data ratio increases. This performance degradation aligns with the baseline's inability to capture pairwise distance relationships, as its graph-based design relies solely on connectivity, neglecting local geometric structure.

Notably, when the percentage of missing data exceeds 20%, the performance of the baseline stabilizes at a low level (approximately 10% in tolerant match rate), suggesting that any reconstruction at this stage is largely due to random overlaps with the original dataset rather than meaningful structural recovery.

Interestingly, across all levels of sparsity, our method maintains high neighborhood accuracy same as the previous experiment (exceeding 80%), highlighting its robustness in preserving local topology. This resilience stems from our design, which leverages co-occurrence frequency to encode proximity, even when full structural information is not available.

However, when the missing data ratio exceeds 70%, performance becomes more volatile. At this point, co-occurrence signals become too sparse for reliable inference, and the gap between our method and the baseline [36] narrows. This limitation underscores a known challenge in high-sparsity scenarios: when neighbor relationships are insufficiently sampled, even distance-aware methods struggle to recover accurate structures.

In summary, the experiment demonstrates that our method is robust and scalable under realistic levels of sparsity (up to 60%), with clear advantages over the state-of-the-art. Its ability to preserve local geometry makes it well-suited for moderately sparse database scenarios, though extreme sparsity remains challenging for all current approaches.

5.3.2. Quantitative Comparison on Real-World 2D Spatial Datasets

In realistic scenarios, attackers often face strict constraints on the number of observable query responses, either due to system-imposed query budgets or privacy-preserving mechanisms. Consequently, a key challenge in reconstruction attacks is recovering as much spatial structure as possible with limited information. To assess the effectiveness of our method under such low-leakage conditions, we evaluate its performance on four real-world 2D datasets, and compare it against state-of-the-art reconstruction baselines [36, 38].

Our primary baseline is the recent method of Markatou et al. [36], which has been employed in prior studies. To simulate realistic query leakage, we uniformly sample a fixed percentage of range queries (from 1% to 5%) under a uniform distribution and use them as the sole input to each reconstruction method. We then evaluate the output using the four metrics mentioned before.

As shown in Fig. 5.5, the performance of all methods generally improves with higher query ratios, with MSE decreasing and accuracy-related metrics increasing. Across all datasets and query budgets, our method consistently outperforms the baselines by a clear margin. Notably, under the lowest query ratio of 1%, our method achieves a tolerant match rate that more than doubles that of Markatou et al. [36]. Furthermore, due to the geometry-aware properties of our t-SNE-based reconstruction, our approach reliably preserves local spatial relationships, achieving neighbor accuracy above 72% across all settings even under minimal information observed.

As mentioned before, one of the most critical concerns is obtaining the highest quality of the reconstructed dataset under minimal query leakage. A practical attack should be capable of extracting meaningful structural information under extremely constrained leakage, while also scaling to higher coverage when available.

To further assess this, we we fix the query ratio at two representative levels—1% and 25% (extremely low and relatively low query coverage)—and compare MSE results across the four real-world datasets. We additionally include the earlier method of Markatou et al. [38], which relies on full dataset ordering as a prerequisite and uses a statistical technique to infer the original data.

The results, summarized in Table 5.2 and 5.3 demonstrate that our method is significantly more query-efficient than the baselines. As expected, the method from [38] fails to operate under these settings, since recovering dataset order requires access to at least 50% of the query responses—far exceeding our evaluation budget. In contrast, As expected, the method from [38] fails to operate under these settings, since recovering dataset order requires access to at least 50% of the query responses—far exceeding our evaluation budget. For instance, under just 1% query coverage, our method achieves up to 80% reduction in MSE relative to the baseline, and in some cases reconstructs geometry with an MSE as low as 1.29.

These findings highlight the robustness of our framework in sparse regimes, demonstrating that even minimal query access can yield accurate geometric reconstructions—reinforcing the threat posed by range query leakage in real-world systems.

Table 5.2: MSE of reconstruction using 1% query ratio.

Method	Paris	Shanghai	AMS	Manhattan
REMIN	42.17	22.81	23.52	14.25
Markatou et al. [38]	-	-	-	-
Markatou et al. [36]	53.07	36.92	37.86	42.97

Table 5.3: MSE of reconstruction using 25% query ratio.

Method	Paris	Shanghai	AMS	Manhattan
REMIN	20.46	21.26	19.04	1.29
Markatou et al. [38]	-	-	-	-
Markatou et al. [36]	26.14	23.02	37.20	10.82

5.4. Reconstruction in Higher Dimensions

To evaluate the scalability and generalizability of our approach, we conduct experiments on synthetic grid-structured datasets with increasing dimensionality, from 2D to 6D. These settings reflect realistic encrypted database scenarios where records are embedded in high-dimensional feature spaces. As dimensionality grows, the reconstruction

task becomes significantly more challenging due to sparser co-occurrence observations and more complex geometric structures.

We compare our method against the leading query-efficient baseline [36] under a fixed query coverage of 1% and uniform query distribution. To manage the exponential growth in data size and computational cost, we use progressively smaller grids (e.g., 16^3 for 3D, 6^5 for 5D). Evaluation is based on MSE and Chamfer distance, both of which capture geometric fidelity. We omit tolerant match and neighbor accuracy in dimensions ≥ 3, as the geometric complexity and uniformity of inter-point distances in high-dimensional spaces render local neighborhood metrics less reliable.

Table 5.4: Performance on higher-dimensional grid datasets with 1% query ratio under uniform distribution.

Dim	Method	MSE	Chamfer Dist
2D (32^2)	REMIN	0.9753	0.4113
	Markatou et al. [36]	24.6954	0.5184
3D (16^3)	REMIN	0.2991	0.5414
	Markatou et al. [36]	0.5968	0.5969
4D (8^4)	REMIN	0.6547	0.6851
	Markatou et al. [36]	0.6814	0.7061
5D (6^5)	REMIN	0.8624	0.9174
	Markatou et al. [36]	1.0639	1.0369
6D (4^6)	REMIN	1.1759	1.0247
	Markatou et al. [36]	1.1888	1.1168

As shown in Table 5.4, our method consistently achieves lower MSE and Chamfer distance across all dimensions. Even in 6D, our attack maintains a clear advantage, demonstrating its robustness to geometric complexity. This suggests that our distance-based co-occurrence representation preserves meaningful spatial structure even when embedding dimensionality increases.

These findings demonstrate the scalability and robustness of our approach: despite extreme query sparsity and dimensional complexity, our attack can still recover the structural layout of the dataset far more effectively than existing methods. This underscores the potential risks of leakage in high-dimensional encrypted systems, where traditional defenses may underestimate the attacker's reconstruction capabilities.

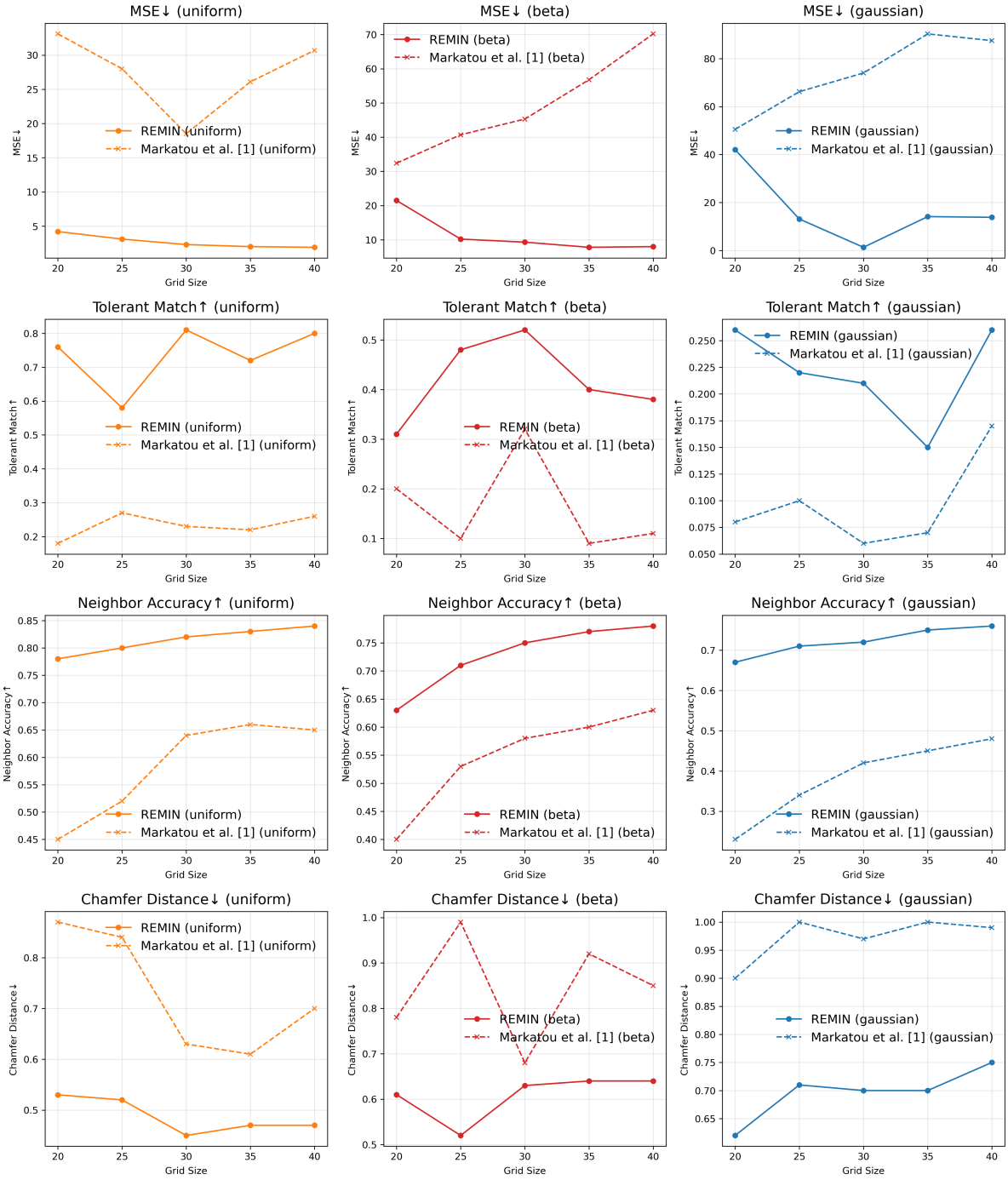


Figure 5.4: Reconstruction performance on varying grid sizes (20×20 to 40×40) and data distributions with 1% query ratio.

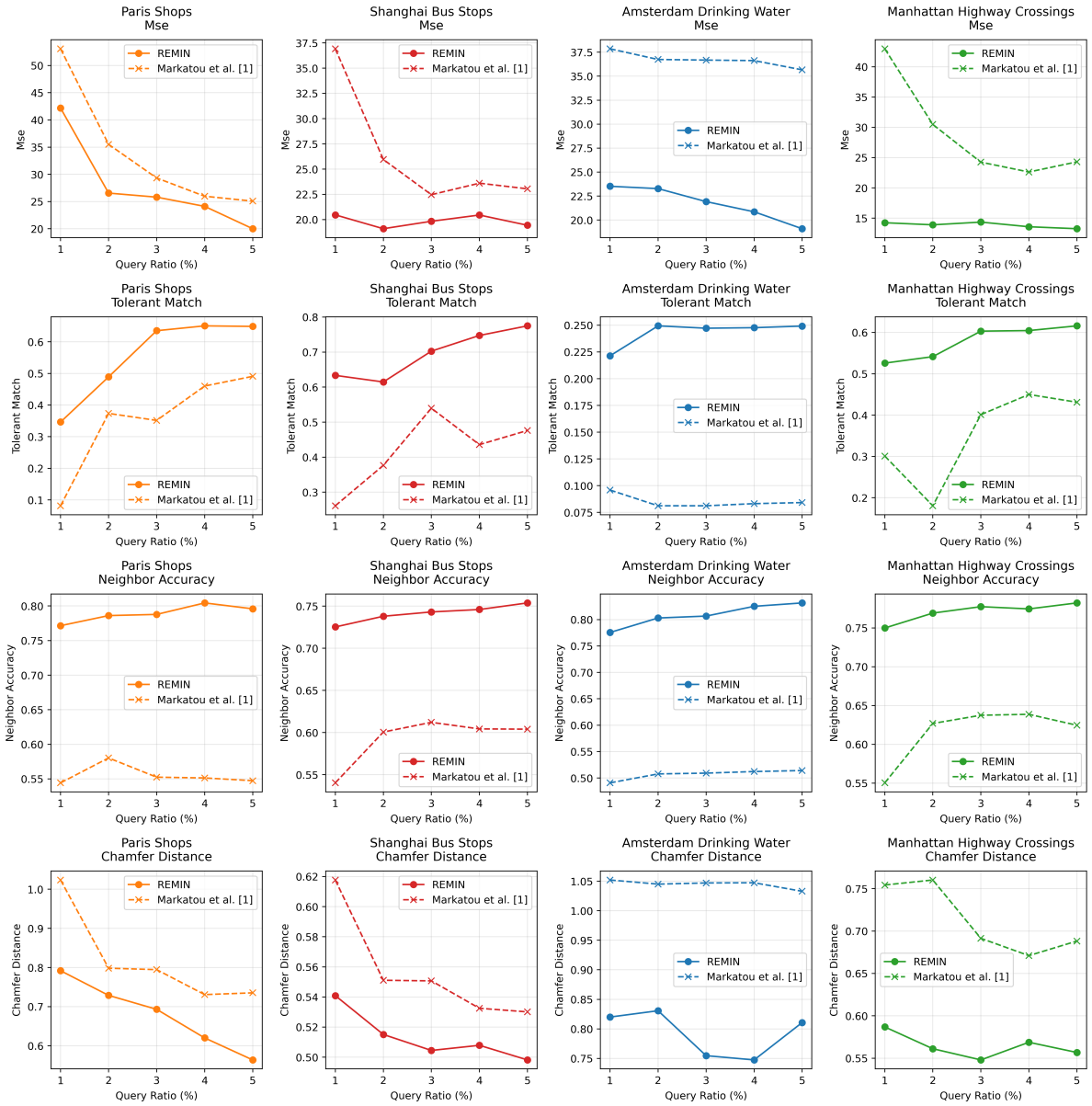


Figure 5.5: Reconstruction performance on real-world datasets under varying query ratios (1% to 5%) with uniform query sampling. Our method consistently outperforms method of Markatou et al. [36] and other baselines in all metrics.

6

REMIN-P Attack: Leveraging Auxiliary Information for Improved Reconstruction

In this section, we explore a new approach for reducing reconstruction error by utilizing auxiliary information. The technique— our poisoning strategy—operates on the output of any reconstruction algorithm and is agnostic to the underlying inference mechanism. As such, it represents a general augmentation layer that can enhance or exploit the structure recovered by any range leakage attack. We describe the new attack, called REMIN-P, and demonstrate its effectiveness in improving the database reconstruction.

6.1. Intention

We explore how to further minimize reconstruction error, especially in scenarios where limited query data leads to imperfect reconstructions. In practical settings, some records in the dataset may be known to the attacker, either due to unavoidable leakage or intentional injections. For instance, in a medical record database, an attacker can create an encrypted record for themselves by visiting a hospital at a specific time, thereby introducing a poisoned record into the system. This knowledge—referred to as poisoning anchor points—can be exploited to refine the reconstruction.

We show that even small amounts of auxiliary knowledge can significantly improve the reconstruction of the underlying data. This section explores how poisoned anchor points can be strategically used to adjust the geometry of the reconstructed dataset and reduce edge distortions, with a focus on the impact of two different injection strategies.

6.2. The REMIN-P Attack Framework

The core idea behind the REMIN-P Attack is to inject auxiliary anchor points into the reconstructed dataset, using known or leaked record positions as spatial references. These anchor points are then used to perform a post-processing adjustment to realign the reconstructed points, improving the overall structure and minimizing distortion.

Such auxiliary information may originate from different sources. We distinguish between two typical threat models:

- **Passive Attacker** gains access to a small, random subset of ground-truth record positions, either through metadata leakage, user-side exposure, or cross-database matching.
- **Active Attacker** deliberately injects or identifies structurally meaningful points (e.g., central locations, synthetic users) with known coordinates.

These scenarios motivate two corresponding injection strategies-random anchor selection and cross-shaped an-

chor placement (Fig. 6.1). The former explores the effect of anchor quantity under realistic, uncontrolled leakage, while the latter demonstrates how even a few strategically placed anchors can serve as structural scaffolds for global alignment.

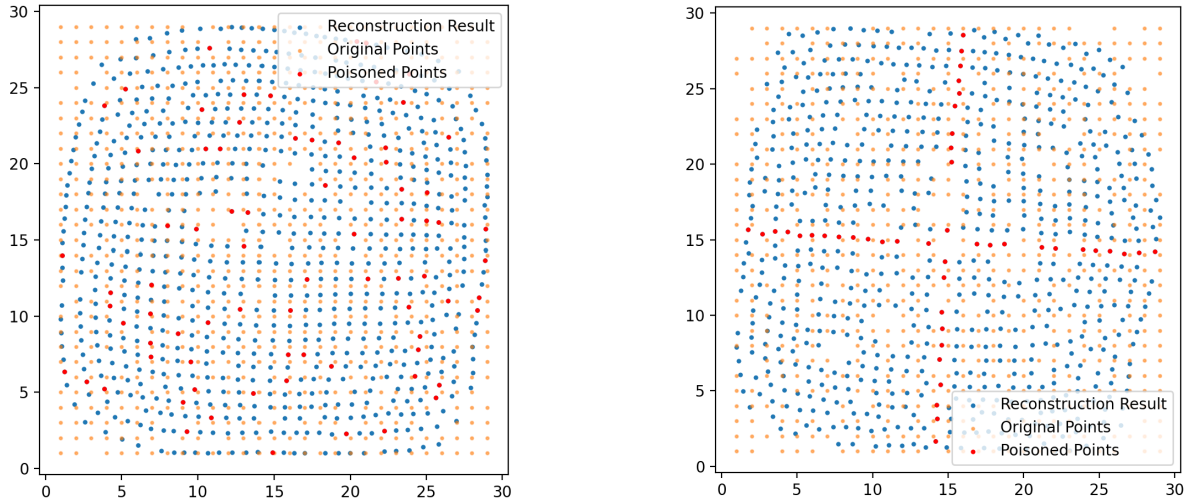


Figure 6.1: Comparison of two anchor injection strategies. The left figure illustrates the random placement of poisoned points, while the right shows the cross-shaped method, highlighting the structural improvement achieved by targeting central axes.

Given that the reconstructed embedding may exhibit mild geometric distortions due to manifold learning effects, we propose a post-processing adjustment step that leverages these known anchor points. Specifically, the attacker uses the identified correspondences between a small set of original and reconstructed coordinates to fit a smooth warping function that aligns the entire structure more closely with its true layout. The full procedure is outlined in Algorithm 4 (See Appendix D).

Since edge points typically have fewer neighboring points in the query responses, the positional information extracted for these points tends to be sparser, resulting in greater distortion in the reconstruction edges. To address this, we aim to align the reconstructed space with the true space based on the radial deviation of poisoned points relative to the layout center. By applying 1D interpolation along each axes, the algorithm efficiently warps the reconstructed layout to better match the ground-truth structure. Unlike global affine transformations, this method adapts to local distortions while preserving the relative neighborhood consistency of the reconstruction.

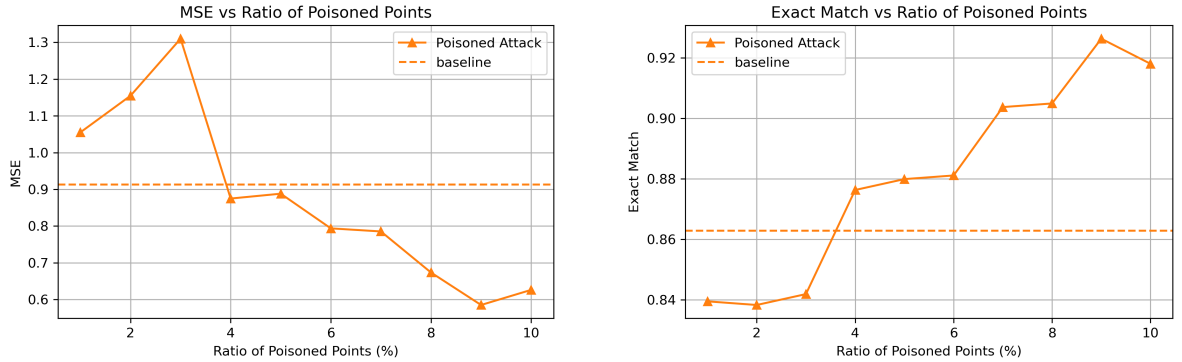
6.3. Experimental Evaluation

To evaluate the effectiveness of poisoning anchor strategies, we conduct experiments on structured grid datasets ranging from 20×20 to 35×35 in size with 5% query ratio. We simulate the attacker having access to the true coordinates of a small subset of records (anchors), which are used to fit a transformation that adjusts the reconstructed layout.

For the random injection strategy, we fix the grid size (30×30) and vary the anchor injection ratio across values 1% to 10%. For each ratio, we randomly sample the corresponding number of poisoned points from the dataset. For the cross-shaped injection, we vary the grid size from 20×20 to 40×40, and insert anchor points along the central row and column (i.e., forming a cross), with the number of poisoned points growing linearly with grid size.

Our results reveal complementary performance trends under the two anchor injection strategies, with the cross-shaped method achieving strong alignment with fewer anchors, and the random method showing gradual improvement as the injection ratio increases.

Random Injection. Fig. 6.2 illustrates the quantitative performance of the random poisoning strategy. When the poisoned point ratio is very low (13%), the adjustment often introduces additional distortion, degrading the reconstruction accuracy. This is attributed to biased or under-constrained interpolation, where the limited number of anchor points leads to unreliable global correction. However, once the poisoned ratio exceeds 3%, the adjustment consistently outperforms the original reconstruction. As the number of anchors increases, the quality



(a) Mean Squared Error↓

(b) Tolerant Match Ratio↑

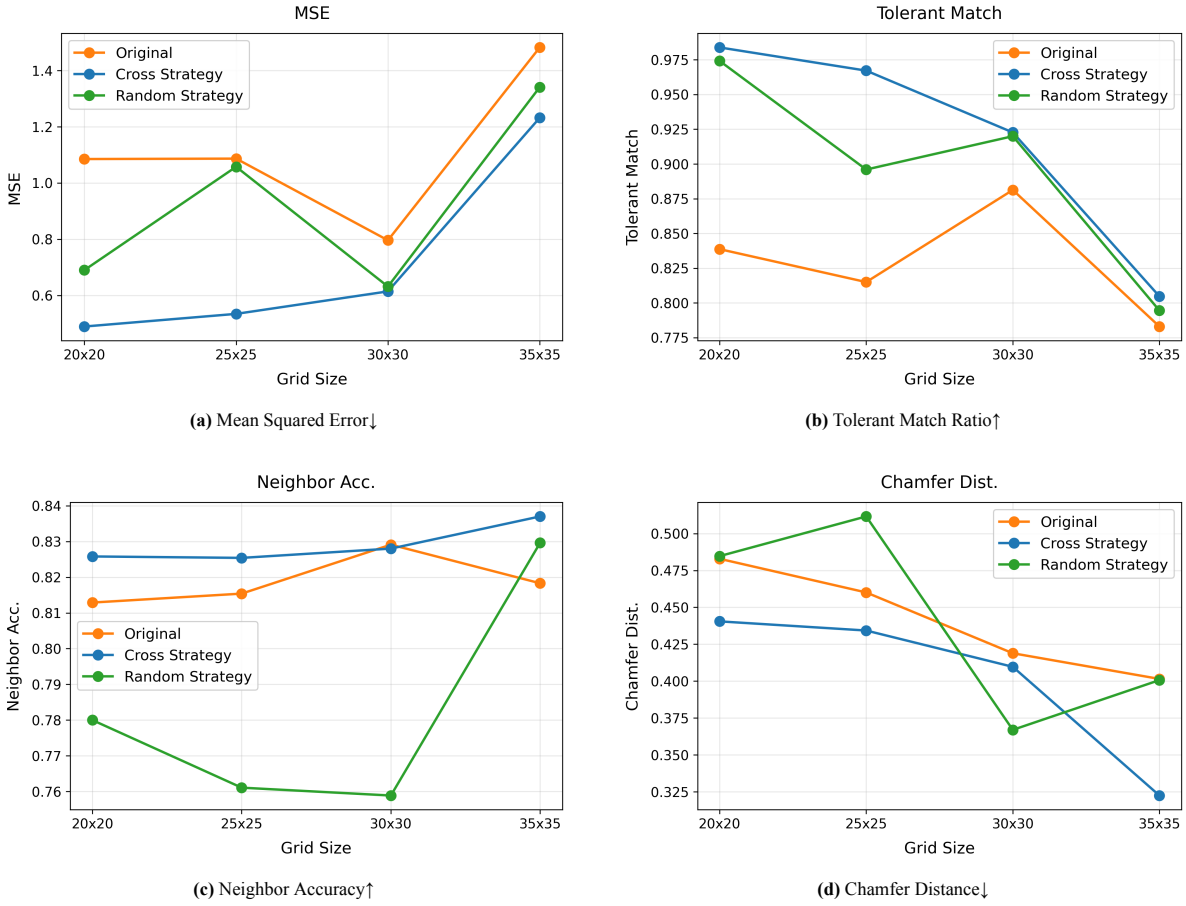
Figure 6.2: (Left) MSE and (Right) Tolerant match ratio of reconstructed coordinates as the ratio of poisoned points increases under the random injection strategy. Results are based on a 30×30 grid dataset with 5% query ratio.

of the correction steadily improves across all evaluation metrics. However, the improvement is gradual, and a large number of anchors are required to reach the performance level achievable by a well-placed set of structural anchors.

Cross-Shaped Injection. In contrast, the cross-shape strategy demonstrates that even limited auxiliary knowledge can be leveraged to significantly improve reconstruction fidelity. As shown in Fig. 6.3, even with only 19 anchors in a 30×30 grid (lower than 2.5%), we observe a reduction of over 22% in MSE and a gain of 4% in tolerant match accuracy compared to the no-anchor baseline. This targeted selection yields significant structural improvement, particularly in correcting edge distortions. The boundary regions, which often exhibit curvature or compression artifacts are realigned more accurately when poisoned points span both axes.

Moreover, as shown in Table 6.3, this method exhibits greater stability and effectiveness across dataset sizes (from 20×20 to 35×35) compared to random poisoning. Notably, while both strategies lead to improvements, the cross-shaped anchors consistently outperform random anchors at low budget levels, especially for correcting global misalignment.

Our experiments confirm that even limited auxiliary knowledge—either leaked or injected—can be leveraged to significantly improve reconstruction fidelity. In practice, it suggests that attackers may benefit from strategically introducing or identifying semantically informative points, thereby refining the geometry of otherwise noisy embeddings. This calls for stronger defense mechanisms beyond traditional leakage profiling, including robustness against minimal anchoring and geometry-aware leakage mitigation.



(a) Mean Squared Error↓

(b) Tolerant Match Ratio↑

(c) Neighbor Accuracy↑

(d) Chamfer Distance↓

Figure 6.3: Comparison of reconstruction metrics before and after post-processing under centerline poisoning attack. Grid sizes range from 20×20 to 35×35 under 5% query ratio.

7

Conclusion and Discussion

7.1. Conclusion

In this work, We address the challenge of reconstructing spatial datasets under severe query leakage constraints, a critical problem in privacy-preserving systems and adversarial settings. We present a novel machine learning-based attack that leverages representation learning from highly limited range query leakage. Our method, based on t-SNE and co-occurrence analysis, is the first to demonstrate effective reconstruction under general datasets, outperforming state-of-the-art baselines by up to 50% in MSE. These results highlight a critical vulnerability in current privacy-preserving systems and show that even minimal query exposure can lead to meaningful data recovery, calling for stronger defenses in real-world deployments.

7.2. Discussion and Future Work

Challenges in Extremely Sparse Datasets. Experimental observations indicate a dramatic decline in reconstruction performance in highly sparse datasets, characterized by over 70% empty regions. In such scenarios, the reduced co-occurrence of distant records severely limits the attacker's observable information. This sparsity also increases the likelihood of queries returning very few or even single records, further weakening the statistical signal needed for reconstruction. Future work could explore incorporating additional signals, such as the frequency of individual point occurrences, to estimate the size of empty neighborhoods, thereby improving reconstruction capabilities in sparse data environments.

Framework Limitations and Generalization. While REMIN shows strong performance under extremely limited query leakage, its scalability under higher leakage ratios (e.g., 50%) remains underexplored. Evaluating whether full database reconstruction is achievable in such settings may further reveal the method's robustness and practical potential. Moreover, the use of t-SNE introduces sensitivity to parameters such as perplexity and stochastic variability, which may affect reconstruction fidelity—especially in non-uniform datasets where isolated points offer limited relational signals. Although our poisoning-based REMIN-P helps mitigate some of these issues, future work could explore embedding algorithms that incorporate structural constraints or density-aware regularization. More broadly, the modular structure of our framework suggests potential for extending to other types of leakage, by transforming diverse observable signals into learnable spatial representations through flexible embedding techniques.

Opportunities and Challenges in Query Poisoning. Our REMIN-P shows that even limited manipulation—through observing or injecting a small number of ground-truth records—can substantially improve reconstruction outcomes. A slightly more active attacker could further exploit this by crafting specific queries or strategically inserting points to guide the embedding process. Future work may explore adaptive poisoning strategies that leverage dataset geometry or embedding sensitivities.

8

Ethics Considerations

This work demonstrates the vulnerability of searchable encryption systems to practical database reconstruction attacks using minimal leakage, which can help inspire stronger security measures and privacy-preserving techniques. By exposing these risks, our research contributes positively to the advancement of secure encrypted database systems. Below we discuss key ethical considerations regarding intellectual property, intended usage, potential misuse, risk control, and human subjects.

Intellectual property. All comparative attacks and defenses, models, datasets and implementation libraries are open-source.

Intended Usage. Our work demonstrates how minimal leakage from encrypted databases can enable approximate data reconstruction, exposing critical vulnerabilities in current systems. These findings aim to drive the development of more secure encryption schemes that better protect sensitive information.

Potential Misuse. Our findings could be misused to reconstruct sensitive location data, infer private financial/medical information, or manipulate data via poisoning attacks. Potential defenses may include padding [49], ORAM [21, 43], and poisoning detection system.

Risk Control. To further mitigate potential risks, we will release the code of this work including algorithms and the implementations in the experiments. We believe that transparency can reduce the risks related to this work, encourage reliable code reuse and promote advancement of database security.

Human Subject. This research does not involve human subjects and any personally identifiable information.

References

[1] Hervé Abdi and Lynne J Williams. “Principal component analysis”. In: *Wiley interdisciplinary reviews: computational statistics* 2.4 (2010), pp. 433–459.

[2] Lucas Ballard, Seny Kamara, and Fabian Monrose. “Achieving efficient conjunctive keyword searches over encrypted data”. In: *International conference on information and communications security*. Springer. 2005, pp. 414–426.

[3] Laura Blackstone, Seny Kamara, and Tarik Moataz. “Revisiting leakage abuse attacks”. In: *Cryptology ePrint Archive* (2019).

[4] Alexandra Boldyreva et al. “Order-preserving symmetric encryption”. In: *Advances in Cryptology-EUROCRYPT 2009: 28th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cologne, Germany, April 26-30, 2009. Proceedings 28*. Springer. 2009, pp. 224–241.

[5] Dan Boneh and Brent Waters. “Conjunctive, subset, and range queries on encrypted data”. In: *Theory of Cryptography: 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, The Netherlands, February 21-24, 2007. Proceedings 4*. Springer. 2007, pp. 535–554.

[6] Raphael Bost. “$\sum o\varphi o\varsigma$: Forward secure searchable encryption”. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. 2016, pp. 1143–1154.

[7] Raphaël Bost, Brice Minaud, and Olga Ohrimenko. “Forward and backward private searchable encryption from constrained cryptographic primitives”. In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. 2017, pp. 1465–1482.

[8] Miguel A Carreira-Perpinán. “A review of dimension reduction techniques”. In: *Department of Computer Science. University of Sheffield. Tech. Rep. CS-96-09* 9 (1997), pp. 1–69.

[9] David Cash et al. “Dynamic searchable encryption in very-large databases: Data structures and implementation”. In: *Cryptology ePrint Archive* (2014).

[10] David Cash et al. “Highly-scalable searchable symmetric encryption with support for boolean queries”. In: *Advances in Cryptology–CRYPTO 2013: 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*. Springer. 2013, pp. 353–373.

[11] David Cash et al. “Leakage-abuse attacks against searchable encryption”. In: *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*. 2015, pp. 668–679.

[12] Melissa Chase and Seny Kamara. “Structured encryption and controlled disclosure”. In: *International conference on the theory and application of cryptology and information security*. Springer. 2010, pp. 577–594.

[13] Jack Cook et al. “There’s always a bigger fish: A clarifying analysis of a machine-learning-assisted side-channel attack”. In: *Proceedings of the 49th Annual International Symposium on Computer Architecture*. 2022, pp. 204–217.

[14] Pádraig Cunningham. “Dimension reduction”. In: *Machine learning techniques for multimedia: Case studies on organization and retrieval*. Springer, 2008, pp. 91–112.

[15] Reza Curtmola et al. “Searchable symmetric encryption: improved definitions and efficient constructions”. In: *Proceedings of the 13th ACM conference on Computer and communications security*. 2006, pp. 79–88.

[16] Francesca Falzon et al. “Full database reconstruction in two dimensions”. In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. 2020, pp. 443–460.

[17] Craig Gentry. “Fully homomorphic encryption using ideal lattices”. In: *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 2009, pp. 169–178.

[18] Marilyn George, Seny Kamara, and Tarik Moataz. “Structured encryption and dynamic leakage suppression”. In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2021, pp. 370–396.

[19] Javad Ghareh Chamani et al. “New constructions for forward and backward private symmetric searchable encryption”. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018, pp. 1038–1055.

[20] Eu-Jin Goh. “Secure indexes”. In: *Cryptology ePrint Archive* (2003).

[21] Oded Goldreich and Rafail Ostrovsky. “Software protection and simulation on oblivious RAMs”. In: *Journal of the ACM (JACM)* 43.3 (1996), pp. 431–473.

[22] Philippe Golle, Jessica Staddon, and Brent Waters. “Secure conjunctive keyword search over encrypted data”. In: *International conference on applied cryptography and network security*. Springer. 2004, pp. 31–45.

[23] Paul Grubbs et al. “Learning to reconstruct: Statistical learning theory and encrypted database attacks”. In: *2019 IEEE symposium on security and privacy (SP)*. IEEE. 2019, pp. 1067–1083.

[24] Paul Grubbs et al. “Pump up the volume: Practical database reconstruction from volume leakage on range queries”. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018, pp. 315–331.

[25] Thomas Hofmann, Bernhard Schölkopf, and Alexander J Smola. “Kernel methods in machine learning”. In: (2008).

[26] Yong Ho Hwang and Pil Joong Lee. “Public key encryption with conjunctive keyword search and its extension to a multi-user system”. In: *International conference on pairing-based cryptography*. Springer. 2007, pp. 2–22.

[27] Mohammad Saiful Islam, Mehmet Kuzu, and Murat Kantarcioglu. “Access pattern disclosure on searchable encryption: ramification, attack and mitigation.” In: *Ndss*. Vol. 20. Citeseer. 2012, p. 12.

[28] Seny Kamara and Charalampos Papamanthou. “Parallel and dynamic searchable symmetric encryption”. In: *Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers 17*. Springer. 2013, pp. 258–274.

[29] Seny Kamara, Charalampos Papamanthou, and Tom Roeder. “Dynamic searchable symmetric encryption”. In: *Proceedings of the 2012 ACM conference on Computer and communications security*. 2012, pp. 965–976.

[30] Georgios Kellaris et al. “Generic attacks on secure outsourced databases”. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. 2016, pp. 1329–1340.

[31] Evgenios M Kornaropoulos, Charalampos Papamanthou, and Roberto Tamassia. “The state of the uniform: Attacks on encrypted databases beyond the uniform query distribution”. In: *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2020, pp. 1223–1240.

[32] Evgenios M Kornaropoulos et al. “Leakage inversion: Towards quantifying privacy in searchable encryption”. In: *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. 2022, pp. 1829–1842.

[33] Marie-Sarah Lacharité, Brice Minaud, and Kenneth G Paterson. “Improved reconstruction attacks on encrypted data using range query leakage”. In: *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2018, pp. 297–314.

[34] Chang Liu et al. “Search pattern leakage in searchable encryption: Attacks and new construction”. In: *Information Sciences* 265 (2014), pp. 176–188.

[35] Evangelia Anna Markatou and Roberto Tamassia. “Full database reconstruction with access and search pattern leakage”. In: *International Conference on Information Security*. Springer. 2019, pp. 25–43.

[36] Evangelia Anna Markatou and Roberto Tamassia. “Reconstructing with Even Less: Amplifying Leakage and Drawing Graphs”. In: *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. 2024, pp. 4777–4791.

[37] Evangelia Anna Markatou et al. “Attacks on encrypted response-hiding range search schemes in multiple dimensions”. In: *Proceedings on Privacy Enhancing Technologies* (2023).

[38] Evangelia Anna Markatou et al. “Reconstructing with less: Leakage abuse attacks in two dimensions”. In: *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*. 2021, pp. 2243–2261.

[39] MongoDB, Inc. *Queryable Encryption: Technical Guide*. Tech. rep. Accessed: 2025-04-20. MongoDB, 2022. url: `https://www.mongodb.com/docs/manual/core/queryable-encryption/`.

[40] Muhammad Naveed, Manoj Prabhakaran, and Carl A Gunter. "Dynamic searchable encryption via blind storage". In: *2014 IEEE Symposium on Security and Privacy*. IEEE. 2014, pp. 639–654.

[41] David Pouliot and Charles V Wright. "The shadow nemesis: Inference attacks on efficiently deployable, efficiently searchable encryption". In: *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. 2016, pp. 1341–1352.

[42] Dawn Xiaoding Song, David Wagner, and Adrian Perrig. "Practical techniques for searches on encrypted data". In: *Proceeding 2000 IEEE symposium on security and privacy. S&P 2000*. IEEE. 2000, pp. 44–55.

[43] Emil Stefanov et al. "Path ORAM: an extremely simple oblivious RAM protocol". In: *Journal of the ACM (JACM)* 65.4 (2018), pp. 1–26.

[44] Gui-Xian Tian, Mei-Jiao Cheng, and Shu-Yu Cui. "The generalized reciprocal distance matrix of graphs". In: *arXiv preprint arXiv:2204.03787* (2022).

[45] Warren S Torgerson. "Multidimensional scaling: I. Theory and method". In: *Psychometrika* 17.4 (1952), pp. 401–419.

[46] Laurens Van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008).

[47] Laurens Van Der Maaten, Eric O Postma, H Jaap Van Den Herik, et al. "Dimensionality reduction: A comparative review". In: *Journal of machine learning research* 10.66-71 (2009), p. 13.

[48] Yauhen Varabei et al. "Intelligent Clustering as a Means to Improve K-means Based Horizontal Attacks". In: *2019 IEEE 30th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC Workshops)*. IEEE. 2019, pp. 1–6.

[49] Serge Vaudenay. "Security flaws induced by CBC padding—applications to SSL, IPSEC, WTLS..." In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2002, pp. 534–545.

[50] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. doi: `10.1038/s41592-019-0686-2`.

[51] Guofeng Wang et al. "Leakage models and inference attacks on searchable encryption for cyber-physical social systems". In: *IEEE Access* 6 (2018), pp. 21828–21839.

[52] Jing Yao et al. "SoK: A systematic study of attacks in efficient encrypted cloud data search". In: *Proceedings of the 8th International Workshop on Security in Blockchain and Cloud Computing*. 2020, pp. 14–20.

A

Parameter Sensitivity

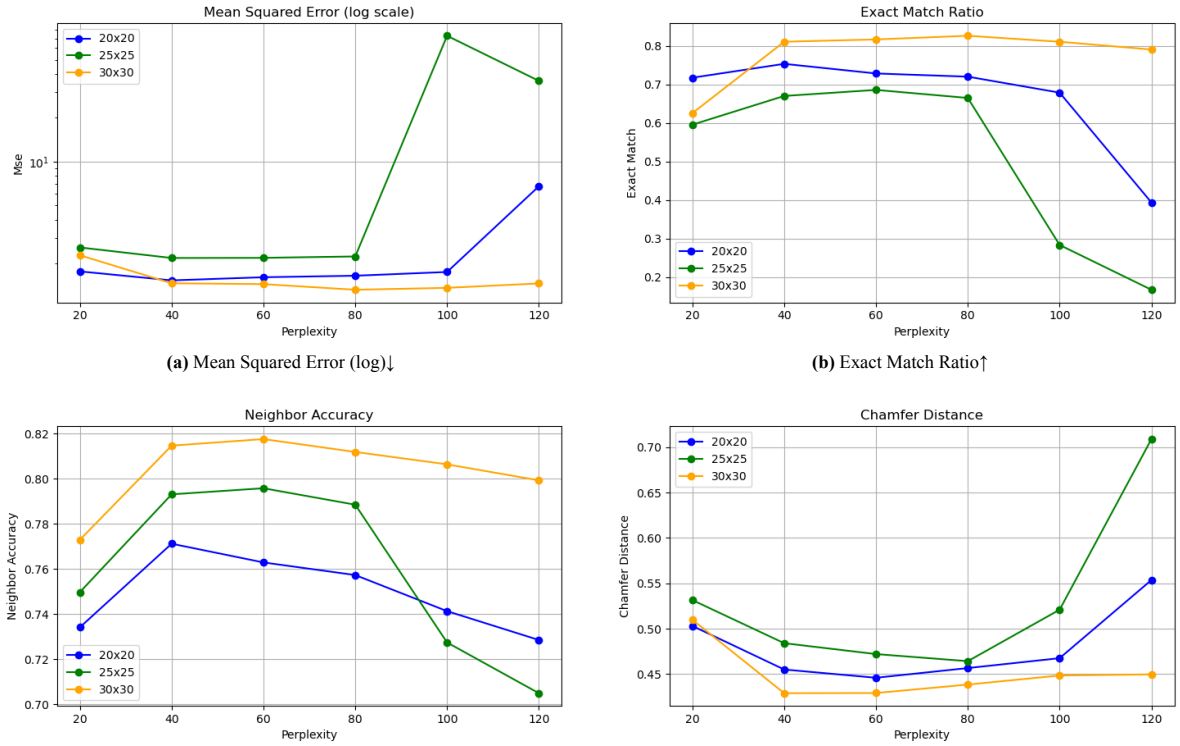
Following the attack design described above, we conduct a detailed investigation into the impact of key parameters on the reconstruction quality. Among them, the perplexity parameter of the t-SNE algorithm, a key parameter that balances the preservation of local and global structures, plays a particularly critical role. This section presents an empirical study of its sensitivity and how to select it appropriately for datasets of varying sizes.

t-SNE has been identified as the most suitable dimensionality reduction algorithm in our attack pipeline due to its strength in preserving local neighborhoods. However, its effectiveness is heavily influenced by the perplexity parameter, which governs the trade-off between local and global structure preservation. Perplexity can be intuitively interpreted as the effective number of nearest neighbors that t-SNE considers when constructing the probability distribution of points. A lower perplexity value forces the algorithm to focus on local neighborhoods, which may result in over-clustering and distortion of the global geometry. Conversely, a higher perplexity encourages better preservation of global structure, but may oversmooth local relationships, especially in datasets with fine-grained variations.

Given that the effective neighborhood size should ideally scale with the total number of points, we hypothesize that the optimal perplexity is not fixed across datasets, but rather dependent on the size of the dataset—i.e., the number of records to be reconstructed. To validate this hypothesis, we conduct a series of experiments measuring reconstruction performance under varying perplexity values across datasets of increasing size (20×20, 25×25, and 30×30 grids, corresponding to 400, 625, and 900 points, respectively). For consistency, all experiments are conducted on structured grid databases with 0% missing records and 1% query coverage under uniform distribution. We evaluate four key metrics—mean squared error (MSE), tolerant exact match ratio, neighborhood accuracy and Chamfer distance—across a wide range of perplexity values (from 20 to 120). Results are shown in Fig. A.1.

Our findings consistently support our hypothesis: as the total number of records increases, the optimal perplexity also shifts upward. For smaller datasets (e.g., 20×20, with 400 points), lower perplexity values (around 40) yield the best performance, as these preserve local structures without over-smoothing. In contrast, for larger datasets (e.g., 30×30, with 900 points), higher perplexity values (around 60-80) are necessary to capture sufficient neighborhood information and maintain global structural consistency.

These results suggest that perplexity should be carefully tuned based on dataset size. For small datasets (≤ 600 points), a perplexity of 30–50 is ideal, while medium-sized datasets (900 points), 60–80 is more effective. Larger datasets (≥ 35) require 80-120 to achieve optimal structure preservation. When perplexity is chosen appropriately, the reconstruction quality improves significantly, as visualized in Fig. 1.1.



(a) Mean Squared Error (log)↓

(b) Exact Match Ratio↑

(c) Neighbor Accuracy↑

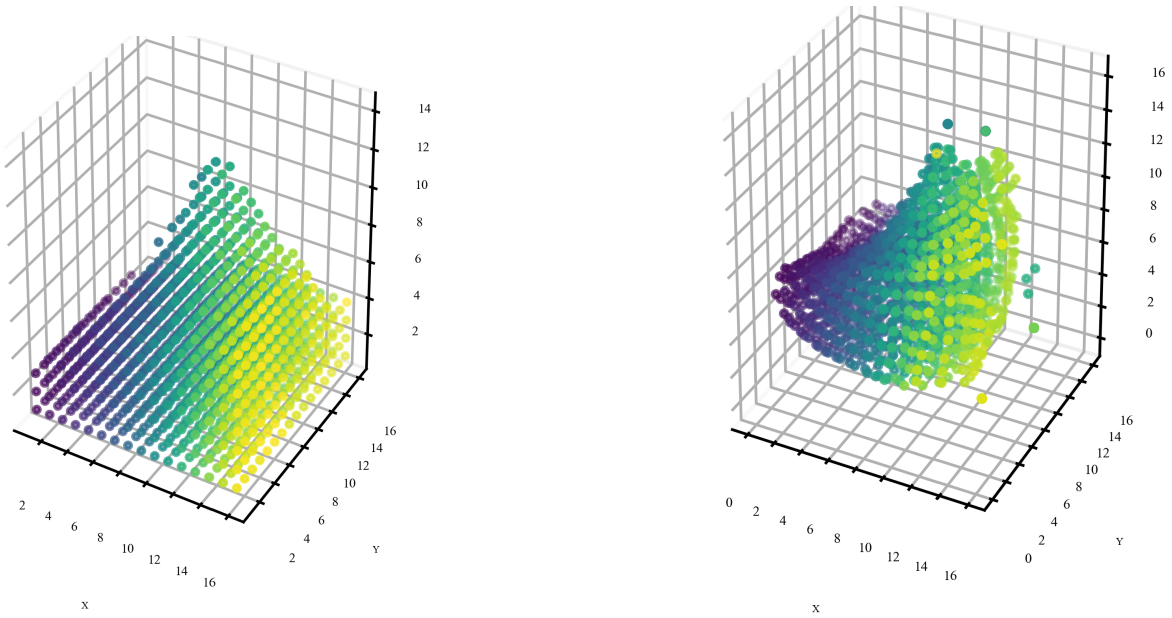
(d) Chamfer Distance↓

Figure A.1: Parameter Sensitivity Analysis: Perplexity vs. Dataset Size. Reconstruction performance (MSE, neighborhood accuracy, etc.) is evaluated across perplexity values (20–120) for grid datasets of increasing size (400–900 points). All tests use 1% query coverage and 0% missing data.

B

Supplementary Experiments

To complement the evaluation presented in the main text, we include additional figures in this appendix. These visualizations provide further evidence of the effectiveness and efficiency of our method across different settings.



(a) Original Coordinates

(b) Reconstruction of REMIN

Figure B.1: (Left) Original NH dataset in 3D coordinates. (Right) Reconstruction by our method REMIN using 1% uniformly sampled queries, achieving 72.38% neighbor accuracy, Chamfer distance of 0.5674, and MSE of 0.6745.

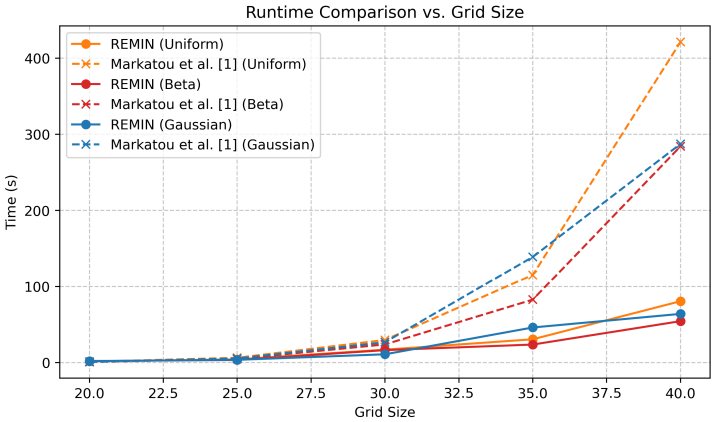


Figure B.2: Runtime across varying grid sizes (20×20 to 40×40) and data distributions with 1% query ratio.

C

Alignment and Refinement Algorithms

The reconstructed coordinates from co-occurrence data lack absolute position and may differ by rotation, scaling, and translation. To correct this, we use Procrustes Analysis (Algorithm 2) to align coordinates to a reference grid. Since Procrustes only handles global alignment, we further refine positions with Simulated Annealing (Algorithm 3) to snap points to integer grid locations, ensuring consistency with the original spatial structure.

Combining these methods yields a well-aligned, discretized embedding.

Algorithm 2 AlignAndScale($coords, grid\_shape$)

Input: $coords$: Input coordinates from previous step to be aligned, represented as an $n \times 2$ matrix
$grid\_shape$: The ranges of the dataset $((x_{min}, x_{max}), (y_{min}, y_{max}))$
Output: $scaled$: Aligned and scaled coordinates in target grid space

*/*Estimate grid size ($g \times g$) based on total points*/*
$n \leftarrow |coords|$, $g \leftarrow \lceil \sqrt{n} \rceil$
/ Generate a reference grid of n evenly spaced points */*
$grid \leftarrow$ first n points of meshgrid(g, g)
/ Align input coordinates to the reference grid using Procrustes analysis. Procrustes is an algorithm to align two point sets via translation, rotation, and scaling. Here we use the implementation from SciPy library [50]. */*
$, aligned \leftarrow$ Procrustes$(grid, coords)$
/ Define rescaling ranges */*
$x_r, y_r \leftarrow [1, grid\_shape[0] - 1], [1, grid\_shape[1] - 1]$
/ Normalize and scale to target range */*
$scaled \leftarrow$ normalize$(aligned) \times (x_r, y_r) + (x_r[0], y_r[0])$
/ Return final coordinates */*
return $scaled$

Algorithm 3 SimulatedAnnealing($coords, grid\_shape$)

Input: $coords$: Input coordinates from previous step to be aligned, represented as an $n \times 2$ matrix
$grid\_shape$: The ranges of the dataset $((x_{min}, x_{max}), (y_{min}, y_{max}))$
Output: S: Snapped coordinates

/ Init temp T, min temp T_{min}, cooling rate α */*
Initialize $T, T_{\text{min}}, \alpha$
/ Round coords and compute initial energy */*
Set $S = \text{round}(coords)$, $E = \text{Energy}(S, coords)$
/ Main annealing loop */*
while $T > T_{\text{min}}$ **do**
/ Generate candidate by perturbation */*
    $S' = \text{Perturb}(S, grid\_shape)$
/ Evaluate new candidate energy */*
    $E' = \text{Energy}(S', coords)$
/ Accept better/worse solution */*
    if $E' < E$ **or** random() $< e^{-(E'-E)/T}$ **then**
/ Update state and energy */*
        $S \leftarrow S', E \leftarrow E'$
    end if
/ Cool down temperature */*
    $T \leftarrow \alpha T$
end while
/ Return optimized snapped coords */*
return S

D

REMIN-P: Radial Correction Using Poisoned Anchors

Algorithm 4 constitutes the fundamental component of REMIN-P, aiming to refine reconstructed coordinates by leveraging a limited set of known poisoned points as anchor references. It first aligns the reconstructed embedding to the true coordinates using Procrustes analysis, then rescales to fit the original domain. Next, it learns radial correction functions mapping displacements from the center in reconstructed space to true space via interpolation on the anchor points. Finally, these mappings are applied to all points to obtain corrected coordinates that better match the true distribution.

Algorithm 4 RadialCorrection(X, $\hat{X}$, I, N_0, N_1)

Input: X: True coordinates of records, $X \in \mathbb{R}^{n\times 2}$
$\hat{X}$: Reconstructed coordinates, $\hat{X} \in \mathbb{R}^{n\times 2}$
I: Indices of poisoned (known) points
N_0, N_1: Original coordinate ranges
Output: $\tilde{X}$: Corrected coordinates

/ Center true/reconstructed coordinates */*
Compute mean μ_X and $\mu_{\hat{X}}$ of X and $\hat{X}$
/ Align shape and scale */*
Apply Procrustes alignment between X and $\hat{X}$ centered at μ_X, $\mu_{\hat{X}}$
/ Ensure domain bounds */*
Rescale aligned $\hat{X}$ to fit within $[0, N_0] \times [0, N_1]$
/ Select anchor points */*
Extract poisoned points: $X_I \leftarrow X[I]$, $\hat{X}_I \leftarrow \hat{X}[I]$
/ Displacement from center */*
Compute offsets: $\Delta x_i = \hat{X}_i^{(1)} - \mu_{\hat{X}}^{(1)}$, $\Delta y_i = \hat{X}_i^{(2)} - \mu_{\hat{X}}^{(2)}$
/ Learn radial mapping */*
Fit 1D interpolators:
$f_x : \Delta x_i \mapsto X_i^{(1)} - \mu_X^{(1)}$, $f_y : \Delta y_i \mapsto X_i^{(2)} - \mu_X^{(2)}$
for each point $j = 1$ to n **do**
/ X displacement */*
$\delta x_j \leftarrow \hat{X}_j^{(1)} - \mu_{\hat{X}}^{(1)}$
/ Y displacement */*
$\delta y_j \leftarrow \hat{X}_j^{(2)} - \mu_{\hat{X}}^{(2)}$
/ Corrected X */*
$\tilde{X}_j^{(1)} \leftarrow f_x(\delta x_j) + \mu_X^{(1)}$
/ Corrected Y */*
$\tilde{X}_j^{(2)} \leftarrow f_y(\delta y_j) + \mu_X^{(2)}$
end for
/ Return corrected coordinates */*
return $\tilde{X}$