Robust de-anonymization of a prototypical Electronic Health Record system

Karnav Popat[1], Swati Waghdhare [5], Sandeep Budhiraja [6], Anurag Agrawal [2, 4], Subhashis Banerjee [1, 3, 4, *]

1 Computer Science, Ashoka University
2 Trivedi School of Biosciences, Ashoka University
3 Centre for Digitalisation, AI and Society
4 Koita Centre for Digital Health
5 Department of Endocrinology, Max Hospital, Saket, New Delhi
6 Institute of Internal Medicine, Max Hospital, Saket, New Delhi

* suban@ashoka.edu.in

Abstract

The rapid adoption of electronic health records (EHRs) by both the public and private sectors of Indian healthcare has presented significant opportunities for healthcare innovation. At the same time, concerns have been raised about the privacy and security of personal healthcare records. This paper investigates the vulnerability of a prototypical EHR used by Max Healthcare, a leading Indian healthcare provider, to de-anonymization attacks. Using a model inspired by Narayanan and Shmatikov [1], we demonstrate that purportedly “anonymized” EHRs are susceptible to robust re-identification when the adversary has partial background information, even a minimal set that is expected to be reasonably available to acquaintances or through social media. We demonstrate that our re-identification attack retains its power even when the adversary has errors in their information or when the dataset has privacy-preserving noise added to it. We find that as few as four known data points on an individual are enough to de-anonymize victims in a personal medical dataset. We further outline a novel neighbourhood attack that exploits familial relationships in EHRs to enable the de-anonymization of entire family networks even when individual family-members’ data is completely secret. We show that the integration of such “anonymized” EHRs into the budding digital health ecosystem without other robust privacy measures exposes participating patients to complete medical and non-medical de-anonymization.

Author summary

With the rapid digitization of healthcare data across the world, a new class of privacy concerns is emerging. A common sharing practice for personal digital health data is to use or release it “anonymized”, i.e., by removing identifiers like name and date of birth and retaining relevant medical data. Our work shows that anonymization without verifiable access control is not sufficient to protect privacy, and that it is possible to uniquely de-anonymize patients from a supposedly anonymous dataset of medical information. This is a grave cause of concern for patient privacy worldwide. Data sharing requires more caution.

1 Introduction

Electronic health records (EHRs) have become integral to modern healthcare systems, offering significant benefits for patient care, research, and policy-making. India's healthcare ecosystem, one of the largest in the world, is at a critical juncture of evolving prototypes and standards for digitalization. However, while the rapid digitalization of health records has ushered in opportunities for innovation and efficiency, it has also heightened concerns about data privacy and security [2].

The Digital Information Security in Healthcare Act (DISHA) and the Digital Personal Data Protection (DPDP) Act in India represent two pivotal legislative efforts aimed at addressing these concerns. DISHA focuses specifically on healthcare, mandating stringent protocols for secure storage, sharing, and processing of health data. It provides a broad framework for Indian Digital Health Data (DHD) to be turned into EHRs and shared with different hospitals and healthcare centres. It requires fiduciaries to provide guarantees of purpose limitation and anonymity. Complementing this, the DPDP Act provides a broader framework for personal data protection, including health data, and introduces key principles such as purpose limitation, data minimization, and the rights of data principals. The Indian medical industry and government machinery have become exceedingly optimistic on EHRs as the pathway to the digitalization of the healthcare system [3] [4], with pioneers from both the private and public sectors [5] [6].

A data sharing model that can protect user privacy is essential for such digitalization endeavours. Data sharing is crucial not only for research, but also for effective intelligence gathering and epidemiology, data analytics and training of machine learning models, and topic discovery using health data. However, privacy preserving data sharing requires effective operational models for purpose limitation – ensuring that no function creep is possible – and neither of the Acts define the standards tightly enough. Data anonymization is by far the most common approach for purpose limitation of data sharing, though it is well known in computer science that it is rarely effective [7].

Despite efforts to anonymize sensitive health data, significant research has shown that anonymization methods are severely vulnerable to de-anonymization attacks, particularly in the context of high-dimensional and sparse datasets [1] [8] [9]. Medical data, by the nature of the information required to be stored, is often exceedingly sparse. Sparse data is especially easy to de-anonymize [10]. Specifically, this sparsity means that when adversaries possess auxiliary datasets or background knowledge, the risk of privacy breaches is further amplified. In the medical context, background knowledge about a medical subject can be easily obtained digitally or through day-to-day interactions. Existing research has shown that medical datasets can be extremely compromising to patient privacy [11] [12] [13].

An absolute notion of privacy might be that no information about an individual should be learnable with access to a statistical database that may not already be learnt without any such access. Indeed, this was the notion that was originally introduced by Dalenius [14] and later termed as *inferential privacy* by Ghosh and Kleinberg [15]. In her celebrated result, Dwork [16] not only proved that absolute inferential privacy is impossible to achieve, but also observed that if the adversary has access to arbitrary auxiliary information, an individual's inferential privacy would be violated even when she doesn't participate in the database, because information about her could be leaked by correlated information of other individuals. This led to the development of the notion of *differential privacy* [16] [17] [7], which measures the difference in the information gained by the adversary when the individual's data is collected vs. when it is not collected; thus it measures the *additional* privacy risk an individual incurs by participating in a database. This framework provides a mathematically rigorous approach to privacy, leveraging calibrated noise to protect against a wide range of adversarial attacks, including those leveraging auxiliary information.

While differential privacy has shown significant promise in protecting aggregate query results, its application to high-dimensional and sparse datasets, such as electronic health records (EHRs), remains a challenge. This is mainly because making people indistinguishable in sparse high-dimensional dimensional datasets requires adding so much noise, that the utility of the data becomes questionable. Research building on Dwork's foundational work has explored constructions like the Laplace mechanism [17] [18] and extensions to interactive and non-interactive settings [19], yet these approaches often struggle to balance privacy guarantees with data utility in sparse domains.

The privacy risks of releasing and using "anonymized" micro-data (such as EHRs) have been widely documented [20]. Landmark studies include the re-identification of a Massachusetts hospital discharge database [21], which demonstrated how voter registration records could be used to identify patients, and privacy breaches involving anonymized AOL search data [22]. Narayanan and Shmatikov [1] extended these findings by demonstrating robust de-anonymization techniques on large, sparse datasets, highlighting that even a small amount of auxiliary information can compromise privacy. Their work underscored the fundamental challenges in anonymizing high-dimensional datasets while preserving their utility for research and analysis, which has only seen success in limited scopes [23] [24] [25].

In this paper, we use the model suggested by Narayanan and Shmatikov [1] to demonstrate that EHRs from a major Indian hospital that are purportedly "anonymized", are vulnerable to robust de-anonymization with even a small amount of background information. We specify both the formal model of privacy breach as well as the practical scenarios of EHR de-anonymization by demonstrating a de-anonymization attack on "de-identified" individual-level data from Max Healthcare using minimal auxiliary information. We outline a novel attack specific to sparse medical data, and validate it on a synthetically derived dataset. We highlight the privacy risks associated with making anonymization guarantees to released sensitive data which is statistically de-anonymizable.

The findings from this study underscore the critical need for alternatives and deeper analysis of data sharing models.

2 Results

2.1 The Max Healthcare Dataset

To test our de-anonymization model on real-world EHRs, we use an anonymized (de-identified) dataset provided by Max Healthcare.

Table 1 contains a summary of the features in the provided data. The dataset is long-format and cross-sectional, containing 2,692 patient records with 629 possible features for each patient. 442 of the 629 features consist of survey responses, relying on self-reporting by the patient or relatives, along with information from past medical records. 187 features consist of measurements and observations including the results of medical tests and information such as height, weight, pallor, etc.

The features in the dataset can also be categorized based on the information they contain. There are 7 features holding personal data such as gender, date of birth, marital status, etc. There are 53 features pertaining to the patient's responses on their lifestyle, such as travel habits, alcohol consumption, etc. There are 382 features containing medical information, such as allergies, patient's history with specific diseases (whether they have had it and when), and patients' relatives' history with the diseases. Finally, there are 164 features representing various tests and measurements, from pulse and temperature to platelet count and various other blood test components.

Type	**Category**	**Feature**	**Count**
Survey	Personal	Personal	7
Survey	Lifestyle	Sleep	1
Survey	Lifestyle	Physical Activity	15
Survey	Lifestyle	Food habits	5
Survey	Lifestyle	Alcohol History	12
Survey	Lifestyle	Travel History	5
Survey	Lifestyle	Social & Personal History	15
Survey	Medical	Allergies	6
Survey	Medical	Current Medications	80
Survey	Medical	Past Medical History	56
Survey	Medical	Past Surgical History	9
Survey	Medical	Family History	231
Observation	Medical	Anthropometric Data	5
Observation	Medical	Investigation Comments	18
Observation	Measurement	Vitals	4
Observation	Measurement	Systemic Examination	8
Observation	Measurement	General Examination	13
Observation	Measurement	Blood Test Report	136
Observation	Measurement	Radiographs/Spirometry	3
		Total	**629**

Table 1. Max Healthcare dataset features

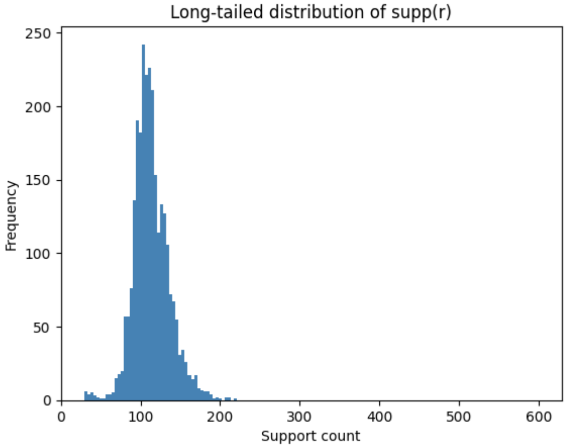

Fig 1. Long-tailed distribution of number of non-null columns per person

This dataset, as expected of an individual-level set of personal health records, is sparse. The majority of records have fewer than 200 out of 629 features with non-null values, and the majority of features have fewer than 100 records with non-null values. A full description of the dataset with detailed features and examples is included in Appendix 2.

Per our formal specification of sparsity, we can show that the Max Healthcare dataset is sparse by evaluating the $Similarity(r, r')$ of each pair of rows in the dataset (refer to section 4.4 for $Similarity$ function). We refer to the set of non-null features of a record r as the *support* of the record $supp(r)$ (similarly $supp(c)$ for a column c). Fig. 1 and Fig. 2 show that the support functions for rows and columns are sharply peaked

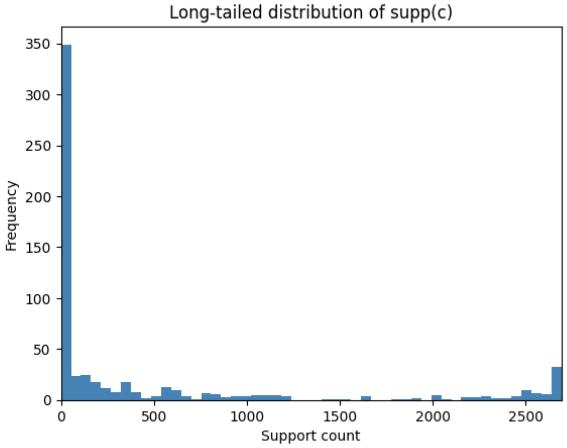

Fig 2. Long-tailed distribution of number of non-null records per feature

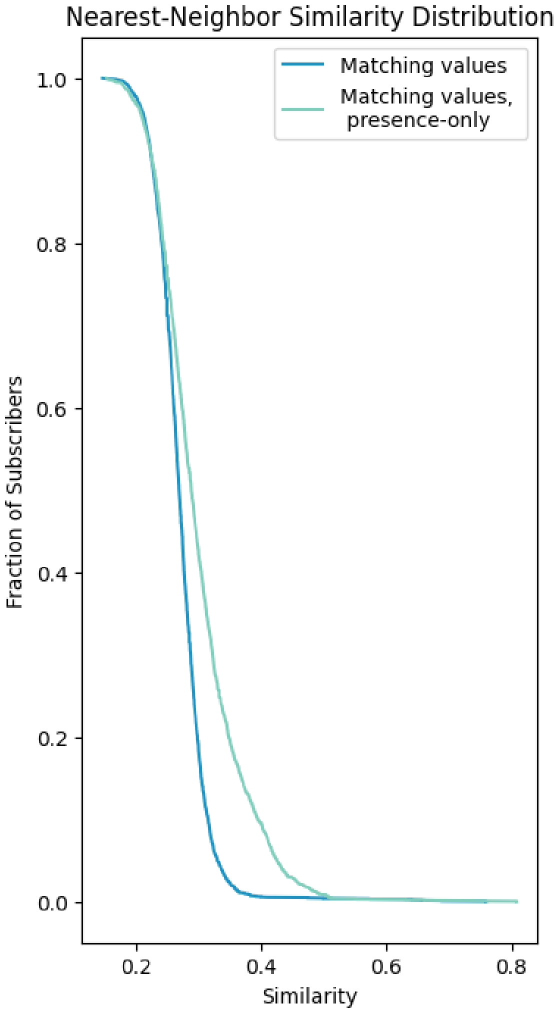

Fig 3. The majority of records in the Max Healthcare dataset have no other record with $Similarity > 0.4$

and long-tailed, as may be expected for sparse datasers.

Moreover, if the dataset is sparse, we would expect the majority of records r to have no other record r' with a high degree of similarity. Fig. 1 demonstrates that this is true for the dataset. Fig. 2 similarly shows that this is equally true when considering the sparsity of the features themselves. As we can see, most of the features are only non-null for a single-digit number of patients; this is why we can be certain that having even a small amount of information about a person's medical history, habits, etc. can be strongly indicative.

Fig. 3 shows that our dataset is sparse by considering each record's most similar record in the dataset. By comparing every feature in every record against every other record, we can show that for the vast majority of records, practically all records are at least 60% dissimilar from every other record; even when considering only the presence of a non-null value, practically all records are at least 50% dissimilar from every other record.

2.2 De-anonymization attack

Figure	$aux(r)$	$Sim(r, r')$
4	Most Informative	Perfect
4	Most Informative	20% Tolerance
4	Most Informative	Presence-only
5	Random	Perfect
5	Random	20% Tolerance
5	Random	Presence-only
6	Acquaintance	Perfect
6	Acquaintance	20% Tolerance
6	Acquaintance	Presence-only
7	Insider	Perfect
7	Insider	20% Tolerance
7	Insider	Presence-only
8	App	Perfect
8	App	20% Tolerance
8	App	Presence-only

Table 2. De-anonymization experiments

To test the degree of privacy in the dataset, we performed a background-information attack, as described in section 4.2, where we modeled an adversary that obtains a small amount of background information $aux(r)$ about a person and uses it to re-identify the person from their anonymized record. We conducted a series of thirty experiments, with different combinations of the auxiliary information selection function and similarity function.

To evaluate the de-anonymization potential, for each record $r \in D$, we extract $aux(r)$ based on the variant of the auxiliary information function we've chosen. We then use our attack algorithm ($A[D, aux(r)]$, defined in section 4.2) to find the output r' ($A[D, aux(r)] \rightarrow r'$). If $r == r'$, i.e., the indices of the source and target record match, we have successfully de-anonymized the record r.

We carried out the experiments summarized in Table 2, with results summarized in Table 3:

Figure	$aux(r)$	$Sim(r, r')$	k for $> 90\%$	k for $> 80\%$	k for $> 60\%$
4	Most Informative	Perfect	4	4	3
4	Most Informative	20% Tolerance	5	4	3
4	Most Informative	Presence-only	-	-	10
5	Random	Perfect	5	4	3
5	Random	20% Tolerance	10	8	6
5	Random	Presence-only	-	16	11
6	Acquaintance	Perfect	4	3	3
6	Acquaintance	20% Tolerance	10	8	6
6	Acquaintance	Presence-only	-	-	-
7	Insider	Perfect	6	4	3
7	Insider	20% Tolerance	7	5	4
7	Insider	Presence-only	-	-	-
8	App	Perfect	4	4	3
8	App	20% Tolerance	11	9	7
8	App	Presence-only	-	-	-

Table 3. De-anonymization results

2.3 Adversary has most informative $aux(r)$

First, we establish the worst-case scenario by providing the adversary with the k most informative features for each record measured by rarity of non-null values. Although this is not a realistic attack scenario, it provides a useful baseline for comparison against different allowed error rates.

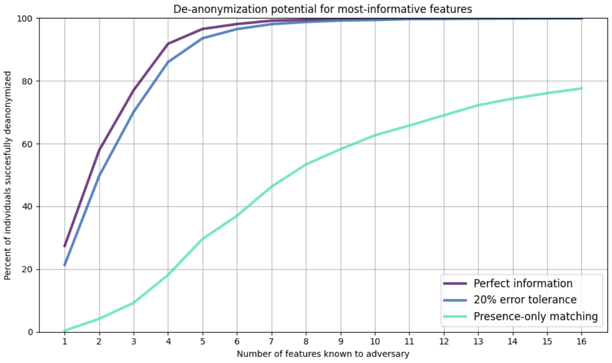

Fig 4. De-anonymization with k most informative features

2.3.1 Perfect information

First, we assume that the adversary has access to auxiliary information that is perfect, i.e. there is no noise added to the database and no imperfections in the adversary's information. In such a case, We find that only 4 features are sufficient to identify $> 90\%$ records in the Max Health database with high certainty (no other record with > 0.8 similarity). With an additional feature (5 in total), this is true of more than 97% of records.

2.3.2 20% error tolerance

Introducing a ±20% error tolerance in matching the adversary's information against the database (equivalent to adding upto 20% of noise in the database) has an equivalent 20% effect on the de-anonymization power, now requiring 5 features to identify records with high certainty.

This is significant because it indicates that, if the adversary has access to some uncommon background information about a victim, even extremely high errors in their knowledge can be compensated for with volume of information. This opens up a new class of "brute force guessing" attacks by which an adversary with the vaguest idea about the patient can repeatedly guess a larger volume of features, thereby successfully finding the patient's record in the database.

2.3.3 Presence-only information

Finally, we test without matching any values, checking only to see whether the record has a non-null value for each feature. This simulates the most robust condition where an adversary might have limited or no information about the details of a patient's medical condition, only that the patient has some history with a disease or lifestyle or whether they underwent the test in question. We find that this reduces the de-anonymization power significantly, requiring 10 features to reach $> 60\%$ success rate for most informative features, while it fails to reach the 90% threshold entirely.

Although it might appear that presence-only matching represents a failure condition for the de-anonymization attack, we must remember that presence-only matching represents the most harmful and dangerous form of background information amplification [1] [31] [8]. In the previous cases, the adversary needed to have some concrete information about the victim in order to de-anonymize them. In the presence-only case, very weak information, vague knowledge or demographic guesses would suffice to compromise their privacy. The most pernicious example would be information about medical tests. Under the presence-only paradigm, simply knowing that a patient has been tested for cancer or some other disease in the last few years would be enough to find out the results of the tests! This is a clear case of severe privacy harm for which a 60% vulnerability is extremely high.

Further, consider the case where the adversary has access to a comprehensive database and is implementing both the de-anonymization attack with presence-only information, as well as the neighbourhood attack that we outline in section 4.5. In this scenario, even though the adversary knows about one patient's test status (say, that a child has recently been tested for some infectious disease), they can use that information to de-anonymize the child as well as the parents and other relatives. In this case, the "low" 60% de-anonymization yielded by the extremely limited background information translates to far more people being harmed.

2.4 Randomly selected features

A more realistic example is an adversary who has access to some arbitrary information set about a person and wants to use that external background knowledge to identify the victim in the database. Our attack remains feasible for a set of k features randomly selected from the set of non-null features of the target record. This exhibits similar performance to the most informative set for perfect and 20% error-tolerated information.

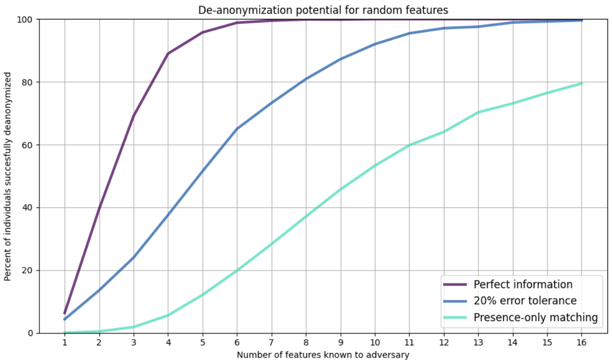

Fig 5. De-anonymization with k randomly selected features

2.4.1 Perfect information

We find that with as few as 5 randomly selected features available to the adversary as $aux(r)$, it is possible to uniquely identify $> 95\%$ of records with high certainty (no other record with > 0.8 similarity).

2.4.2 20% error tolerance

With randomly selected features and an error margin $(+-20\%)$, we require 10 features to de-anonymize $> 90\%$ of records in the database. This falls to 8 features for a tolerance of 80%. This shows that when the adversary only has access to more common (less sparse) features, it becomes tougher to tolerate "guesses" or errors in their knowledge.

2.4.3 Presence-only information

With presence-only matching, the percentage of de-anonymized records stays below the 90% threshold value. With a full 16 presence-only features available to the adversary, it is possible to de-anonymize 80% of the patients in the dataset, and 60% with 11 features. Note that these still represent extremely achievable volumes of information, since 11 features of presence-only data can be accumulated even solely through background knowledge of the victim's lifestyle and recent medical tests, both of which can be collected through the public-facing internet or for any of the adversary's many acquaintances.

Finally, we test our attack using different subsets of auxiliary information representing the information realistically available to different adversaries. To demonstrate vulnerability to specific adversaries, we demonstrate three different realistic adversaries who acquire background information through plausible means and are able to successfully de-anonymize the patient.

2.5 Malicious colleague's information

2.5.1 Perfect information

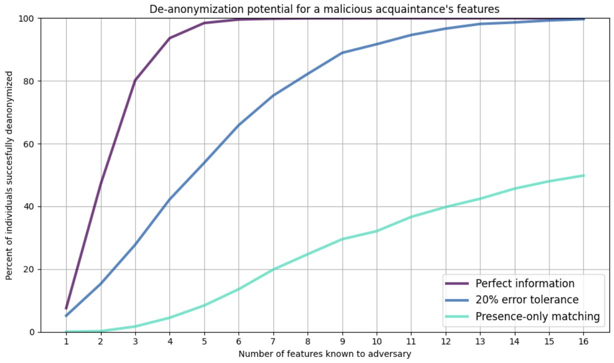

Fig 6. De-anonymization with malicious colleague's information

Under the assumption that a malicious colleague has perfect information about a subset of features, we find that 4 features are sufficient to identify $> 90\%$ of records in the Max Healthcare database with high certainty, and 3 features for $> 80\%$ certainty. This is comparable to the performance observed with the most informative features, which is particularly concerning since this is the class of information that is most accessible not just to adversaries in contact with the victim but also those observing the victim's internet presence.

2.5.2 20% error tolerance

Introducing a $+-20\%$ error tolerance requires 10 features to identify $> 90\%$ of records and 8 features for $> 80\%$ certainty. This shows greater impact from noise compared to perfect information scenarios, reflecting that these features are more tightly bound. That is, because the information in this subset is more likely to be correlated together (because lifestyle traits tend to cluster), it is less likely to represent statistically unlikely combinations that can be exploited to identify the patient.

2.5.3 Presence-only information

Under the presence-only information paradigm, the attack only de-anonymizes upto 50% of patients. For the other subsets of quasi-public information, it is even weaker - 10-30%. While this means fewer patients are immediately identified, the danger that the attack represents is still extremely high, as indicated in 4.1.3. The large amount of additional information that the presence-only attack reveals, combined with the potential for further amplification through the neighbourhood attack, represents stark privacy risks.

2.6 Medical insider's information

2.6.1 Perfect information

A medical insider, such as a nurse or lab technician, may have access to a broader range of features. Similarly, an adversary within the medical system or who has gained access to data shared within EHRs or an ABDM-like ecosystem, has access to many features. With perfect information, we find that 6 of these features are sufficient to identify $> 90\%$ of records with high certainty, and 4 features for $> 80\%$ certainty.

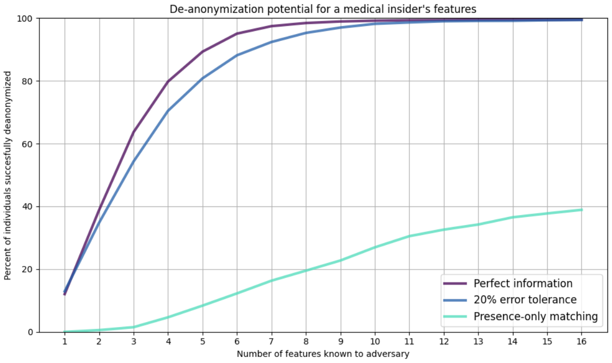

Fig 7. De-anonymization with medical insider's information

2.6.2 20% error tolerance

With a $+-20\%$ error tolerance, the medical insider requires 7 features to de-anonymize $> 90\%$ of records and 5 features for $> 80\%$ certainty. This demonstrates moderate robustness against noise and reveals that there is an even higher risk from a medical insider (who could be a doctor, nurse, medical technician, or anyone embedded into a medical data ecosystem such as ABDM), compared to an ordinary acquaintance.

2.7 Smart medical service's information

2.7.1 Perfect information

A smart medical service, such as a health tracking app, may have access to a curated set of health-relevant features. With perfect information, we find that 4 features are sufficient to identify $> 90\%$ of records with high certainty. This is again comparable to the performance of the most-informative features.

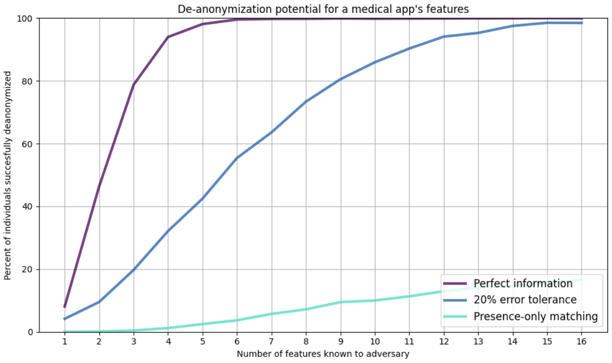

Fig 8. De-anonymization with smart medical service's information

2.7.2 20% error tolerance

With a $+-20\%$ error tolerance, the smart medical service requires 11 features to reach $> 90\%$ de-anonymization and 9 features to reach $> 80\%$ with 9 features. This is again similar to the performance of a malicious colleague or acquaintance.

3 Discussion

Our findings reveal significant privacy vulnerabilities in EHRs, particularly when realistic adversaries possess even small amounts of external information. Our de-anonymization experiments demonstrated that many of the stakeholders in the EHR ecosystem, when able to "guess" as few as 4 to 11 features, could potentially re-identify a substantial proportion of patients in the dataset. Any malicious acquaintance, or any online medical service, when given seemingly innocuous details such as lifestyle habits or basic medical history, could exploit the sparsity of EHRs to compromise privacy. This vulnerability is robust to various constraints, including the extent of the adversary's knowledge, random errors in their "guesses", and limitations based on the nature of information plausibly available to them. The ease with which realistic adversaries can breach privacy highlights the urgent need for caution when introducing personal and sensitive data into the digital healthcare ecosystem.

One caveat of our attack model is that, in the case of adversaries using semi-public or "guessable" information, the adversary must guess up to 10 uncorrelated or semi-correlated data points to successfully execute the attack. For example, information about a person's smoking history might be encoded in multiple columns, but each of those columns is highly correlated to one another; as a result, the adversary must obtain information on more independent characteristics such as smoking history, travel habits, etc. to create a uniquely-identifiable combination of features. However, the weak correlation that still exists between such features means that there are still many potential avenues of de-anonymization.

The risk of escalation is further increased by the integration of these EHRs into the Indian health ecosystem through the government's Ayushman Bharat Digital Mission (ABDM). The inter-connectedness of datasets means that the de-anonymization of a single EHR can have far-reaching consequences. Under the ABDM model, each record in the EHR is embedded with a unique identifier (the Ayushman Bharat Health Account (ABHA) ID), which links to the patient's record with every other healthcare provider and which is in turn linked to the patient's government identity (Aadhaar ID). This means that a successful de-anonymization attack on one EHR could potentially expose an individual's identity across the entire Indian digital ecosystem. This breach could reveal not only sensitive medical information but also other critical personal data, including financial identities and linked demographic details.

The harms that we outline are therefore severe. The simple de-anonymization of an EHR that we have demonstrated can be escalated in both breadth (to the victim's family) and in depth (to the victim's non-medical data). Once an adversary collects the small amount of background information needed to identify the victim in the "weakest link" healthcare provider's EHR, they can then use that to identify the victim across the healthcare ecosystem, escalate the de-anonymization to their family members, and obtain critical private details for every family member with a digital health record. Such a scenario would not only compromise individual privacy but also erode public trust in the digital healthcare system.

When the adversary is restricted to knowing only whether there is a non-null feature (rather than having some actual, perfect or error-prone, information about the patient's data), the absolute success rate of the de-anonymization falls. However, the presence-only de-anonymization results actually represent extreme risk, with possibly far more privacy harm than the other cases.

Consider the case of a woman who smokes. All women above the age of 40 are recommended to get regular screenings for breast cancer [37] [38]. Consider an adversary with only three pieces of information about a victim X: "X is female", "X is above the age of 40", and "X smokes". All of this information can be trivially gathered or guessed about millions of people from their internet presences, including public social media profiles, public academic or professional profiles, or broader personal databases. Using this information, and the knowledge of early cancer screening guidelines, the adversary can make informed guesses about more than 20 features for the patient's record in our dataset (smoking habits, lifestyle choices, medical history including cancer screenings and other routine tests). This would be more than sufficient to de-anonymize her with high certainty. This creates a scenario where an adversary who has basic background information and knows a patient recently had a cancer screening, can find out the results of her cancer diagnosis! Consequently, a database where $> 60\%$ of patients can be de-anonymized with presence-only information represents an extremely insecure database where $> 60\%$ of patients have functionally no privacy.

As Cynthia Dwork's [31] work establishes, the inherent sparsity of personal datasets makes meaningful privacy practically unattainable. Even when an individual's data is not explicitly included in a dataset, their privacy can still be compromised through correlations and auxiliary information available to adversaries. This is particularly concerning in the context of Indian healthcare, where the scale and critical nature of medical data exacerbate these vulnerabilities. With one of the largest healthcare ecosystems in the world, India's rapid digitalization of health records creates vast datasets that are both high-dimensional and sparse, making them prime targets for de-anonymization attacks. The sensitivity of medical data, combined with the potential for widespread privacy breaches, poses significant risks to patient trust and the integrity of the healthcare system.

The only viable way to safeguard sensitive health data in the face of these vulnerabilities is through the implementation of strongly restrictive access paradigms and strong purpose limitation mechanisms. Access controls must ensure that only authorized entities can access specific subsets of data, while purpose limitation must strictly define and enforce the permissible uses of the data, preventing function creep and unauthorized secondary uses. However, this is not currently the case in Indian EHR systems such as Max's, nor in the government's electronic health project, the Ayushman Bharat Digital Mission (ABDM). To preserve privacy in the digital healthcare ecosystem, it is imperative to adopt stricter access controls, implement purpose limitation at both the policy and technical levels, and ensure compliance through rigorous auditing and accountability measures. Without these changes, the promise of digital health in India will remain overshadowed by significant privacy risks.

3.1 Neighbourhood attack

The implications of the neighbourhood attack specified in section 4.5 further amplify the privacy risks associated with EHRs, as we demonstrate that an individual's data can be compromised even if their own information remains entirely private from any adversary. To investigate the feasibility of using one known record's information to identify relatives' records, we conducted a series of experiments representing two attacks: one attempting to identify the parents' records, and one attempting to identify any siblings. Unfortunately, we were unable to validate the neighbourhood attack algorithm on the Max Healthcare dataset because the dataset contains no pairs of closely related patients. As a result, the attack (correctly) yields no related records for each of the de-identified records. We discuss the neighbourhood attack experiments and their results here to demonstrate how the attack can be conducted and how the results establish the absence of any related individuals.

To attempt each attack, we conducted the experiments summarized in Table 4:

For each experiment, we isolated the features from the identified record r^* which represent a relative's history, and compared them to the corresponding features in each candidate record $r' \in D$.

Note that the dataset of personal health records provided by Max Healthcare

Figure	Attack	Features	Target Features
9	Parents	r^* Father's history	r''s history
9	Parents	r^* Mother's history	r''s history
10	Parents	r^* Grandfather's history	r' Father's history
10	Parents	r^* Grandmother's history	r' Mother's history
11	Siblings	r^* Siblings' history	r''s history
12	Siblings	r^* Father's history	r' Father's history
12	Siblings	r^* Mother's history	r' Mother's history

Table 4. Neighbourhood attack experiments

contained no parent-child or sibling pairs. As a result, the desired outcome is for our attack algorithm to return no matching pairs, with the deviating features for each possibly related pair of records.

3.1.1 Identifying Parents

With the foreknowledge that our dataset contained no parent-child pairs, we conducted the attacks outlined in the Methods section, with the goal of establishing that the attack would currently return no valid parent records which could be de-anonymized using the de-anonymized child record.

We found that, individually, each comparison (as listed in Table 4) yielded a small number of matches (which were necessarily false positives). Since each attack contained multiple feature-set comparisons, combining them together correctly yielded no matches in our dataset.

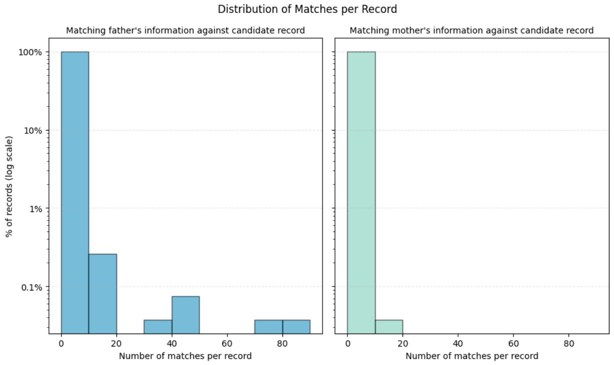

Fig 9. The number of matches for the identified record against candidate parents' records using parent history

As illustrated in Fig.9, the attack correctly returned no positive matches using the parents' medical history. 88.2% of records returned no matches for possible father records, and 96% of records returned no matches for possible mother records. No pair of records matched on the features for both parents, confirming that there were no parent-child pairs in our dataset.

Similarly, as illustrated in Fig. 10, the attack correctly returned no positive matches using the grandparents' medical history. 99.2% of records returned no matches for possible father records, and 99.5% of records returned no matches for possible mother records. No pair of records matched on the grandparents' features for both parents.

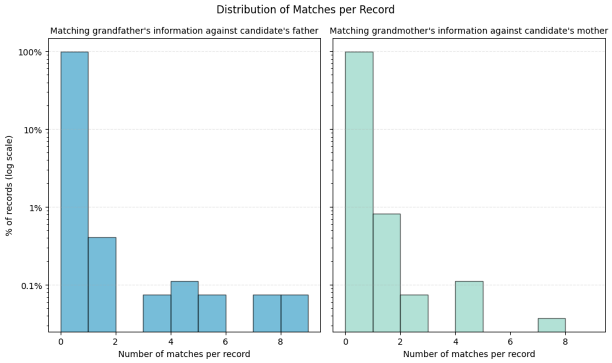

Fig 10. The number of matches for the identified record against candidate parents' records using grandparent history

3.1.2 Identifying Siblings

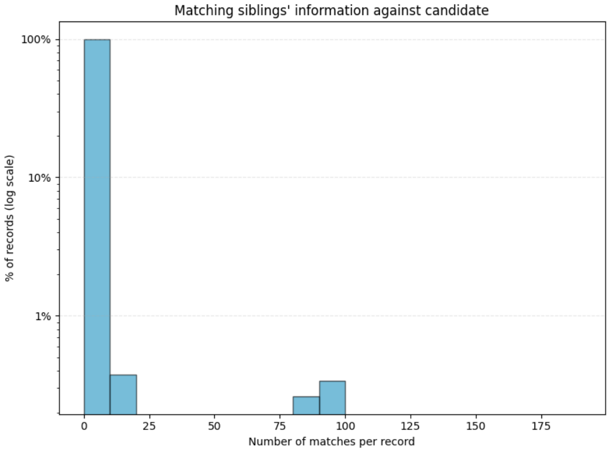

Fig 11. The number of matches for the identified record against candidate siblings' records using siblings' history

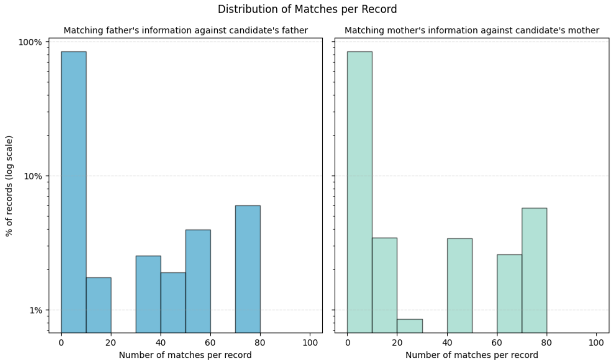

Fig 12. The number of matches for the identified record against candidate siblings' records using mutual parents' history

As illustrated in Fig. 11, 79.6% of records returned no matches for possible sibling records using the siblings' medical history. As illustrated in Fig. 12, 52.2% of records returned no matches for possible sibling records using the father's history, and 52.04% using the mother's history. No pair of records matched when considering the siblings', father's and mother's history together, leading to the correct conclusion that there are no sibling pairs in the dataset.

These results indicate that by leveraging familial relationships encoded in EHRs, an adversary can de-anonymize not only the target patient but also their relatives, creating a cascading effect that extends the breach across entire family networks. This attack is particularly concerning in the Indian context because families often share medical services and personnel, and because familial medical history is often too critical a component of patient records to remove.

4 Methods

Define a database D consisting of n records and m features. We are interested in Electronic Health Records in the context of personal health data, i.e. information about individual patients and treatments. We thus take each $i \in \{1, 2, \ldots, n\}$ to represent the index of one patient and each $j \in \{1, 2, \ldots, m\}$ to represent the index of one medical feature. For simplicity, we will primarily refer to a "record" r as one patient's row of data consisting of $(r_0, \ldots, r_{m-1})$, the adversarial attack's victim.

Each observation r_j can be of many types, but we can simplistically represent each observation as one of four types:

- A date, representing personal information or the date of a medical procedure
- A numeric value, representing the quantification of some personal or medical information
- A boolean value
- A string observation containing rich information about the patient

Even more simplistically, we can reduce the majority of features in a medical database to a simple binary. For example, the string *gender* feature can be reduced to is_female holding $True$ for a female patient or $False$ otherwise. Similarly, a numeric medical information can be reduced to *is_abnormal* representing whether the value is outside of the normal range or not. This helps us model a weaker form of attack where an adversary might not directly know information in the database but has background knowledge that informs general assumptions.

To compare two records, we need a similarity function that compares each feature of the two records and returns a similarity score between 1 (identical) and 0 (completely distinct).

Following the convention of Narayanan and Shmatikov, we refer to the set of non-null features of a record as the *support* of the record $supp(r)$. The support of two records is $supp(r) \cup supp(r')$. Similarly, the support of a column c, denoted $supp(c)$, is the set of non-null features in a particular feature column

We then specify our similarity function as:

$$Similar(r, r') = \frac{\sum_j Sim(r_j, r'_j)}{|supp(r) \cup supp(r')|}$$

where $Sim(r_j, r'_j)$ is a comparison function that checks the distance of the values of r_j and r'_j in the case of numeric values, or checks if they are within some tolerance of each other otherwise. We therefore call two records "0.5 similar" if half their features satisfy the similarity condition of $Sim(r_j, r'_j)$, and so on. $\sum Sim(r_j, r'_j)$ is therefore the total similarity of two records on each of their common features. We outline the variations of the similarity measure that we experiment with in Section 4.4. More detailed definitions of the functions $aux(r)$, $supp(r)$, $Sim(r_j, r'_j)$ and $Similar(r, r')$ are provided in Appendix 1.

4.1 Dataset Sparsity

A pre-requisite for our auxiliary information deanonymization attack is that the dataset must be *sparse*. A *sparse* dataset is one where:

- A majority of records have non-null values for only a small proportion of features
- A majority of features have non-null values for only a small proportion of records

More formally, we can say that a dataset is sparse if the probability of two unique records having the same non-null values is low (below some threshold δ) [1]. Given a suitable small value ϵ, we can then call a dataset sparse if

$$Prob[Similar(r, r') > \epsilon \quad \text{for all} \quad r' \neq r] < \delta$$

We know that personal EHR data is sparse because the vast majority of patients only suffer from, and are tested for, specific ailments. As a result, any database of electronic health records of patients will necessarily be sparse. We show this for our dataset in Section 4.1 and Fig. 1 and 2.

4.2 De-anonymization algorithm

To demonstrate the lack of privacy in the dataset, we demonstrate a background-information attack on a prototypical EHR modelled as our database D.

We model an adversary that obtains a small amount of background information $aux(r)$ about a person representing record r in D. This $aux(r)$ is a subset of the information about the person available in record r, with some error tolerance and/or limitations. We outline the different variations of $aux(r)$ that we will experiment with in Section 4.3. We can quantify the certainty requirement as ϵ, the minimum similarity score between the auxiliary information and the record above which we consider it re-identified.

A de-anonymization attack is thus successful if, given some background information about a patient, an adversarial algorithm A given the database and $aux(r)$, can return the patient's record r' with a high enough similarity measure ϵ. Formally:

$$A[D, aux(r)] \to r' \in D \text{ such that } Similar(aux(r), r') > \epsilon$$

We can define the simplest adversarial algorithm A as:

- Calculate the similarity $Scores$ of r using $Similar(r, r') = \sum Sim(r_j, r'_j) \;\; \forall \; r_j \in aux(r)$ for every $r' \in D$. That is, we compare our information on our victim to the each patient in the database and calculate how much they match.
- If $max(Scores) \geq \epsilon$ and the record with the maximum score is unique, then return the record with the maximum score. That is, if we find a record that matches our information closely enough, and no other record matches it as closely, we have found our victim.

- Otherwise, return the l records in $max(Scores)$ (we consider this a failure to identify the record). If we aren't able to pick out the matching patient from the database, we haven't been able to identify them.

This algorithm simply matches the auxiliary information against every record in D and calculates the similarity score. If the record r' with the highest similarity is unique, it returns r'. If the best match is not unique, it returns all the records which have the highest similarity score. It may be noted, however, that it may still be considered a successful attack if k is small.

4.3 Auxiliary information selection $aux(r)$

For each record r that we attempt to de-anonymize, we have to select the background information $aux(r)$ which will be available to the adversary. We implement three variations in the way this auxiliary information is selected.

4.3.1 k most informative features

To establish a lower-bound on the number of features required to de-anonymize a record, we first grant the adversary the k most informative features that are present for that record. This represents the worst-case scenario where the adversary happens to have the information which would be most useful in de-anonymizing the victim.

For this purpose, We measure informativeness by the rarity of the feature. Specifically, we select the k features in $supp(r) \cup supp(r')$ with the lowest $supp(c)$ (lowest support in a column), indicating that those features least often have non-null values for any patient record.

4.3.2 Randomly selected features

We attempt to de-anonymize each record using k features selected randomly from the non-null features of the target record r. In cases where we must select k features but $supp(r) < k$, we calculate the similarity score on the size of $supp(r)$ features instead.

To reduce the variance of results, we take the average of the similarity scores from ten instances of the random selection of features for each record. Note that this has no impact for records where $supp(r) < k$ or $supp(r) \approx k$.

4.3.3 Adversary-accessible features

However, neither of the above two approaches truly represent any realistic scenario. Some features, by virtue of being less common, are much more informative for de-anonymization than others; consequently, random selection can be misleading. It is not very likely that an adversary would have access to specific information such as the patient's sodium serum or squamous epithelial cell blood test values. However, there is a pool of more "public" features that we can guess an adversary would be more likely to observe or gain access to: height, allergies, clubbing, major medical history, etc.

Malicious colleague/acquaintance: In our day-to-day activities, we interact with a large number of acquaintances of uncertain intent. Any of these individuals would possess some knowledge of our personal information and lifestyle (travel habits, allergies, etc.) or could acquire it if determined. For example, a malicious colleague would know the answers to:

- Do you have any skin allergies?

- How old were you when you started consuming alcohol regularly?

- How often do you Travel or go on Vacation?
- Have you travelled domestically in last one Year?

Note that this is not just information that a colleague would know; it is information that many people, including public-facing individuals, publish on the internet for anyone to see. Instagram and Twitter accounts reveal detailed social activity including data points that reveal or are highly correlated to relevant medical information. Many professions, which can be easily identified through legitimate means, are similarly highly correlated with medically relevant information. Many of these datapoints can also be inferred from one other because of strong correlation, such as between the consumption of alcohol and nicotine. The bar for obtaining this information, as a result, is quite low.

Malicious medical insider: EHRs are designed to be used and shared by medical institutions and professionals to make treatment simpler. Insider attacks or improprieties in the handling of this data can cause severe vulnerabilities. The minimum information that such a vulnerability would expose would be the patient's current medical state and the medical history required to diagnose them. This exposure includes features like:

- Diabetes Mellitus (Y/N)
- Year of Start
- Year of End
- Duration
- Surgery Name (Y/N)
- Year of Surgery
- Place of Surgery

Once again, note that this doesn't just include the example case of a regular general practitioner or doctor who attended to the patient personally. While that is certainly one vector to obtain the information, it is also possible for this information to be provided to the adversary through an EHR system with doubtful purpose-limitation.

Smart medical app: There are several health-related apps and services which collect medical data from consumers who voluntarily offer it to the operators for utility benefits. The subset of features collected by smartwatches and such services is not nearly enough (since they would only have access to vitals and lifestyle/travel habits). However, there are several smart home blood test services and similar services which collect medical information, sometimes with vulnerable security and obscure terms and conditions. This includes:

- Blood pressure
- Temperature
- Absolute monocyte count
- Drug name (Y/N)
- Strength
- Frequency
- Duration

By attempting de-anonymization on the database using auxiliary information drawn from different subsets of features representing realistic adversaries, we can show that the database is vulnerable to real-world malicious actors under more realistic assumptions.

4.4 Similarity scoring function $Sim(r_j, r'_j)$

To simulate different ranges of inaccuracy in the adversary's knowledge, we can add error tolerances to our comparison function, i.e., reduce the precision with which the adversary knows a feature.

4.4.1 Perfect information:

The most simple model of an attack is an adversary that directly obtains k features of auxiliary information about a particular record r from the database D. However, especially in the medical context, only a limited range of adversaries would have perfect information from the database. Given that features include numeric values like blood test results, it's more likely that an adversary such as an insider or an eavesdropper would have less specific information about certain features.

4.4.2 20% error tolerance:

In the event of the adversary "guessing" or extrapolating features such as medical measurements, or deriving them from a different source, there would be errors in their information. To investigate this scenario, we allow an error tolerance of 20% in all numeric features. If $r_j \in aux(r)$ is within 20% of $r'_j \in r'$, $Sim()$ returns a 1, otherwise 0.

Note that there is significant literature on adding privacy-preserving noise to a database as a means of ensuring that statistical attacks cannot be conducted on it [17] [26]. As Narayanan and Shmatikov [1] established, there is no formal difference between an error rate in the adversary's information and statistical noise added to D for security. As a result, this scenario models both the cases where either the adversary has made slightly incorrect "guesses" about the information or the EHR has statistical noise added to it to protect against adversarial attack. Both represent cases where the adversary's information, when compared to the information stored in the database, will be mismatched by a small amount, and it equivalent whether the mismatch exists on the part of the adversary (error) or on the part of the database (noise).

4.4.3 Presence of non-null features:

A result of the severe sparsity of the database is that the mere presence of a non-null feature for a record r is revealing, regardless of its value. Thus, we now ignore the values altogether and only check whether or not the feature is non-null. If both r_j and r'_j contain a non-null value (for eg. $is\_abnormal$ is $True$ for both), $Sim()$ returns a 1, otherwise 0. This allows us to model a scenario with much lower requirements on the attacker, who need only know general information such as whether a patient has any allergy at all, or has ever been tested for a particular disease.

To implement this comparison, we select every feature m in the Max Healthcare database where $count(mode(m)) > \delta\% * count(m)$. For $\delta = 50$, this is true for 89% of the 629 features in the database (558 features). For $\delta = 66$, this is true for 83% of the database (521 features). For each of these, we replace the modal value with $False$, and any other value with $True$. For the "presence-only" versions of the attack, we use only this modified subset of the database, without using the remaining (less sparse) features.

This approach has the consequence of showing that our attack is robust to protection schemes such as bucketization [27] which prevent raw numeric values from being released. It further shows that the presence-only version of the attack can be conducted even when the adversary does not have permissions to access the database. If the adversary only has access to the database interface, they can make use of side channel leaks of information to potentially reconstruct presence-based information about the database [28] [29] [30].

4.5 Neighbourhood attack algorithm

There is a significant body of literature on using social network graph structures to extract information from a personal database [31] [32] [33]. These graph neighbourhood attacks make use of an extra source of background information in addition to the adversary's background information: their knowledge of the relationships between different people in the database [34] [35]. Since social media data inherently contains information about who you follow or interact with, information about the activity of a victim's connections could be used to get information on the victim.

Consider that medical history data can also be framed as a social network graph containing information on relationships between patients. A personal medical record contains features on the person's parents, siblings, and often grandparents. That same parent's information will also be included in the parent's medical record. Consequently, gaining access to a person's record also gives the adversary access to part of the parent's record.

We can therefore outline a novel attack made possible by the nature of personal EHRs: the neighbourhood re-identification attack. Literature on privacy in social network data has found that "neighbourhood attacks" can reveal information about a victim even if the victim themselves is not present in the database. Even more easily, neighbourhood attacks can be used to take a single de-anonymized record and extract additional information from the database without any incremental auxiliary information [31]. Since many families in India consult the same doctor or visit the same hospital when needed, this represents a significant attack vector at the household level. While there is literature on preventing these attacks, much of it focuses on changing the disclosure of data or obscuring relationships, which is not practical for medical datasets [36].

This attack is possible because personal EHRs such as the Max Healthcare dataset require significant amounts of information on a person's family medical history to help identify hereditary and environmental disorders. Our dataset includes 231 features relating to relatives' medical history, as illustrated in Fig. 13. In a dataset that is sufficiently sparse, it is likely that two records with corresponding medical histories of at least two relatives, will be from the same family tree.

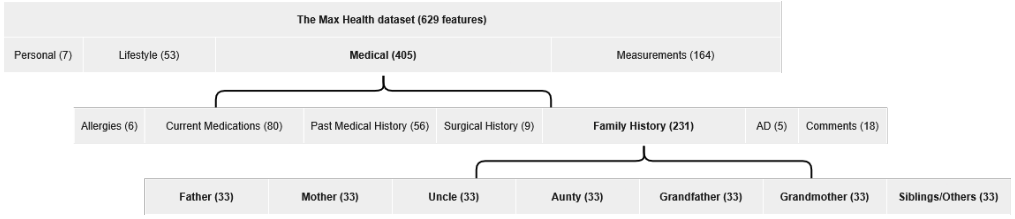

Fig 13. The hierarchy of family-related features in the Max Healthcare EHR

We model the attack using the same database D as before. We add the condition that each $r \in D$ must contain some n features which contain information about the individual's relatives. We denote these features as r_{mi} through $r_{(m+n)i}$, where m denotes the index of the feature in r, and i is the index of the relationship to the individual represented by r. We take an adversary that uses background information $aux(r)$ to successfully identify the record $r' = r \in D$ using the algorithm A from the previous section.

We denote this de-anonymized record as $r*$, and the records of the relatives of patient $r*$ as $r*_i$. We know that there are some features $r*_{mi}$ through $r*_{(m+n)i}$ which hold medical information on the relative i of the patient represented by $r*$.

By treating this information as the auxiliary information $aux(r*_i)$ available to the

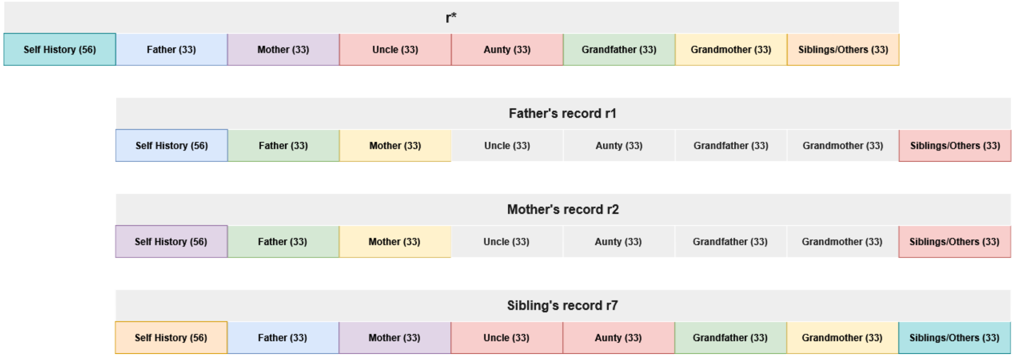

Fig 14. A family of records in D. Corresponding colours indicate fields that can be used as $aux(r*_i)$ once $r*$ is identified.

adversary to de-identify the relative, we can construct an adversarial algorithm similar to A which would allow the adversary to de-identify every relative of the patient whose information is present in $r*$.

This neighbourhood attack is thus successful if

$$A'[D, r*] \rightarrow r*_i \in D \text{ such that } Similar(r*_{mi},\ r*_i) > \epsilon$$

We can define the neighbourhood attack's adversarial algorithm A' as:

- Use algorithm A with $aux(r)$ to successfully de-anonymize a record $r*$
- For any i where $r*_{mi}$ is non-null, assign $aux(r*_i) = r*_{mi}$
- Use this auxiliary information with A and D to find the record $r*_i$ representing the relative of $r*$. If no such record exists in D, return that the relatives have not been recorded in the database.
- Recursively repeat steps 2 and 3 for all $r*_{imj}$ in $r*_i$, representing the relatives of the relatives.
- Repeat steps 2, 3 and 4 for all values of i in $r*$

Though this attack makes it possible to set off a chain reaction which de-anonymizes every member of their immediate and extended family who have a medical history, it has two limitations:

- In cases where the medical history of the relative is not present in the original de-anonymized record $r*$, either because of the relative has no medical history or because of incomplete self-reporting, it is not possible to de-anonymize the relative. It is still possible, however, to accomplish the lesser de-anonymization of revealing whether the relative's record is in the database D or not.
- The attack can only be propagated upwards. Due to the nature of heredity, in our dataset and in the majority of EHRs described in literature, medical information on the patients' descendants is never included. As a result, while the attack can be used to de-anonymize the patient's family tree, it cannot be used iteratively to de-anonymize the entire dataset.

However, these limitations do not take away from the primary advantage of this attack; it severely reduces the background information necessary to de-anonymize multiple people in a dataset by making use of networks within the records.

5 Conclusion

This paper demonstrates a viable avenue of attack to uniquely identify individuals in datasets which claim to be securely anonymized. It creates a formal method of specification to show that this is a quality inherent to the nature of personal data, especially in the medical domain. It demonstrates that this attack is effective in a real-world scenario for real data from a leading medical provider. It further demonstrates that the attack is robust to a variety of assumptions about the quality and quantity of information available to the de-anonymizing adversary. It further specifies methods to propagate the de-anonymization to more data belonging to the individual and their relatives, demonstrating that privacy harms exist even for patients with no publicly available information.

While we successfully show that there are privacy risks with the existing approach to medical anonymization, and suggest mitigation techniques from a systems perspective, there are no algorithmic methods of securing privacy on large, sparse, personal datasets. This has implications on the vulnerability of EHRs in the Indian case as well as globally, particularly in a rapidly innovative AI-driven world. In the Indian case, the assumptions of privacy made by the new, digital medical paradigm led by the Ayushman Bharat Digital Mission are extremely vulnerable to attacks similar to ours. In the global case, the increasing propagation of interoperable digital records and their use to train AI services creates new vectors for such attacks.

Further investigation is necessary to establish the true risks to privacy when EHRs are used in tightly linked and cross-referenced systems such as ABDM. We hope that this paper highlights the need for a new approach to the storage of sensitive personal data at scale.

6 Acknowledgements

The authors would like to thank the Centre for Digitalization, AI & Society and the Trivedi School of Biosciences at Ashoka University for facilitating the research process. We are sincerely grateful to Dr. Abhaya Indrayan for constructive comments and the team at Max Healthcare for providing the prototypical EHR.

References

1. Narayanan A, Shmatikov V. Robust De-anonymization of Large Sparse Datasets. In: Proceedings of the 2008 IEEE Symposium on Security and Privacy. IEEE; 2008. p. 111–125.

2. Price II WN, Cohen IG. Privacy in the age of big medical data. Nature Medicine. 2019;25:44–56.

3. Wadhwa M. Towards a New Indian Model of Information and Communications Technology-Led Growth and Development. CSD Working Paper Series; March 2020. Available from: `https://csd.columbia.edu/sites/default/files/content/docs/ICT%20India/Papers/ICT_India_Working_Paper_25.pdf`.

4. Srivastava SK. Adoption of Electronic Health Records: A Roadmap for India. Healthcare Informatics Research. 2016;22(4):261–269. doi:10.4258/hir.2016.22.4.261.

5. Honavar SG. Electronic Medical Records – The Good, the Bad and the Ugly. Indian Journal of Ophthalmology. 2020;68(3):417–418. doi:10.4103/ijo.IJO_278_20.

6. Venkat A. Max Healthcare improves patient safety with e-health record system; 2017. How-To. Available from: `https://www.cio.com/article/218500/max-healthcare-improves-patient-safety-with-e-health-record-system.html`.

7. Dwork C, Roth A. The Algorithmic Foundations of Differential Privacy. Found Trends Theor Comput Sci. 2014;9(3–4):211–407. doi:10.1561/0400000042.

8. de Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: The privacy bounds of human mobility. Scientific Reports. 2013;3:1376.

9. de Montjoye YA, Radaelli L, Singh VK, Pentland A. Unique in the shopping mall: On the reidentifiability of credit card metadata. Science. 2015;347(6221):536–539.

10. Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences. 2013;110(15):5802–5805.

11. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PloS one. 2011;6(12):e28071. doi:10.1371/journal.pone.0028071.

12. El Emam K, Yakovlev A, Neisa A, Jonker E. Evaluating the privacy risks of de-identified data. Journal of the American Medical Informatics Association. 2008;15(5):627–637. doi:10.1197/jamia.M2716.

13. Gariépy-Saper K, Decarie N. Privacy of Electronic Health Records: A Review of the Literature. Journal of the Canadian Health Libraries Association. 2021;42(1):74–84. doi:10.29173/jchla29496.

14. Dalenius T. Towards a methodology for statistical disclosure control. Statistik Tidskrift. 1977;15(429-444):2–1.

15. Ghosh A, Kleinberg R. Inferential Privacy Guarantees for Differentially Private Mechanisms. CoRR. 2016;abs/1603.01508.

16. Dwork C. Differential Privacy. In: Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II. ICALP'06. Berlin, Heidelberg: Springer-Verlag; 2006. p. 1–12. Available from: `http://dx.doi.org/10.1007/11787006_1`.

17. Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis. In: Proceedings of the Third Conference on Theory of Cryptography. TCC'06. Berlin, Heidelberg: Springer-Verlag; 2006. p. 265–284. Available from: `https://doi.org/10.1007/11681878_14`.

18. Dong J, Roth M, Su WJ. Gaussian differential privacy. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2022;84(1):3–37.

19. Machanavajjhala A, Kifer D, Gehrke J, Venkitasubramaniam M. L-diversity: Privacy beyond k-anonymity. ACM Transactions on Knowledge Discovery from Data (TKDD). 2007;1(1):3–es.

20. Xu L, Jiang C, Wang J, Yuan J, Ren Y. Information security in big data: Privacy and data mining. IEEE Access. 2014;2:1149–1176.

21. Sweeney L. Weaving Technology and Policy Together to Maintain Confidentiality. Journal of Law, Medicine and Ethics. 1997;25(2-3):98–110.

22. Barbaro M, Zeller T. A Face Is Exposed for AOL Searcher No. 4417749; 2006. Available from: `https://www.nytimes.com/2006/08/09/technology/09aol.html`.

23. McSherry F, Mironov I. Differentially private recommender systems: building privacy into the Netflix prize contenders. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM; 2009. p. 627–636.

24. Dwork C, Kohli N, Mulligan DK. Differential privacy in practice: Expose your epsilons! Journal of Privacy and Confidentiality. 2019;9(2):1–30. doi:10.29012/jpc.689.

25. Gadotti A, Rocher L, Houssiau F, Creţu AM, de Montjoye YA. Anonymization: The imperfect science of using data while preserving privacy. Science Advances. 2024;10(29):eadn7053. doi:10.1126/sciadv.adn7053.

26. Kargupta H, Datta S, Wang Q, Sivakumar K. On the privacy preserving properties of random data perturbation techniques. In: Third IEEE International Conference on Data Mining; 2003. p. 99–106.

27. Jayapradha J, Prakash M, Alotaibi Y, Khalaf OI, Alghamdi SA. Heap Bucketization Anonymity—An Efficient Privacy-Preserving Data Publishing Model for Multiple Sensitive Attributes. IEEE Access. 2022;10:28773–28791. doi:10.1109/ACCESS.2022.3158312.

28. Shahverdi A, Shirinov M, Dachman-Soled D. Database Reconstruction from Noisy Volumes: A Cache Side-Channel Attack on SQLite. In: 30th USENIX Security Symposium. USENIX Association; 2021. p. 1019–1035. Available from: `https://www.usenix.org/conference/usenixsecurity21/presentation/shahverdi`.

29. Wang L, Grubbs P, Lu J, Bindschaedler V, Cash D, Ristenpart T. Side-Channel Attacks on Shared Search Indexes. In: IEEE Symposium on Security and Privacy (SP); 2017. p. 673–692.

30. Chen S, Wang R, Wang X, Zhang K. Side-Channel Leaks in Web Applications: A Reality Today, a Challenge Tomorrow. In: IEEE Symposium on Security and Privacy; 2010. p. 191–206.

31. Backstrom L, Dwork C, Kleinberg J. Wherefore Art Thou R3579X? Anonymized Social Networks, Hidden Patterns, and Structural Steganography. In: WWW; 2007. p. 181–190.

32. Hongyan Zhang LLXW Li Xu. De-anonymizing Social Networks with Edge-Neighborhood Graph Attacks. In: Security and Privacy in Digital Economy. Springer; 2020. p. 726–737.

33. Fang J, Li A, Jiang Q, Li S, Han W. A Structure-Based De-Anonymization Attack on Graph Data Using Weighted Neighbor Match. In: IEEE Fourth International Conference on Data Science in Cyberspace; 2019. p. 480–486.

34. Qian J, Li XY, Zhang C, Chen L, Jung T, Han J. Social Network De-Anonymization and Privacy Inference with Knowledge Graph Model. IEEE Transactions on Dependable and Secure Computing. 2019;16(4):679–692. doi:10.1109/TDSC.2017.2697854.

35. Korolova A, Motwani R, Nabar SU, Xu Y. Link privacy in social networks. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management. CIKM '08. New York, NY, USA: Association for Computing Machinery; 2008. p. 289–298. Available from: https://doi.org/10.1145/1458082.1458123.

36. Zhou B, Pei J. Preserving Privacy in Social Networks Against Neighborhood Attacks. In: IEEE 24th International Conference on Data Engineering; 2008. p. 506–515.

37. Centers for Disease Control and Prevention. What Is Breast Cancer Screening?; 2023. Available from: https://www.cdc.gov/breast-cancer/screening/index.html.

38. American Cancer Society. American Cancer Society Recommendations for the Early Detection of Breast Cancer; 2023. Available from: https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast html.