iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84236

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Privacy Preservation System for Multiple Attributes

Bonumaddi Kumari[1], Chikkala Soujanya[2]
[1]Associate Professor, [2]MCA Final Semester, Master of Computer Applications, Sanketika Vidya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, India

Abstract: Privacy preservation and data utility are two important key factors for publishing the data of multiple sensitive attributes. Traditional anonymization methods often fail to protect against re-identification risks in complex datasets. The proposed system uses semantic L-diversity technique to partition the data into the different buckets. Slicing technique is used to partition the data into multiple sensitive tables along with the quasi table. This approach also uses a bucket id to keep the associations among the sensitive tables and quasi table. Additionally, the framework provides a mechanism to generate and securely re-identify summary tables derived from sensitive datasets, ensuring a balance between accessibility and confidentiality. The system proposes a framework that ensures data utility while minimizing privacy risks. The system aims to provide a solution for organizations to publish valuable data responsibly without compromising individual privacy.

I. INTRODUCTION

Privacy-preserving data publishing (PPDP) is the process of sharing data while ensuring that sensitive or personal information remains protected. Many organizations, such as hospitals, government agencies, and businesses, need to share data for research, policy-making, or analysis. However, publishing raw data can expose individuals' private details, leading to privacy risks. To prevent this, PPDP uses various techniques to anonymize or modify the data while keeping it useful. PPDP is crucial because privacy laws, such as GDPR and HIPAA, require organizations to protect personal information. Without proper privacy protection, data breaches can lead to identity theft, discrimination, or misuse of information. Some popular techniques in PPDP include k-anonymity, which ensures that each individual's data is indistinguishable from at least k-1 others, l-diversity, which ensures diversity in sensitive attributes, and differential privacy, which adds statistical noise to data to prevent re-identification.

PPDP is widely used in healthcare, finance, social sciences, and cybersecurity. For example, hospitals may share patient data for medical research while ensuring that no individual's health records are exposed. The goal is to balance privacy with data utility, allowing organizations to gain insights from data while protecting individuals' confidentiality.

II. EXISTING SYSTEM

K-anonymity groups records to ensure each is indistinguishable from at least K-1 others based on quasi-identifiers. However, it may not fully protect sensitive information and can over-generalize data, reducing its usefulness. While it enhances privacy, it struggles with complex data and safeguarding sensitive attributes effectively. L-Diversity enhances privacy by ensuring each group of records contains at least L different sensitive values. This prevents attackers from deducing personal information by maintaining variety in sensitive attributes. While L-Diversity simply ensures distinct sensitive.

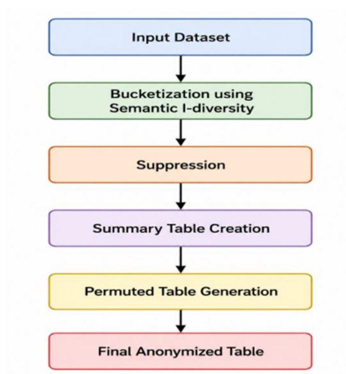


Figure : Flow Chat

A. Challenges

The major challenges that are faced in the existing system are

1) Insufficient Protection for Multiple Sensitive Attributes: Traditional privacy-preserving techniques such as K-Anonymity and basic L-Diversity are primarily designed to protect datasets containing a single sensitive attribute. When multiple sensitive attributes exist in the same dataset, these methods fail to provide comprehensive privacy protection. This increases the possibility of exposing confidential information and limits their applicability in real-world datasets such as healthcare and financial records.
2) Vulnerability to Inference and Background Knowledge Attacks: Existing systems are vulnerable to inference attacks, where attackers combine anonymized data with publicly available or background information to identify individuals. Even if direct identifiers are removed, quasi-identifiers can still reveal sensitive information when correlated with external datasets. As a result, the privacy of individuals cannot be fully guaranteed using traditional approaches.
3) High Information Loss and Reduced Data Utility: Most existing anonymization techniques rely heavily on generalization and suppression to protect privacy. While these methods improve confidentiality, they also remove important details from the dataset. This reduces the accuracy and usefulness of the published data for research, data analysis, and decision-making. Achieving a balance between privacy and utility remains a significant challenge.
4) Semantic Similarity Between Sensitive Values is Ignored: Traditional L-Diversity ensures that different sensitive values exist within a group but does not consider their semantic similarity. For example, diseases such as "Flu" and "Viral Fever" are counted as different values even though they convey nearly the same information. This allows attackers to make accurate guesses about an individual's sensitive information, weakening the overall privacy protection.
5) Scalability and Performance Issues with Large Datasets: As the size of the dataset increases, traditional privacy-preserving techniques require more computational time and memory. Processing large volumes of records using repeated generalization and anonymization operations becomes inefficient and difficult to scale. This makes existing systems less suitable for handling modern big data applications where fast processing and efficient privacy preservation are essential.

III. PROPOSED SYSTEM

The proposed system improves data privacy by using a combination of Semantic L-Diversity, Slicing, and Bucketization. Semantic L-Diversity ensures that each group of records contains different and meaningful sensitive values, making it difficult for attackers to identify an individual's private information. Slicing divides the dataset into separate columns so that sensitive and non-sensitive information are stored independently, reducing the chances of linking personal details while still keeping the data useful for analysis. Bucketization further enhances privacy by grouping similar records into buckets and randomly shuffling sensitive values within each bucket, making it difficult to trace information back to a specific person. Together, these techniques provide stronger privacy protection while maintaining the usefulness of the data for research and analysis.

IV. ALGORITHM

The proposed system uses privacy-preserving techniques to protect multiple sensitive attributes while maintaining data utility. It processes the input dataset by identifying quasi-identifiers and sensitive attributes before applying Semantic L-Diversity, Bucketization, Slicing, and Suppression. These techniques work together to reduce the risk of privacy breaches while ensuring that the published data remains useful for analysis. The overall algorithm focuses on achieving a balance between data privacy and data utility.

A. Input Data Processing

The system begins by loading the input dataset and identifying quasi-identifiers and multiple sensitive attributes. It analyzes the dataset to determine the number of records, unique sensitive values, and the required bucket size. This preprocessing step prepares the data for further privacy-preserving operations.

B. Bucket Formation

The system divides the dataset into multiple buckets based on the calculated bucket size. Each bucket is created in such a way that it contains a diverse set of sensitive values. This grouping helps prevent attackers from identifying individuals while maintaining meaningful data for analysis.

C. Semantic L-Diversity
Semantic L-Diversity ensures that every bucket contains multiple distinct and semantically different sensitive values. Instead of considering only the number of different values, it also considers their meaning. This makes it much harder for attackers to infer sensitive information from the published data.

D. Suppression Process
The suppression process checks whether the occurrence of any sensitive value exceeds its predefined threshold within a bucket. If the threshold is exceeded, additional occurrences are replaced with placeholder values to reduce disclosure risks. This improves privacy while minimizing unnecessary information loss.

E. Correlated Table Generation
The system creates separate correlated tables for quasi-identifiers and sensitive attributes. Each sensitive record is assigned a unique identifier, allowing relationships to be maintained without directly exposing confidential information. This approach strengthens privacy while preserving important data associations.

F. Summary Table and Permutation
Finally, the system combines the processed information into a summary table using the assigned identifiers. The records within each bucket are randomly permuted to prevent direct linkage between individuals and their sensitive information. This final step improves resistance against inference attacks while preserving the overall usefulness of the dataset.

V. TECHNIQUES
The proposed system uses several privacy-preserving techniques to secure sensitive information while maintaining the usefulness of the published data. These techniques work together to prevent unauthorized disclosure of personal information, reduce the risk of inference attacks, and improve data quality for research and analysis. By combining multiple anonymization methods, the system provides stronger privacy protection compared to traditional approaches.

A. Semantic L-Diversity
Semantic L-Diversity enhances privacy by ensuring that each bucket contains multiple distinct and semantically different sensitive values. Unlike traditional L-Diversity, it considers both the diversity and the meaning of sensitive attributes, making inference attacks much more difficult.

B. Bucketization
Bucketization divides the dataset into multiple groups based on quasi-identifiers. Sensitive values within each bucket are separated and randomly arranged, reducing the possibility of linking individuals with their confidential information while preserving data utility.

C. Slicing
Slicing partitions the dataset into multiple columns by separating quasi-identifiers from sensitive attributes. This reduces the relationship between identifying and confidential information while maintaining useful attribute correlations for analysis.

D. Suppression
Suppression protects highly sensitive information by replacing selected values that exceed predefined threshold limits with placeholder values. This minimizes the disclosure of sensitive information while maintaining an acceptable balance between privacy and data utility.

VI. METHODOLOGY
The methodology of the proposed system follows a systematic process to anonymize datasets containing multiple sensitive attributes. Initially, the input dataset is collected and analyzed to identify quasi-identifiers and sensitive attributes. The system then calculates the bucket size and determines the appropriate L-value required for Semantic L-Diversity.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

After preprocessing, the dataset is divided into buckets while ensuring that each bucket satisfies the Semantic L-Diversity requirement. Threshold values are applied to control the frequency of sensitive attributes, and suppression is performed whenever these limits are exceeded. The system then creates correlated tables for quasi-identifiers and sensitive attributes, followed by the generation of a summary table using unique identifiers.

Finally, the records within each bucket are randomly permuted to prevent direct association between sensitive information and individuals. The anonymized dataset is then evaluated using privacy and utility metrics to ensure that the proposed approach provides strong privacy protection while preserving the usefulness of the published data.

A. Input

1	age	education	occupatioi	relationsh	gender	hours	state	Salary	Zip Code	Disease	Disease Ca	Disease Sensitivity
2	25	11th	Machine-c	Own-child	Male	40	United-Sta	50115	40469	Tooth Sen	Oral & Dei	0
3	38	HS-grad	Farming-fi	Husband	Male	50	United-Sta	59837	42068	Pneumoni	Respirator	1
4	28	Assoc-acd	Protective	Husband	Male	40	United-Sta	90892	93510	Diverticuli	Digestive [	0
104	22	Some-coll	Transport-	Own-child	Male	39	United-Sta	59969	77489	Lung Canc	Cancer	1
105	58	9th	Other-serv	Not-in-fan	Female	40	United-Sta	59823	49526	Acne	Skin Cond	0
106	52	HS-grad	Adm-cleri	Unmarriec	Female	28	United-Sta	54053	65076	Bipolar Di:	Mental He	1
107	36	10th	Machine-c	Not-in-fan	Female	40	United-Sta	51620	65999	Multiple S	Neurologi	1
108	41	HS-grad	Protective	Husband	Male	40	United-Sta	46495	96869	Conjunctiv	Common I	0
109	28	HS-grad	Transport-	Own-child	Male	40	United-Sta	52837	50282	Multiple S	Neurologi	1
1783	46	HS-grad	Craft-repa	Husband	Male	40	United-Sta	57417	52550	Parkinson	Neurologi	1
1784	41	HS-grad	Transport-	Husband	Male	60	United-Sta	40062	97113	Rhinitis	Mild Respi	0
1785	48	Bachelors	Exec-mana	Not-in-fan	Female	35	United-Sta	50860	97099	Liver Failu	Liver Disea	1
1786	71	9th	Protective	Husband	Male	39	Cuba	58674	53067	Common (	Common I	0
31997	38	Assoc-acd	Other-serv	Husband	Male	40	El-Salvado	58568	17658	Breast Car	Cancer	1
31998	38	HS-grad	Craft-repa	Husband	Male	40	United-Sta	57466	90937	Anxiety	Mild Ment	0
31999	46	Some-coll	Sales	Unmarriec	Female	40	United-Sta	59339	49222	Acid Reflu	Digestive [	0
32000	37	HS-grad	Machine-c	Unmarriec	Female	20	United-Sta	40743	38355	Bipolar Di:	Mental He	1
32001	62	HS-grad	Exec-mana	Husband	Male	40	United-Sta	94318	97380	TB	Respirator	1

Figure 1: Orginal Dataset

1	age	education	occupatioi	relationsh	gender	hours	state	Salary	Zip Code	Disease	Disease Ca	Disease Se	bucket
2	32	Some-coll	Adm-cleri	Not-in-fan	Male	38	United-Sta	50935	95880	Myopia	Eye Condi	0	1
3	19	HS-grad	Exec-mana	Not-in-fan	Female	35	United-Sta	53340	68163	Breast Car	Cancer	1	1
4	43	HS-grad	Machine-c	Not-in-fan	Female	40	United-Sta	44443	31160	Tooth Sen	Oral & Dei	0	1
104	33	Bachelors	Exec-mana	Husband	Male	40	United-Sta	113380	55624	Tooth Sen	Oral & Dei	0	1
105	21	HS-grad	Adm-cleri	Wife	Female	43	United-Sta	98499	89800	Gastritis	Digestive [	0	1
106	55	Assoc-voc	Other-serv	Wife	Female	42	United-Sta	58844	25269	Bipolar Di:	Mental He	1	1
107	47	HS-grad	Tech-supp	Unmarriec	Female	55	United-Sta	42235	26252	Cirrhosis	Liver Disea	1	1
108	39	Assoc-acd	Prof-speci	Not-in-fan	Male	59	United-Sta	41247	27287	Mild Hype	Mild Cardi	0	1
109	51	HS-grad	Machine-c	Not-in-fan	Male	50	United-Sta	57370	15398	Type 1 Dia	Autoimmu	1	1
1783	57	7th-8th	Craft-repa	Husband	Male	40	United-Sta	110409	39538	Bronchitis	Respirator	1	6
1784	57	Masters	Exec-mana	Husband	Male	45	United-Sta	109399	85068	Athlete's F	Skin Cond	0	6
1785	44	Doctorate	Exec-mana	Wife	Female	32	United-Sta	91759	86655	Osteoarth	Joint & Bo	0	6
1786	30	Some-coll	Sales	Not-in-fan	Female	45	United-Sta	55156	19472	Mild Hype	Mild Cardi	0	6
31997	47	Bachelors	Exec-mana	Not-in-fan	Male	40	United-Sta	81632	73920	Bronchitis	Respirator	1	101
31998	53	Some-coll	Exec-mana	Wife	Female	30	United-Sta	47319	59520	Gastritis	Digestive [	0	101
31999	46	Assoc-voc	Adm-cleri	Unmarriec	Female	40	United-Sta	45656	48046	High Chole	Mild Cardi	0	101
32000	59	Masters	Prof-speci	Husband	Male	7	United-Sta	109067	41650	Lung Canc	Cancer	1	101
32001	35	Some-coll	Exec-mana	Husband	Male	58	United-Sta	52983	31785	Tooth Sen	Oral & Dei	0	101

Figure 2 : Bucketized using Semantic L-diversity

1	age	education	occupatioi	relationsh	gender	hours	state	salary	zip code	disease	disease ca	disease se	bucket
2	32	Some-coll	Adm-cleri	Not-in-fan	Male	38	United-Sta	50935	95880	Myopia	Eye Condi	0	1
3	19	HS-grad	Exec-mana	Not-in-fan	Female	35	United-Sta	53340	68163	Breast Car	Cancer	1	1
4	43	HS-grad	Machine-c	Not-in-fan	Female	40	United-Sta	44443	31160	Tooth Sen	Oral & Dei	0	1
104	33	Bachelors	xxxx	xxxx	Male	40	United-Sta	113380	55624	xxxx	Oral & Dei	0	1
105	21	HS-grad	xxxx	Wife	Female	43	United-Sta	98499	89800	Gastritis	Digestive [	0	1
106	55	Assoc-voc	xxxx	Wife	Female	42	United-Sta	58844	25269	xxxx	Mental He	1	1
107	47	xxxx	Tech-supp	Unmarriec	Female	55	United-Sta	42235	26252	Cirrhosis	Liver Disea	1	1
108	39	Assoc-acd	xxxx	xxxx	Male	59	United-Sta	41247	27287	Mild Hype	Mild Cardi	0	1
109	51	xxxx	xxxx	xxxx	Male	50	United-Sta	57370	15398	xxxx	Autoimmu	1	1
17983	40	Bachelors	xxxx	xxxx	Female	60	United-Sta	95475	27480	xxxx	Autoimmu	1	57
17984	39	Assoc-voc	xxxx	xxxx	Male	38	United-Sta	82225	12062	xxxx	Kidney Dis	1	57
17985	36	xxxx	xxxx	xxxx	Female	40	United-Sta	55693	32347	xxxx	Liver Disea	1	57
17986	34	xxxx	xxxx	xxxx	Male	40	United-Sta	53411	12593	xxxx	Skin Cond	0	57
31997	47	xxxx	xxxx	xxxx	Male	40	United-Sta	81632	73920	xxxx	Respirator	1	101
31998	53	xxxx	xxxx	Wife	Female	30	United-Sta	47319	59520	xxxx	Digestive [	0	101
31999	46	Assoc-voc	xxxx	xxxx	Female	40	United-Sta	45656	48046	xxxx	Mild Cardi	0	101
32000	59	Masters	xxxx	xxxx	Male	7	United-Sta	109067	41650	Lung Canc	Cancer	1	101
32001	35	xxxx	xxxx	xxxx	Male	58	United-Sta	52983	31785	xxxx	Oral & Dei	0	101

Figure 3 :Suppressed Sensitive Attributes

1	Hrs/salary	Age	Gender	Zip code/State	Bucket
2	38,50,935	32	Male	95880,United-States	1
3	35,53,340	19	Female	68163,United-States	1
4	40,44,443	43	Female	31160,United-States	1
5	20,41,027	19	Male	66616,United-States	1
6	38,53,192	43	Female	68833,United-States	1
7	40,49,582	34	Male	76904,United-States	1
23561	40,81,514	55	Female	57662,Canada	75
23562	40,41,159	27	Male	23470,United-States	75
23563	45,52,467	36	Male	25255,United-States	75
23564	40,56,870	44	Male	62255,United-States	75
23565	50,49,926	39	Male	73344,United-States	75
23566	40,42,074	54	Female	22319,United-States	75
31995	40,40,775	64	Male	23013,United-States	101
31996	45,46,938	77	Male	58564,United-States	101
31997	40,81,632	47	Male	73920,United-States	101
31998	30,47,319	53	Female	59520,United-States	101
31999	40,45,656	46	Female	48046,United-States	101
32000	71,09,067	59	Male	41650,United-States	101
32001	58,52,983	35	Male	31785,United-States	101

Figure 4 : Correlated Quasi Attributes

1	sid	SAs
2	1	xxxx(124), Some-college(30), Bachelors(30), HS-grad(30), Assoc-acdm(14), 11th(14), 10th(13), Masters(13), Prof-school(12), Assoc-voc(12), 7th-8th(6), 9th(5), 1st-4th(5), 12th(4), Doctorate(3), 5th-6th(2)
3	2	xxxx(251), Exec-managerial(5), Adm-clerical(5), Other-service(5), Craft-repair(5), Prof-specialty(5), Machine-op-inspct(5), Handlers-cleaners(5), Farming-fishing(5), Transport-moving(5), Sales(5), Tech-support(5), Protective-serv(5), ?(5), Priv-house-serv(1)
4	3	xxxx(135), HS-grad(30), Bachelors(30), Some-college(30), Masters(17), Assoc-acdm(14), Assoc-voc(13), 11th(12), 10th(10), 7th-8th(6), 5th-6th(5), 12th(4), Doctorate(4), Prof-school(3), 9th(3), 1st-4th(1)
5	4	xxxx(249), Other-service(5), Sales(5), Craft-repair(5), Machine-op-inspct(5), Adm-clerical(5), Prof-specialty(5), Protective-serv(5), Farming-fishing(5), Tech-support(5), Transport-moving(5), ?(5), Handlers-cleaners(5), Exec-managerial(5), Priv-house-serv(2), Armed-Forces(1)
6	5	xxxx(130), Some-college(30), Bachelors(30), HS-grad(30), Masters(16), Assoc-acdm(14), 10th(13), 11th(11), Assoc-voc(9), 7th-8th(9), 12th(8), Doctorate(6), 9th(5), Prof-school(4), 5th-6th(1), 1st-4th(1)
7	6	xxxx(248), ?(5), Exec-managerial(5), Farming-fishing(5), Other-service(5), Sales(5), Craft-repair(5), Machine-op-inspct(5), Transport-moving(5), Adm-clerical(5), Handlers-cleaners(5), Protective-serv(5), Prof-specialty(5), Tech-support(5), Priv-house-serv(4)
200	199	xxxx(125), Bachelors(30), Some-college(30), HS-grad(30), Masters(19), 11th(17), Assoc-voc(14), 5th-6th(8), 12th(8), 9th(7), 10th(7), Prof-school(6), 7th-8th(5), Doctorate(5), Assoc-acdm(5), 1st-4th(1)
201	200	xxxx(251), Exec-managerial(5), Machine-op-inspct(5), ?(5), Other-service(5), Prof-specialty(5), Adm-clerical(5), Handlers-cleaners(5), Protective-serv(5), Transport-moving(5), Craft-repair(5), Tech-support(5), Sales(5), Farming-fishing(5), Priv-house-serv(1)
202	201	xxxx(130), Bachelors(30), HS-grad(30), Some-college(30), 11th(21), Masters(15), Assoc-acdm(10), 10th(7), Assoc-voc(7), Prof-school(4), 9th(4), 12th(3), 1st-4th(3), 7th-8th(2), Preschool(2), Doctorate(1), 5th-6th(1)
203	202	xxxx(230), Tech-support(5), Prof-specialty(5), Exec-managerial(5), Adm-clerical(5), Transport-moving(5), Sales(5), Craft-repair(5), Other-service(5), ?(5), Handlers-cleaners(5), Machine-op-inspct(5), Protective-serv(5), Farming-fishing(5), Priv-house-serv(4), Armed-Forces(1)

Figure 5 : Correlated Sensitive Attributes (Education & Qualification)

1	sid	SAs
2	101	xxxx(194), Breast Cancer(2), Myopia(2), Tendonitis(2), Sickle Cell Disease(2), Cushingâ€™s Syndrome(2), Tooth Sensitivity(2), Pancreatic Cancer(2), Ulcerative Colitis(2), Diverticulitis(2), Liver Failure(2), Osteoarthritis(2), Common Cold(2), HIV/AIDS(2), Sinusiti
3	102	xxxx(174), Own-child(27), Not-in-family(27), Unmarried(27), Husband(27), Wife(19), Other-relative(16)
4	103	xxxx(196), Conjunctivitis(2), Sickle Cell Disease(2), Alzheimerâ€™s(2), Parkinsonâ€™s(2), Cystic Fibrosis(2), Eczema(2), Cirrhosis(2), Mild Asthma(2), Hemophilia(2), Pneumonia(2), Bipolar Disorder(2), High Cholesterol(2), Tooth Sensitivity(2), Psoriasis(2), Myc
5	104	xxxx(182), Husband(27), Own-child(27), Not-in-family(27), Unmarried(27), Wife(16), Other-relative(11)
6	105	xxxx(193), Chronic Kidney Disease(2), Kidney Failure(2), Migraines(2), Schizophrenia(2), Athleteâ€™s Foot(2), Osteoarthritis(2), Irritable Bowel Syndrome(2), Insomnia(2), Bronchitis(2), Liver Failure(2), Tendonitis(2), Parkinsonâ€™s(2), Eczema(2), HIV/AIDS(2)
7	106	xxxx(189), Own-child(27), Not-in-family(27), Husband(27), Unmarried(27), Wife(15), Other-relative(5)
200	299	xxxx(194), Eczema(2), CataractsÂ (EarlyÂ Stage)(2), Cirrhosis(2), Hemophilia(2), Irritable Bowel Syndrome(2), Leukemia(2), Cushingâ€™s Syndrome(2), Myopia(2), Rhinitis(2), Bronchitis(2), Chronic Kidney Disease(2), Breast Cancer(2), Ulcerative Colitis(2), HIV/
201	300	xxxx(182), Own-child(27), Husband(27), Not-in-family(27), Unmarried(22), Other-relative(17), Wife(15)
202	301	xxxx(181), Down Syndrome(2), Tooth Sensitivity(2), Cystic Fibrosis(2), Pancreatic Cancer(2), Mild Depression(2), Crohnâ€™s Disease(2), Psoriasis(2), Kidney Failure(2), Stroke(2), Athleteâ€™s Foot(2), Alzheimerâ€™s(2), TB(2), Obesity(2), Acid Reflux(2), Cirrhos
203	302	xxxx(172), Not-in-family(27), Husband(27), Own-child(27), Unmarried(27), Wife(15), Other-relative(5)

Figure 6: Correlated Sensitive Attributes(Disease & Relationship)

1	hours_salary	age	gender	zipcode_state
2	3,85,09,35,001	32,002	Male,101	95880,United-States,102
3	3,55,33,40,001	19,002	Female,10	68163,United-States,102
4	4,04,44,43,001	43,002	Female,10	31160,United-States,102
5	2,04,10,27,001	19,002	Male,101	66616,United-States,102
6	3.85.31.92.001	43.002	Female.10	68833.United-States.102
17017	2,04,78,04,107	19,108	Male,207	69567,United-States,208
17018	4,04,72,03,107	24,108	Male,207	34882,United-States,208
17019	5,54,46,80,107	50,108	Male,207	93333,United-States,208
17020	4,05,57,86,107	29,108	Male,207	60235,United-States,208
17021	2,45,82,28,107	44,108	Male,207	39242,Iran,208
31997	4,08,16,32,201	47,202	Male,301	73920,United-States,302
31998	3,04,73,19,201	53,202	Female,30	59520,United-States,302
31999	4,04,56,56,201	46,202	Female,30	48046,United-States,302
32000	7,10,90,67,201	59,202	Male,301	41650,United-States,302
32001	5,85,29,83,201	35,202	Male,301	31785,United-States,302

Figure 7: Summary table

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

1	hours_salary	age	gender	zipcode_state
2	5,24,82,81,001	45,002	Male,101	26898,United-States,102
3	4,04,82,84,001	39,002	Male,101	59976,United-States,102
4	4,08,10,39,001	50,002	Male,101	45658,United-States,102
5	3,05,89,99,001	20,002	Male,101	51854,Mexico,102
6	2,04,14,22,001	22,002	Male,101	63228,England,102
17017	4,04,31,47,107	27,108	Male,207	11393,United-States,208
17018	3,44,82,01,107	33,108	Male,207	28601,Germany,208
17019	5,55,80,99,107	35,108	Male,207	81623,United-States,208
17020	4,09,40,24,107	28,108	Male,207	29500,United-States,208
17021	3,04,72,48,107	21,108	Male,207	64072,United-States,208
31997	4,55,91,07,201	29,202	Male,301	95503,United-States,302
31998	3,25,06,35,201	30,202	Female,30	32737,United-States,302
31999	4,05,16,66,201	47,202	Male,301	10601,United-States,302
32000	4,05,11,27,201	57,202	Male,301	47066,United-States,302
32001	4,05,67,24,201	33,202	Female,30	73595,United-States,302

Figure 8: Permuted Summary Table

VII. CONCLUSION

Privacy preservation using **bucketization**, **semantic L-diversity**, and **slicing** helps protect sensitive data while keeping it useful. Since bucketization is based on **disease category**, applying semantic L-diversity ensures a good mix of diseases in each bucket, making it harder to identify individuals. However, when different **thresholds** are used, the level of suppression in disease categories changes significantly, while other sensitive attributes remain less affected. This shows that diseases need stricter privacy controls. At the same time, **slicing** keeps quasi-identifiers separate from sensitive data, improving privacy. Together, these methods ensure that data can be shared safely while still being useful for analysis.

REFERENCES

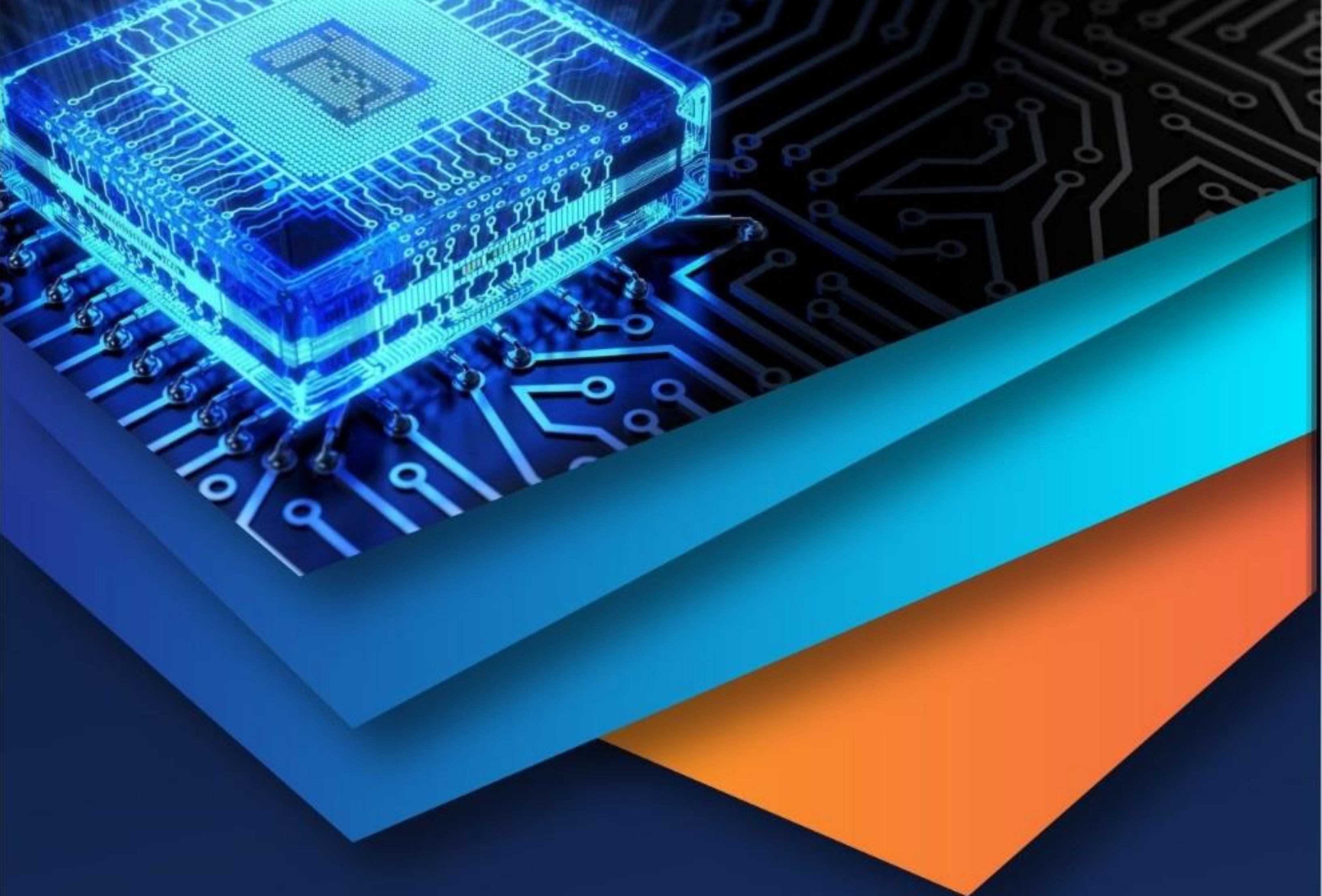
[1] Xiao, Yuelei, and Haiqi Li. "Privacy preserving data publishing for multiple sensitive attributes based on security level." Information 11.3 (2020): 166.
[2] Wang et al. "A Privacy-Preserving Method for Publishing Data with Multiple Sensitive Attributes." 2022.
[3] Liu et al. "Privacy Preserving Data Publishing for Heterogeneous Multiple Sensitive Attributes." 2023.
[4] Zhang et al. "Privacy Preserving Data Publishing with Multiple Sensitive Attributes Using Overlapped Slicing for Microdata Anonymization." 2022.
[5] Li et al. "Privacy Preserving Data Publishing with Multiple Sensitive Attributes Using a Differential Privacy-Based Mechanism." 2020.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

BIBLOGRAPHY

Bonumaddi Kumari working as an Assistant Professor in computer Science and Engineering at Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh. With 3 years of experience in Computer Science and Engineering (B.tech), accredited by NAAC, she has a membership in IAENG with her area of interest in Java Full Stack, DBMS, Data Structures, and C Language.

Chikkala Soujanya is currently pursuing his final semester of Master of Computer Applications (MCA) at Sanketika Vidya Parishad Engineering College, which is accredited with an 'A' grade by NAAC, affiliated to Andhra University, and approved by AICTE. With a keen interest in Privacy preservation and data publishing(PPDP), has undertaken her postgraduate project titled "Privacy Preservation System for Multiple Attributes." The project is carried out under the guidance of Bonumaddi Kumari, Assistant Professor , SVPE



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429