2023 26th International Conference on Computer and Information Technology (ICCIT)13-15 December, Cox's Bazar, Bangladesh

Secure Land Purchasing using Different Multi-Party Skyline Queries with Anonymous Information

Dola Das, Md. Jahid Hasan, Sk. Nahid Hasan, Arifur Rahman and Kazi Md. Rokibul Alam
Department of Computer Science and Engineering
Khulna University of Engineering & Technology, Bangladesh
Email: dola.das@cse.kuet.ac.bd, jahid.cse.kuet@gmail.com, hasan1707080@stud.kuet.ac.bd, rarifkhan652@gmail.com, rokib@cse.kuet.ac.bd

***Abstract*—A large number of available information causes data overload for users to choose the optimal options. In case of online land purchasing, it is challenging for a user to choose an optimal option from a vast number of lands with different areas and prices. Again, datasets of the seller (*i. e.* land owner) are sensitive since it is a business issue. Therefore, ensuring the privacy of the information, anonymity of the seller and their information etc. are crucial. This paper proposes a framework to find out the dominant (*i. e.* optimal) datasets using several skyline query algorithms namely block nested loop (BNL), bitmap, index, and nearest neighbour (NN). Besides, to ensure data privacy, separate decryption keys are allocated among multiple authorities, to ensure data anonymity, mixed network-based ElGamal cryptosystem with re-encryption and shuffling techniques are used, to ensure anonymity of the seller, an anonymous identity after registration phase is assigned. Lastly, The selection processes, privacy, result analyses etc. shows the efficiency of the framework and makes easier for land buyers to choose the optimal options. This paper is a preliminary report of the proposed system.**

Index Terms*—*Multi-party skyline query, ElGamal cryptosystem, Mixed network, Re-encryption, Data privacy, Anonymity.

I. Introduction

Online land purchasing is growing in popularity as it provides buyers with a convenient way to view a range of lands from distances [18]. They have full access to details about the property, such as its address, size, price, and any added features or amenities. But it may be challenging to select the best lands from among thousands with authentic information. In order to overcome this difficulty, the skyline algorithms can be helpful. A skyline query can be used to help online land buyers find lands with several preferable aspects. The skyline query retrieves lands that are the best in at least one criterion and not worse than others in any other criterion [1]. This helps buyers identify lands that meet their preferences across multiple dimensions simultaneously. For example, a larger land may come at a higher price, or a land closer to the city center may be small in size. In this work, by using skyline queries, buyers can explore the trade-offs between different criteria and find lands that strike the right balance for their needs. Additionally, to secure personal information, avoid fraud, maintain confidentiality, encourage fairness, and foster confidence between buyers and online platforms, data security and anonymity of both data and land owner are essential in online land purchasing.

Recently, a variety of approaches have been used in this field. For the purpose of protecting land information, some researchers put various security algorithms, such as blockchain and elliptic curve cryptography [17]. Again, some researchers merely use the skyline technique and choose the most dominant point from large datasets, but they don't guarantee the security of the data. Data security was implemented in [18] but there was no registration process discussed clearly. However, due to these reasons, the existing frameworks in this field are not efficient enough.

In this paper, a new framework is proposed where land selection includes data privacy, data and owner's anonymity using ElGamal cryptosystem [2], re-encryption and Fisher Yate shuffling, unique anonymous identity etc. Thus, this framework proposes some new features such as:

1) Making a user-friendly interface that allows land owners to register themselves with relevant information for selling.

2) Providing data security by dividing decryption keys among 2 servers.

3) Confirming data anonymity by ElGamal re-encryption and Fisher Yate shuffling of encrypted datasets by 2 servers.

4) Ensuring the data owners' anonymity with an unique id.

5) Based on the request from buyers, displaying the results on the website with comparisons of the skyline query results.

Following sections provide more details about secure land purchasing framework. Section II explains some of the contemporary works in this domain. The necessary tools, the overall architecture of the proposed framework, and description of the models used in this work is addressed in section III. Section IV speculates the framework. Finally, section V concludes the overall processes and future plans for this work.

II. Literature Review

Online land purchasing has become popular in recent years. However, the privacy of database and the land owners must be provided since any coercers may modify the known owner's data intentionally. Several researchers have recently developed a number of encryption algorithms to provide data privacy for land purchasing schemes. But none of them did not consider data privacy, data anonymity, owner anonymity, result revelation, etc. altogether. Also, it is also difficult to select the best lands from the many which were not considered by almost anyone in the past.

979-8-3503-5901-5/23/$31.00

The Skyline operator is introduced in [1], with the goal of extending relational database systems with skyline queries to select the best from the many. The Skyline result is determined by the shortest path distance, which varies depending on the algorithm [4] introduced aggregate skyline queries. The nearest neighbor algorithm is proposed by [23]. Again, in [10], they created a block nested loop (BNL) algorithm that is similar to a naive nested loop. They made a comparison of all tuples in memory and it determines whether the tuple is chosen or not. It is unsuitable for large databases because of its multiple scanning so index techniques are developed in [15]. In [11], they introduced a bitwise representation technique called bitmap, which is completely non-blocking and has a lower initial response time than other methods. In the paper [12], a new skyline algorithm is introduced and used on $2D$ data. The algorithm in [13] is the first skyline algorithm to use nearest neighbor (NN) techniques based on the concept of R-trees [14]. Besides many algorithms like top-k dominating [21], [22] exist to handle incomplete data set.

Again, many security protocols for data privacy have recently been developed in this field. Blockchain application is used for property buying and selling in [16]. They used Angular to build the front end and Solidity Contract for backend support. A method for securing land registers using blockchain and the SHA256 hash function is demonstrated in paper [17]. But using brute force to recover the original data from blockchain technology is difficult and time-consuming in their works. In [7], the problem of secure spatio-textual skyline query processing is defined in the cloud environment, and two secure spatio-textual skyline query is proposed. But the authors of [7] and [8] did not consider the registration process. The developers have proposed several skyline algorithms. However, the majority of them do not address owner authentication, privacy and anonymity issues together concerning multi-party skyline queries [3].

In the proposed work, the registration process is demonstrated, where users can register with their names and a random id is assigned to each user. Also, the ElGamal cryptosystem is used to provide data privacy and anonymity. Finally, the optimal results with desired aspects are formed and evaluated.

III. Methodology

This section describes the overview of the proposed framework and individual stages of this work. Fig. 1 depicts the overall workflow diagram of the proposed work.

A. Overview of the Proposed Franework

The system is developed with some entities namely 2 mix servers *COH*1 (Cryptographic Operation Handler 1) and *COH*2 (Cryptographic Operation Handler 2), an assistant *ICSP* (Information collector and service provider), $(N > 1)$ land owners LO_1, ..., LO_n $(n \in (1, ..., N))$ and buyers (*i. e* users) etc. Each LO_n receives an anonymous unique identity ID_n from *ICSP* after completing online registration. Each LO_n can have more than 1 land to sell *i. e* DA_{n1}, ..., DA_{nj} dataset where l_n is the maximum number of LO_n's dataset, every data has

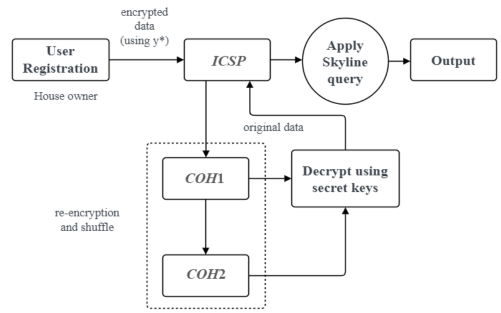

Fig. 1: Overview of the proposed framework

same D dimensions, the uniform data type (*i. e* numeric) and $(u, v) \in (1, ..., D)$. COH_1 and COH_2 have their own key pair (x_1, y_1) and (x_2, y_2), respectively and a combined encryption key y^* is generated from (y_1, y_2). y^* is open to all and each LO_n needs to encrypt $DA_n = DA_{n1}$, ..., DA_{nj} where $j \approx l_n$ and send them to *ICSP*. Now, *ICSP* gives all the datasets DA_1, ..., DA_n of all LO_1, ..., LO_n to COH_1 and COH_2 for re-encryption, shuffling and decryption of the datasets. Finally, *ICSP* computes the skyline queries based on the decrypted results and lists the best lands $w.r.t.$ the ID_n. Here, the desired criteria to develop the farmework are as follows:

1) Owner's anonymity: In order to conceal the link between the owner and it's dataset, ensuring owner's anonymity is must.

2) Data privacy: To eliminate the unwanted modification of the datasets, privacy mechanism implementations is needed.

3) Data anonymity: In order to remove links between the submitted dataset and the outputs, anonymization of datasets are important.

4) Auditable deeds: To ensure the authenticity of the results, showing the results publicly is very important.

Besides, there are some assumptions to implement the proposed framework as follows:

- No entity can know LO_n's dataset in plain and encrypted forms except LO_n itself.
- No owner can know the number of dominant datasets plus their owners about the dominated datasets.

B. Individual Stages

This section comprises of user registration, data encryption and submission, re-encryption and shuffling, skyline queries etc. They are described as follows:

1) User Resistration: Registration of LO_n's will be conducted under *ICSP* and it inherits the mechanism of anonymous credential from [20] for owner's anonymity. The process is as follows:

- Each LO_n shows it's legimitacy by name, email, address etc. personal information privately to *ICSP*.
- If *ICSP* is convienced, it provides a random unique anonymous credential (along with required attributes) ID_n to LO_n. Later on, LO_n can appear anonymously to everyone.

2) Data Encryption and Submission: This stage comprises of encryption of datasets by the owners with the ElGamal cryptosystem [5], submission of the encrypted datasets, approval of the posted datasets on the website. To do these, the interactions between LO_ns and *ICSP* are as follows:

- The combined encryption key $y*$ is calculated as $y* = y_1.y_2 \pmod p = g^{x1+x2} \pmod p$ where p is a large prime number and g is a generator of the multiplicative group Z_p* of the integers modulo p. Later on, mod(p) will be omitted.
- Now, each LO_n encrypts its DA_1, ..., DA_n as $\underline{DA}$ using $y*$ based on the commutative cryptosystem [2] and submits them to *ICSP* with their anonymous identity. Here, For each DA_{nj}, the encrypted result will be as $c_1 = g^{x1+x2}$ and $c_2 = DA_{nj}.y*^{x1+x2}$. Thus, other datasets are also encrypted.
- When LO_n finds its dataset on the website, it puts an unique confirmation code on the specific place of the website.

3) Re-encryption and Shuffling: In the proposed framework, re-encryption and shuffling of the encrypted data is used to anonymize the datasets [6]. The procedures between *COH*1, *COH*2 and *ICSP* are as follows:

- The encrypted datasets $\underline{DA}$ are collected by *ICSP*.
- Now, *ICSP* sends the $\underline{DA}$ to *COH*1 and *COH*2 to re-encrypt and shuffle the encrypted datasets sequentially. For each encrypted DA_{nj} *i. e.* c_1 and c_2, the re-encrypted form is $c_1' = g^{x1+x2}.g^{x1+x2}$ and $c_2' = DA_{nj}.y*^{x1+x2}.y*^{x1+x2}$. Thus, other datasets are also re-encrypted.
- After re-encrypting each DA_{nj}, *COH*1 and *COH*2 shuflle the re-encrypted results sequentially using Fisher Yate shuffling algorithm [19].
- Now, using the secret key x_1 and x_2, *COH*1 and *COH*2 decrypt the re-encrypted and shuffled datasets for skyline queries. Here, for a single data, the decryption process will be $DA_{nj} = \frac{C_2'}{C_1'^{(x1+x2)}} = \frac{DA_{nj}.y_*^{(x1+x2)}.y_*^{(x1+x2)}}{(g^{x1+x2})^{(x1+x2)}.(g^{x1+x2})^{(x1+x2)}}$. Thus, all the re-encrypted datasets are decrypted to form the original datasets.
- These decrypted datasets are posted on the website with unique identity for all.

4) Skyline Queries: While getting back the original data, different skyline algorithms are applied to get the desired results. Generally, skyline query returns the dominant datasets $i.\ e.$ the winners against all other dataset. For example, for two different data DA_{nj} and DA_{oq} of LO_n and LO_o, respectively, if $(1 <= (u,v) <= D)$, $(n,o) \in (1,...,N)$ and $(j,q) \approx (l_n,l_o)$, DA_{nj} is said to be dominant $i.\ e.$ (DA_{nj} $< DA_{oq}$) if ($DA_{nj}[u] < DA_{oq}[u]$) for at least one u and ($DA_{nj}[u] <= DA_{oq}[u]$) for all u. Again, DA_{nj} and DA_{oq} are both said to be non-dominated or, winners if ($DA_{nj}[u]$ $< DA_{oq}[u]$) and ($DA_{nj}[v] > DA_{oq}[v]$) [2]. In the proposed framework, to gain the dominant datasets, 4 skyline algorithms namely block nested loop (BNL), bitmap, nearest neighbor (NN) and index are used and discussed below. Finaly, a comparison is stated among them to verify the correctness of the results.

TABLE I: Computation Time for Various Stages

Individual Stage	Processing Time
Registration	10.2 ms/owner
Data submission	145 ms/owner
Re-encryption and decryption	18.9 sec
BNL/Bitmap/NN/Index	0.09/0.78/0.12/0.15 sec

1) Block Nested Loop (BNL): It compares data points with other ones by dividing large datasets into blocks [9].
2) Bitmap: The bitmap [9] precomputes a bitmap index by representing each data point as a binary vector. If (Q) is the set of all data points and (p) is a point in the skyline then Skyline(Q) $= p$ in Q for all q in Q, there does not exist q' in Q such that q' dominates p and $q'\ ! = q$
3) Nearest Neighbor (NN): The NN algorithm [9] finds the nearest point that dominates a given query point. If q is a query point, this algorithm finds the point p in DS that minimizes the Euclidean distance to q with the condition p dominates q $i.\ e.$ p is a skyline point $w.\ r.\ t.\ q$.
4) Index: The indexing algorithm [9] quickly locates skyline points in a dataset by minimizing the amount of required pairwise comparisons.

After calculating the skyline results, a comparison is stated as in section IV-*B*.

IV. Model Evaluation

For the execution of the prototype of proposed framework, python 3.8.5 was used under 1.6-1.8 GHz, 8 GB RAM, and 64-bit operating system. A synthetic dataset of 12 LO_n each having up to 25 dataset was used for the prototype. It was computed on a single computer. Here, 4 skyline algorithms were applied and the output of each skyline result as well as their comparisons are displayed. Initially, for the registration system, Xampp local server with version 7.2.28 was used where the control panel version was 3.2.4.

A. Experimental Results

Here, if there are 12 LO_n each having up to 25 data, the time computation for each stage is as Table I. Here, in the registration stage, for providing anonymous identity, $ICSP$ needs 10.2 ms/LO_n. After that, for encrypting dataset of LO_n with $y*$ and submit them to $ICSP$, each LO_n needs 145 ms. To anonymize and decrypt the re-encrypted data, $COH1$, $COH2$ and $ICSP$ need 18.9 seconds. Finaly, for BNL, bitmap, NN and index algorithm, it takes 0.09 second, 0.78 second, 0.12 second and 0.15 second, respectively for 300 datasets. When a land buyer will want to know the best lands with some specific aspects, the output procedure will start from the skyline calculation which makes the framework more practical and less time consuming.

Now, the target is to maximize the area and minimize the price. For simplification, a small dataset of 1 lands of 12 LO_ns with location, area and price is considered and shown in Table II and the skyline results from Table II are as below:

TABLE II: Sample input data of land owners

City	Area(sq feet)	Price(tk)
Dhaka	209	735887
khulna	262	912408
Rajshahi	264	1003829
Barishal	279	1065733
Satkhira	244	1282450
Dhaka	234	1247070
Bogura	207	1032238
Sylhet	245	1235467
Mymensing	139	441778
Khulna	295	964161
Barishal	183	680018
Rajshahi	129	439993

Among the 12 data stated above, the encryption, re-encryption and shuffling and getting back the original data by decryption for 6 data is shown in Fig. 2.

Index	Id	Area	Price
0	542	209	735887
1	653	262	912408
2	354	264	1003829
3	111	279	1065733
4	985	244	1282450
5	254	234	1247070

a) Input Data

Id	Area	Price
329536	127072	447419296
397024	159296	554744064
215232	160512	610328032
67488	169632	647965664
598880	148352	779729600
154432	142272	758218560

a) User Encrypted Data

Id	Area	Price
42615936	31781376	120844950336
13362624	33587136	128297201472
30577536	28169856	150127274880
118578240	29373696	154386460880
78610752	31540688	109839324672
65248128	25160256	88589020608

a) Re-encryption and shuffling by COH1

Id	Area	Price
102581760	257840640	984907809280
327152640	243978240	927698608640
603476480	242129920	843210977280
910297600	225495040	1185188992000
500894720	193149440	680077329920
234736640	216253440	1152492211200

d) Re-encryption and shuffling by COH2

Id	Area	Price
111	279	1065733
354	264	1003829
653	262	912408
985	244	1282450
542	209	735887
254	234	1247070

e) Decrypted Original Data

Fig. 2: The processings of datasets

1) BNL: The output from BNL algorithm for the dataset in Table II are described in Table III. The algorithm works like the following steps:

1) (209,735887) is not dominated by any point so it entered the skyline list.
2) Similarly (262,912408), (264,1003829), (279, 1065733), (244,1282), (234, 1247070), (245,1235467), (139,441778) are not dominated so they are inserted. (207, 1032238) is dominated by (209,735887) so it is discarded.
3) (295,964161) dominates the points (264,1003829), (279, 1065733), (244,1282), (234, 1247070), (245,1235467). So all are discarded from the list.
4) (183,680018) and (129,439993) is not dominated nor dominate any point so it is inserted into the list. So the skyline points are (209,735887), (262,912408), (295,964161), (139,441778), (183,680018), and (129,439993)

TABLE III: Output data for BNL and Bitmap algorithms

City	Area(sq feet)	Price(tk)
Dhaka	209	735887
Khulna	262	912408
Mymensing	139	441778
Khulna	295	964161
Barishal	183	680018
Rajshahi	129	439993

2) Bitmap: Here, the area is represented in the X coordinate and the price is in the Y coordinate. All the data and its corresponding bitmap representation are shown in Table IV. Here co-ordinate to (209,735887) is (8, 4) means if we sort all 12 areas in descending order then 209 is the 8th maximum area and if we sort all 12 prices in ascending order then 735887 is the 4th minimum price. The bit representation of 8 is (12-8)+1 = 5 bits from the left is 1 and the remaining is 0. Similarly for 4 is (12-4)+1 = 9 bits from the left is 1 and the remaining is 0. Then the coefficient of x dimension, Cx and y dimension, Cy have to be found out.

Cx = 011100000100 (4th bit from the right of all data points in x co-ordinate marked in green color) and Cy = 110000001011(5th bit from the right of all data points in y co-ordinate marked in green color). After Cx $\wedge$ Cy, it becomes 010000000000. Here, the decisions about skyline are:

- If it has only one 1 then it is included in the skyline point.
- If more than one 1 is in the operation then it will not be in the skyline point.

Therefore, (209,735887) is a skyline point.

Let's consider another point (245, 1235467). Its coordinate point is (5,10). Cx = 011100010100 and Cy = 011100010100 and Cx & Cy = 011100010100. Here, no of 1 is more than one so it is not included in the skyline point.

TABLE IV: Bitmap representation of input data

(area,price)	Co-ordinate (x, y)	Bitmap representation (x, y)
(209, 735887)	(8,4)	(111110000000, 111111111000)
(262, 912408)	(4,5)	(111111111000, 111111110000)
(264, 1003829)	(3,7)	(111111111100, 111111000000)
(279, 1065733)	(2,9)	(111111111110, 111100000000)
(244, 1282450)	(6,12)	(111111100000, 100000000000)
(234, 1247070)	(7,11)	(111111000000, 110000000000)
(207, 1032238)	(9,8)	(111100000000, 111110000000)
(245, 1235467)	(5,10)	(111111110000, 111000000000)
(139, 441778)	(11,2)	(110000000000, 111111111110)
(295, 964161)	(1,6)	(111111111111, 111111100000)
(183, 680018)	(10,3)	(111000000000, 111111111100)
(129, 439993)	(12,1)	(100000000000, 111111111111)

Same as the output of BNL algorithm, Table III shows the dominant objects calculated from bitmap algorithm.

3) Nearest Neighbour: At first, using the modulus distance, the nearest points were found out. Then, considering a logical vertical and horizontal line, the points were dropped right side of the vertical line and top side of the horizontal line.

From 2nd column of Table IV, the first nearest point from the origin is (1, 6) which is represented in Fig. 3(a). Similarly, the next nearest point from the origin is (4, 5) as shown in Fig 3(b). Thus, the horizontal and vertical lines cover the skyline points inside it. Then it recursively makes space and any point between these spaces is discarded and so on. The overall output for Table II is shown in Table V.

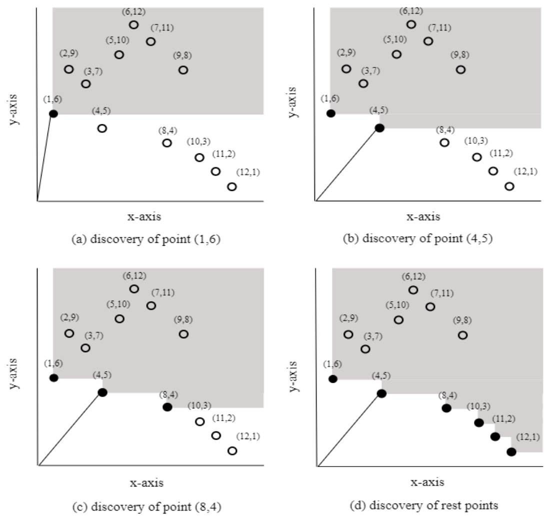

Fig. 3: Nearest neighbor algorithm working procedure

TABLE V: Output for NN Algorithm

City	Area	Price
Dhaka	295	964161
Khulna	262	912408
Mymensing	209	735887
Khulna	139	441778
Barishal	183	680018
Rajshahi	129	439993

TABLE VI: Output for Index Algorithm

City	Area	Price
Dhaka	295	964161
Rajshahi	129	439993
Khulna	139	441778
Mymensing	183	680018
Khulna	209	735887
Barishal	262	912408

4) Index: The Index scheme exploits a transformation mechanism that maps high-dimensional points into single dimensional space and a B+ tree structure is used to index the transformed points. Here, the index listing is shown in Table VII and it is stated that the algorithm is terminated at cmin = 4 because (4,5) in list 1 are less than or, equal to 5 and in list 2, (8,4) is less than or, equal to 8. Therefore, there is no need to further proceed.

TABLE VII: Sorted Index Listing to find out skyline points

list 1		list 2	
(1, 6)	cmin=1	(12, 1)	cmin=1
(2, 9)	cmin=2	(11, 2)	cmin=2
(3, 7)	cmin=3	(10, 3)	cmin=3
(4, 5)	cmin=4	(8, 4)	cmin=4
(5, 10)	cmin=5	(9, 8)	cmin=8
(6, 12)	cmin=6		
(7, 11)	cmin=7		

Thus, the final output from index algorithm is shown in Table VI.

B. Comparison among skyline algorithms

After applying the skyline algorithm on the small dataset in Table II, for all the cases the same result is formed. After

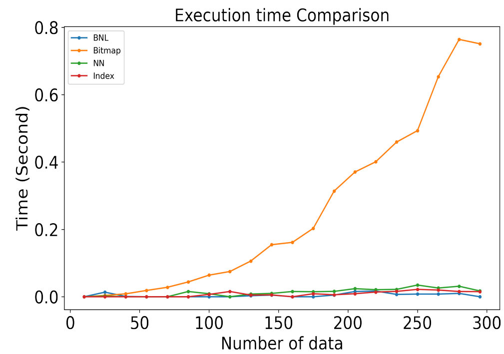

Fig. 4: Execution time comparison of all algorithms

applying the BNL, and bitmap, the same output with no indexed changes is found out. In these cases, there are total of 6 output skyline points as in Table III. In the case of the NN and Index algorithm, the output points are the same as the previous 2 algorithm's outputs but their order is different.

The execution times for all 4 algorithms are compared in Fig 4. Now, according to the execution time from Fig. 4, the performances of algorithms can be measured. For BNL and NN, both perform similarly. BNL perform slightly better than NN. Again, it can be stated that BNL takes almost constant time, but the NN algorithm needs extra time with the increment of size of input data. For 300 data, NN needs around 0.12 seconds and BNL needs 0.09 seconds whereas for 10000 data, NN needs around 3 seconds and BNL needs only around 0.7 seconds. Therefore, it can be stated that, in this experiment, for a huge number of data, BNL performs better.

Again, in case of bitmap, index and NN, bitmap takes a huge amount of time compared with the other 2 algorithms. For 300 data, bitmap needs around 0.78 seconds where index and NN need 0.15 seconds and 0.12 seconds, respectively. Here, when the number of data increases, the length of the bitmap representation increases, therefore the time for bitmap algorithm also increases. In the case of the index algorithm, it performs better than the bitmap algorithm but compared to BNL and NN, it is not a well performed algorithm. Therefore, for a large number of datasets, the BNL performs the best and the bitmap algorithm should be avoided.

The proposed framework maintains the desired criteria stated in section III-A as follows:

1) Owner's anonymity: In the proposed framework, the anonymity of the LO_n is maintained by the anonymous identity discussed in section III-B-1.

2) Data privacy: To ensure the privacy of the datasets, encryption using $y*$ by LO_n own is stated in section III-B-2. Again, since the decryption keys are divided among 2 entities, data privacy can not be breached while atleast 1 entity is honest.

3) Data anonymity: In order to anonymize the encrypted datasets, re-encryption and shuffling of the encrypted datasets are implemented by $COH1$ and $COH2$ in section III-B-3.

TABLE VIII: A Comparison Based on the Proposed Cryptographic Aspects

Aspects	Proposed Framework	[3]
Authentication of owner	by registration	not mentioned
Data storage	a public website	third party server
Encryption key generation	by combining 2 keys	single key of each owner
Decryption key generation	by 2 distinct keys of 2 servers	single key of each owner
Data anonymization	by Fisher Yate shuffling Algorithm	XOR and permutation
Skyline queries conduction	by specific entity	owners themselves
Re-encryption conduction	by 2 specific servers	not used

4) Auditable deeds: Since the datasets are posted on the website along with the unique identity, each LO_n can identify it's own datasets. Also, all of the publicly disclosed results confirm the authenticity of the results.

A comparison based on the aspects considered in this work and run time are represented in Table VIII and Table IX. These tables stated that the proposed framework is more practical and secured according to the considered features.

TABLE IX: A Comparison Based on the Number of Operations Required for Proposed Security Aspects

Aspects	Proposed Framework	[3]
Registration	1 exponentiation and 1 modulus	Not considered
Encryption key	1 from 2 cryptographic handler	1 for each owner
Decryption key	2 distinct key	1 for each owner
Re-encryption	2	not considered
Shuffling	1	1

V. Conclusion and Future work

The proposed system for multi-party land selling and buying platform using skyline query adopts 2 server to perform the cryptographic operations and an assistant to complete non-cryptographic and other major tasks. Here the decryption key division between 2 server ensures the data privacy, re-encryption and shuffling ensures the anonymity of datasets and the anonymous identity of the authenticate land owner provides the owner's anonymity. Finally, the 4 skyline algorithm individually provides the best lands within specific criteria. Here, the algorithms perform differently according to the datasets. In the modern world, online land purchasing is more practical and useful. With the increment of the land owner, the data increases rapidly. So skyline algorithm is useful then. In future, skyline queries on dynamic datasets with more dimensions, volumes of computations and communications etc. must be evaluated.

References

[1] S. Borzsony, D. Kossmann, and K. Stocker, "The skyline operator," in 17th Int. Conf. on data engineering. IEEE, pp. 421–430, 2001.

[2] D. Das, K. M. R. Alam and Y. Morimoto, "A Framework for Multi-party Skyline Query Maintaining Privacy and Data Integrity," 2021 24th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2021, pp. 1-6, doi: 10.1109/ICCIT54785.2021.9689854.

[3] M. Qaosar, K. M R. Alam, A. Zaman, C. Li, S. Ahmed, M. A. Siddique, and Y. Morimoto, "A framework for privacy-preserving multi-party skyline query based on homomorphic encryption," IEEE Access, Vol. 7, pp. 167481–167496, 2019.

[4] Bency, A.C., 2014, "A Study on Dynamic Skyline Queries". Published in International Journal for Research in Applied Science and Engineering Technology (IJRASET).

[5] T. ElGamal, "A Public-Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms," IEEE Trans. on Information Theory, Vol. 31, No. 4, pp. 469.472, 1985.

[6] D. Das, K. M. R. Alam and Y. Morimoto, "An Anonymity Retaining Framework for Multi-party Skyline Queries Based on Unique Tags," in IEEE Transactions on Dependable and Secure Computing, doi: 10.1109/TDSC.2023.3323961.

[7] Kossmann, D., Ramsak, F. and Rost, S., 2002, January, "An online algorithm for skyline queries". In VLDB'02: Proceedings of the 28th International Conference on Very Large Databases (pp. 275-286). Morgan Kaufmann.

[8] Y. Teng, D. Liu, X. Liu, W. Zhao, H. Liu and C. Fan, "Secure Spatio-textual Skyline Queries on Cloud Platform," 2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), 2020, pp. 251-259.

[9] Papadias D, Tao Y, Fu G, Seeger B. An optimal and progressive algorithm for skyline queries. InProceedings of the 2003 ACM SIGMOD international conference on Management of data 2003 Jun 9 (pp. 467-478).

[10] S. Zhang, S. Ray, R. Lu, Y. Guan, Y. Zheng and J. Shao, "Toward Privacy-Preserving Aggregate Reverse Skyline Query With Strong Security," in IEEE Transactions on Information Forensics and Security, vol. 17, pp. 2538-2552, 2022, doi: 10.1109/TIFS.2022.3188147.

[11] Bavirthi, S.S., Supreethi, K.P. An efficient framework for spatio-textual skyline querying and minimizing search space using R+ tree indexing technique. Int. j. inf. tecnol. 14, 1263–1271 (2022).

[12] Tan, K., Eng, P. Ooi, B. "Efficient Progressive Skyline Computation. Very Large Data Bases" Conference (VLDB), 301-310, Rome, Italy, September 11-14, 2001.

[13] Kossmann, D., Ramsak, F. and Rost, S. (2002) 'Shooting stars in the Sky', VLDB '02: Proceedings of the 28th International Conference on Very Large Databases, pp. 275–286. doi:10.1016/b978-155860869-6/50032-9.

[14] Hjaltason, G.R. and Samet, H. (1999) 'Distance browsing in spatial databases', ACM Transactions on Database Systems, 24(2), pp. 265–318. doi:10.1145/320248.320255.

[15] N. Beckmann, H. Kriegel, R. Schneider, B. Seeger. "The R*-tree: An Efficient and Robust Access Method for Points and Rectangles", SIGMOD, 1990.

[16] K.L. Tan, P.K. Eng and B.C. Ooi. "Efficient Progressive Skyline Computation", VLDB, 2001.

[17] D. Bhanushali, A. Koul, S. Sharma and B. Shaikh, "BlockChain to Prevent Fraudulent Activities: Buying and Selling Property Using BlockChain," 2020 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2020, pp. 705-709.

[18] S, Krishnapriya & Sarath, Greeshma. (2020). "Securing Land Registration using Blockchain. Procedia Computer Science". 171. 1708-1715. 10.1016/j.procs.2020.04.183.

[19] Yusfrizal, D. Adhar, U. Indriani, E. Panggabean, A. Sabir and H. Kurniawan, "Application of the Fisher-Yates Shuffle Algorithm in the Game Matching the World Monument Picture," 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), Manado, Indonesia, 2020, pp. 1-6, doi: 10.1109/ICORIS50180.2020.9320766.

[20] S. Tamura and S. Taniguchi, "Enhancement of anonymous tag based credentials," in Information Security Comput. Fraud, Vol. 2, No. 1, pp. 10–20, 2014.

[21] H. M. A. Fattah, K. M. A. Hasan and T. Tsuji, "Weighted top-k dominating queries on highly incomplete data", Inf. Syst., vol. 107, Jul. 2022.

[22] Fattah, H. A., Hasan, K. A., & Tsuji, T. Indexed top-k dominating queries on highly incomplete data. In Proceedings of the international conference on big data, IoT, and machine learning (p. 231). Springer.

[23] Kung, H.T., Luccio, F. and Preparata, F.P., 1975. "On finding the maxima of a set of vectors. Journal of the ACM (JACM)", 22(4), pp.469-476.