**CHAPTER 4**

**4.1. Overview**

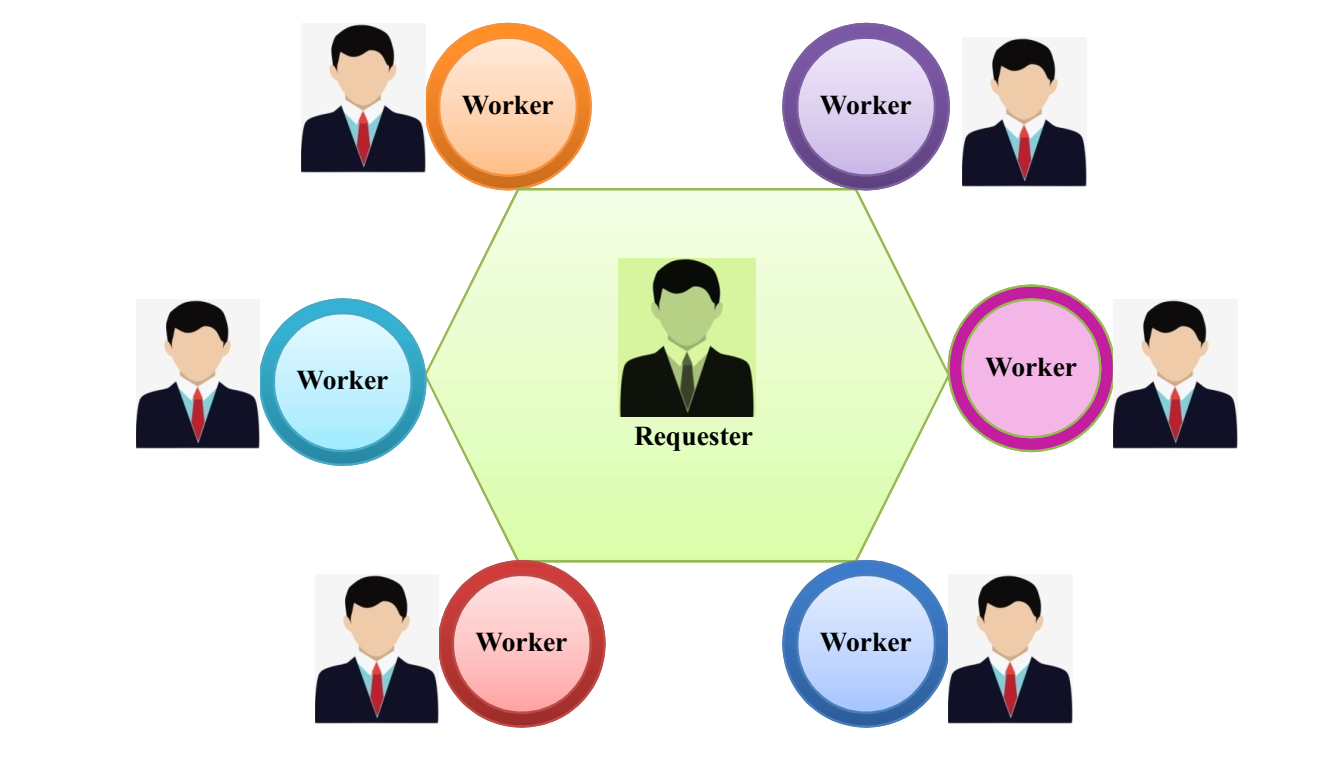
In recent times, the crowdsourcing pattern has grabbed more interest in real-time applications such as creation, research, web research, tagging, and processing of training because of its effectiveness (Chawla et al.2019). The main phases of the crowdsourcing system are task handling, task location, and expert task reformation. Query optimization refers to the process of improving the performance of a database query by selecting the most efficient query execution plan (Balaji et al.2021). The basic function of query optimization is to manage the continuous relational database. Although queries are sent to the database server simultaneously, the query optimizer does not communicate with the user (Panahi et al. 2019). The analyzer is the one who initially analyzes the query and achieves the optimization at some point. In some cases, problems occur when analyzing the database structure which adds additional complexities to query processing (Qu et al.2019). The different characteristics of the data architecture are used to collect the queries and the correct queries must be gathered in the database. The procedure path is selected based on different handling times for every query and the time difference varies in terms of hours and minutes. Based on the intention of the query it is evaluated using a query optimizer. An effective query optimizer selects the optimal query (Zhang et al. 2020). For a high-quality query, the objective functions are computed in terms of latency and cost. This chapter aims in selecting the optimal KNN query in crowdsourcing system. To overcome an optimal query selection problem, an optimal KNN query selection utilizing the Enhanced Horse Optimization Algorithm-based Intelligent Query Optimization in Crowdsourcing Systems (EHOA-IQOCSS) is implemented, and the cost and latency are the two objectives that need to be reduced by the algorithm. Optimal query selection combined with query sorting can be done with an inexpensive technique. Each query with the best query location accomplishment is selected for optimization. The representation value is easily chosen with the help of the improved optimization function. In this way, the determination issue associated with the increased cost is tackled along with increased accuracy. Besides standardization of the downside ordering section is used to limit the accuracy and improve the intellectual batch efficiency. The query improver needs to evaluate the query before its submission and the task is now to select the best query produced. Since an excellent query identification system comprises many consequences, an optimal crowdsourcing system is used to enhance these consequences. This EHOA-IQOCSS method accomplishes trade-off among latency as well as costs in the query of process. The SELECT operation and JOIN operation are used to mimic the human form of choosing certain objects based on a specific condition and to merge objects from two places in compliance with certain conditions respectively. This method’s efficiency is computed based on different consequences namely Point of Interest (POI) information, space information, gang dimension, ranges, as well as KNN query.

**4.2. Proposed Methodology**

For the query optimization process, a novel EHOA-IQOCSS method is proposed. The proposed method majorly uses chaotic concepts depending on an improved genre of HOA. In the crowdsourcing technique, the consumer plans to form a group based on Service Location (SL). Therefore, members of a group accept the query assessment protocol and store the details about the mobile location. Based on the location provided to consumers, K-nearest neighbor applications or query thresholds are scaled. Group formation helps the client identify group members for the question assessment task. The participant in the whole group is evaluated with queries from their limited database. A query gets back with its associated spatial data and the point of interest. A protocol for consecutive or coincident query ranking has a trade-off between transmission cost and processing times. The KNN and query ranges are entranced depending on the location of the user. Group formation assists prioritization of users of the group members, during query evaluation. Based on the total count of group members, the queries are accessed from a specific database. An authenticated query is returned with its associated spatial data and a point of interest is provided. There is a trade-off between processing time and transmission cost based on whether a sequential or parallel queries ranking process is implemented.

**4.2.1 Parallel Processing in Crowdsourcing**

Depending on parallel processing, the user posts the KNN and the state of the span query to each member of the group. The concept of crowdsourcing based on parallel processing is exhibited in Figure 4.1. The local database queries are estimated by group members and the query data location, and the POI information is returned. Thus, a reliability level calculation query response and data selection are applied to the query submitted evaluation for the last response, before the data location and the POI recipient of the group member.

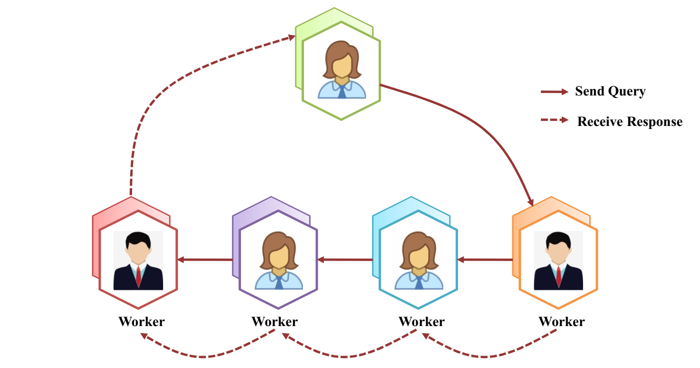
****

**Figure 4.1. Parallel processing**

Various group members use their scalability in the query-processing environment which is a basic dignity of parallel processing operations. The number of contributors can be multiplied without any extra processing time based on the query computation process. The members of the group can be evaluated in parallel. Otherwise, the additional storage locality and POI associated with the parallel processing of multiple group members enhances the communication overhead. Processing time typically increases when group members are included in query-processing task, but parallel processing helps control timing needs.

**4.2.2 Serial Processing in Crowdsourcing**

The applicants select members from the group and assign an ID to each participant who visited the flags (Mahin Mehnaz Tabassum et al. 2017). These moves are only applicable to group members participating in query estimation. Serial processing in a crowdsourcing environment is given in Figure 4.2.

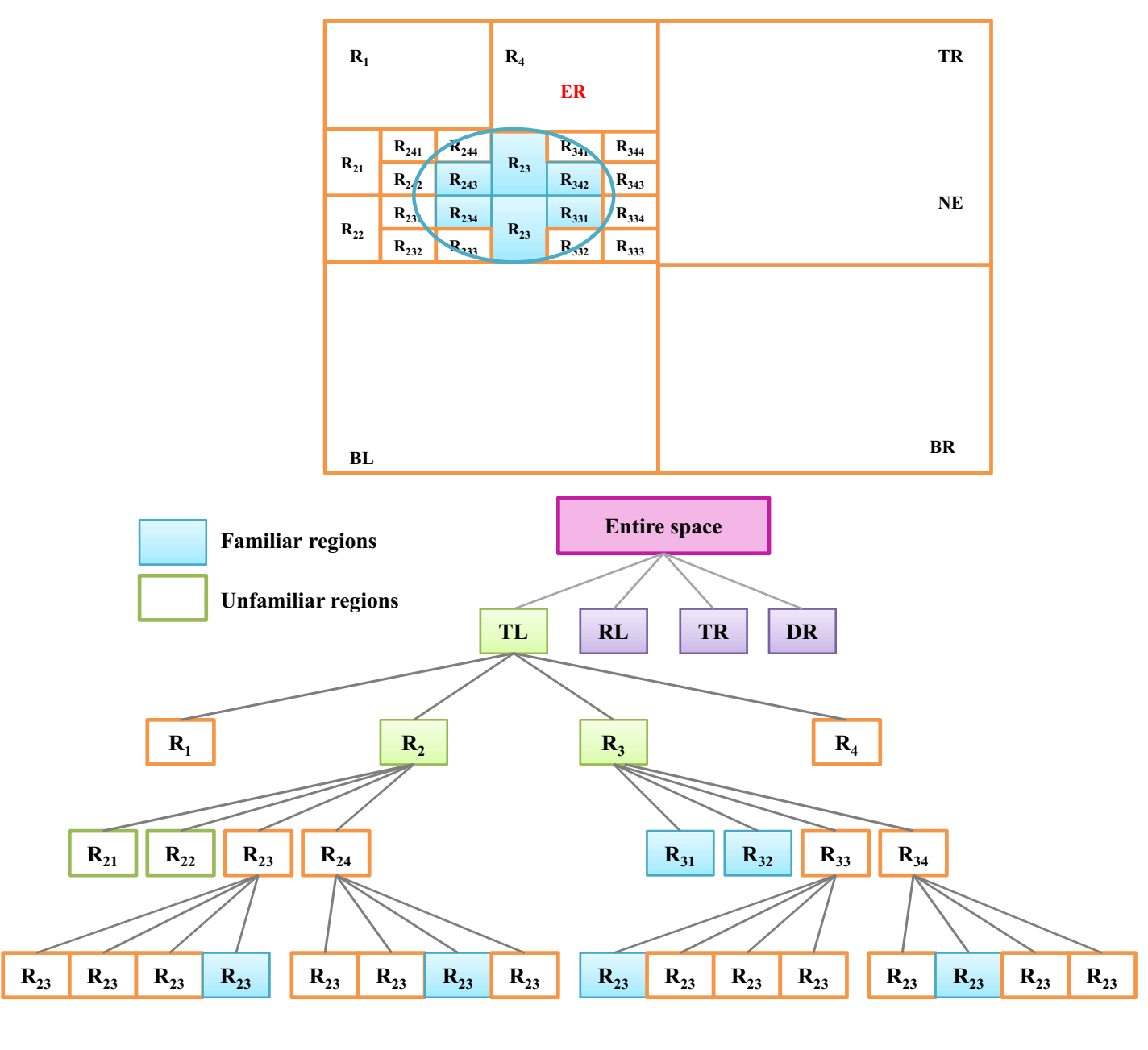


**Figure.4.2. Sequential processing**

A local database visit is used with space data and the derived POI and ID are set to verifiable for visited flags. Then the member selects another member with an incorrect ID and the selected person's details are provided. The optimized query message is updated, and the visited flag ID is set to true if the user has provided an optimized query from their database. A reordered message is sent to the next member with an incorrect ID. This method is applied to all visited flags. Finally, the query requester receives this message. Finally, depending on the overall results, the query requester returns the correct result, and the query comment is discovered using the confidence level. The POI will not be retrieved by the query candidate, even if multiple participants use the same data space. Individuals opting for limited communication bandwidth will desire serial processing.

**4.2.3 Space information index**

The hierarchical information architecture (Hashem et al. 2018) of the space tree is broadly utilized and is also called the quad-tree. The whole space is thus represented as the source node for the space tree. The segmentation task is terminated when the known and unknown region is accepted or when the space tree achieves the peak boundary area. The tree node value is set to 1 if the node area is displayed within the close region. Otherwise, the strange field is not concatenated, and the value becomes 0. The region covers the well-known area, then sets the node value back to 0, and the area is classified into 4 segments. Figure 4.3 shows the easy and increasing regulation for partitioning the space trees. Typically, a POI tree or KNN query evaluation in the query limit is used to discover the query answer by using the POI user system.



**Figure 4.3 Space Tree Structure**

 is expressed in terms of a parallel technique query candidate with all group members, which gives the function . The candidate establishes the k-neighbor POI from the purchased POI to the KNN query. The POI does not combine more than one query response due to the query range. Then, the accuracy is resolved by using the confidence level of the query response. While providing the same locality and POI the detail is transferred from one member to the other. The query applicant obtains the POI of the correct information location after serial processing, and the query response is determined by the confidence level.

**4.2.4 Trust stage**

Two types of trust stages namely distance and area are estimated utilizing the query answer standard. The best response quality determines the trust stage, and the query range of POI depends entirely on the region. The Query response contains all the POI for the query instance of the KNN whose confidence level is taken from the Query Response using the average  confidence value. Here  is the  nearest neighbor from  in the answer set . A circular segment  has a center  and  radius at a mutual distance 

The LCR group is denoted  with a circular area cross-section . The LCR near region  is denoted as and the entire region near  is represented as . As denoted below equation, the confidence levels are evaluated in accordance with the area for the POI value .

 (4.1)

The query response obtained by using the area-based confidence level is provided in the following equation.

 (4.2)

For a set of well-known LCRs, the query range  is denoted as. Now the region is denoted as  and the entire region is denoted as. Now the query range is obtained based on the trust stage.

(4.3)

Here  and  elect the trust stage calculated based on the area. For distance-level, trust-based KNN query estimation, a specific distance-constrained range POI information is not available. The  nearest POI value in the distance-based trust level calculation is 0.5. Then the distance  is not less than half the distances among  and . Here  is taken from a well-known location, and its distance is denoted as , which is the nearest boundary of the known location. The distance between  and  is denoted as C2 . As shown in Equation (4.4), the calculated distance-based trust stage for the POI is represented in the response to the 

 (4.4)

For each POI, if the distance-based trust stage is zero, then the query location  does not include the identified location. Using the equation below, the KNN is established based on the query response distance.

 (4.5)

Here  and  represent the trust stages calculated based on distance. Some POI detail has a logical distance, and the distance between POI is estimated by reducing the value of the distance . The closest trust stage POI has the confidence level value, while the nearest  POI has a minimum trust stage value. Based on the trust level calculations, the threshold query and KNN are evaluated.

**4.2.5 Trust Stage Computation-KNN Query using EHOA-IQOCSS**

The query processing is conducted by using the Enhanced Horse Optimization Algorithm based Intelligent Query Optimization in Crowdsourcing Systems (EHOA-IQOCSS). By employing an optimization technique, the cost, communication time, and communication overheads are reduced. To determine the query’s cost, latency, as well as communication overhead the KNN query optimization takes either parallel or serial query processing. Besides, the EHOA-IQOCSS is used to select the best KNN query for the customer with less latency, cost, and communication overhead. A novel bio-inspired system was introduced with its key resources being the features of the horse herd. The EHOA-IQOCSS includes parameters for configuration and instruction flow. Few variables include horse distribution rate (HDR), horse memory pool (HMP), solitary stallion percent (SSP), and dominant stallion percent (DSP). Simultaneously, few configurable HOA parameters were typical to other bio-inspired methods such as the reorganization frequency, dimensions, size of population  , and the number of iterations. The stages involved in EHOA-IQOCSS are depicted in Figure 4.4. The EHOA’s major stages are mentioned below.

Step 1.  horses were randomly generated in  dimension penetrating spaces.

Step 2. The method computes Fitness values using objective function as well as update values.

Step 3. The iteration value is set to 0. When it is less than 1, the processes proceed to step 4, otherwise it proceeds to step 8.

Step 4. The optimal horse in accordance with FV is the herd leader from the set  of recently initialized herds thus .

Step 5. Based on Fitness Value the best horse is a single stallion thereby forming set .

Step 6. The remaining horses are assigned randomly to the herd from .

Step 7. The worst horse in terms of fitness values are randomly distributed in the  dimension searching space

Step 8. For each herd  of size  from, rank is from  so that the maximum rank value is assigned to the horse with the best Fitness value, Step 9. The gait was upgraded from [1,2] to an arbitrary number. When the horse was single then the velocity was upgraded in step 10, otherwise, velocity was upgraded in step 11.

Step 10. The following equations was employed in updating the location of single stallion.

(4.6)

So that , are the velocity values in  and  signifies arbitrary number, indicates the gait horse, represents the center of the adjacent herd, and  signifies the location of the stallion.

Step 11. When the horse belongs to the herd,

(4.7)

So that and signifies the rank of  and center of herd.

Step 12. The location update equation are.

(4.8)

In equation 4.3, and  denotes the position in  and respectively.

(4.9)

The memory update equation is expressed as .

 (4.10)

A picture containing text, screenshot, diagram, line

Description automatically generated

**Figure 4.4. Steps involved in EHOA-IQOCSS**

So that  denotes a standard distributed number with a standard deviation of 1 and a mean of 0.

Step 14. For every herd of size from, the ranks are from. The horse with the best FV receives the highest rank value, which is the maximum value in the case of a maximization problem or the minimum value in the case of a minimization problem. When two horses with indices  from have a similar FV, the horse with the  index has the greatest rank when compared to the horse with the  index. The herd center is calculated as a weighted arithmetic mean of the position of the horses, where the weight is the rank.

Step 15. The nearby herd was defined for each stallion from .

Step 16. The iteration value was improved by 1.

Step 17. The processes return the  best solution value.

**4.2.6 Process Involved in EHOA-IQOCSS Model**

The EHOA-IQOCSS model optimizes selectivity and coupled queries with low delay and expense. In this section, the objective functions and query optimization of the EHOA-IQOCSS model are obtained. The objective or fitness function for crowdsourcing follows based on latency, data privacy, as well as reduced cost as follows.

 (4.11)

The constraint derivation including latency, cost, and data privacy are discussed below.

**4.2.6.1 Homomorphic Encryption Model**

Requester privacy as well as Employee privacy for homomorphic encryption model are developed in this section. This technique is well-organized Employee for encryption bit–wise information as well as the computational result of decryption (Renukadevi et al. 2021). By using the model provided, plaintext 0 or 1, they find a plaintext and reject them easily.

**Key generation:**

After, suppose that  and  mutually independent large prime numbers. Thus, the Legendre symbols  selection-residues.

Consequently, the private and public keys are classified as as well as , whereis the Jacobi symbol (Gong et al. 2014)

**Encryption:**

The encrypted text is taken as  and non-zero random number like is chosen. Hence, the cipher-text  is indicated as:

(4.12)

**Decryption:**

If the cipher text is  and ‘*’* is a quadratic residue, then t is verified by the prime factor . First, compute  and,  denote the quadratic residue hold and.

## 4.2.6.2 Monetary Cost (CRWDcost )

## Now, a query plans with cost is generated . It refers to the reward given to a crowdsourcing operator for activation . The cost of executing the operator is 0. In accordance with the crowdsourcing task computing conditions by unit, the cost is . The unit cost is. When the operator has an additional join condition the larger unit cost obtains Join operator

(4.13)

In Eq. (4.13), the query plans is the query  where the cost value is minimized.

## 4.2.6.3. Latency(CRWDlatency )

## The query evaluation speeds are scaled by delay. The optimization as well as delay estimation is nontrivial. The general crowd structure consists of a limited number of crowd workers.

 (4.14)

In Eq. (4.14), the cost budget containing query. It defines the minimum cost one where there are multiple projects with low latency.

**4.3 Summary**

A novel EHOA-IQOCSS algorithm is proposed to design the query optimization processes. The EHOA-IQOCSS design is based on an improved genre of HOA employing chaotic concept. It aims to achieve an optimal tradeoff among latency as well as cost in the query optimization processes. Besides for the data encryption process, homomorphic encryption technique is used. Also, the EHOA-IQOCSS model optimizes selecting and joining of queries with minimum cost and latency. The experimental evaluation of the EHOA-IQOCSS design is conducted under a series of simulations.