CHAPTER 2

LITERATURE REVIEW

2.1 SPATIAL DATABASES

Recently, attention has been focused on spatial databases, which combine conventional and spatially related data, such as Geographic Information Systems, CAD/CAM, or VLSI. A language has been developed to query such spatial databases. It recognizes the significantly different requirements of spatial data handling and overcomes the inherent problems of the application of conventional database query languages.

Spatial databases, addressing the growing data management and analysis needs of spatial applications such as Geographic Information Systems, have been an active area of research for more than two decades. Shashi Shekhar (2003) has produced taxonomy of models for space, spatial data types and operators, spatial query languages and processing strategies, as well as spatial indexes and clustering techniques. However, more research is needed to improve support for network and field data, as well as query processing.

The recent advances in database technology have enabled the development of a new generation of spatial databases, where the DBMS is able to manage spatial and non-spatial data types together. Most spatial databases can deal with vector geometries (e.g., polygons, lines and points), but have limited facilities for handling image data. However, the widespread availability of high-resolution remote sensing images has improved
considerably the application of images to environmental monitoring and urban management. Therefore, it is increasingly important to build databases capable of dealing with images together with other spatial and non-spatial data types. Lubia Vinhas (2003) describes a solution for efficient handling of large image data sets in a standard object-relational database management system. By means of adequate indexing, compression and retrieval techniques, satisfactory performances can be achieved using a standard DBMS, even for very large satellite images.

Selection and join queries are fundamental operations in Data Base Management Systems (DBMS). Support for nontraditional data, including spatial objects, in an efficient manner is of ongoing interest in database research. Toward this goal, access methods and cost models for spatial queries are necessary tools for spatial query processing and optimization. Yannis Theodoridis (2000) presents analytical models that estimate the cost (in terms of node and disk accesses) of selection and join queries using R-tree based structures. The proposed formulae need no knowledge of the underlying R-tree structure(s) and are applicable to uniform-like and non-uniform data distributions. In addition, experimental results are presented which show the accuracy of the analytical estimations when compared to actual runs on both synthetic and real data sets.

Much of the data that we encounter has a spatial aspect. Yet this has not been readily exploited by traditional RDBMS. Over the past five years there has been a confluence of Geographic Information System technology, RDBMS architecture and SQL standards that has fostered the implementation of spatial processing within the RDBMS. Adler (2001) discusses the overview of spatial processing and the evolution of technology leading to the development of the IBM DB2 spatial Extender that exploits the IBM DB2 universal Database object relational support to implement a standard based SQL spatial capability.
Path based Range Query (PRQ) for Two dimensional spatial database can be defined as follows: Given a sequence of query point, $P = \{p_1, p_2, \ldots, p_n\}$, and a search distance $d$, want to report all points in the spatial database that are within a distance $d$ of some point $p_i$ in $P$. This query arises from traveler information system where it is often a feature to report events that lie nearby a planned route. The simple method of performing Repeated Range Query (RRQ), i.e. the standard range query for each query point $p_i$ and combining the results is inefficient as it involves multiple searches on the database. Hoong Kee Ng (2004) presented an algorithm for the PRQ that uses only one pass of the R-tree which simultaneously processes all the points in the query path $P$. Generalize pruning rules for the standard range query and also present new ones for efficient processing of PRQ. An experiment shows that the algorithm for PRQ outperforms RRQ significantly and that this is consistent across the various problem parameters studied.

Observing that networks are ubiquitous in applications for spatial databases, a new data model and query language is defined that especially supports graph structures. This model integrates concepts of functional data modeling with order sorted algebra. Besides object and data type hierarchies, graphs are available as an explicit modeling tool, and graph operations are part of the query language. Graphs have three classes of components, namely, nodes, edges, and explicit paths. These are at the same time object types within the object type hierarchy and can be used like any other type. Explicit paths are useful because real world objects often correspond to paths in a network. Furthermore, a dynamic generalization concept is introduced to handle heterogeneous collection of objects in a query. In connection with spatial data types, this leads to powerful modeling and querying capabilities for spatial databases, in particular for spatially embedded network such as highways, rivers, public transport and so forth. Martin Erwig (1994) used multilevel order sorted algebra as a formal framework for the specification of
this model. First level algebra defines types and operations of the query language and the second level algebra defines kinds and type constructors as function between kinds, and provides the types that can be used at the first level.

Earth science (ES) applications handle very large geospatial data sets and interactive response time is required by its query processing. Spatial selection is one of the very important basic operations for geo-spatial databases. It retrieves all the objects that intersect with a given point or rectangle. Keiichi Tamura (2001) presented a novel approach for the parallel processing of spatial selection of very large geospatial databases using partitioned parallelism. To evaluate this approach, the Extended Sequoia 2000 benchmark is used, which has real world data and real queries. The experimental results of parallel processing of spatial selection show good speed-up.

K-Nearest Neighbor (k-NN) queries are used in GIS and CAD/CAM applications to find the k spatial objects closest to some given query points. Most previous k-NN research has assumed that the spatial databases to be queried are local, and that the query processing algorithms have direct access to their spatial indices; e.g., R-trees. Clearly, this assumption does not hold when k-NN queries are directed at remote spatial databases that operate autonomously. While it is possible to replicate some or all the spatial objects from the remote databases in a local database and build a separate index structure for them, such an alternative is infeasible when the database is huge, or there are large number of spatial databases to be queried. Danzhou Liu (2002) proposed a k-NN query processing algorithm that uses one or more window queries to retrieve the nearest neighbors of a given query point and also proposed two different methods to estimate the ranges to be used by the window queries. Each range estimation method requires different
statistical knowledge about the spatial databases. Apart from not requiring
direct access to the spatial indices, the window queries used in the proposed
algorithm can be easily supported by non-spatial database systems containing
spatial objects.

Similarity based retrieval of images is an important task in many
image database applications. A major class of users’ requests requires
retrieving those images in the database that are spatially similar to the query
image. Gudivada (2004) proposed an algorithm for computing the spatial
similarity between two symbolic images. A symbolic image is a logical
representation of the original image where the image objects are uniquely
labeled with symbolic names. Spatial relationships in a symbolic image are
represented as edges in a weighted graph referred to as spatial orientation
graph. Spatial similarity is then quantified in terms of the number of as well
as the extent to which the edges of spatial orientation graph of the database
image conform to the corresponding edges of the spatial orientation graph of
the query image. The algorithm is robust in the sense that it can deal with
translation, scale, and rotation variances in images. The algorithm has
quadratic time complexity in terms of the total number of objects in both the
database and query images.

2.2 INTELLIGENT TRANSPORTATION SYSTEMS

Drivers wish to travel in the best conditions in terms of time, safety,
economy and comfort. Currently the information (travel time and forecast)
related to journey is collected by different agencies at different frequency and
at different levels of detail. Little effort has been made to integrate and utilize
the information efficiently. A prototype Road Information System (RIS) is
being developed at The University of Calgary. A RIS is a special-purpose
Geographic Information System (GIS) that collects, organizes and
disseminates information related to the street and highway networks of a city
or a region. The basic structure of the RIS is the road and street network which is made up of a series of nodes and links. Attributes from different applications can be referenced by these nodes and links. The structure, if properly designed, can serve many applications such as route determination and route guidance, commercial and service information provision, dispatching and monitoring fleet vehicles, as well as road inventory and traffic management. Lee (1989) research effort of the prototype RIS concentrates on data modeling, spatial data base management, communications and real-time applications for the many Automatic Vehicle Location, Navigation and Guidance (AVLNG) Systems presently being developed.

Sudip Misra (2005) developed the first Learning Automatonbased solution to the dynamic single source shortest path problem. It involves finding the shortest path in a single-source stochastic graph topology where there are continuous probabilistic updates in the edge-weights. The algorithm is significantly more efficient than the existing solutions, and can be used to find the “statistical” shortest path tree in the “average” graph topology. It converges to this solution irrespective of whether there are new changes in edge-weights taking place or not.

In such random settings, the new learning automata solution converges to the set of shortest paths. The algorithm can be applicable in domains ranging from ground transportation to aerospace, from civilian applications to military, from spatial database applications to telecommunications networking.

On the other hand, the existing algorithms will fail to exhibit such a behavior, and would recalculate the affected shortest paths after each weight-change. The important contribution of the algorithm is that all the edges in a stochastic graph are not probed, and even if they are, they are not all probed equally often. Indeed, the algorithm attempts to almost always probe only
those edges that will be included in the shortest path graph, while probing the other edges minimally. This increases the performance of the algorithm. All the algorithms are tested in environments where edge-weights change stochastically, and where the graph topologies undergo multiple simultaneous edge-weight updates. It is superior in terms of the average number of processed nodes, scanned edges and the time per update operation, when compared with the existing algorithms.

Jagadeesh (2002) suggested that the performance of conventional route computation algorithms tends to deteriorate as the size of the network increases. A series of experiments were conducted on a real city road network to evaluate a heuristic technique before incorporating it into a hierarchical route-finding algorithm based on road types. The improved hierarchical algorithm computes a near optimal route in a fast and efficient manner by restricting the route computation to small subnetworks. The solutions provided by the algorithm are comparable to the optimal solutions.

One of the major focus areas of ITS is the development of dynamic route guidance systems for road vehicles. The basic purpose of a route guidance system is to provide the drivers with help to plan a route according to their preference and to guide them through the planned route to the destination. The system helps the driver to avoid congested roads and also to find the best route in an unfamiliar locality. Apart from benefiting the individual commuters, route guidance systems could aid the transportation industry to make substantial savings by reducing unwanted travel time and fuel wastage.

With the implementation of Intelligent Transportation System for system management purpose, there is now the ability to extract archived data that can be used to evaluate the implementation of new operational strategies. In recognition of the need to provide feedback to decision-makers, efforts are
underway to provide rigorous documentation of ITS benefits and cost. Bertini (2005) described how Advanced Traffic Management System (ATMS) data are being used to contribute toward these evaluations. Case examples are described in the areas of freeway management, incident management, arterial management and transit management. Building a complete ITS system requires collaboration in time, funding and institutional arrangements. ITS components that are integrated can result in synergistic effects when considered as an entire system. It is shown that in some cases it is possible to build upon national level statistics describing ITS benefits by using data collected from the systems themselves. It is hoped that further effort to integrate transportation planning with evaluation methodologies will incorporate the necessary empirical result from a wide variety of studies. In this way, better databases can be developed, and heightened accountability will be more pervasive in the evaluation of ITS improvement.

2.3 GENETIC ALGORITHM

A physical distribution system has a number of optimization problems. Most of them belong to a combinatorial problem, to which conventional mathematical programming methods may hardly be applied. Watanabe (2001) reports on two applications of the genetic algorithm (GA) to physical distribution scheduling problems, which arise at real physical distribution centers. The developed GA schedulers took the place of conventional schedulers, which were coded by rule-based technologies. Advantages of the introduction of GA schedulers into the physical distribution system are as follows: (1) the GA becomes a general problem-solver engine. Once we develop this engine, we only have to develop interfaces for the applications; and (2) fitness functions necessary for the GA force the physical distribution schedulers to have approximate performance estimation. This was not taken into consideration when the rule-based scheduler was used. Two
applications of the discussed schedulers were implemented with real
distribution centers, and they brought much efficiency to their management.

Liang ZOU (2007) suggested that the application of random A*
algorithm, clears out the biggest obstruction between the genetic algorithm
and dynamic route guidance of how to get the initial generation of genetic
algorithm. The developed models and algorithms are implemented with
Guangzhou electronic map and their computational performance is analyzed
experimentally. The results indicate that dynamic route guidance A*
algorithm is suitable for route guidance problem in FIFO dynamic network,
dynamic route guidance Q-learning algorithm is suited for route guidance
problem in steady non-FIFO dynamic network, and dynamic route guidance
genetic algorithm is suitable for route guidance problem in non-FIFO
dynamic network.

The route guidance system, which provides driving advice based on
traffic information about an origin and a destination, has become very popular
along with the advancement of handheld devices and the global position
system. Since the accuracy and efficiency of route guidance depend on the
accuracy of the traffic conditions, the route guidance system needs to include
more variables in calculation, such as real time traffic flows and allowable
vehicle speeds. As variables considered by the route guidance system
increase, the cost to compute multiplies. As handheld devices have limited
resources, it is not feasible to use them to compute the exact optimal solutions
by some well-known algorithm, such as the Dijkstra’s algorithm, which is
usually used to find the shortest path with a map of reasonable numbers of
vertices. To solve this problem, Chu Hsing Lin (2008) proposed to use the
genetic algorithm to alleviate the rising computational cost. By the use of
genetic algorithm to find the shortest time in driving with diverse scenarios of
real traffic conditions and varying vehicle speeds. The effectiveness of the
genetic algorithm is clearly demonstrated when applied on a real map of modern city with very large vertex numbers.

Heydar Toossian Shandiz (2008) proposed a method for controlling traffic lights in order to have maximum flow in the route which results in a moving traffic. As decision is made based on stochastic data, the method improves the decision in practice. In this method sensor send information on a computer, then based on genetic algorithm timing of green in traffic light is adjusted. Simulation result based on real data shows the full capacity of route is reached.

Francisco (2001) suggested a method to obtain the best possible route between two points on a real road map as optimization module in a fleet management system and implemented based on evolutionary computation techniques. Chakraborty (2005) proposed a GA-based algorithm to find out several alternative routes depending on different criteria, according to the driver's choice, such as shortest path by distance, the path which has minimum number of turns, passing through mountains or by the side of a river. The algorithm is quite efficient in finding alternative non overlapping routes with different characteristics.

2.4 HIGH PERFORMANCE CLUSTER

The terms "grid computing" and "cluster computing" have been used almost interchangeably to describe networked computers that run distributed applications and share resources. They have been used to describe such a diverse set of distributed computing solutions that their meanings have become ambiguous. Both technologies improve application performance by executing parallelizable computations simultaneously on different machines, and both technologies enable the shared use of distributed resources.
However, cluster and grid computing represent different approaches to solving performance problems; although their technologies and infrastructure differ, their features and benefits complement each other. A cluster and a grid can run on the same network at the same time, and a cluster can even contribute resources to a grid. Figure 2.1 shows Computer cluster topology.

![Figure 2.1 Computer cluster topology](image)

Mark Baker (2000) states that cluster computing are not a new area of computing. It is, however, evident that there is a growing interest in its usage in all areas where applications have traditionally used parallel or distributed computing platforms. The mounting interest has been fuelled in part by the availability of powerful microprocessors and high-speed networks as off-the-shelf commodity components as well as in part by the rapidly maturing software components available to support high performance and high availability applications.

Lohn (2002) presented results from a study comparing a recently developed co-evolutionary genetic algorithm (CGA) against a set of evolutionary algorithms using a suite of multiobjective optimization benchmarks. The CGA embodies competitive co-evolution and employs a simple, straightforward target population representation and fitness calculation based on developmental theory of learning. Because of these
properties, setting up the additional population is trivial, making implementation no more difficult than using a standard GA. Empirical results using a suite of two-objective test functions indicate that this CGA performs well at finding solutions on convex, non convex, discrete, and deceptive Pareto-optimal fronts, while giving respectable results on a non uniform optimization. On a multimodal Pareto front, the CGA yields poor coverage across the Pareto front, yet finds a solution that dominates all the solutions produced by the eight other algorithms.

Lu (2004) proposed a High performance neural network for image segmentation which utilized a PC cluster to implement back propagation training on large data sets. Back propagation training of the very large neural network described here requires substantial computing resources that impact the time required for training.

2.5 MULTI OBJECT GENETIC ALGORITHM

Genetic algorithms (GAs) are powerful, general purpose adaptive search techniques which have been used successfully in a variety of learning systems. In the standard formulation, GAs maintain a set of alternative knowledge structures for the task to be learned, and improved knowledge structures are formed through a combination of competition and knowledge sharing among the alternative knowledge structures. David Schaffer (1987) extended the GA paradigm by allowing multidimensional feedback concerning the performance of the alternative structures. The modified GA is shown to solve a multiclass pattern discrimination task which could not be solved by the unmodified GA.

Multi-objective evolutionary algorithms which use non-dominated sorting and sharing have been mainly criticized for their (i) $O(MN s)$ computational complexity (where M is the number of objectives and N is the
population size), (ii) non-elitism approach, and (iii) the need for specifying a sharing parameter. Kalyanmoy Deb (2000) suggested a non-dominated sorting based multi-objective evolutionary algorithm (we called it the Non-dominated Sorting GA-II or NSGA-II) which alleviates all the above three difficulties. Specifically, a fast non-dominated sorting approach with O(MN 2) computational complexity is presented. Second, a selection operator is presented which creates a mating pool by combining the parent and child populations and selecting the best (with respect to fitness and spread) N solutions. Simulation results on a number of difficult test problems show that the proposed NSGA-II, in most problems, is able to find much better spread of solutions and better convergence near the true Pareto-optimal front compared to Pareto Archived Evolution Strategy (PAES) and Strength Pareto Evolutionary Algorithm (SPEA) two other elitist multi-objective EAs which pay special attention towards creating a diverse Pareto-optimal front. Moreover, the definition of dominance is modifying in order to solve constrained multi-objective problems efficiently. Simulation results of the constrained NSGA-II on a number of test problems, including a five-objective, seven-constraint non-linear problem, are compared with another constrained multi-objective optimizer and much better performance of NSGA-II is observed. Because of NSGA-II's low computational requirements, elitist approach, parameter-less niching approach, and simple constraint-handling strategy, NSGA-II should find increasing applications in the coming years.

Presence of multiple objectives in a problem, in principle, gives rise to a set of optimal solutions (largely known as Pareto-optimal solution), instead of single optimal solution. This type of problem is known as multi-objective optimization problem (MOP). In general, a MOP has been solved using weighted sums or decision-making schemes. An alternative way is to look for the Pareto-optimal front. Many evolutionary algorithms (EAs) like
genetic algorithm (GA) have been suggested to solve a MOP, hence termed as multi-objective evolutionary algorithms (MOEAs). Nondominated sorting genetic algorithm (NSGA-II) is one such MOEA which demonstrates the ability to identify a Pareto-optimal front efficiently. Thus, it provides the decision maker (DM) a complete picture of the optimal solution space. Amar Kishor (2004) presented an application of NSGA-II in order to solve a multi-objective series system reliability optimization problem. Here, conflicting objectives such as maximization of system reliability and minimization of the system cost have been considered.

In the field of multi-objective optimization using evolutionary algorithms conventionally different objectives are aggregated and combined into one objective function using a fixed weight when more than one objective needs to be optimized. With such a weighted aggregation, only one solution can be obtained in one run. One method is to assign uniformly distributed weight to each individual in the population of the evolutionary algorithm. The other method is to change the weight periodically when the evolution proceeds. In this way a full set of Pareto solutions can be obtained in one single run.

On-Line Analytical Processing (OLAP) tools are frequently used in business, science and health to extract useful knowledge from massive databases. An important and hard optimization problem in OLAP data warehouses is the view selection problem, consisting of selecting a set of aggregate views of the data for speeding up future query processing. A common variant of the view selection problem addressed in the literature minimizes the sum of maintenance cost and query time on the view set. Converting what is inherently an optimization problem with multiple conflicting objectives into one with a single objective ignores the need and value of a variety of solutions offering various levels of trade-off between the
objectives. Michael Lawrence (2006) applied two non-elitist multiobjective evolutionary algorithms (MOEAs) to view selection under a size constraint. Emphasis is to determine the suitability of the combination of MOEAs with constraint handling to the view selection problem, compared to a widely used greedy algorithm. We observe that the evolutionary process mimics that of the greedy in terms of the convergence process in the population. The MOEAs are competitive with the greedy on a variety of problem instances, often finding solutions dominating it in a reasonable amount of time.

Miroslav Kulich (2003) presented an application of three soft computing techniques – ant colony optimisation, genetic algorithm, and neural networks to rescue operation planning. It considers the task as the multiple traveling salesmen problem and proposes suitable heuristics in order to improve the performance of the selected techniques. Then it applies the implemented solutions to a real data.

Masaya Yoshikawa (2005) proposed a new hybrid routing algorithm which combines Tabu search with Ant Colony Optimization. The proposed hybrid technique enables to find the shortest route including the blind alley. Experiments prove the effectiveness in comparison with conventional routing algorithm such as Dijkstra algorithm.