2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Spatiotemporal Aggregate Functions for Spatiotemporal Data

HyunHo Shin, SangHo Kim, KeunHo Ryu
Database Laboratory, Chungbuk National University, Cheongju Chungbuk, Korea
e-mail : { shinhh, shkim, khryu } @dblab.cbu.ac.kr

***Abstract* – Aggregate operator which belongs to query operations are important in specialized systems such as geographic information system(GIS) and spatial database system. Most of data describing objects in the real world are characterized by space and time attributes. Till now, however, works on aggregate operations have only dealt with spatial or temporal aspect of object. The current demand of aggregate operations relates to spatiotemporal data which are contained both spatial and temporal data concurrently. Therefore, work on spatiotemporal operations is focused on database area. In this paper, we propose spatiotemporal aggregate functions that operate on spatiotemporal data. Above all, we support spatiotemporal aggregate functions on the basis of three dimensional spatiotemporal models that are defined with the linear one dimensional temporal domain. The proposed algorithms are evaluated through some implementation results. We are sure that the achievement of our work is useful and efficient.**

I. Introduction

The interest on temporal and spatial databases has been growing increasingly, many works have been performed up to now. Early works focused on spatiotemporal modeling, indexing and recording[1]. In recent times, works have been going on query languages by appending temporal and spatial concepts to existing query languages like SQL in databases. Related works have been actively studied, such as functions are extended by adding temporal or spatial concepts, and process method on function. One of the query operations, aggregation, is applied to entire relation or tuples which form part of relation. It is an important operation that computes all of the values or select a representative value in query language[2].

In real world applications, user queries on spatiotemporal data, such as "What is the average rainfall in 'A' region from August, 2001 to September, 2001", have both temporal and spatial constraints and focus on the past and present time. However, previous works for aggregation are only attached to temporal or spatial aggregation. So, those are not directly available in real world because the characteristic of the real world data have both spatial and temporal condition.

In this paper, on the basis of the spatiotemporal databases and models, we propose spatiotemporal aggregate functions that are able to statistically analyze data by generating additional information from computation on past and current condition, as well as to process historical and spatial management on spatiotemporal data. Spatiotemporal aggregate functions can represent query briefly, so user can conveniently query on spatiotemporal data by extending query to support both temporal and spatial concurrently.

The rest of this paper is organized as follows: In section 2, we describe the related works on the spatial and temporal aggregate functions. Section 3 defines spatiotemporal models and spatiotemporal topological operators that extract spatiotemporal information. Section 4 describes the proposed spatiotemporal aggregate functions. In section 5, the proposed spatiotemporal aggregate functions will be implemented and evaluated with the estate management application. Finally, Section 6 present conclusion of our work and future directions.

II. Related Works

Differ from the query on conventional databases, the aggregate functions return the result value which are computed on the data satisfying a some condition[3]. Also, it is used to improve the system performance and to simplify the complex query that is made by user.

Temporal aggregation is performed on the historical information of time-varying data. In connection with this characteristic, many works are progressed on new aggregate function to add the temporal dimension in aggregation of conventional databases. Also, there are many researches on the aggregation tree[4] to process the aggregate operation in the temporal databases. This approach shows superior performance than temporal aggregate operation which is extended from the previous natural aggregate operation. At the same time, the approach of processing for temporal aggregate operation is proposed, that is based on the B^+-tree to store the aggregate value into the disk[5,6].

In order to offer new aggregate function for query, many researches are progressed on the aggregate for spatial. For efficient query processing, the MRA-tree (Multi-Resolution Aggregation Tree) handle the approximate aggregate query which is used in the popular multidimensional index[7]. The approach of processing for aggregate operation, such as MIN and MAX[8], is also researched on the spatial aggregate functions.

However, the researches for spatial and temporal aggregation just support the lopsided information into the temporal or spatial. Therefore, in this paper, we propose the

This work was supported by University IT Research Center Project and KOSEF RRC Project(Cheongju Univ. ICRC) in Korea

spatiotemporal aggregate function which is directly applied to the spatiotemporal data to solve the above problems.

III. SPATOTEMPORAL DATA

A. Spatiotemporal model

The spatiotemporal database store both spatial and non-spatial attributes, and offer efficient management and services on spatiotemporal attributes of existing objects on space. To make it possible spatiotemporal database must be able to represent information of spatial and non-spatial object, support retrieval and update of data, and offer historical information on change of position and topological relationship between spatial objects over time. Therefore, spatiotemporal data model should be needed to store efficiently spatiotemporal objects in database.

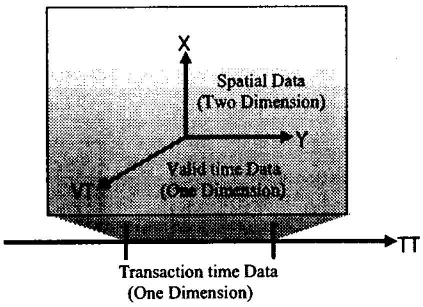

Fig. 1. Linear Three Dimensional Structures

In this paper, we support spatiotemporal aggregate functions on the basis of three dimensional spatiotemporal models that are defined with the linear one dimensional temporal domain, as figure1. That is to say, domain of linear three dimensional spatiotemporal data hieratically present any three spatiotemporal factor that is defined by domain of one dimensional linear transaction time. Factors which organize spatiotemporal domain are made up two dimensional temporal domains that consist of valid time and transaction time, as well as two dimensional spatial domains. In order to represent combined structures, spatiotemporal domains have four dimensional the structure at least. However it is not able to construct over the three dimensions in the theory, so that use linear three dimensional structures.

B. Storage of spatiotemporal model

In this paper, we apply to relational spatiotemporal data model to recording data and aggregation, which consider spatiotemporal property of spatial attribute. A relational spatiotemporal data model consists of the spatial table which add the temporal attribute on the any objects, and a non-spatial table. In order to store the defined spatiotemporal model, database structures consist of two relations "spatial historical relation" and "spatial attribute historical relation". The spatial historical relation stores spatial information of each region that has been changed with temporal evolution. And the spatial attribute historical relation stores information related to each spatial attribute that is changed over time.

Table 1. Spatial Historical Relation Structure

FID	Name	Size	X1	Y1	X2	Y2	VTs	VTe	Geo
int	varchar	int	float	float	float	float	Date	date	Binary

As see in table 1, the spatial historical relation stores the changed spatial information according to the interval of valid time corresponding to each spatial object.

Table 2. Spatial Attribute Historical Relation Structure

FID	Name	VTs	VTe	Value_1	...	Value_n
int	Varchar	date	date	float	...	float

As see in table 2, the attribute historical relation stores the changed attribute information according to the interval of valid time corresponding to each spatial object.

C. Spatiotemporal topological operators

The important and fundamental functions of database are to store the data and to retrieve the stored data. Accordingly we should consider about not only a representation method of data but also an operator which can retrieve the data efficiently. In this paper, proposed spatiotemporal aggregate functions execute aggregation process with obtained spatiotemporal data through spatiotemporal topological operators. Users also can execute various aggregate query by combining spatiotemporal aggregate functions and topological operators[9].

Table 3. Classification of Spatiotemporal Operators

Operator	Type of Operators
Spatiotemporal Topological Operator	tEqual(), tDisjoints(), tIntersects(), tTouches(), tCrosses(), tWithins(), tContains(), tOverlaps()

The table 3 shows classified spatiotemporal operators that are used in this paper.

IV. SPATIOTEMPORAL AGGREGATE FUNCTIONS

In this section, on the basis of spatiotemporal data model, we define spatiotemporal aggregate functions that are able to analyze data statistically to generate optional information through computation on the past and current information. We also introduce algorithm of stSUM among the proposed algorithms. Table 4 shows the notations to use definition for spatiotemporal aggregate functions.

Table 4. Notations for Spatiotemporal Aggregate Functions Definition

Notation	Descriptions
A	an attribute of spatial object
SC	a construction factor of spatial object(MBR).
W_L	a layer of given window area
F_L	a retrieved layer of spatial object
CVT	a common valid time
′	a current relation

[*Definition 1*] stCOUNT is the total of F_L which is contained within the W_L in spatial and within the valid time of given W_L. It is defined as follows:

Spatial (F_L) ***stCOUNT Spatial*** (W_L)
$= [COUNT(\forall F_L) \textit{ with } \exists SC'(F_L) \cap \exists SC'(W_L) \wedge CVT(F_L, W_L)]$

[*Definition 2*] stSUM is sum of attributes value of F_L which is contained within the W_L in spatial and within the valid time of given W_L. It is defined as follows:

Spatial (F_L) ***stSUM Spatial*** (W_L)
$= [SUM(\forall A'(F_L)) \textit{ with } \exists SC'(F_L) \cap \exists SC'(W_L) \wedge CVT(F_L, W_L)]$

[*Definition 3*] stAVG is attribute value of stSUM/ stCOUNT of F_L which is contained within the W_L in spatial and within the valid time of given W_L. It is defined as follows:

Spatial (F_L) ***stAVG Spatial*** (W_L)
$= [AVG(SUM(A'(F_L)) / COUNT(F_L))$
$\textit{with } \exists SC'(F_L) \cap \exists SC'(W_L) \wedge CVT(F_L, W_L)]$

[*Definition 4*] stMAX is max attribute value among attributes of F_L which is contained within the W_L in spatial and within the valid time of given W_L. It is defined as follows:

Spatial (F_L) ***stMAX Spatial*** (W_L)
$= [MAX(F_L) \textit{ with } \exists SC'(F_L) \cap \exists SC'(W_L) \wedge CVT(F_L, W_L)]$

[*Definition 5*] stMAX ⇔ stMIN

The table 5 shows the algorithm of stSUM which returns value of spatiotemporal attribute among the defined spatiotemporal aggregate functions. In the above, we proposed aggregate functions, such as stCOUNT, stSUM, stAVG, stMAX, stMIN, for spatiotemporal data.

There are two types of aggregate functions. One is the computational aggregate functions which yield the results through computation during given valid time interval, such as stCOUNT, stSUM, and stAVG, and the other is the selective aggregate functions which select the representative value of satisfied some conditions during given valid time interval, such as stMAX and stMIN. Spatiotemporal aggregate functions execute aggregate operation by comparing both spatial information and valid time.

Table 5. Algorithm of stSUM

```
Spatiotemporal Aggregate Function stSUM
Input : feature fid, spatial object type
Output : sum of feature attributes
Method :
Identify input features from the window;     /* query window */
While(Check current feature schema) {
/*extract spatiotemporal satisfying spatiotemporal conditions */
 If (feature fid in current feature schema equal feature fid in input
  tuple){
        if(valid time of feature fid in current feature schema <=
          valid time of feature fid in input tuple) {
             for(Retrieve all tuple != NULL) {
                 stSUM+ = aggregated attributes;
                     }
                }
          }
    return stSUM
  }
End stSUM   /* result of stSUM */
```

V. Implementation And Evaluation

In this section we implement the proposed spatiotemporal aggregate functions that apply to the estate management application, and evaluate the proposed algorithms.

A. Implementation platform

Our implementation platform is follows. The operating system is windows 2000 sever, database sever is Microsoft SQLServer2000 and language is Visual C++ 6.0. We have the experiment with assumption that experimental data generate the 13648 random data for the price of estate in Jung-Gu in Seoul, have the uniform distribution in the implementation, and used parameters are size of window, number of spatiotemporal data and interval of temporal query and so on. Figure 2 shows the result which applied a stAVG spatiotemporal aggregate function to obtained spatiotemporal data using tIntersects topological operator.

[*Query example*] "What is the average of estate from January, 2000 to December, 2002 with query window area(W)? "

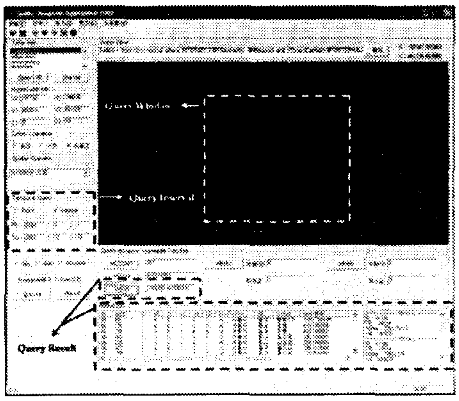

Fig. 2. Result of Execution on Window Query

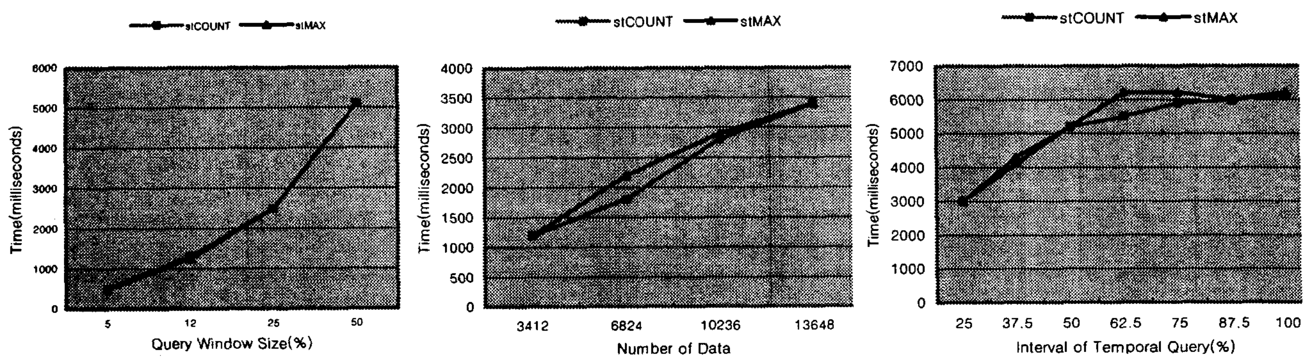

Fig. 3. Result of Performance Evaluation

B. Experimental Evaluation

We test the performance for the proposed algorithms of spatiotemporal aggregate functions in this paper. Figure 3 shows that size of query window, number of data and interval of temporal query execute performance evaluation on the stCOUNT of computational aggregate functions and stMAX of selective aggregate functions evaluate a performance. Therefore, proposed aggregate function can be applied to spatiotemporal data which existing aggregate functions can not support. We also show the validity of performance of algorithm from the evaluation of proposed spatiotemporal aggregate functions. In the experiments, the query processing time have O(nlogn) of time complexity by the size of query window, and O(n) of time complexity by the number of data.

VI. Conclusion

Recently, many works have been investigating query operations regarding spatiotemporal data analysis which change over time. In particular, one of the query operations, aggregation, is an operation that returns result values through the processing of computation on the data which satisfy some conditions. It is also an important operation that analyzes the used data in various applications. However, previous works of aggregation are attached only to temporal or spatial aggregation. So those are not directly available for the real world data because the characteristic of the real world data has both spatial and temporal conditions.

In order to solve these problems, in this paper, we described the spatiotemporal topological operators which extract spatiotemporal data and database structures representing spatiotemporal objects. On the basis of the database structures and spatiotemporal topological operators, we also proposed the aggregate functions, such as stCOUNT, stSUM, stAVG, stMAX, stMIN, and described algorithms. The proposed aggregate functions are suitable for statistical analysis on temporal and spatial information of various spatiotemporal data which are stored in spatiotemporal database. With such work, we allow more various spatiotemporal query operations to be employed by using combination of spatiotemporal aggregate functions and topological operators.

Besides, we also showed the convenience and improvement of query representation conforming to application systems, and facility of analysis on spatiotemporal applications which previous temporal or spatial aggregate functions can not analyze by applying to the estate management system. Then, we showed the validity of performance of proposed algorithms from the evaluation of spatiotemporal aggregate functions.

In the future we plan to improve aggregate algorithms and index for process of aggregate query. We also will propose extended spatiotemporal aggregate functions which can aggregate analysis in detail.

References

[1] T.K.Sellis, "Research Issues in Spatio-temporal Database System," SSD, pp.5-11, 1999.

[2] J.Gray, "The Benchmark Handbook for Database and Transaction Processing Systems," Morgan Kaumann, 1991.

[3] DongHo Kim, InHong Lee and KeunHo Ryu, "The Design and Implementation of Aggregate Function of Temporal Databases in Main Memory," Journal of the KISS, Vol.21, No.8, Aug, 1994

[4] N.Kline, and R.T.Snodgrass, "Computing Temporal Aggregates," International Conference on Database Engineering, p222-231, 1995.

[5] D. Zhang, A. Markowetz, V. Tsotras, D. Gunopulos and .Seeger, "Efficient Computation of Temporal Aggregates with Range Predicates," in PODS 01, Santa Barbara, CA, 2001.

[6] J.Yang and J.Widom, "Incremental Computation and Maintenance of Temporal Aggregtes," International Conference on Data Engineering, pp.51-60, 2001.

[7] I.Lazaridis and S.Mehrotra. "Progressive approximate aggregate queries with a multi-resolution tree structure," SIGMOD, 2001.

[8] Donhui Zhang and Vassilis J.Tsotras, "Improving Min/Max Aggregation over Spatial Object," ACM-GIS, 2001.

[9] SangHo Kim, Goo Kang and KeunHo Ryu, " Design of Three Dimensional Spatial Topological Relational Operators", Journal of the KISS, Vol.10, No.2, pp.211-220, 2003.

[10] DongHo Kim and KeunHo Ryu, "A Relational Spatiotemporal Database Query Language and Their Operation," Journal of the KISS, Vol.5, No.10, Aug, 1998.