

Reunión Espacio-Temporal en escenarios donde sólo un conjunto de entrada está indexado

Anibal Jorge Diaz

Departamento de Informática

Facultad Tecnológica Nacional Regional Concepción Del Uruguay

Universidad Tecnológica Nacional, Entre Ríos Argentina

anibaljdiaz@yahoo.es

Edilma Olinda Gagliardi

Departamento de Informática

Facultad de Ciencias Físico, Matemáticas y Naturales

Universidad Nacional de San Luis, Argentina

oli@unsl.edu.ar

Gilberto A. Gutiérrez Retamal

Facultad de Ciencias Empresariales

Departamento de Auditoría e Informática

Universidad del Bío-Bío, Chillán, Chile

ggutierr@ubiobio.cl

Resumen.

Los sistemas de bases de datos que administran objetos espaciales y temporales, llamadas bases de datos espacio-temporales (BDET), han recibido creciente interés en los últimos años. Uno de los tipos de consulta espacio-temporal que se caracteriza por analizar la historia espacial de los objetos es la *reunión espacio-temporal (RET)*, en la cual se requiere combinar dos conjuntos de objetos espacio-temporales de acuerdo con algún predicado que involucre tanto atributos espaciales como temporales. Entre las distintas combinaciones que se pueden presentar para reunir dos conjuntos, consideramos en este trabajo el caso donde sólo uno de los conjuntos de entrada se encuentra indexado, con el fin de estudiar las estrategias que nos permitan con el menor costo pasar de este escenario inicial a uno donde contemos con dos conjuntos indexados, y realizar la reunión espacio-temporal de estos conjuntos.

Palabras claves: Bases de datos espacio-temporales (BDET); Reunión espacio-temporal (RET); 3DR-tree, 3DRp-tree (3DR-tree podado).

1. Introducción.

Las BDET requieren administrar datos de objetos cuya posición y/o forma cambian en distintos instantes. Dentro de este dominio, son muchas las aplicaciones que se benefician del

procesamiento eficiente de consultas espacio-temporales, tales como sistemas de comunicación móvil, sistemas de control de tráfico, sistemas de información geográfica asociadas con fenómenos donde la identidad de las componentes geográficas puede cambiar en el tiempo, y aplicaciones multimedia. Un aspecto importante al momento de procesar los datos de entrada es si los mismos están o no indexados. Uno de estos escenarios es aquel donde del par de conjuntos de entrada, uno de ellos se encuentra indexado y el otro conjunto no.

La organización del presente artículo es la siguiente: en la Sección 2, nos referimos a contextos de movimiento de objetos; en la Sección 3, describimos los distintos estados de indexación que pueden presentar los conjuntos de entrada junto con los métodos de indexación desarrollados sobre los cuales implementamos los algoritmos de consulta; en la Sección 4, presentamos los distintos algoritmos de reunión espacio-temporal; y por último en la Sección 5, mostramos el trabajo actual y futuro.

2. Contextos según el movimiento de los objetos

Movilidad es una palabra clave que se tiene en cuenta debido a todas las aplicaciones y servicios que ocurren en la realidad. Estas aplicaciones necesitan de una disponibilidad de información que sólo puede ser provista por una BDET, capaz de almacenar y recuperar los datos de manera eficiente. Para ello se requieren de métodos de indexación que consideren tanto el almacenamiento eficiente, sobre todo de los atributos dinámicos del objeto móvil, más una actualización constante de esos atributos, y de una respuesta adecuada frente a los diversos tipos de consultas.

Para categorizar varios enfoques de indexación se presentan tres contextos diferentes: *movimiento sin restricciones*, *movimiento restringido* y *movimiento restringido a redes*.

El movimiento sin restricciones es aquel en donde los objetos se mueven libremente dentro de la región en la que se encuentran. Ejemplo de este contexto es, entre otros, el monitoreo de animales que luego de permanecer en cautiverio son reubicados en sus hábitats naturales. Existen diferentes métodos de indexación que se diferencian de acuerdo a la información que indexan: métodos que indexan el pasado; métodos que indexan el presente y métodos que indexan el presente y el futuro.

El movimiento restringido es aquel en el que el contexto se presenta limitado para los objetos que se muevan. Esta limitación tiene que ver con obstáculos que obligan al objeto a tomar caminos alternativos ya que éstos no pueden ser sorteados. Un ejemplo de este contexto es el de un robot autónomo, al cual se le indica qué hacer, pero no cómo hacerlo; es decir, planifica sus movimientos en base al conocimiento que tiene de un determinado ambiente y sus obstáculos.

En algunas otras aplicaciones de la realidad, los movimientos quedan restringidos a una red, porque no interesa el espacio subyacente a ella para moverse, es decir, al ver una carretera y un automóvil moviéndose en ella, no interesa el espacio o el tamaño de la banquina adyacente a la misma, como así tampoco la posición del objeto en términos de coordenadas x e y , sino que interesa la posición relativa respecto de la red, es decir, en el kilómetro 24 de la ruta 3, por ejemplo [10]. A menudo, una red dirige el tráfico de los objetos que se mueven en ella. Es decir, al tener una red predefinida, los objetos que se mueven en ella tienden a moverse siguiendo los diferentes caminos que la misma presenta. La red puede ser tanto artificial, creada basándose en un criterio para el movimiento de los objetos, como real, basada en un mundo real.

3. Métodos de indexación

Las consultas de reunión espacio-temporal requieren que se combinen dos conjuntos de objetos espacio-temporales de acuerdo con algún predicado que involucre tanto atributos espaciales como temporales. Entre las distintas combinaciones que se pueden presentar para reunir dos conjuntos, se presentan los escenarios donde ambos conjuntos pueden estar indexados, o donde sólo uno puede estar indexado, o escenarios donde ninguno de los conjuntos de entrada están indexados. Consideraremos en este trabajo el caso donde sólo uno de los conjuntos de entrada, de objetos espacio-temporales con movimiento sin restricción, se encuentra indexado.

Los métodos de acceso tradicionales no soportan datos espacio-temporales; en consecuencia, se han desarrollado diferentes propuestas para mantener simultáneamente información sobre tiempo y espacio, las cuales indexan objetos que se mueven en un espacio bidimensional. La mayoría de estos métodos extienden los métodos de acceso espacial para incluir componentes temporales. Estos métodos se pueden clasificar basándose en el tipo de dato, acerca de objetos en movimiento, con el que tratan. Algunos métodos se enfocan en la recuperación de datos históricos y otros en la recuperación de posiciones futuras de objetos en movimiento basadas en la posición actual y en patrones de movimiento [1,8]. En este trabajo nos enfocamos en la historia de los objetos con movimiento sin restricciones y analizamos el comportamiento de las estructuras MVR-tree, 3DR-tree y 3DRp-tree que aplicamos en nuestra investigación.

El índice espacio-temporal 3DR-tree [12], es un método de acceso espacio-temporal que considera el tiempo como una dimensión adicional a las coordenadas espaciales. De esta manera cuando un objeto, en un espacio bidimensional cambia su posición o forma geométrica, se inserta un rectángulo mínimo envolvente (MBR), formado por dos puntos en tres dimensiones. Este MBR es una aproximación geométrica de un objeto real el cual es capaz de contener los atributos espaciales del objeto.

La estructura 3DR-tree es eficiente en el uso del almacenamiento y en el procesamiento de consultas tipo time-interval. Sin embargo, una de sus principales desventajas es su ineficiencia para consultas de tipo time-slice debido a la gran altura que pueden alcanzar los MBR de objetos con poca movilidad [11]. Esta deficiencia se reduce considerablemente gracias a la técnica de particionado que se aplica a esta estructura obteniéndose estructuras 3DR-tree de menor altura [7], llamadas 3DRp-Tree (3DR-Tree podado) [3].

El índice espacio-temporal surge al partir la estructura 3DR-tree de indexación de un conjunto de datos, la cual tiene por objeto reducir la altura del árbol. Esta idea fue propuesta por [7] llamada APR-tree y obtenida en función de la carga de trabajo. En esta estructura el intervalo temporal que administra cada 3DR-tree particionado depende de la carga de trabajo de la estructura 3DR-tree. Esta estructura no comparte nodos, pero el particionado del árbol principal 3DR-tree genera duplicados en los 3DRp-tree, lo cual debe tenerse en cuenta en el algoritmo de búsqueda a través de la estructura.

La estructura Multiversión R-tree [11] es una estructura formada por múltiples árboles R-tree, la cual se aplica para responder consultas tipo timeslice e intervalos cortos. Optamos por usar esta estructura porque permite trabajar con datos históricos y porque su rendimiento se mantiene al aumentar el número de objetos indexados [5].

4. Algoritmos de reunión espacio-temporal

Administrar datos espaciales que cambian su posición en el tiempo requiere de un procesamiento de datos espacio-temporales efectivo y eficiente y ha dado lugar a un amplio campo de investigación en el tema. En este sentido, nuestra línea de trabajo pretende contribuir a la comprensión de la reunión espacio-temporal introduciendo y evaluar estrategias para el procesamiento de consultas de este tipo. Debido a la componente de tiempo, las BDET necesitan administrar grandes volúmenes de datos acumulados a través de un gran período de tiempo. Una consulta sobre estos datos puede ser resuelta por medio de una búsqueda exhaustiva, es decir, accediendo a todos los objetos de la base y retornando aquellos que pertenecen a la respuesta. Esta aproximación es ineficiente debido al tamaño de estas bases de datos. Una solución más apropiada consiste en construir índices sobre los datos de manera tal que las consultas puedan responderse accediendo solamente a una pequeña parte de la base de datos.

La recuperación de datos y la forma en que ésta se realiza en las bases de datos espacio-temporales es nuestra principal área de interés. Un objeto espacio-temporal es una abstracción de una entidad que tiene un identificador, o clave primaria, una localización, una forma espacial y una propiedad temporal junto con alguna otra característica que lo describe. En consecuencia, es posible una indexación de este tipo de objetos.

Una de las primeras estimaciones de la consulta de reunión espacio-temporal se aplicó usando el método de acceso espacio-temporal 3DR-tree, y su implementación es prácticamente la misma que la propuesta en [2].

Otra de las estimaciones para la reunión espacio-temporal es la propuesta en [3] donde se divide la estructura 3DR-tree de los conjuntos de entrada, en varios árboles 3DR-tree podados, con el objeto de reducir su altura y por lo tanto reducir el costo de búsqueda.

El otro enfoque que encaramos es a través de las estructuras multiversión *MVR-tree* la cual es una suma de R-tree en el que cada árbol mantiene las posiciones de los objetos que evolucionaron en un intervalo de tiempo pasado. De tal manera que, si dada una consulta de reunión espacio temporal con una condición espacial y temporal prefijada, con uno de los conjuntos de datos indexados por estructuras *MVR-tree*, podemos aplicar estrategias de indexación del conjunto de entrada no codificado y luego en la reunión propiamente dicha aplicar técnicas de recorrido de árbol que nos permitan obtener una respuesta tan o más eficiente que las conocidas para responder esta consulta. La propuesta es aplicar la estrategia de filtrar los intervalos de tiempo de las estructura que se intersectan con la condición temporal dada en sendos arreglos; posteriormente, continuar con una política de recorrida de árbol en profundidad, procesando en cada nivel todos los objetos que pertenecen al intervalo temporal y que se intersectan espacialmente a efectos de filtrar los mismos.

Los únicos algoritmos comparables de reunión espacio-temporal los encontramos en los trabajos de [2,4] y por lo tanto no existen otros para determinar mejores diseños para resolver la RET en diversas aplicaciones.

5. Trabajo Futuro.

Como trabajo futuro se planifica trabajar explorando esta consulta aplicando el método de acceso espacio-temporal *MV3R-tree* para casos en que uno de los conjuntos este indexado y el otro no. También, investigar los casos que se presentan en escenarios donde los conjuntos de entrada no vienen indexados. Estos trabajos tienen como objetivo desarrollar

algoritmos que respondan eficientemente la consulta de reunión espacio-temporal y realizar experimentos que nos permitan comparar su efectividad con los métodos existentes.

6. Bibliografía consultada.

- [1] Agarwal, P., Arge, L. & Erickson, J. (Indexing moving points. Symposium on Principles of Database Systems (pp. 175-186) ACM Press. (2000).
- [2] Brinkhoff, T., Kriegel, H. And B. Seeger. "Efficient Processing of Spatial Join Using R-Trees". Proc. ACM SIGMOD, 1993.
- [3] Diaz, A., Gutierrez, G., Gagliardi, O. "Algoritmo de reunión espacio-temporal usando estructura 3DR-Tree podada". CACIC2007. Octubre 2007-Corrientes, Argentina.
- [4] Gutierrez, Gilberto. Tesis Doctoral : "Métodos de Acceso y Procesamiento de Consultas Espacio-Temporales". Universidad de Chile. Santiago. Abril de 2007.
- [5] Hadjieleftheriou, M., Kollios, G., Tsotras, V., Gunopoulos, D. "Indexing Spatio-Temporal archives". The VLDB Journal, Vol. 15, Issue 2, June 2006.
- [6] Hadjieleftheriou, M., Kollios, G., Tsotras, V., Gunopoulos, D. "Efficient Indexing of Spatio-Temporal objects". In Extending Database Technology (pp. 251-268). 2002
- [7] Hyung-ju Cho, Jun-Ki Min, Chin-Wan Chung. "An Adaptive Indexing Technique Using Spatio-Temporal Query Workloads". Korean Advanced Institute of Science. Information and Software Technology. Vol 46, pp. 229-241. March 2004.
- [8] Kollios, G., Gunopoulos, D, Tsotras, V. "On Indexing Mobile Objects". ACM Symposium on Principles of Database Systems (pp. 261-272). ACM Press 1999.
- [9] Mokbel, M., Ghanem, T., Aref, W. "Spatio-temporal Acces Methods". IEEE Data Engineering Bulleti (pp. 40-49) 2003.
- [10] Psofer, D. "Indexing the Trajectories of Moving Objects". Bulletin of the IEEE Computer Society Technical Commitee on Data Engineering. (2002)
- [11] Tao, Yufei and Papadias, Dimitris. "The MV3R-Tree: A Spatio-Temporal Access Method for Timestamp and Interval Queries". Proceedings of the 27th VLDB Conference, Roma, Italy, 2001
- [12] Varzirgiannis, M., Theodoridis, Y., Sellis, T. "Spatio-Temporal Composition and Indexing for Large Multimedia Applications". ACM Multimedia Systems. (1998).