Una Extensión del FHQT Temporal para Distancias Continuas

Andrés Pascal[1], Anabella De Battista[1], Norma Edith Herrera[2], Gilberto Gutierrez[3]

[1] Dpto. de Sistemas de Información, Universidad Tecnológica Nacional, Entre Ríos, Argentina,
{pascala, debattistaa }@frcu.utn.edu.ar
[2] Dpto. de Informática, Universidad Nacional de San Luis, Argentina,
nherrera@unsl.edu.ar
[3] Universidad del Bio Bio, Facultad de Ciencias Empresariales, Chillán, Chile,
ggutierr@ubiobio.cl

Resumen El modelo de bases de datos métrico-temporal permite abordar aquellas situaciones en las que resulta necesario realizar búsquedas por similitud teniendo en cuenta también la componente temporal. En este artículo presentamos una mejora al índice métrico-temporal *FHQT-Temporal*, que soporta valores continuos de la función de distancia, manteniendo la eficiencia ante cambios de valores del radio de búsqueda e incrementos de los intervalos de tiempo. Además se muestran resultados de la verificación experimental de esta estructura para un conjunto de datos determinado.
Palabras Claves: Espacios Métricos, Bases de Datos Temporales, Índices, Bases de Datos Métrico-Temporales.

1. Introducción

Las bases de datos clásicas se organizan bajo el concepto de búsqueda exacta sobre datos estructurados. Esto significa que la información se organiza en registros los cuales se dividen en campos que contienen valores completamente comparables. Una consulta retorna todos aquellos registros cuyos campos coinciden con los aportados en la consulta (búsqueda exacta). Una característica importante de las bases de datos clásicas es que capturan sólo un estado de la realidad modelizada, usualmente el más reciente. Por medio de las transacciones, la base de datos evoluciona de un estado al siguiente descartando el estado previo.

Actualmente las bases de datos han incluido la capacidad de almacenar datos no estructurados tales como imágenes, sonido, texto, video, datos geométricos, etc. La problemática de almacenamiento y búsqueda en estos tipos de base de datos difiere de las bases de datos clásicas en varios aspectos: los datos no son estructurados por lo que no es posible organizarlos en registros y campos; la búsqueda exacta carece de interés; resulta de interés mantener todos los estados de la base de datos y no sólo el más reciente para poder consultar el intervalo de tiempo de vigencia de los objetos. Es en este contexto donde surgen nuevos modelos de bases de datos.

El modelo de ***espacios métricos*** [3], permite trabajar con objetos no estructurados y realizar búsquedas por similitud sobre los mismos. Un espacio métrico es un par (U, d) donde U es un universo de objetos y $d : U \times U \rightarrow R^+$ es una función de distancia

definida entre los elementos de U que mide la similitud entre ellos. Una de las consultas típicas en este modelo es la búsqueda por rango, denotado por $(q, r)_d$, que consiste en recuperar los objetos de la base de datos que se encuentren como máximo a distancia r de un elemento q dado.

El modelo de ***bases de datos temporales*** [7] incorpora al tiempo como una dimensión, por lo que permite asociar tiempos a los datos almacenados y consultar por los objetos vigentes en un intervalo o en un instante de tiempo dado.

Existen aplicaciones donde resulta de interés realizar búsquedas por similitud teniendo en cuenta también la componente temporal. Es en este ámbito donde surge el ***modelo métrico-temporal***. En este modelo se puede trabajar con objetos no estructurados con tiempos de vigencia asociados y realizar consultas por similitud y por tiempo en forma simultánea.

Varios índices han sido diseñados para resolver consultas métrico-temporales [1,2,5,6]. Todos ellos están basados en el *Fixed Height Queries Tree*, un índice para espacios métricos, y asumen que la función de distancia d devuelve valores discretos. En este trabajo presentamos una extensión del *Fixed Height Queries Tree Temporal(FHQT Temporal)* para distancias continuas.

Este artículo está organizado de la siguiente manera. En la Sección 2 se expone el trabajo relacionado definiendo los conceptos necesarios para la comprensión de este artículo. En la Sección 3 presentamos nuestro aporte definiendo la extensión del *FHQT-Temporal* para distancias continuas. En la Sección 4 presentamos la evaluación experimental y finalizamos en la Sección 5 dando las conclusiones y trabajo futuro.

2. Trabajo Relacionado

2.1. El Modelo Métrico-Temporal

El modelo *métrico-temporal* está orientado a satisfacer búsquedas sobre objetos no estructurados que poseen uno (una sola dimesión temporal) o dos (bitemporal) instantes o intervalos de tiempo asociados y que además no pueden ser recuperados a través de un atributo clave por medio de una búsqueda exacta. Sea U un universo de objetos válidos, se define un Espacio Métrico-Temporal mediante el par (X, d), donde $X = U \times N \times N \times N \times N$ y d es la función de distancia $d : U \times U \rightarrow R^+$. Cada elemento $x \in X$ es una 5-upla $(o, t_{vi}, t_{vf}, t_{ti}, t_{tf})$, donde o es un objeto (una huella digital, una imagen, un sonido, etc), $[t_{vi}, t_{vf}]$ es el intervalo de validez de o en la realidad y $[t_{ti}, t_{tf}]$ el intervalo de tiempo transaccional asociado. Por simplicidad, se definen todos los tiempos como valores pertenecientes al conjunto N. Estos valores pueden ser fechas, horas, etc., pero en cualquier caso se pueden representar mediante números naturales. La función de distancia d mide la disimilitud entre dos objetos y cumple con las propiedades de toda métrica, es decir, positividad, simetría, reflexividad y desigualdad triangular [3].

Formalmente una *consulta métrico-temporal* se define como una 4-upla $(q, r, t_{iq}, t_{fq})_d$, tal que:

$$(q, r, t_{iq}, t_{fq})_d = \{o/(o, t_{io}, t_{fo}) \in X \wedge d(q, o) \leq r \wedge (t_{io} \leq t_{fq}) \wedge (t_{iq} \leq t_{fo})\}$$

Una forma trivial de resolver una consulta métrico-temporal, sin realizar una búsqueda exhaustiva en la bases de datos, es construir un índice métrico agregándole a cada

objeto un intervalo temporal que represente la vigencia del mismo. Luego, ante una consulta $(q, r, t_{iq}, t_{fq})_d$ en primer lugar se utilizará el índice métrico para descartar los objetos obj que están a distancia mayor que r de q; y posteriormente se realizará un recorrido del conjunto de elementos no descartados en el primer paso para determinar qué objetos conforman la respuesta a la consulta, que serán aquellos cuyo intervalo de vigencia que se superpone con $[t_{iq}, t_{fq}]$.

Varios índices métrico-temporales se han propuesto en este ámbito: el *Pivot-FHQT* [1], el *Historical-FHQT* [2], el *Event-FHQT* [5] y el *FHQT-Temporal* [6]; todos ellos han tomado como base el Fixed Height Queries Tree[3], un índice para espacios métricos. Todos estos índices asumen que la función de distancia retorna valores discretos.

2.2. FHQT-Temporal

El *FHQT-Temporal* es un FHQT al cual se le agrega un intervalo de tiempo en cada nodo del árbol. Este intervalo representa el período de tiempo de vigencia de todos los objetos del subárbol cuya raíz es dicho nodo. En cada nodo hoja, este intervalo es el período total de vigencia de los objetos que contiene. Para un nodo interior, el intervalo se calcula tomando el tiempo inicial mínimo y el tiempo final máximo de sus hijos.

Este índice permite resolver consultas por similitud puras, temporales puras y métrico-temporales. Como se basa en un FHQT, solo permite funciones de distancia discretas.

Formalmente, un *FHQT-Temporal* es un árbol en el cual un nodo interior V es una 3-upla (t_{ini}, t_{fin}, $\{(d_1, h_1), (d_2, h_2),..., (d_m, h_m)\}$) donde:

- $h_1..h_m$ son los m hijos del nodo V,
- las d_i , para $i = 1..m$, son las distancias entre el pivote correspondiente al nivel de V y los objetos contenidos en las hojas de los subárboles de h_i.
- los dos primeros componentes de la 3-upla, t_{ini} y t_{fin}, se definen de la siguiente manera: t_{ini} = $min_{j=1..m}(t_{ini}(h_j))$, y t_{fin} = $max_{j=1..m}$ $(t_{fin}(h_j))$.

Las hojas del *FHQT-Temporal* tienen una estructura similar; están representadas por una 3-upla (t_{ini}, t_{fin}, $\{e_1, e_2,..., e_l\}$), donde:

- los e_i para $i = 1..l$ son los l elementos que contiene la hoja, que a su vez están formados por tres componentes: el objeto o, el tiempo inicial del mismo t_{io}, y su tiempo final t_{fo}.
- los valores t_{ini}, t_{fin} poseen el mismo significado que para los nodos interiores, pero aplicados a los elementos e_i.

En la Figura 1 se muestra el esquema genérico del *FHQT-Temporal*. La estructura es dinámica, permitiendo tanto altas como bajas ya sea de instantes o intervalos contenidos en el intervalo que el índice posee hasta el momento, como de objetos con tiempos fuera de éste.

Cuando se realiza una consulta métrico-temporal, se procede de la siguiente manera: en cada nivel del árbol se seleccionan los subárboles hijos del nodo que se está procesando, cuyos intervalos temporales se intersectan con el intervalo o instante de la consulta. De éstos, posteriormente se eligen los que cumplen con la restricción de similitud tomando en cuenta la firma de la consulta y el radio de búsqueda. Este procedimiento

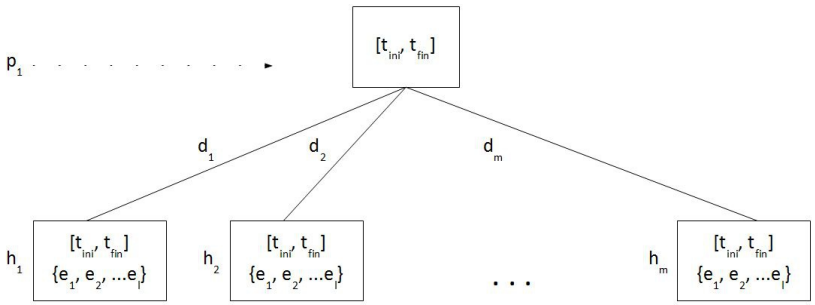

Figura 1. *Esquema del FHQT-Temporal*

se repite hasta llegar al último nivel. Para cada hoja no descartada, luego de verificar la superposición temporal, se realiza una búsqueda secuencial sobre todos los elementos contenidos en las mismas, comparando tanto el aspecto temporal como la distancia de cada elemento a la consulta.

3. Extensión del FHQT-Temporal para Distancias Continuas

El *FHQT*$^+$*-Temporal* es una variante del *FHQT-Temporal* generalizada que soporta valores continuos de la función de distancia. Para ello, en lugar de asociar un número natural a cada hijo de un nodo, se asocian dos intervalos de valores de distancias. El primer intervalo representa el rango máximo correspondiente a la rama, mientras que el segundo (incluido en el anterior) constituye el rango actual de valores, es decir, el intervalo formado por el mínino y el máximo valor de distancia del pivote a los objetos contenidos en las hojas del subárbol. Los intervalos máximos son constantes y se calculan cuando se construye el árbol en base al histograma de distancias, de tal manera de que el árbol tenga una alta probabilidad de quedar balanceado. Los intervalos actuales son variables y se van actualizando de acuerdo a los objetos que se insertan en la estructura. Para determinar en qué rama se agrega un nuevo elemento, se utilizan los intervalos máximos y ante una consulta sólo se usan los actuales. Al utilizar estos últimos intervalos en las búsquedas, se incrementa la capacidad de filtrado por similitud, ya que los intervalos son más pequeños y quedan espacios vacíos entre intervalos consecutivos, como veremos más adelante.

Un *FHQT*$^+$*-Temporal* es un árbol r-ario donde el valor de r es un parámetro que se define en forma previa a su construcción, normalmente en base a la distribución del histograma de distancias. Formalmente, es un árbol donde cada nodo interior V es una 3-upla (t_{ini}, t_{fin}, $\{(int_{x1}, int_{a1}, h_1), (int_{x2}, int_{a2}, h_2),..., (int_{xm}, int_{am}, h_m)\}$) donde:

- $h_1..h_m$ son los m hijos del nodo V,
- los int_{xi} , para $i = 1..m$, son los intervalos *máximos* de distancias entre el pivote correspondiente al nivel de V y los objetos que pueden pertenecer a las hojas de los subárboles de h_i.

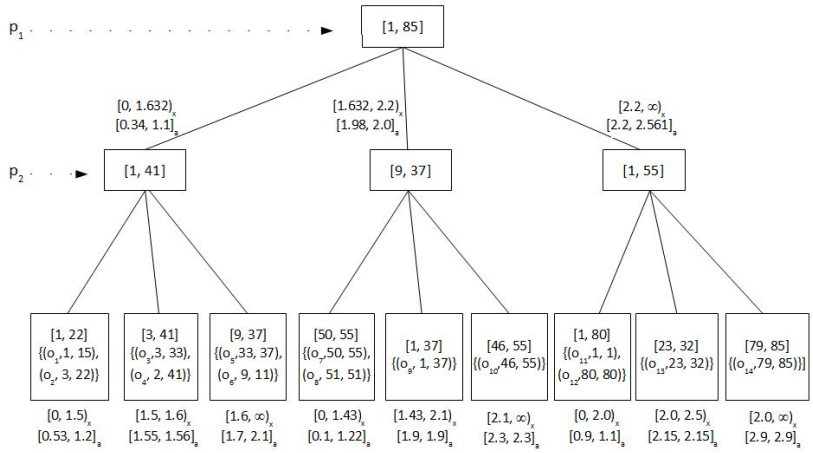

Figura 2. *Ejemplo de un FHQT$^+$-Temporal*

- los int_{ai} , para $i = 1..m$, son los intervalos *actuales* de distancias entre el pivote correspondiente al nivel de V y los objetos contenidos actualmente en las hojas de los subárboles de h_i.
- los dos primeros componentes de la 3-upla, t_{ini} y t_{fin}, se definen de la siguiente manera: t_{ini} = $min_{j=1..m}(t_{ini}(h_j))$, y t_{fin} = $max_{j=1..m}\ (t_{fin}(h_j))$.

Las hojas del *FHQT$^+$-Temporal* se definen de la misma forma que para el *FHQT-Temporal*

Para calcular los intervalos máximos correspondientes al nodo raíz del árbol, se toma una muestra de la base de datos, se calcula el histograma de distancias y se divide el espacio en r intervalos, de tal manera de que cada uno de ellos posea la misma cantidad (± 1) de elementos. Luego para cada nodo hijo se procede de la misma manera, pero considerando solo los elementos de la muestra que fueron asignados a dicho nodo. De esta manera todos los nodos interiores tendrán exactamente r hijos.

Una vez definidos los rangos, se realiza la inserción de los elementos, actualizando los intervalos actuales. Sea o el objeto a insertar, v el nodo donde se quiere insertar el objeto, $[d_{xi}, d_{xf})$ y $[d_{ai}, d_{af}]$ los intervalos máximo y actual asociados al nodo, y p el pivote del nivel, primero se verifica que $d(p, o) \in [d_{xi}, d_{xf})$ y si esto se cumple, se actualiza el intervalo actual haciendo $d_{ai} := min(d_{ai}, d(p, o))$ y $d_{af} := max(d_{af}, d(p, o))$. El aspecto temporal se procesa de la misma manera que en el *FHQT-Temporal*.

Ante una consulta, para visitar un nodo se comprueba que la distancia de la consulta al pivote del nivel pertenezca al intervalo actual asociado al nodo mas menos el radio de búsqueda, es decir, que $f_n \in [d_{ai} - r, d_{af} + r]$. En la Figura 2 se muestra un ejemplo del *FHQT$^+$-Temporal*. Es interesante notar que los intervalos actuales correspondientes a los nodos de un mismo nivel, usualmente no cubren todo el espacio posible. Al reducir el tamaño de estos rangos, aumenta la probabilidad de que una rama se descarte ya que la comprobación anterior se realiza sobre un intervalo más pequeño. Por ejemplo, si el radio de búsqueda es $0{,}3$ y la distancia entre la consulta y el pivote del primer nivel es

1,5, si se utilizan los rangos máximos se deben procesar el primer y segundo hijo del nodo raíz, ya que $1{,}5 \in [0-0{,}3\ ,\ 1{,}632+0{,}3)$ y $1{,}5 \in [1{,}632-0{,}3\ ,\ 2{,}2+0{,}3)$, mientras que usando los rangos actuales ambos se descartan porque $1{,}5 \notin [0-0{,}34\ ,\ 1{,}1+0{,}3]$ y $1{,}5 \notin [1{,}98-0{,}3\ ,\ 2{,}0+0{,}3]$.

Este índice permite resolver consultas por similitud puras, temporales puras y métrico-temporales con funciones de distancia tanto continuas como discretas.

4. Evaluación Experimental

Debido a que el modelo métrico-temporal es relativamente reciente, no existen aún bases de datos disponibles para realizar experimentos, por lo que se optó por adaptar la base $NASA$ [4] que es frecuentemente utilizada por investigadores del área de espacios métricos (disponible para su descarga en *http://www.sisap.org/library/dbs/vectors/*), para la determinación experimental de la eficiencia de esta nueva estructura ante consultas métrico-temporales.

La base de datos $NASA$ es un conjunto de 40.150 vectores 20-dimensionales de números reales, que representan características de imágenes obtenidas por la $NASA$.

Partiendo de este conjunto de datos se generó la base de datos métrico-temporal *NASA*^{MT} asignando a cada vector un identificador y un intervalo de vigencia, que indica el período de validez del objeto. El intervalo total considerado fue [1, 1000]. Luego, mediante un proceso aleatorio se generaron lotes de 1.000, 5.000, 10.000, 20.000 y 30.000 elementos.

Una vez construido el índice, se seleccionaron al azar 100 objetos de la base de datos *NASA*^{MT} y se generaron cuatro lotes de consultas métrico-temporales mediante la asignación en forma aleatoria de intervalos/instantes de tiempo. Uno de los lotes para cada base de datos se compuso solamente de consultas instantáneas y los demás fueron construídos asociándoles intervalos correspondientes al 10 %, 25 % y 50 % del intervalo total ($[1, 1000]$).

Para completar los parámetros requeridos en las consultas, se definieron tres radios de búsquedas distintos para cada base de datos, que devuelven en promedio aproximadamente el 1 %, 5 % y 10 % de los objetos contenidos ante las consultas por similitud de los lotes definidos anteriormente. Estos radios fueron calculados experimentalmente y son los siguientes: 0,453; 0,69855 y 0,8275.

Como función de distancia se utilizó la distancia euclidiana que es la medida utilizada usualmente sobre esta bases de datos para realizar pruebas por similitud. En estas pruebas sólo se tomó en cuenta como variable de costo la cantidad de evaluaciones de la función de distancia ya que la estructura se mantuvo en memoria principal.

4.1. FHQT$^+$-Temporal: Análisis de los Resultados Obtenidos

En esta sección se presentan, grafican y analizan los resultados obtenidos para el *FHQT*$^+$*-Temporal* en comparación con la solución trivial que utiliza un FHQT como índice métrico donde cada objeto tiene su intervalo de vigencia asociado. En la solución triial, primero se busca por similitud y para cada objeto resultante de esta búsqueda, se usa su intevalo de vigencia asociado para determinar si forma o no parte de la respuesta.

Cabe notar que cualquiera de dichas soluciones tiene, como mínimo, el costo correspondiente a un índice métrico.

Variación del Costo en Función del Tamaño de la Base de Datos En los gráficos de la Figura 3 se muestran las curvas de costos correspondientes a los lotes de consultas instantáneas y 50 % del intervalo respectivamente, en comparación con la solución trivial. El eje x indica la cantidad de elementos de la base de datos NASA^{MT} y el eje y, el promedio de evaluaciones de la función de distancia.

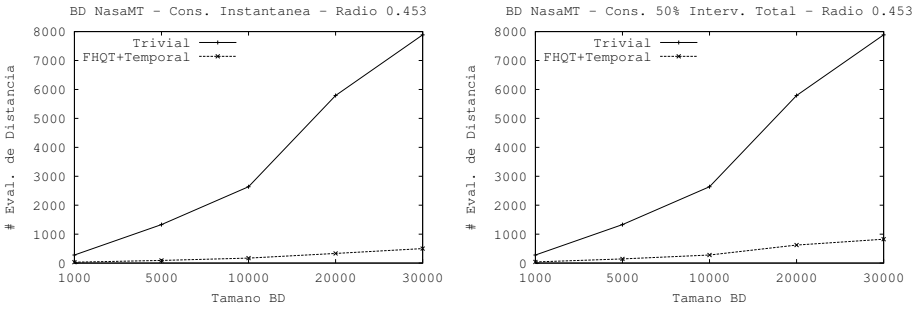

Figura 3. *Costo consultas métrico-temporales mediante un FHQT$^+$-Temporal a la base de datos NASA^{MT}, en función del tamaño)*

Las curvas muestran en ambos casos que ante el aumento de la cantidad de elementos del conjunto de datos consultado, el costo se incrementa de forma similar a una curva lineal aunque los factores por los cuales se multiplica son muy distintos. Claramente se ve que el *FHQT$^+$-Temporal* supera ampliamente la performance de la solución trivial. En el mejor de los casos –cuando el tamaño es el mayor y las consultas son instantáneas–, su costo es de sólo el 7,89 % del correspondiente a la solución trivial, y en el peor de los casos, un 12,6 %. Es importante notar que el costo de la solución trivial no varía en función del tamaño del intervalo de tiempo, por lo cual en ambos gráficos la curva es la misma.

El *porcentaje de evaluaciones*, medido como la cantidad de objetos resultantes sobre el costo promedio de las consultas, fue del $0{,}1$ % al $1{,}0$ % para la solución trivial, es decir que por cada elemento resultante tuvieron que ser evaluados 1000 objetos en el peor de los casos y 100 en el mejor caso. Para el *FHQT$^+$-Temporal* este porcentaje se eleva a $0{,}7$ % y $9{,}4$ % respectivamente, correspondientes a 143 y 11 evaluaciones de la función de distancia por cada objeto resultante. En ambos casos, las mejoras no son producidas por la variación de la cantidad de elementos, sino por la modificación del radio e intervalo de búsqueda.

El *FHQT$^+$-Temporal* supera claramente en eficiencia a la solución trivial en todos los casos, por lo cual en los apartados siguientes sólo se analizan las variaciones de este índice respecto a los distintos parámetros.

Variación del Costo en Función del Radio de Búsqueda En la Figura 4 se presentan dos gráficas de costo del *FHQT*$^+$*-Temporal* en función del radio de búsqueda. La primera corresponde a consultas instantáneas y la segunda a intervalos de tiempo con tamaño promedio igual al 50 % del total.

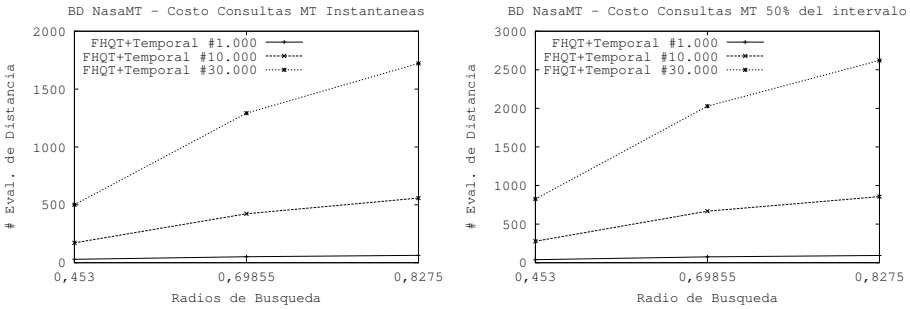

Figura 4. *Costo consultas métrico-temporales mediante un FHQT*$^+$*-Temporal a la base de datos NASA*^{MT}*, en función del radio*

Como es lógico, la cantidad de evaluaciones de la función de distancia se incrementa considerablemente al aumentar el radio de búsqueda. Sin embargo, el porcentaje de evaluaciones mejora sustancialmente. Este porcentaje varía entre $0{,}7$ y $0{,}9$ para el radio menor, y es alrededor de 9 veces más grande para el mayor radio. Esto significa que la eficiencia del *FHQT*$^+$*-Temporal* aumenta cuando se incrementa el radio de búsqueda. En todo índice métrico, es natural que el porcentaje de evaluaciones en algún momento aumente al consultar con mayores radios debido a que la cantidad de resultados también es mayor. En el caso extremo, cuando se consulta con un radio que incluye a todos los elementos de la base de datos, este porcentaje es cercano a 100. Sin embargo, en una base métrico-temporal, si sólo se aumenta el radio y el intervalo de tiempo de la consulta permanece constante, puede ser que la cantidad de resultados sea la misma ya que no todos los elementos cumplirán la restricción temporal.

En el *FHQT*$^+$*-Temporal*, cuando se aumenta el radio de búsqueda las restricciones temporales se hacen más importantes para el proceso de descarte de elementos, es decir que la cantidad de elementos que cumplen con la condición de similitud es mayor, pero la cantidad de elementos que cumplen con la restricción temporal se mantiene igual. Esta es una de las causas del aumento del porcentaje de evaluaciones.

Variación del Costo en Función de la Amplitud del Intervalo de Tiempo Para evaluar la influencia de las variaciones de la amplitud del intervalo temporal sobre el costo de las consultas se presentan los gráficos de la Figura 5. En el eje X se ubican en primer lugar las consultas instantáneas y a continuación los intervalos temporales correspondientes al $10\,\%$, $25\,\%$ y $50\,\%$ del intervalo total. Los valores del eje Y representan las cantidades promedio de evaluaciones de la función de distancia para los lotes de 100

consultas. El radio de consulta correspondiente al primer gráfico es $0{,}453$ y el del segundo $0{,}8275$.

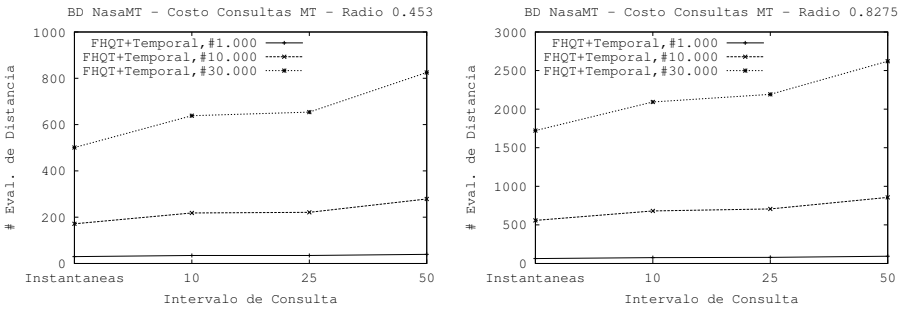

Figura 5. *Costo consultas métrico-temporales mediante un FHQT$^+$-Temporal a la base de datos NASA^{MT}, en función del intervalo temporal*

Como se ve en ambos gráficos, el costo de las consultas que toman el 50 % del intervalo total es entre un 50 % y un 70 % mayor que el de las consultas instantáneas. Por otro lado, la cantidad de elementos resultantes es alrededor de dos veces mayor (lo que es coherente con el aumento del intervalo temporal, ya que los objetos se encuentran uniformemente distribuidos en cuanto al tiempo). Por esta razón el porcentaje de evaluaciones aumenta también, entre un 30 % y un 40 %.

5. Conclusiones y Trabajo Futuro

En este trabajo presentamos una extensión de un método de acceso métrico-temporal que soporta valores continuos de la función de distancia, el *Fixed Height Queries Tree$^+$-Temporal(FHQT$^+$-Temporal)*, que permite resolver eficientemente consultas métrico-temporales. Este índice permite resolver consultas por similitud puras, temporales puras y métrico-temporales con funciones de distancia tanto continuas como discretas. Los experimentos han mostrado que el FHQT$^+$-Temporal tiene en la totalidad de los casos un costo menor que la solución trivial, y son notables las mejoras respecto al porcentaje de evaluaciones del índice cuando se incrementan los radios y los intervalos de consulta.

Actualmente estamos trabajando en la extensión de otros índices métrico-temporales para distancias continuas.

Referencias

1. A. De Battista, A. Pascal, N. Herrera, and G. Gutierrez. Metric-temporal access methods. *Journal of Computer Science & Technology,*, 10(2):54–60, 2010.

2. De Battista, A. Pascal, G. Gutierrez, and N. Herrera. Un nuevo índice métrico-temporal: el historical fhqt. In *Actas del XIII Congreso Argentino de Ciencias de la Computación*, Corrientes, Agentina, 2007.
3. Edgar Chávez, Gonzalo Navarro, Ricardo Baeza-Yates, and José Luis Marroquín. Searching in metric spaces. *ACM Comput. Surv.*, 33(3):273–321, 2001.
4. K. Figueroa, G. Navarro, and E. Chávez. Metric spaces library, 2007. Available at `http://www.sisap.org/Metric_Space_Library.html`.
5. A. Pascal, A. De Battista, G. Gutierrez, and N. Herrera. Indice metrico-temporal event-fhqt. In *Actas del XIIII Congreso Argentino de Ciencias de la Computación*, La Rioja, Argentina, 2008.
6. A. Pascal, De Battista, G. Gutierrez, and N. Herrera. Procesamiento de consultas métrico-temporales. In *XXIII Conferencia Latinoamericana de Informática*, pages 133–144, Costa Rica, 2007.
7. B. Salzberg and V. J. Tsotras. A comparison of access methods for temporal data. *ACM Computing Surveys*, 31(2), 1999.