

Capítulo 1

Introducción

1.1 Ambiente móvil

Nuestro trabajo de tesis se centra en la evaluación de consultas en ambiente móviles. Un ambiente móvil consiste en un conjunto de productores y consumidores de datos que cambian de posición geográfica a lo largo del tiempo e interactúan entre sí a través de una infraestructura de comunicación, por ejemplo una red inter-vehicular [CDI+08].

Los principales aspectos que se consideran para la evaluación de consultas en un ambiente móvil son: el cambio de posición geográfica de productores y consumidores y la producción y consumo de datos que puede ser discreta (“*one shot*”) o continua. Tener una producción continua de datos implica que dada una consulta se obtiene un flujo de datos. Por otra parte, el hecho de tener una producción discreta de datos quiere decir que ante una consulta sólo se va a tener la respuesta. Finalmente la movilidad de productores y consumidores implica que las respuestas obtenidas tienen cierta validez espacio-temporal.

Consideremos un escenario donde los productores y consumidores son los vehículos que se comunican a través de una red-inalámbrica *ad-hoc* de corto alcance (con estándares como IEEE 802.11 o *Ultra Wide Banda UWB*) [CDI+08]. La producción discreta de datos se hace por servicios que proveen información sobre ubicación de hoteles o gasolineras y sobre eventos como accidentes viales. Por otro lado, los datos continuos representan variables que se observan al monitorear el medio ambiente, por ejemplo el tráfico de la ciudad. Como se puede observar en la figura 1.1, la movilidad ocasiona que la validez y pertinencia de los datos producidos dependan de la localización de productores y consumidores. Por ejemplo, el valor de la variable estado del tráfico en un tiempo t_1 no es el mismo que en un tiempo t_2 debido a que ya hubo un cambio de posición geográfica.

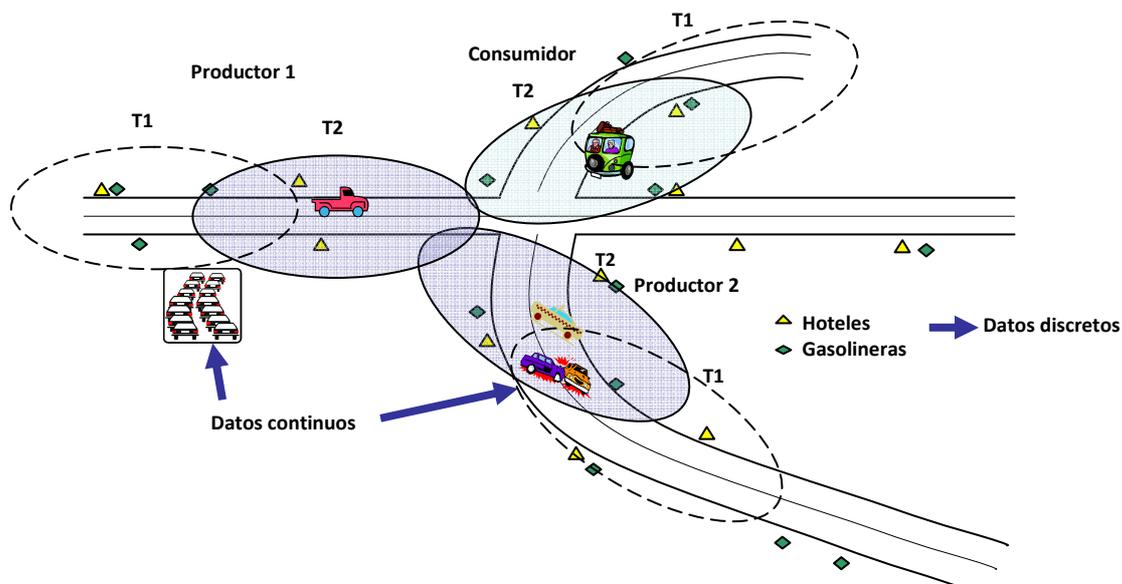


Figura 1.1 Ambiente móvil.

Se podrían resolver consultas como Q1: ¿Cuáles son las áreas de descanso cercanas a mi ubicación dentro de dos horas y que tengan gasolinera y hotel con cuartos disponibles para dos personas y que puedan ser reservados por Internet?

En un ambiente móvil no existe una base de datos única que concentre los datos y se ponga a disposición de los usuarios. Pero si existen un conjunto de productores móviles que generan datos (discretos y continuos) a cerca de la región

que van recorriendo. Por lo tanto el consumo de datos en un ambiente móvil se realiza sobre estos productores que varía en una región en un intervalo de tiempo. Así, la finalidad del consumo de datos es tener una vista sobre el estado del ambiente en ciertos momentos y regiones. El estado del ambiente es construido a partir del conocimiento que aportan los productores.

El reto consiste en proponer estrategias de evaluación de consultas para realizar la mediación de datos entre productores y consumidores en este tipo de ambiente (evaluación de consultas), para que un consumidor pueda en todo momento obtener datos validos a pesar de su movilidad y la de los productores.

1.2 Consulta de datos en un ambiente móvil

El problema que se aborda en esta tesis se enuncia de la siguiente manera:

Dado un ambiente donde existen productores y consumidores móviles, se requiere un proceso de mediación que en todo momento pueda recuperar los datos (continuos y discretos) a partir de los productores y entregarlos a los consumidores, basándose en un conjunto de requerimientos expresados en forma de una consulta.

Existen tres principales aspectos a considerar: el cambio continuo o frecuente de posición de productores y consumidores y el procesamiento de los datos producidos en forma discreta o continua.

Considerando el aspecto de la movilidad en la evaluación de consultas, se presentan los siguientes casos:

1. Que la consulta no se pueda llevar a cabo debido a que el productor cambió de posición y por lo tanto ya no hay quien provea los datos. O bien, el consumidor se movió y debido a esto ya no hay quien reciba los datos.

2. En caso de que la consulta se pueda llevar a cabo, los resultados pueden inválidos en caso de que la consulta tenga que ver con la posición de productor y/o consumidor.

En el escenario de la primera situación (ver figura 1.2) ocurre que: dado un instante de tiempo t_1 en que coinciden C (consumidor) y P (productor), C puede realizar un consulta sobre P, pero dado que ambos van a gran velocidad y el procesamiento por parte del productor es lento, la respuesta por parte de P ocurre en otro instante de tiempo t_2 en donde C ya se encuentra fuera de su alcance. Por lo tanto, la consulta no se lleva a cabo y se pierde debido a la ausencia de productor/consumidor.

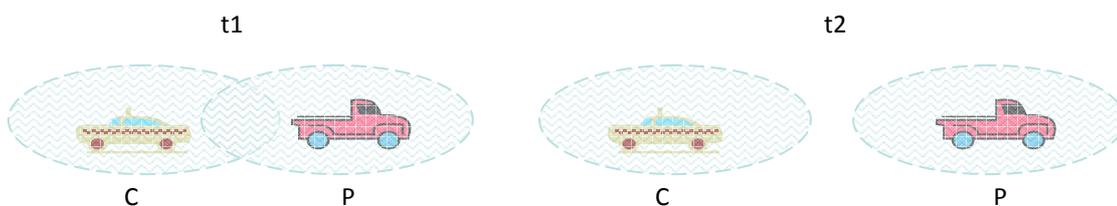


Figura 1.2. Movilidad de productores y consumidores.

En cuanto a la validez de los resultados consideremos que ya se tiene identificada el área de interés que satisface una consulta. Se pueden presentar tres casos:

- 1) Que el resultado producido por P en un tiempo t_1 sea de total interés para el consumidor (ver figura 1.3 A).
- 2) P se ha desplazado y el resultado producido en un tiempo t_2 es parcial, es decir, que la información sea sobre sólo una parte del área de interés y que incluya además información sobre otra área (ver figura 1.3 B).
- 3) La respuesta se genera en un tiempo t_3 en donde el resultado sea completamente fuera del área de interés, es decir no válido (ver figura 1.3 C).

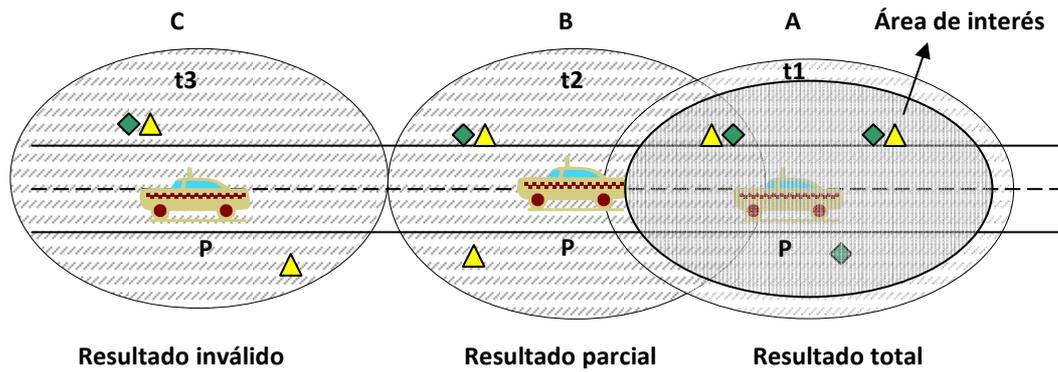


Figura 1.3. Validez de la información.

El resultado parcial presenta el mayor reto ya que implica enviar información que no es de interés para el consumidor y por lo tanto, saturar la red con información que no es útil; o bien filtrar los datos y solo notificar datos relevantes al consumidor.

Por otro lado, el aspecto de la presencia de datos continuos en la evaluación de consultas implica administrar flujos de datos considerando su orden y velocidad de producción [BW01]. Para procesar flujos de datos se requiere de técnicas adecuadas como: ventanas (para procesar el flujo por partes), *queues* (para almacenar de manera temporal tuplas que deben ser procesadas) y *synopses* (que mantienen el estado de los operadores) entre otros [ABB04]. Para nuestra investigación nosotros consideramos que los flujos tienen validez espacio-temporal debido a la movilidad de productores y consumidores.

El objetivo principal de esa tesis es construir un evaluador de consultas que implemente estrategias para recuperar y notificar datos (continuos y discretos) entre consumidores y productores móviles. Para lograr esto nosotros proponemos:

- Un modelo de datos que considera la continuidad de la información producida y su validez espacial con respecto al instante en que son producidos.
- Operadores espacio-temporales que se basan en el modelo de datos definido, para expresar el filtrado espacio-temporal de los flujos de datos.
- Una arquitectura para el evaluador de consultas móviles.

1.3. Evaluador de consultas en ambientes móviles

Con la finalidad de identificar las implicaciones asociadas el problema de consumo de datos en ambiente móviles hemos propuesto una taxonomía de consultas que considera los principales aspectos de una ambiente móvil: cambio de posición espacio-temporal de consumidores y productores y producción de datos (continua y discreta). La taxonomía de consultas nos permitió clasificar los trabajos relacionados que atacan los diferentes tipos de consultas móviles.

A partir de la taxonomía de consultas se identifican dos familias de operadores de filtrado: ventanas temporales y ventanas espaciales. El operador de ventana temporal es similar a la utilizada por los DSMS [BBDM02]. Sin embargo hemos propuesto diversos tipos de ventanas temporales: *snapshot*, *landmark*, *sliding overlap* y *sliding disjoint*. El operador de ventana espacial es la principal aportación de nuestro trabajo de tesis. La semántica de este operador consiste en consumir un área de interés particular asociada a la posición actual de un consumidor móvil.

Nuestro principio de solución se basa en la especificación de los operadores de ventana (temporal y espacial) y sus subclases (*landmark*, *sliding overlap* y *sliding disjoint*). A partir de la especificación construimos un prototipo implementado en lenguaje JAVA.

Para validar nuestro prototipo realizamos un experimento que consiste en tomar como datos de entrada datos espacio-temporales generados por un simulador de tráfico (*Network-based Generator* [B00, B002]) y aplicar los operadores propuestos a partir de un conjunto de consultas. A partir de la experimentación realizada pudimos obtener una serie de resultados que se discuten en las siguientes secciones.

1.4 Organización de documento

El resto de la tesis esta organizada de la siguiente manera:

- **Capítulo 2** propone una taxonomía de consultas en un ambiente móvil. Además presenta el estado del arte sobre el procesamiento de consultas en ambientes móviles tomando como base la taxonomía propuesta.
- **Capítulo 3** describe nuestro principio de solución para el procesamiento de consultas de datos en ambientes móviles. El capítulo describe el modelo de datos continuos espaciales (datos espacio-temporales) que usamos y define una serie de operadores de ventana temporal y espacial que hemos propuesto.
- **Capítulo 4** describe la implementación de los operadores propuestos y una arquitectura del evaluador de consultas.
- **Capítulo 5** presenta la validación de nuestra implementación a través de un experimento. El capítulo describe las características del generador de datos espacio-temporales que simula el comportamiento del tráfico en una ciudad.
- **Capítulo 6** concluye la tesis, sintetiza nuestro trabajo, presenta nuestras contribuciones y discute las perspectivas.