



**Universidad**  
Zaragoza

# IDENTIFICACIÓN DE PATRONES Y ALGORITMOS DE CONSOLIDACIÓN EN BASES DE DATOS DE POSICIONAMIENTO

---

Pilar Barbero Iriarte  
9 de diciembre de 2015

Universidad de Zaragoza

1. Introducción
2. Análisis de los datos
3. ¿Cómo abordar el problema?
4. Comparativa casos estudiados
5. Conclusiones
6. Demostración
7. Preguntas

# INTRODUCCIÓN

---

Contexto:

- Empresa Zaragozana de telecomunicaciones.
- Almacenamiento de posiciones GPS de sujetos.
- Capacidad de guardado de posiciones limitada.

Problemas:

- Exceso de éstas.
- No existe preprocesado antes de la inserción.
- No existe postprocesado después de la inserción.
- No todas aportan información.

Objetivo:

- Eliminar posiciones repetidas.
- Eliminar posiciones que no aporten información.

## ANÁLISIS DE LOS DATOS

---

- **Id:** Identificador numérico



- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición
- **Latitud:** real que representa la latitud GPS

- **Id**: Identificador numérico
- **IdServidor**: Identificador numérico del servidor que realiza la inserción
- **Recurso**: identificador del sujeto que transfiere la posición
- **Latitud**: real que representa la latitud GPS
- **Longitud**: real que representa la longitud GPS

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición
- **Latitud:** real que representa la latitud GPS
- **Longitud:** real que representa la longitud GPS
- **Velocidad:** entero que representa la velocidad instantánea

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición
- **Latitud:** real que representa la latitud GPS
- **Longitud:** real que representa la longitud GPS
- **Velocidad:** entero que representa la velocidad instantánea
- **Orientación:** entero que representa la orientación respecto al norte en grados

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición
- **Latitud:** real que representa la latitud GPS
- **Longitud:** real que representa la longitud GPS
- **Velocidad:** entero que representa la velocidad instantánea
- **Orientación:** entero que representa la orientación respecto al norte en grados
- **Cobertura:** booleano que indica si tiene cobertura (n. satélites)

- **Id:** Identificador numérico
- **IdServidor:** Identificador numérico del servidor que realiza la inserción
- **Recurso:** identificador del sujeto que transfiere la posición
- **Latitud:** real que representa la latitud GPS
- **Longitud:** real que representa la longitud GPS
- **Velocidad:** entero que representa la velocidad instantánea
- **Orientación:** entero que representa la orientación respecto al norte en grados
- **Cobertura:** booleano que indica si tiene cobertura (n. satélites)
- **Error:** error en la toma de posición



¿CÓMO ABORDAR EL PROBLEMA?

---

## ¿CÓMO ABORDAR EL PROBLEMA?

- Desarrollo de algoritmos de consolidación a través de nociones de distancia y tiempo.
- Uso de algoritmos de *clustering* con el fin de identificar varias posiciones con su centro del clúster y consolidarlas en ésta.















## COMPARATIVA CASOS ESTUDIADOS

---

| Método     | Tiempo        | N.  | Iteraciones |
|------------|---------------|-----|-------------|
| K-means    | 0.69 secs     | 500 | 9           |
| DBSCAN     | 2 min 30 secs | 9   | 111         |
| DJ-Cluster | 0.37 secs     | 22  | 11          |

- DBSCAN más lento de todos.
- DBSCAN consolidación mayor.
- K-means se queda en 500 clústers.
- DJ-Clúster baja de los 500.
- DJ-Clúster menor tiempo de ejecución.

| Método                                | Tiempo    | N.   |
|---------------------------------------|-----------|------|
| Cons. por adelgazamiento              | <0.01 sec | 800  |
| Cons. por distancia simple            | <0.01 sec | 507  |
| Cons. por distancia $t_0$ –alcanzable | <0.01 sec | 21   |
| Cons. por tiempo                      | <0.01 sec | 1786 |

- Alternar distintos tipos de consolidación en función del espacio crítico en ese momento.

## CONCLUSIONES

---

- Algoritmos de consolidación simples son *simples*, pero eficaces.

- Algoritmos de consolidación simples son *simples*, pero eficaces.
- Noción de vecindario distinto al euclídeo implementada en algoritmos de consolidación simple.

- Algoritmos de consolidación simples son *simples*, pero eficaces.
- Noción de vecindario distinto al euclídeo implementada en algoritmos de consolidación simple.
- Algoritmos de *clustering* más avanzados, pero más complejos a la hora de implementar.

- Algoritmos de consolidación simples son *simples*, pero eficaces.
- Noción de vecindario distinto al euclídeo implementada en algoritmos de consolidación simple.
- Algoritmos de *clustering* más avanzados, pero más complejos a la hora de implementar.
- Importante un procesamiento previo.



- Algoritmos de consolidación simples son *simples*, pero eficaces.
- Noción de vecindario distinto al euclídeo implementada en algoritmos de consolidación simple.
- Algoritmos de *clustering* más avanzados, pero más complejos a la hora de implementar.
- Importante un procesado previo.
- A la hora de recuperar una traza con los datos borrados, mejor *DJ-Clúster*

- Algoritmos de consolidación simples son *simples*, pero eficaces.
- Noción de vecindario distinto al euclídeo implementada en algoritmos de consolidación simple.
- Algoritmos de *clustering* más avanzados, pero más complejos a la hora de implementar.
- Importante un procesado previo.
- A la hora de recuperar una traza con los datos borrados, mejor *DJ-Clúster*
- Noción de ruido de DBSCAN importante, tanto DJ-Clúster como K-means sólo encuentra clústers de tamaño 1.

# DEMOSTRACIÓN

---

## PREGUNTAS

---

http://github.com/pbarbero/TFM

pbarbero / TFM

Watch 1 Star 0 Fork 0

Code Issues Pull requests Wiki Pulse Graphs Settings

Trabajo Fin de Máster Modelización e Investigación Matemática, Estadística y Computación — Edit

88 commits 2 branches 0 releases 1 contributor

Branch: master New pull request New file Find file HTTPS https://github.com/pbarbe Download ZIP

|                     |                                   |                                   |
|---------------------|-----------------------------------|-----------------------------------|
| pbarbero            | frame.comparatiya                 | Latest commit deebfea 5 hours ago |
| consolidationExpert | changes in checkpoints            | 7 days ago                        |
| consolidationSimple | changes in checkpoints            | 7 days ago                        |
| data                | reorder files                     | 7 days ago                        |
| dataAnalysis        | reorder files                     | 7 days ago                        |
| defensa             | frame comparativa                 | 5 hours ago                       |
| papeles             | reorder files                     | 7 days ago                        |
| report              | lasts changes to report           | 7 days ago                        |
| reuniones           | add own algorithms to comparative | 9 days ago                        |
| .gitignore          | update gitignore                  | 7 days ago                        |
| License.md          | Create License.md                 | 9 days ago                        |
| README.md           | change readme                     | a day ago                         |

README.md

## Identificación de patrones y algoritmos de consolidación en bases de datos de posicionamiento