

Creación de Mapas de Riesgo Ferroviario a partir de mapas de calor de incidencias

Creating Railroad Risk Maps from incident knowledge base heatmaps

José Gómez Castaño*

^a Jefatura de Normalización y Nuevos Desarrollos de la Dirección de Gestión de Red e Innovación. ADIF

^b Consultor Sistemas de Información Ferroviaria. STRUCTURALIA

Resumen

Uno de las tareas más importantes a las que se enfrenta cualquier Administrador de Infraestructuras es la de anticiparse a las posibles incidencias que se puedan producir en las mismas. Para ello, lo más efectivo es contar con procedimientos preventivos que se puedan anticipar a las posibles situaciones.

Para conseguir este objetivo, contar con una base de conocimiento es algo fundamental. Es necesario conocer qué tipo de incidencias se han producido y dónde. En este artículo se describe cómo sacar partido a la componente geoespacial de esta base de conocimiento utilizando la técnica de representar las incidencias como mapas de calor o "heatmaps".

Los mapas de calor son representaciones gráficas en las que se muestran las zonas de mayor actividad de cierta variable. En este caso se presentan las zonas con más incidencias, atendiendo a sus diversos atributos. El resultado es la generación de Mapas de Riesgos.

En este artículo se muestra cómo desarrollar algunos programas que responden a un procedimiento para generar los mapas de calor a partir de los datos de incidencias básicos, y una aplicación web funcional (<http://www.astroide.es/trenes/heatmap.php>).

Palabras clave: sig, mapa de calor, kml, heatmap, incidencias, mapa de riesgo, python, google maps, openlayers

Abstract

One of the most important tasks facing any Infrastructure Manager is to anticipate possible problems that may occur in them. To do this, the most effective is to have preventive procedures can anticipate possible situations.

To achieve this goal, having a knowledge base is essential. You need to know what type of incidents have occurred and where. This article describes how to leverage the geospatial component of this knowledge base using the technique of representing incidents such as heat maps or "heatmaps".

The heat maps are graphical representations shown in the areas of greatest activity of a certain variable. In this case there are areas with more incidents, according to various attributes. The result is the generation of risk maps.

This article shows how to develop some programs that respond to a procedure to generate heat maps from the basic event data.

keywords: gis, heatmap, kml, events, risk maps, python, openlayers, google maps

* jgomezc@adif.es

1. Situación actual

Para el análisis de las incidencias, partimos de los registros que se hacen cada vez que se produce alguna en la red ferroviaria. En todos los casos queda constancia de diversos atributos ligados a la misma. Desde el momento en que se produce, elementos de la infraestructura afectados, personal que ha intervenido, tipo de material que la ha producido, si se han visto afectados trenes y cuáles.

Para el trabajo que nos ocupa, el dato más interesante es el lugar donde se ha producido. En la operación ferroviaria, el lugar está ligado a la línea y punto kilométrico dentro de ella, o a una instalación ferroviaria en concreto. Este término es bastante amplio y puede tratarse de una estación, bifurcación, puesto, agujas, instalaciones de seguridad, de electrificación, etc.

Por lo tanto, no hay una relación directa entre la denominación o codificación donde ha tenido lugar la incidencia y su posición geográfica en un mapa. Aunque en el Anexo I de la Directiva INSPIRE [1] se encuentra recogido un modelo relativo a redes de transporte ferroviario, este no ha sido adoptado todavía por ningún Administrador de la Infraestructura. Esta situación será subsanada en el futuro, dado que también se encuentra contemplado la LISIGE [2].

Por este motivo se parte de ubicaciones sin geolocalizar. Las bases de datos que se utilizan tradicionalmente para almacenar este tipo de información no tienen extensiones espaciales, lo que complica la aplicación de una solución puramente geográfica.

2. Metodología propuesta

Vamos a crear una base de conocimiento a partir de las incidencias que se van produciendo. Esta se encontrará almacenada en un gestor de bases de datos con extensiones espaciales. El uso de este tipo de bases de datos nos permitirá hacer interrogaciones a la misma en lenguaje SFSQL[3]. Los atributos alfanuméricos seguirán registrándose bajo el mismo modelo que hasta ahora, pero podremos tener la información geográfica en 2 tablas al efecto.

La primera contendrá los puntos relevantes de red ferroviaria, con un campo con la identificación tradicionalmente utilizada para ellos. Su posición geográfica se almacenará en un campo de tipo Geometría, y puede representarse por un punto o un polígono. La segunda almacenará la geometría de las líneas.

El objetivo de estas dos tablas es poder geolocalizar la posición de cada incidencia. Podemos tener dos casos. Que se haya producido en una instalación. Su ubicación se obtendrá cruzando la tabla de alfanumérica con la de ubicaciones. En un segundo caso, la incidencia puede haberse producido en plena vía. En este caso solo disponemos del punto kilométrico y la línea. Para conocer su posición habrá que calcularla aplicando un algoritmo de segmentación dinámica a partir del km en el que se registra.

En este punto hay que llamar la atención sobre una particularidad que existe en cualquier red viaria que cambia con el tiempo. Las señales indicativas de los puntos kilométricos, no equivalen a la posición geométrica. Esto quiere decir que los puntos kilométricos indicados en el terreno, no se encuentran en la misma posición que resulta si tomamos la geometría de la línea férrea. Es lo que produce que en la realidad haya kilómetros con “más de 1000 metros”. Esto se resuelve tomándolos como fenómenos individuales, como si se tratara de un elemento más de la infraestructura, de la misma forma que se codifica en la BTN25 del IGN. Además se necesita una tabla que relacione el punto kilométrico del terreno con el real geométrico.

Si bien el tener en el mismo gestor de bases de datos toda la información le proporciona mucha eficiencia al sistema, se puede optar por otra solución que proporcione opciones de geolocalización a cualquier otra aplicación. La información geográfica se puede exponer a través de un servicio web y ser consumida desde

el resto de aplicaciones. En la actualidad, estos dos métodos se utilizan en la aplicación de gestión de incidencias de forma habitual.

Sea cual sea el sistema elegido para obtener las coordenadas de las ubicaciones de las incidencias, el resultado será que tendremos la posición de cada una de ellas. Conjugando los atributos alfanuméricos y la información geográfica, el resultado de cada consulta proporcionará el conjunto de ubicaciones que cumplan con la Query SFSQL solicitada.

Siguiendo este esquema se pueden hacer consultas para conocer las ubicaciones (los registros) que hayan tenido una incidencia de cierta categoría, o que haya sido producida por cierto tipo de locomotora, tipo de vagón, o cualquier otro tipo de condición a investigar. A continuación se muestra un ejemplo de consulta en la que obtendremos las ubicaciones de las incidencias con codificación 001 y creadas por el tipo de locomotora 333.

```
SELECT incidencias.codUbicacion, ST_AsText(ubicaciones.ubicacion_geom) AS geom,  
ubicaciones.nombreUbicacion FROM incidencias  
inner join ubicaciones on ubicaciones.codUbicacion = incidencias.codUbicacion  
where incidencias.codIncidencia = "001" and incidencias.tipoLocotora = '333';
```

El resultado sería un conjunto de fenómenos ST_Point o ST_Polygon, dependiendo del tipo de datos almacenados, y en el SRID correspondiente al sistema de referencia utilizado en la tabla. En el caso de tratarse de un punto kilométrico, se puede construir una consulta similar para obtener las coordenadas del punto buscado como fenómeno ST_Point.

A partir de este momento es cuando podremos representarlas sobre un mapa para crear un heatmap. Explotaremos la base de conocimiento que se ha ido creando para representar en el mapa de calor la distribución geográfica de las incidencias de acuerdo a un código de colores. Este mostrará las zonas de mayor densidad de las incidencias filtradas según los criterios que se busquen (figura 1).



Figura 1. Ejemplo de mapa de calor generado

El resultado es un mapa que muestra dónde se han producido las incidencias con mayor frecuencia. Para convertir este en un Mapa de Riesgos, bastará con seleccionar aquellos tipos de incidencia que tienen que ver con aspectos de la infraestructura. Por ejemplo con las instalaciones de seguridad, señalización, puentes, pasos a nivel. Con esta visualización es fácil detectar cuales serían las instalaciones con más riesgo de sufrir un problema o zonas que acumulan puntos negros.

Existen otras fuentes de datos que se pueden utilizar para generar los mapas de riesgo. Podemos aplicar la misma técnica a determinar las zonas con mayor riesgo de sismos. Para ello hemos seleccionado de los últimos 10000 terremotos registrados en España, los que han registrado una magnitud mayor a 3 y se han geolocalizado con datos proporcionados a través del Servicio de Información Sísmica [4]. De ellos obtenemos el mapa de riesgos sísmicos en función de los terremotos registrados, que se muestra en la figura 2. A estos se le puede añadir más información relativa a la geología de las diferentes zonas.

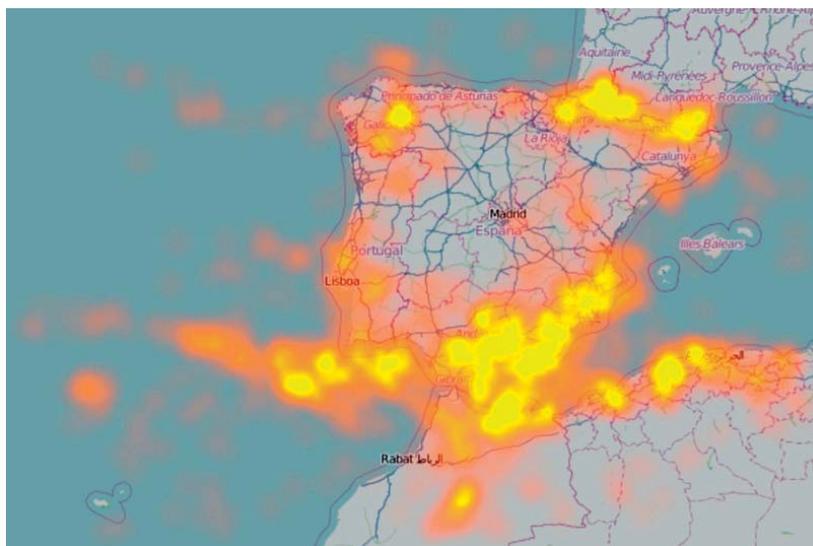


Figura 2. Mapa de los últimos 10000 sismos

La combinación de datos procedentes de diferentes fuentes, conforman el conjunto de Mapas de Riesgo. La codificación de las incidencias puede refinarse para asociarlas a cierto elemento de seguridad, tipo de señal o instalación para que se destaquen en los mapas resultantes los más afectados por las mismas.

Los mapas resultantes pueden ser generados como gráficos y posteriormente añadidos como una capa más en cualquier Sistema de Información Geográfica (SIG). En el ejemplo mostrado más abajo se ha generado un gráfico georreferenciado, una capa KML y un gráfico que automáticamente añade el mapa de calor sobre una cartografía de base de Open Street Map.

3. Métodos de creación del mapa

Son muchos los métodos que se pueden utilizar para la creación de un heatmap [5]

Para construir un gráfico de este tipo hay que representar las agrupaciones de las incidencias, dando un color diferente para cada densidad. Para ilustrar el resultado se han utilizado varios métodos.

La forma más sencilla de construir un gráfico de este tipo es utilizando la librería numpy de python. Los datos se almacenan en un csv

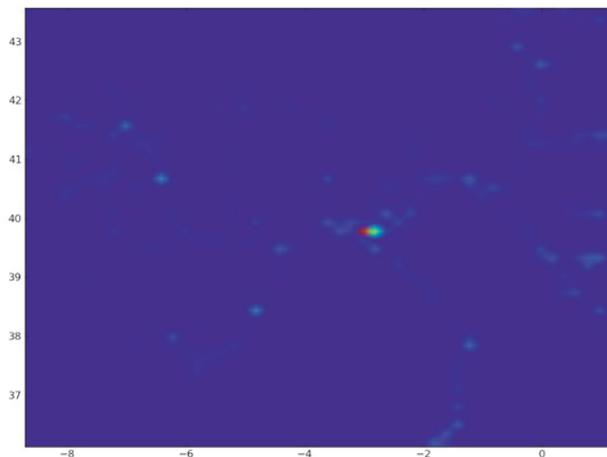
```
#!/usr/bin/python

import numpy as np
import csv
import matplotlib.pyplot as plt

reader = csv.reader(open('datos.csv',
'rb'))
xs = []
ys = []
for index, row in enumerate(reader):
    xs.append(float(row[ 1] ))
    ys.append(float(row[ 0] ))

heatmap, xedges, yedges =
np.histogram2d(xs, ys, bins=50)
extent = [ xedges[ 0], xedges[ -1], yed-
ges[ 0], yedges[ -1]]
plt.imshow(heatmap, extent=extent)

plt.show()
```



Esta función no proporciona una imagen georreferenciada, aunque si pueden usarse los ejes X e Y para determinar la posición de las zonas con mayor acumulación de incidencias.

Para conseguir imágenes de gran calidad podemos utilizar la librería heatmap [6]. Esta permite incluirla en un script y generar una imagen en la que se puede controlar el tamaño y escala e colores de las zonas. Esta librería presenta la ventaja de poder generar un fichero en formato KML usando como base el gráfico generado, lo que permite posicionarlo sobre cartografía

```
#!/usr/bin/python

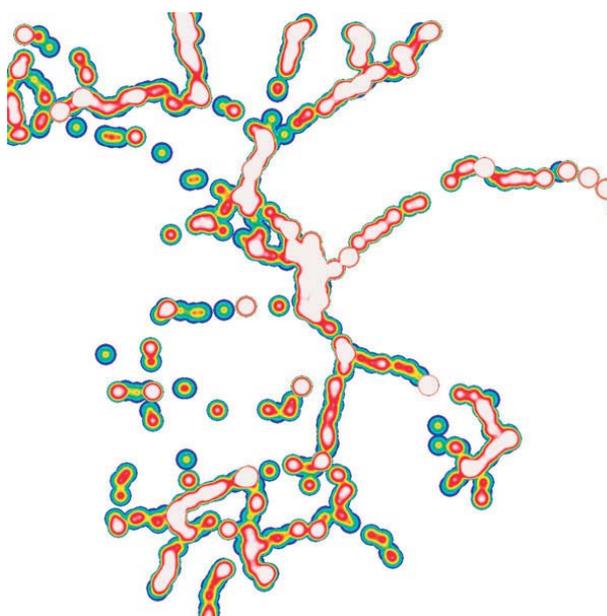
import heatmap
import numpy as np
import csv
import matplotlib.pyplot as plt

hm = heatmap.Heatmap()

reader = csv.reader(open('datos.csv',
'rb'))
dat = []
for index, row in enumerate(reader):
    x =(float(row[ 1] ))
    y = (float(row[ 0] ))
    dat.append([ x,y ])

hm.heatmap(dat, "classic.png", dotsize=50)

hm.saveKML("incidengeorreferenciadas.kml")
```



Por último nos quedaría poder representar el heatmap directamente sobre la cartografía. Para ello podemos utilizar la librería heatmap.py. Partiendo de los datos filtrados mediante SFSQL permite generar un mapa base con Open Street Map y el mapa de calor sobre él.

```
heatmap.py -p datosIncidencias -o riesgos.png --  
width 1000 --osm -B 0.9
```



Para permitir un tratamiento más sencillo en la web se puede usar una librería JavaScript, que trate los datos en una aplicación web en el navegador cliente. La más extendida es heatmap.js [7].

En todos los casos podemos cambiar la resolución espacial, el zoom, densidad de puntos o escalas de colores de representación, por lo que nos permite diversos tipos de representación en función de las necesidades

4. Ejemplo práctico

Para nuestro caso práctico hemos elegido una base de datos PostGIS en la que se ha creado el modelo de datos mostrado en la figura 3. Los datos alfanuméricos se muestran a título de ejemplo. Para la tabla de ubicaciones se han elegido objetos de tipo Point para representarlos. Las coordenadas geográficas se han almacenado en EPSG:4258.

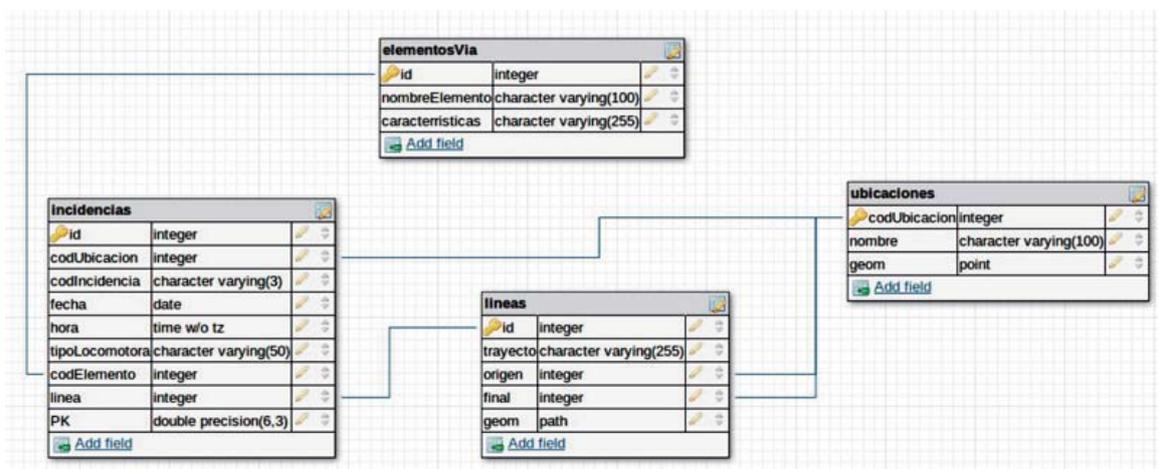


Figura 3. Modelo de datos

Se han generado incidencias de prueba de diverso tipo para todo un año distribuidas por todas las líneas, hasta un total de 50000, para dar una imagen lo más ilustrativo posible, y se han almacenado de acuerdo al modelo anterior en la base de datos.

Para poder generar el mapa de calor, se han desarrollado varios programas en Python a medida basándose en los scripts descritos más arriba. A partir de los datos almacenados estos programas proporcionan varios salidas:

- Imágenes del mapa de calor sin georreferenciar
- Imágenes del mapa de calor georreferenciadas en formato KML(figura 4)
- Mapas de calor sobre cartografía de Open Street Map
- Mapas generados dinámicamente mediante JavaScript

Se ha desarrollado un aplicación web [8] que muestra el resultado completo. Esto permite generar el mapa de riesgos pudiendo filtrarlo por meses y tipología (figura 5). Estos mapas muestran las zonas de riesgo de sufrir un tipo de incidencia determinado.

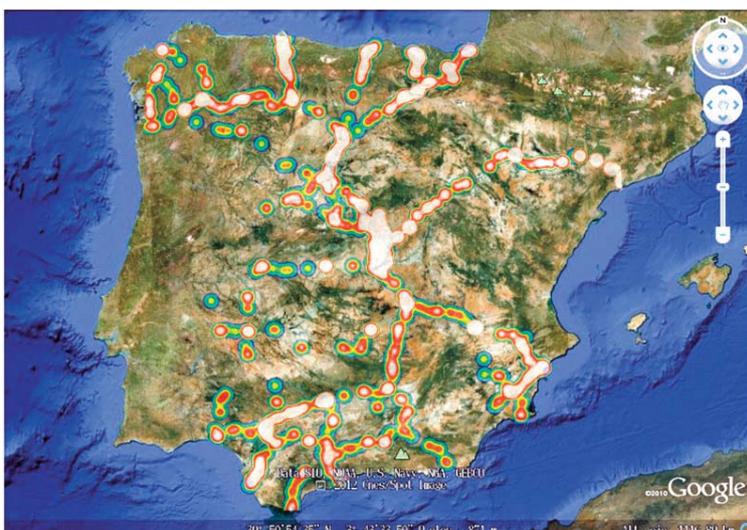


Figura 4. Superposición de capa KML sobre Google Earth

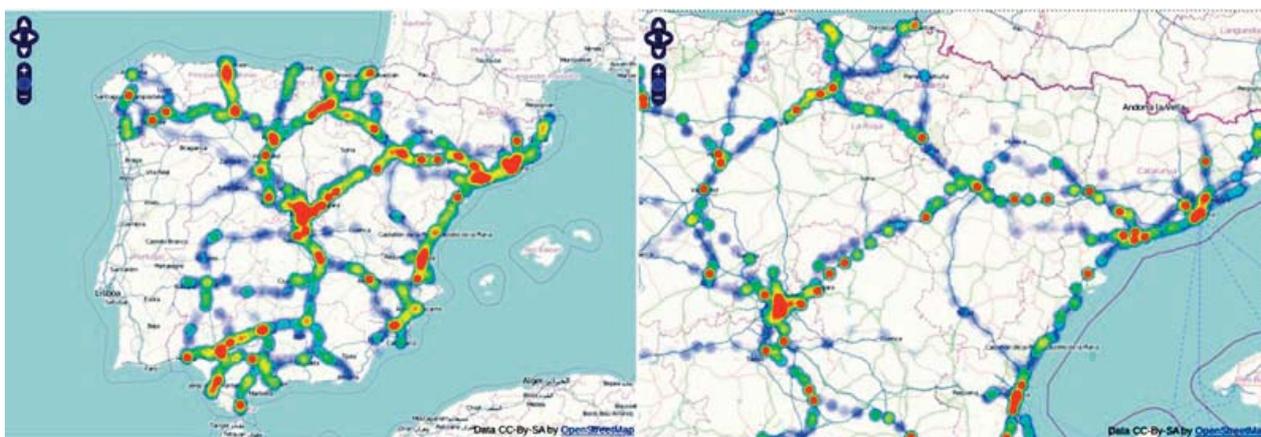


Figura 5. Generación dinámica en Aplicación Web

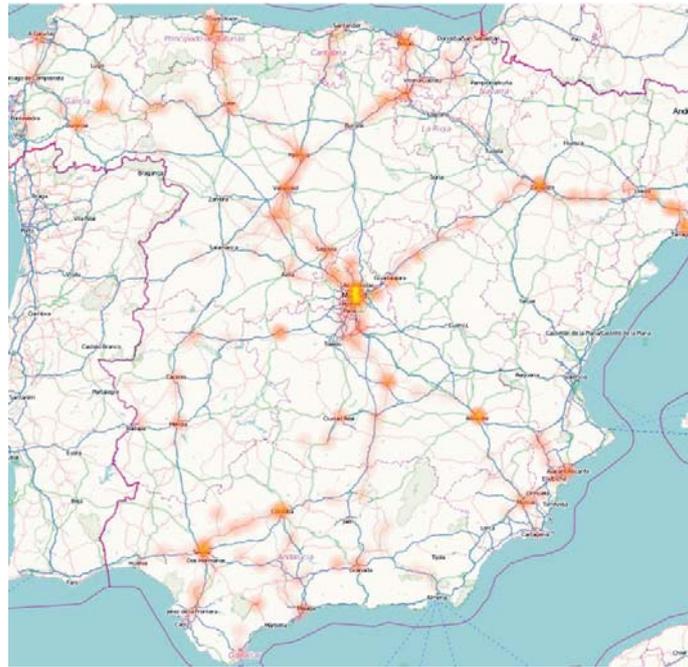


Figura 6. Generación dinámica en Aplicación Web

5. Conclusión

Los mapas obtenidos a partir de una base de conocimiento de incidencias, pueden utilizarse como Mapas de Riesgo. Estos pueden utilizarse para localizar dónde efectuar medidas preventivas sobre la infraestructura o localizar los puntos negros de la misma. Para identificar las zonas afectadas los mapas deben estar georreferenciados o poderse visualizar directamente con la cartografía. Un ejemplo implementado en una aplicación web puede verse en <http://www.astroide.es/trenes/heatmap.php>

[1] <http://inspire.jrc.ec.europa.eu/>

[2] <http://www.boe.es/boe/dias/2010/07/06/pdfs/BOE-A-2010-10707.pdf>

[3] <http://www.opengeospatial.org/standards/sfs>

[4] <http://www.ign.es/ign/layout/sismo.do>

[5] “The History of the Cluster Heat Map”, Leland Wilkinson and Michael Friendly, The American Statistician, Noviembre, 2008

[6] <http://jguy.com/heatmap/>

[7] <http://www.patrick-wied.at/static/heatmapjs/>

[8] <http://www.astroide.es/trenes/heatmap.php>