

Universidad de Costa Rica



Facultad de Ingeniería

Escuela de Ciencias de la Computación e Informática

CI1314 - Bases de Datos II

Profesora: Elzbieta Malinowski G.

Documentación del Proyecto

Versión 4

Oscar Esquivel Oviedo - B22410

Pablo Sauma Chacón - B16195

Leonardo Villalobos Arias - B27219

16 de Octubre del 2014

OBJETIVO DE LA APLICACIÓN

La aplicación tiene como objetivo principal el reducir los daños que se dan por incendios en el país, ya que éstos pueden llegar a ser muy frecuentes se plantea una manera de obtener información acerca de qué tan peligroso puede ser un lugar en cuanto a riesgo de incendio y en caso de que ocurra uno, a quién acudir.

Para cumplir con su objetivo la aplicación cuenta con ciertos objetivos específicos, el primero es poder dar información sobre las estaciones de bomberos más cercanas a un punto cualquiera en el país, para lo cual no sólo toma en cuenta las distancias entre ese punto y las estaciones, sino también las distancias reales a lo largo de caminos que se deben recorrer para llegar desde la estación hasta el lugar. Esto para informar sobre cuáles estaciones deben contactarse en caso de darse un incendio, y que éstas no sólo estén cerca, sino que también sea posible para los bomberos trasladarse.

Como funcionalidad adicional, la aplicación también puede determinar el nivel de peligro general de incendio en cualquier punto del país, tomando en cuenta si éste se encuentra dentro de una zona de riesgo y si es así, de cuánto riesgo, y qué tan lejos están las estaciones de bomberos más cercanas.

DESCRIPCIÓN DE LOS ARCHIVOS DE DATOS

Para clarificar los contenidos de los archivos tipo shapefile y hojas electrónicas de los cuales se importará sus datos a la base de datos, se hará una descripción de sus contenidos y observaciones sobre estos: Principalmente por qué no se encuentran en tercera forma normal.

ARCHIVOS SHAPEFILE

Como origen de los datos espaciales, y algunos datos relacionados a estos, se utilizarán archivos shapefile. El contenido de estos será analizado, y se expondrá por qué algunos de estos archivos no están normalizados en la Tercera Forma Normal (3FN).

provincias2008crtm05	Provincia	Num_Canto	Cod_Prov	POB_1950_H	POB_1950_M	POB_1963_H	POB_1963_M	POB_1973_H	POB_1973_M
Breve descripción de su significado	Nombre de la provincia	Cantidad de cantones que posee la provincia	Código identificador de la provincia	Población masculina en la provincia en 1950	Población femenina en la provincia en 1950	Población masculina en la provincia en 1963	Población femenina en la provincia en 1963	Población masculina en la provincia en 1973	Población femenina en la provincia en 1973
Tipo de dato	Texto	Número	Número	Número	Número	Número	Número	Número	Número
Observación es sobre valores	Único para cada provincia	Derivado de sus provincias	1 dígito, determina el nombre	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo
Violación de 3FN	No	No	No	No	No	No	No	No	No
Razones de violación de 3FN									

provincias2008crtm05	POB_1984_H	POB_1984_M	POB_2000_H	POB_2000_M	Viviendas_	Viviendas_	Viviendas_	Hectares
Breve descripción de su significado	Población masculina en la provincia en 1984	Población femenina en la provincia en 1984	Población masculina en la provincia en 2000	Población femenina en la provincia en 2000	Cantidad de viviendas ocupadas? en la provincia	Cantidad de viviendas desocupadas? en la provincia	Cantidad de viviendas colectivas? en la provincia	Hectáreas correspondientes
Tipo de dato	Número	Número	Número	Número	Número	Número	Número	Número flotante
Observaciones sobre valores	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Campo calculado, datos incorrectos
Violación de 3FN	No	No	No	No	No	No	No	No
Razones de violación de 3FN								

Tabla 1: Datos de provincias2008crtm05

cantones2008crtm05	Ncanton	Nprovincia	Prov	Canto	Pob_2000_H
Breve descripción de su significado	Nombre del cantón	Nombre de la provincia a la que pertenece el cantón	Número de la provincia a la que pertenece el cantón	Número del cantón	Población masculina en la provincia en 2000
Tipo de dato	Texto	Texto	Número	Número	Número
Observaciones sobre valores	Único para cada cantón	Puede repetirse para varios cantones	1 dígito, puede repetirse entre cantones	2 dígitos, no es único para cada cantón	Recolectados mediante censo
Violación de 3FN	Sí	Sí	Sí	Sí	No
Razones de violación de 3FN	Está determinado por CodNum	Está determinado por Ncanton, CodNum o Prov	Está determinado por Ncanton, CodNum o Nprovincia	Está determinado por CodNum	

cantones2008crtm05	Pob_2000_M	Vivienda_O	Vivienda_D	Vivienda_C	CodNum	Hectares
Breve descripción de su significado	Población femenina en la provincia en 2000	Cantidad de viviendas ocupadas en el cantón	Cantidad de viviendas desocupadas en el cantón	Cantidad de viviendas colectivas en el cantón	Código numérico que identifica cada provincia	Hectáreas correspondientes
Tipo de dato	Número	Número	Número	Número	Número	Número flotante
Observaciones sobre valores	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	3 dígitos, identifica a cada cantón individualmente	Campo calculado, datos incorrectos
Violación de 3FN	No	No	No	No	Sí	No
Razones de violación de 3FN					Está determinado por Prov y Canto	

Tabla 2: Datos de cantones2008crtm05

distritos2008crtm05	Provincia	Ncanton	Ndistrito	Prov	Canton	Distrito	Coddist
Breve descripción de su significado	Nombre de la provincia a la que pertenece el distrito	Nombre del cantón al que pertenece el distrito	Nombre del distrito	Número de la provincia a la que pertenece el distrito	Número del cantón al que pertenece el distrito	Número del distrito	Código identificador del distrito
Tipo de dato	Texto	Texto	Texto	Número	Número	Número	Número
Observaciones sobre valores	Puede repetirse para varios distritos	Puede repetirse para varios distritos	Puede repetirse para varios distritos	1 dígito, puede repetirse entre distritos	2 dígitos, puede repetirse entre distritos	2 dígitos, no es único para cada distrito	5 dígitos, único para cada distrito
Violación de 3FN	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Razones de violación de 3FN	Está determinado por Ncanton, Prov, Coddist, o Prov, Canton y Distrito	Está determinado por Coddist o Prov, Canton y Distrito	Está determinado por Coddist o Prov, Canton y Distrito	Está determinado por Provincia, Canton, Ncanton o Coddist	Está determinado por Ncanton o Coddist	Está determinado por Coddist	Determinado por Prov, Canton y Distrito, o Ncanton y Ndistrito

distritos2008crtm05	Pob_1997_M	Pob_1997_H	Pob_2000_H	Pob_2000_M	Vivienda_O	Vivienda_D	Vivienda_C
Breve descripción de su significado	Población femenina en la provincia en 1997	Población masculina en la provincia en 1997	Población masculina en la provincia en 2000	Población femenina en la provincia en 2000	Cantidad de viviendas ocupadas en el distrito	Cantidad de viviendas desocupadas en el distrito	Cantidad de viviendas colectivas en el distrito
Tipo de dato	Número	Número	Número	Número	Número	Número	Número
Observaciones sobre valores	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo	Recolectados mediante censo
Violación de 3FN	No	No	No	No	No	No	No
Razones de violación de 3FN							

Tabla 3: Datos de distritos2008crtm05

bomberos2008crtm05	Nombre	Direccion	Infraestructu	Extintoras	Rescate	Forestales	Provincia	Canton	Distrito
Breve descripción de su significado	Nombre de la estación de bomberos	Dirección de la estación de bomberos	Describe la cantidad de pisos que tiene la estación	Número de unidades extintoras	Número de unidades de rescate	Número de unidades de control forestal	Nombre de la provincia a la que pertenece la estación	Nombre del cantón al que pertenece la estación	Nombre del distrito al que pertenece la estación
Tipo de dato	Texto	Texto	Texto	Texto	Texto	Texto	Texto	Texto	Texto
Observaciones sobre valores	Único para cada estación	Único para cada estación, descripción informal	Formato: Edificio/Estructura de Una/Dos/Cuatro plantas	Formato: No/# Unidad(es)	Formato: No/# Unidad(es)	Formato: No/# Unidad(es)	Puede repetirse para varias estaciones	Puede repetirse para varias estaciones	Puede repetirse para varias estaciones
Violación de 3FN	No	No	No	No	No	No	Sí	No	No
Razones de violación de 3FN							Está determinado por Canton y Distrito		

Tabla 4: Datos de bomberos2008crtm05

redcamino2008crtm05	Ruta	Tipo	Longitud
Breve descripción de su significado	Número de la ruta	Tipo de la ruta	Longitud de la ruta
Tipo de dato	Texto (Puesto que incluye guiones)	Texto	Número flotante
Observaciones sobre valores	Muchas carreteras tienen el valor ND	Enumeración	Campo calculado
Violación de 3FN	No	No	No
Razones de violación de 3FN			

Tabla 5: Datos de redcamino2008crtm05

riesgoincendio2008crtm05	Messec	Clasificac	Riesgo
Breve descripción de su significado	Meses que el área está seca	Clasificación de la velocidad del viento	Nivel de riesgo de la zona
Tipo de dato	Número	Texto	Texto
Observaciones sobre valores	Varían del 1 al 6	Tres categorías: 3-5 m/s, 5-7 m/s y >7 m/s	Bajo, medio, alto o muy alto
Violación de 3FN	No	No	Sí
Razones de violación de 3FN			La combinación de Messec y Clasificac la determinan

Tabla 6: Datos de riesgoincendio2008crtm05

ARCHIVOS DE HOJAS ELECTRÓNICAS

Como otro origen de datos para importar a la base de datos, se tienen datos reunidos en hojas electrónicas en formato Excel. La composición de estas es la siguiente:

	Población				
	Provincia, cantón y distrito	Total	Hombres	Mujeres	Total
Descripción de su significado	Contiene el nombre de la provincia, cantón o distrito	Contiene el valor total de pobladores de la zona	Contiene el valor de la población masculina de la zona	Contiene el valor de la población femenina de la zona	Contiene el valor total de viviendas de la zona
Tipo de Dato	Texto	Entero	Entero	Entero	Entero
Observaciones	Se agrupan de acuerdo a pertenencia, no a si son provincias, cantones o distritos	Es la suma de 'Hombres' y 'Mujeres' en cada tupla respectiva, y a su vez, diferentes tuplas son sumas de sus componentes (regiones pertenecientes a ella)	Algunas tuplas tienen por valor la suma de sus 'componentes' (regiones pertenecientes a estas)	Algunas tuplas tienen por valor la suma de sus 'componentes' (regiones pertenecientes a estas)	Es el resultado de la suma del total de viviendas individuales y colectivas adentro de su misma tupla, y además de la suma de las 'regiones' pertenecientes a ella.

	Viviendas			
	Total	Individuales		Colectivas
Descripción de su significado	Contiene el valor total de viviendas individuales de la zona	Ocupadas Contiene el valor total de viviendas individuales ocupadas de la zona	Desocupadas Contiene el valor total de viviendas individuales desocupadas de la zona	Contiene el valor total de viviendas colectivas de la zona
Tipo de Dato	Entero	Entero	Entero	Entero
Observaciones	Es el resultado de la suma de las viviendas ocupadas y desocupadas dentro de su misma tupla, y además de la suma de las 'regiones' pertenecientes a ella.	Algunas tuplas son el resultado de las suma de 'regiones' pertenecientes a ellas.	Algunas tuplas son el resultado de las suma de 'regiones' pertenecientes a ellas.	Algunas tuplas son el resultado de las suma de 'regiones' pertenecientes a ellas.

Tabla 7: Datos del archivo Población total por sexo y total de viviendas por ocupación, según provincia, cantón y distrito.xls

	Kilómetros			
	Cantón	2003	2006	2007
Descripción de su significado	Contiene el nombre del cantón	Número de kilómetros de carreteras y caminos asfaltados en el cantón para el año	Número de kilómetros de carreteras y caminos asfaltados en el cantón para el año	Número de kilómetros de carreteras y caminos asfaltados en el cantón para el año
Tipo de Dato	Texto	Número	Número	Número
Observaciones	No trae distritos, solo cantones	Se podría calcular en base a las geometrías	Se podría calcular en base a las geometrías	Se podría calcular en base a las geometrías

Tabla 7: Datos de los archivos de datos de carreteras

Los archivos que contienen datos de carreteras siguen todo este mismo formato. La lista de estos archivos es:

- Longitud_de_carreteras_y_caminos_asfaltadasCantones03_08.xls
- Longitud_de_carreteras_y_caminos_de_concretoCantones03_08.xls
- Longitud_de_carreteras_y_caminos_de_lastre_o_tierraCantones03_08.xls
- Longitud_de_carreteras_y_caminosCantones01_08.xls

ESQUEMA CONCEPTUAL

Con base en los objetivos de la aplicación, y los datos disponibles para su uso, se diseñó un esquema conceptual para la base de datos utilizando el modelo Entidad-Relación extendido para datos espaciales. A continuación se presentará el diagrama, seguido de explicaciones y justificaciones de su construcción. Una versión más grande del diagrama, como así sus versiones anteriores se pueden encontrar en los anexos al documento.

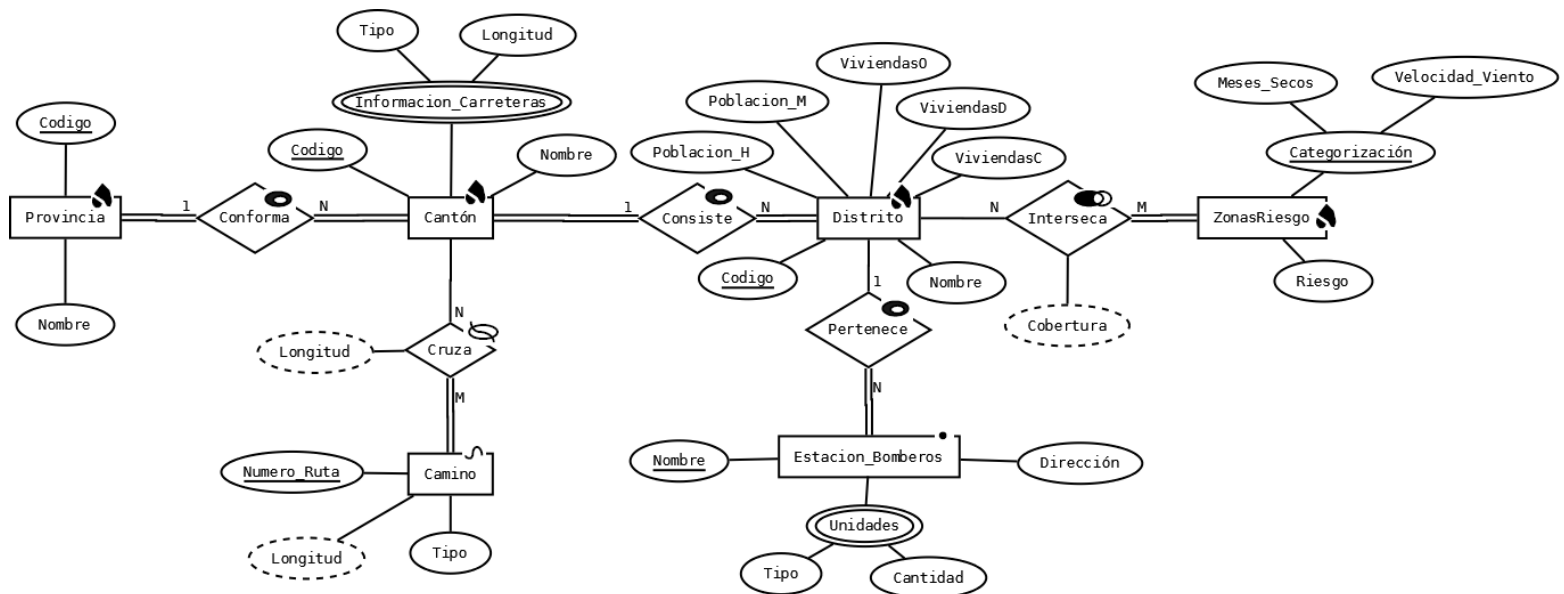


Imagen 1: Diagrama Entidad-Relación que corresponde al Diseño Conceptual

LEYENDA DE ELEMENTOS ESPACIALES

Como no existe un estándar para el uso de símbolos que representan elementos de una base de datos con extensión espacial, se utilizó la siguiente leyenda para dar significado a cada símbolo.

Geometrías	
•	Punto
↷	Línea
🔺	Multipolígono
Relaciones Topológicas	
🔄	Cruza
🕸	Dentro
🔗	Interseca

Donde las relaciones topológicas significan lo siguiente:

- Cruza: Intersección entre una geometría de polígono o multi-polígono con una línea
- Dentro: Contención de una geometría de polígono o multi-polígono dentro de otra, de tal forma que ni el borde ni el interior de la geometría contenida tocan el exterior de la geometría contenedora.
- Interseca: Intersección entre una geometría de polígono o multi-polígono con otra.

ORIGEN DE DATOS

Como el sistema de base de datos tiene el origen de sus datos en archivos shapefile y hojas electrónicas, se explicará a continuación aquellas entidades para las cuales se importarán datos, y de qué archivo, y aquellos datos que serán calculados.

PROVINCIA

Todos los datos de provincia serán importados desde el archivo provincias2008crtm05.shp, incluyendo su geometría.

CANTÓN

Todos los atributos de cantón serán importados del archivo cantones2008ctm05.shp, salvo Información_Carreteras, que será tomado de los datos más actualizados de los archivos de Transporte acerca de longitud de carreteras por cantón. La relación con Provincia se importará también del archivo shapefile.

DISTRITO

Los atributos Código, Nombre, su geometría y su relación con Cantón, serán importados del archivo distritos2008crtm05.shp. El resto de atributos, acerca de población y viviendas, serán importados del archivo de hoja electrónica que corresponde a estos datos, tomando únicamente los datos más recientes.

CAMINO

El número de la ruta, las geometrías de caminos como así su tipo serán importados directamente del archivo redcamino2008crtm05.shp. Sin embargo, a este es necesario aplicarle un proceso de limpieza de datos, dado que no para todos los caminos existe un número de ruta, por lo que este, junto con la longitud del camino, será calculado.

ESTACION_BOMBEROS

Todos los datos de esta entidad serán importados desde bomberos2008crtm05.shp.

ZONASRIESGO

Nuevamente, todos sus atributos provendrán de riesgoincendiocrtm05.shp. Se decidió que Meses_Secos y Velocidad_Viento serían la llave de esta entidad, dado que estos determinan Riesgo en casi todos los casos. En los casos que no se determina se hará una combinación de los dos valores de Riesgo. Como en el shapefile los datos están dispersos, al momento de importarlos se hará una unión de las geometrías, agrupándolas por el campo llave.

CRUZA

Para la relación entre Cantón y Camino, se calculará que caminos atraviesan qué cantones y la longitud de esta intersección, y se almacenarán estos datos, mediante una consulta en la base de datos.

INTERSECA

Similarmente como la relación anterior, se calculará qué zonas intersecan con qué distritos, y el área de la intersección.

ESQUEMA LÓGICO

Con el motivo de facilitar la implementación, típicamente se realiza un mapeo del esquema conceptual a un esquema lógico. En este caso, se realiza un mapeo al modelo relacional con extensión de datos espaciales. Las tablas relacionales producto de este mapeo, con sus justificaciones de estar en tercera forma normal, son las siguientes:

Provincia			
Nombre	Codigo	Nombre	Frontera
Tipo	String	String	geometria
Long	1	10	multipoligono
Restriccion	PK	Not null	

Tabla 8: Relación Provincia

La tabla está normalizada en tercera forma normal dado que la única dependencia funcional que se da, aparte de las de la llave primaria, es que Nombre puede determinar Código, pero no ocurre un problema dado que la llave es un solo atributo.

Canton				
Nombre	Codigo	Nombre	Frontera	Codigo_Provincia
Tipo	String	String	geometria	String
Long	3	20	multipoligono	1
Restriccion	PK	Not null		FK1

Tabla 9: Relación Canton

La tabla está normalizada en tercera forma normal dado que la única dependencia funcional que se da, aparte de las de la llave primaria, es que Nombre puede determinar Código, pero no ocurre un problema dado que la llave es un solo atributo.

FK1 = Provincia

Distrito				
Nombre	Codigo	Nombre	Poblacion_H	Poblacion_M
Tipo	String	String	Integer	Integer
Long	5	20	Integer	Integer
Restriccion	PK	Not null	> 0	> 0

Nombre	ViviendasO	ViviendasD	ViviendasC	Frontera	Codigo_Canton
Tipo	Integer	Integer	Integer	geometria	String
Long	Integer	Integer	Integer	multipoligono	3
Restriccion	> 0	> 0	> 0		FK1

Tabla 10: Relación Distrito

La tabla está en tercera forma normal dado que la llave primaria es de un solo atributo y que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Canton

Camino				
Nombre	Numero_Ruta	Tipo	Longitud	Forma
Tipo	String	String	Double	geometria
Long	15	15	Double	linea
Restriccion	PK	Not null	Calculado	

Tabla 11: Relación Camino

La tabla está en tercera forma normal dado que la llave primaria es de un solo atributo y que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

Estacion_Bomberos				
Nombre	Nombre	Direccion	Posicion	Codigo_Canton
Tipo	String	String	geometria	String
Long	25	100	punto	5
Restriccion	PK			FK1

Tabla 12: Relación Estacion_Bomberos

La tabla está en tercera forma normal dado que la llave primaria es de un solo atributo y que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Canton

Zonas_Riesgo

Nombre	Meses_Secos	Velocidad_Viento	Riesgo	Frontera
Tipo	Integer	String	String	geometria
Long	Integer	10	10	multipoligono
Restriccion	PK	PK		

Tabla 13: Relación Zonas_Riesgo

La tabla está en tercera forma normal dado que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

Interseca

Nombre	Codigo_Distrito	Meses_Secos	Velocidad_Viento	Cobertura
Tipo	String	Integer	String	Double
Long	5	Integer	10	Double
Restriccion	PK FK1	PK FK2	PK FK2	> 0

Tabla 14: Relación Interseca

La tabla está en tercera forma normal dado que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Distrito

FK2 = Zonas_Riesgo

Cruza

Nombre	Codigo_Canton	Numero_Ruta_Camino	Longitud
Tipo	String	String	Double
Long	3	15	Double
Restriccion	PK FK1	PK FK2	> 0

Tabla 15: Relación Cruza

La tabla está en tercera forma normal dado que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Canton

FK2 = Camino

Unidades_Estacion_Bomberos

Nombre	Nombre_Estacion	Cantidad	Tipo
Tipo	String	Integer	String
Long	25	Integer	10
Restriccion	PK-FK1	> 0	PK

Tabla 16: Relación Unidades_Estacion_Bomberos

La tabla está en tercera forma normal dado que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Estacion_Bomberos

Informacion_Carreteras_Canton

Nombre	Codigo_Canton	Longitud	Tipo
Tipo	String	Double	String
Long	3	Double	10
Restriccion	PK-FK1	> 0	PK

Tabla 17: Relación Informacion_Carreteras_Canton

La tabla está en tercera forma normal dado que no existen dependencias funcionales a partir de otros atributos que no sean la llave primaria.

FK1 = Canton

ANEXOS

Como anexos se incluirá las tres versiones del diagrama Entidad-Relación revisado por la profesora Elzbieta Malinowski G.