



UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Iztapalapa

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA
DEPARTAMENTO DE MATEMÁTICAS

**Diseño de un modelo para identificar patrones de
conducta en los niveles de Trata de Personas en el país**

Tesis que presenta:

Paulina Martínez Rosas

Para obtener el grado de:

Maestra en Ciencias

(Matemáticas Aplicadas e Industriales)

Dirigida por:

Dra. Blanca Rosa Pérez Salvador

Jurado calificador:

Presidente:

Secretaria:

Vocal:

México, D.F., a 24 de noviembre del 2015

A mis papás

Agradecimientos

Agradezco a la Universidad Autónoma Metropolitana Unidad Iztapalapa, que me abrió sus puertas ofreciéndome la oportunidad de desarrollarme y estudiar, en un marco de prestigio, calidad y confianza.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por brindarme la oportunidad de participar, como becaria, en el Programa Nacional de Posgrados de Calidad (PNPC) durante la realización de mi posgrado.

Así mismo, quiero expresar mi agradecimiento a mi directora de tesis: Dra. Blanca Rosa Pérez Salvador, por la confianza que depositó en mí, por su paciencia y por todo su apoyo a lo largo de la realización de mi posgrado.

De manera muy especial quiero dar las gracias a mis papás: Pablo Martínez Rivera y Alicia Rosas Téllez, por brindarme su comprensión, cariño y apoyo. Porque gracia a sus consejos y sacrificios no hubiese logrado llegar hasta este punto de culminar mi carrera. Los amo.

Finalmente, quiero agradecer a todos mis amigos y a todas las personas que me han acompañado y ayudado en este camino.

Índice general

Introducción	ix
1. El fenómeno de la trata de personas en México	1
1.1. Antecedentes	1
1.2. Formas de explotación	2
1.3. El perfil de las víctimas en México	4
1.4. Perfil de los Tratantes y de los grupos delictivos vinculados con la Trata de personas	5
1.5. Modus Operandi	7
1.5.1. Captación o reclutamiento	7
1.5.2. Traslado	8
1.5.3. Explotación	9
1.6. Las ganancias económicas	9
1.7. Rutas	9
1.8. Estadísticas de incidencia	12
1.9. Dificultades del estudio de la trata de personas	14
1.10. La Experiencia de la investigación documental	14
2. Aprendizaje no supervisado	17
Introducción	17
2.1. Regla de asociación	18
2.1.1. Análisis de la canasta de mercado	19
2.1.2. Algoritmo Apriori	20
2.2. Análisis de Conceptos Formales	23
2.2.1. Contexto y concepto formal, concepto de retícula	23

Índice general

3. Uso de modelos no supervisados para identificar sospechosos de trata de personas	29
3.1. Identificación de sospechosos de trata en los países bajos	29
3.2. Método de Monte-Carlo para analizar la Trata en México	46
3.2.1. Variables	46
3.2.2. Simulación	48
4. Resultados y conclusiones	59
A. Apéndice	63
A.1. Catálogo de Riesgo por estado relacionado con la trata de personas en México	63
A.2. Investigación Documental	65
Bibliografía	69

Introducción

La Comisión Nacional de los Derechos Humanos (CNDH, 2013) menciona que la trata de personas es un fenómeno delictivo que se encuentra extendido por todo el mundo; miles de personas son víctimas de este delito, particularmente mujeres, niños y niñas son capturados, trasladados, vendidos y comprados con fines de explotación. La forma más conocida de la trata de persona es la explotación sexual, cientos de miles de víctimas también son objeto de trata con fines de trabajo forzoso, servidumbre doméstica o mendicidad infantil. El delito de trata de personas afecta casi a todos los países de todas las regiones del mundo. Entre 2010 y 2012 se identificaron víctimas de 152 nacionalidades diferentes en 124 países de todo el mundo. La trata de personas es un delito de carácter regional e internacional, ya que las víctimas son objetos de trata dentro de su propio país o son trasladadas a países vecinos o llevadas a otros continentes. La mayoría de las víctimas son extranjeras en el país donde se identifican como víctimas. Es decir, esas víctimas (6 de cada 10) son trasladadas cruzando al menos una frontera nacional.

México es un país de origen, tránsito, destino y recepción de migrantes, lo cual contribuye su situación geográfica. Esta situación ha contribuido a aumentar la vulnerabilidad de millares de personas, sobre todo mujeres, niños, niñas y adolescentes, que están expuestos a ser víctimas de la trata de personas, principalmente con fines de explotación sexual y laboral, y es considerado como una modalidad contemporánea de esclavitud.

UNODC (2013) menciona que, se cree que de todos los casos de trata registrados a nivel mundial en 2006, el 79 % correspondió a trata con fines de explotación sexual, en tanto que, al parecer, el 21 % restante de las víctimas fue objeto de trata con fines de trabajos forzoso u otras formas de explotación.

Motivación

No es una tarea fácil analizar el fenómeno de trata de persona debido a la complejidad y a la evolución que ha tenido con la aparición de nuevas tecnologías. Para estudiar este fenómeno se requiere de modelos no supervisados con el fin de poder identificar a presuntos tratantes y sus atributos.

En la literatura podemos encontrar numerosas investigaciones desde el punto de vista de las Ciencias Sociales. Sin embargo, la mayoría sólo describe el problema sin intentar dar un modelo que ayude a resolver este problema. La única investigación que encontramos que propone una metodología matemática para analizar el problema fue el artículo *Semi-automated Knowledge discovery: identifying and profiling human trafficking* (Poelmans et al, 2012). En esta investigación se utilizaron datos de la policía de Amsterdam para identificar posibles sospechosos que con una investigación más profunda les llevo hacer arresto reales, siguiendo esta metodología, en esta tesis pretendemos adaptar el modelo para ser utilizado con datos de algunas instituciones mexicanas.

En México, la trata de personas es uno de los delitos de mayor impacto humano y social, por lo que, la Procuraduría General de la República (PGR), o la Secretaría de Hacienda, o las asociaciones dedicadas a perseguir este delito, tiene como fin disminuirlo y evitar que las víctimas sufrán una situación psicológica y que pueda llegarles ha provocar la muerte.

Nuestro trabajo pretender apoyar en este sentido a las organizaciones que lo requieran.

Objetivos

El propósito de este trabajo es analizar el fenómeno de la trata de personas y proponer una metodología que ayude a identificar sujetos con probables elementos relacionados con este delito en el país. Este trabajo persigue 3 objetivos principales.

1. Proponer un modelo conceptual que describa las conductas de trata de personas con base en casos conocidos y documentados que aplique a nivel nacional. Aquí se va a revisar información documental, tanto nacional como internacional.
2. Construir bases de datos e indicadores relacionados con la actividad tomando en cuenta información de acceso público y de acceso restringido o privado con datos reales o simulados.

3. Obtener un modelo matemático estadístico en función de variables explicativas de sujetos con posibles elementos relacionados al delito de trata de personas y el movimiento de sus ganancias ilícitas. Los datos serán analizados mediante un modelo no supervisado.

Estructura

La estructura de esta tesis consta de 4 capítulos. En el capítulo 1 presentamos el problema de Trata de personas en México, así como, los perfiles de las víctimas y de los tratantes. También citamos el modus de los tratantes que ejecutan este delito. De igual manera, mencionamos las rutas utilizadas por estos sujetos. En el capítulo 2 nos enfocamos en el estudio de dos tipos de modelos del aprendizaje no supervisado: la canasta de mercado y el análisis de conceptos formales (FCA, por sus siglas en inglés), para identificar objetos que presentan atributos comunes (tratantes y sus atributos). En el capítulo 3 revisaremos a detalle el artículo *Semi-automated Knowledge discovery: identifying and profiling human trafficking* escrito por Poelmans et al (2012). Luego implementaremos computacionalmente el modelo con datos simulados usando la información para la identificación de operaciones con recursos de procedencia ilícita que está a disposición de las instituciones gubernamentales pero que no es de acceso público por ser información de carácter confidencial, por lo que simularemos sus valores. En el capítulo 4 mostraremos algunos resultados concluyentes y las conclusiones obtenidas a lo largo de la elaboración de este trabajo.

Finalmente, se incluye en la tesis un Apéndice, en el que se presenta un Catálogo de Riesgo por Estados relacionado con la trata de personas en México, y así mismo, se presentan los resultados obtenidos de casos reales que han sido publicados en los periódicos y los noticieros.

Capítulo 1

El fenómeno de la trata de personas en México

1.1. Antecedentes

La trata de personas es una violación de derechos humanos que actualmente se manifiesta como un grave problema de carácter internacional, tiene lugar en todas las regiones del mundo y muchos países se ven afectados, ya sea como países de origen, de tránsito o de destino, el movimiento de migración es una problemática que también propicia a este delito. Así mismo, es el tercer negocio ilícito más lucrativo tras el tráfico de drogas y de armas.

La Organización de las Naciones Unidas (ONU, 2002) define la trata de personas como: la captación, el transporte, el traslado, la acogida o la recepción de personas, recurriendo a la amenaza o al uso de la fuerza u otras formas de coacción, al rapto, al fraude, al engaño, al abuso de poder o de una situación de vulnerabilidad o a la concesión o recepción de pagos o beneficios para obtener el consentimiento de una persona que tenga autoridad sobre otra, con fines de explotación. Esa explotación incluirá, como mínimo, la explotación de la prostitución ajena u otras formas de explotación sexual, los trabajos o servicios forzados, la esclavitud o las prácticas análogas a la esclavitud, la servidumbre o la extracción de órganos.

El fenómeno de la trata de personas responde a diversos factores relacionados, principalmente, con la oferta y demanda de mano de obra barata, de las desigualdades económicas y de los problemas socioculturales aún no resueltos, como la discriminación y la violencia de género. Otras cuestiones, como la ubicación

La Trata de personas en México

geográfica o el contexto migratorio general, influyen también en la incidencia del delito.

Fondo de las Naciones Unidas para la Infancia (UNICEF, 2010) menciona que, la trata de personas es una práctica muy antigua en la historia de la humanidad, pero a fines del siglo XIX y principios del siglo XX que se reconoció a la trata de personas como un problema social producto de secuestros, engaños o coacciones sobre mujeres vulnerables y que se denominó como Trata de Blancas. En este mundo que actualmente vivimos, la trata de personas se ha convertido en una nueva forma de esclavitud del siglo XXI.

1.2. Formas de explotación

En el delito de la trata de personas implica una relación entre la víctima y el tratante con ciertos elementos que son fundamentales.

Según Women's Link Worldwide (2010) cita que la trata de personas tiene tres elementos:

1. Captación, transporte, traslado, acogida o recepción de una persona.
2. Control de la víctima por medio de: la amenaza o el uso de la fuerza u otras formas de coacción, se aprovechan de la vulnerabilidad de la víctima, o la concesión o recepción por pagos y preventas de la persona que tenga autoridad sobre la víctima.
3. Finalidad de explotación prostitución ajena u otras formas de explotación sexual, trabajos o servicios forzados, esclavitud o prácticas análogas a la esclavitud, servidumbre, o extracción de órganos, entre muchas otras.

Cuando la víctima es menor de edad, basta con que se dé el acto del punto 3, para calificar el hecho como trata de personas, sin que se requiera que este presente los puntos 1 y 2 señalados anteriormente.

La base de este fenómeno son las diferentes formas de explotación. INM, OIM y CIM-OEA (2006) cita que las formas y mecanismo de trata son:

Formas o Mecanismo	Sectores
Laboral	Fábricas, maquiladoras. Trabajo agrícola, plantaciones. Minas, construcciones, pesca. Mendicidad. Trabajo domestico. Vientres de aquiler.
Sexual	Prostitución Forzada. Pornografía (película, fotos, internet). Pedofilia. Turismo sexual. Agencia matrimoniales. Embarazos forzados.
Falsas Adopciones	Venta de niños.
Servidumbre	Prácticas religiosas y culturales. Matrimonios serviles.
Militar	Soldados Cautivos. Niños soldados.
Tráfico de órganos	Sustracción ilícita de órganos, tejidos o componentes (pulmón, riñon, cornea, hígado, corazón, etc..) para ser vendidos en el mercado negro.
Práctica esclavista	Captura, adquisición o cesión de un individuo para explotación o servilismo.

En forma más general, hay dos tipos de trata de personas, las cuales son: trata externa, en que se caracteriza por la captación que ocurre en el país de origen o residencia de la víctima y el traslado a otro país para su explotación. En el proceso de traslado hay cruces de fronteras marítimas, aéreas o terrestres por parte de las víctimas. Trata interna, lo cual se caracteriza por el proceso de captación, traslado y explotación que se da dentro de las fronteras del país.

1.3. El perfil de las víctimas en México

Según UNODC (2013), en México la trata de personas es mayormente para explotación sexual y las víctimas se caracterizan por:

- Belleza física: el cuerpo como condición de vulnerabilidad cuando reúne los criterios establecidos por la demanda de determinado mercado sexual.
- Sexo femenino: el cuerpo femenino es mucho más utilizado que el masculino para fines de explotación sexual.
- Edad mínima: se prefieren a las niñas, adolescentes y mujeres jóvenes.
- Clase social baja y media: los bajos ingresos personales o familiares orillan a las personas a aceptar cualquier tipo de empleo o que las puedan venderlas.
- Escolaridad: la falta de estudios y la escolaridad básica las ponen en riesgos de caer en la trata de personas.
- Estado civil: las mujeres solteras pueden estar en mayor situación de vulnerabilidad.
- Contar con hijos puede aumentar la urgencia de conseguir un trabajo para mantenerlos económicamente.
- Ocupación: estudiantes, empleadas domésticas, meseras y obreras suelen ser reclutadas por tratantes.
- Estatus legal: su condición migratoria los ponen en situación de vulnerabilidad.
- Grupo étnico: las niñas, adolescentes y mujeres adultas indígenas pueden estar en una mayor situación de vulnerabilidad.
- Adicciones: las personas que sufren algún tipo de adicción a drogas o alcohol pueden entrar en situaciones de vulnerabilidad.
- Región: suelen estar en posición de vulnerabilidad las mujeres que habitan comunidades rurales donde el Estado está ausente en términos de seguridad, empleo y protección social, así como mujeres que habitan contextos urbanos marginales.
- Antecedentes de violencia: la presencia de violencia familiar puede ser un factor de expulsión de la familia y colocar a las mujeres en situación de vulnerabilidad.

1.4. Perfil de los Tratantes y de los grupos delictivos vinculados con la Trata de personas

La trata de personas es un negocio muy rentable económicamente, de acuerdo con la Convención de las Naciones Unidas contra la Delincuencia Organizada Transnacional y muchos casos se cometen por redes bien organizadas de delincuentes que en algunos casos tienen agencias de viajes o de empleos.

La expresión “tratante” se refiere a: quienes se dediquen a la captación y el transporte de personas, quienes ejerzan control sobre las víctimas de la trata, quienes las trasladen o mantengan en situaciones de explotación, quienes participen en delitos conexos y quienes obtengan un lucro directo o indirecto de la trata, sus actos constitutivos y los delitos conexos (ONU, 2002).

Funciones y roles de los tratantes de personas en las distintas fases de delito de Trata de personas (UNODC, 2013):

RECLUTAMIENTO	<p>Personas independientes, a veces sin tener ninguna conexión con las víctimas. Ejemplos de esta forma de reclutamiento son falsas agencias de modelos, agencias de viajes, oficinas de empleo que dan falsas promesas como niñeras, trabajadores, <i>au-pairs</i>, etc.</p> <p>Personas de la comunidad de origen de las víctimas (amigos, familiares).</p> <p>En algunos casos son los propios “padrotes” los que participan en la etapa de reclutamiento.</p>	<p>Enganchan a sus víctimas haciéndoles creer que se puede conseguir trabajo en otros lugares.</p>
TRASLADO	<p>Individuos vinculados con redes o que trabajan a título personal (guías).</p> <p>Este tipo de función no requiere una gran especialización.</p> <p>Pueden ser profesionales que trabajan en sector del transporte de mercancías, del correo, etc.</p> <p>En el caso de operaciones transnacionales, el cruce de fronteras implica tareas más específicas y sofisticadas.</p>	<p>Llevan a las víctimas de trata desde y hacia ciertos puntos o en lugares distintos de la explotación. En el caso del componente transnacional, ayudan al traslado antes y después de cruzar fronteras.</p> <p>En este caso, las “guías” acompañan a las víctimas a través de fronteras, ya sea a pie, en barco o por cualquier otro medio de transporte. Conocen las rutas y caminos, los puntos de patrulla y la cantidad correcta de soborno de las autoridades.</p>

La Trata de personas en México

EXPLOTACIÓN	Concurren varias figuras. Pueden ser los propietarios de un burdel, dueños de clubes nocturnos, dueños de una plantación, un abusador de niños que ha comprado a un niño o niña, un médico que le roba los órganos de las víctimas, etc.	Explotar y controlar a la víctima
OTRAS ACTIVIDADES Y ACTORES TRANSVERSALES	<p>Falsificación de documentos Profesionistas, mediadores, ser oficiales de gobierno, ser funcionarios de migración o por corrupción, funcionarios aun en servicio.</p> <p>Halcones, guardias, controladores: figuras de la red delincencial que intervienen en el control de las víctimas.</p>	

Es citado en UNODC (2013), que el delito de trata de personas, generalmente, se lleva a cabo, por las redes de grupos delictivos del primer nivel y, en el caso de la falsificación de documentos, con grupos muy organizados (segundo y tercer nivel) que con frecuencia involucran a las conexiones familiares y de negocios entre los mexicanos y los contactos en el exterior. Con respecto a los delincuentes a pequeña escala, reclutan a una o un pequeño número de las víctimas a través de sus redes sociales y familiares.

Bajo este esquema, se puede decir que los tratantes en México se han identificado en tres niveles diferentes. El primer nivel incluye a familias en situación de vulnerabilidad que puedan vender, rentar o prestar a sus hijas u hijos para conseguir dinero más fácilmente. En este primer nivel participan las redes de familiares, vecinos o parientes que viven cerca de la víctima. El segundo nivel participan grupos locales, miembros de pequeños grupos de delincuentes y criminales que operan individualmente o por grupos más establecidos (en el caso Tlaxcala). El tercer nivel participan los miembros de grandes grupos delictivos organizados como son los cárteles del narcotráfico.

Los tratantes son principalmente hombres, aunque las mujeres también participan en este delito y su participación está en aumento y en algunas regiones pueden desempeñar un papel importante en la cadena delictiva.

La mayoría de los grupos delictivos también están vinculados con choferes de camiones o taxistas, los cuales pueden llevar a las víctimas desde el lugar de origen hasta el destino de explotación.

La Procuraduría General de la República (PGR) establece que la trata de personas está ligada a grupos delictivos, principalmente a organizaciones relacionadas con el narcotráfico.

1.5. Modus Operandi

Se menciona en UNODC (2013) que, la trata de personas en México se explica por diversos factores, como las condiciones de vulnerabilidad socioeconómica de las víctimas, así como del contexto político, sociocultural y económico de los lugares en los cuales se lleva a cabo este delito. Estas condiciones de vulnerabilidad que se manifiestan en espacios socioterritoriales bien definidos son aprovechadas por los tratantes y constituyen una explicación de su modus operandi.

1.5.1. Captación o reclutamiento

Las formas de reclutamiento que se han identificado en el país incluyen principalmente:

- El engaño
- La seducción o enamoramiento
- La captación con ayuda de otras víctimas
- Internet
- Secuestro
- La venta (particularmente en el caso de niños y niñas)
- La adopción ilegal.

Otras formas de enganche

Otras formas de reclutamiento son las personas migrantes, algunas de estas personas que son objeto de tráfico ilícito de migrantes son obligadas a permanecer en países de tránsito, lo cual las hace vulnerables a la trata de personas.

1.5.2. Traslado

Después de que la víctima es reclutada será trasladada al lugar de destino para su explotación. Estos traslados se pueden hacer con o sin pasaportes, visas y documentos de identidad oficiales, pero mayormente se utiliza documentación falsa. En muchos casos, las víctimas son obligadas a llevar credenciales de elector falsas, lo que dificulta la identificación de menores de edad al momento de los operativos e inspecciones de las autoridades. Las víctimas normalmente son transportadas por algún medio público de transporte, entre la etapa de captación y traslado, cuando aún existe la confianza entre la víctima y el enganchador (en el caso de la trata con fines de explotación sexual), el enganchador puede mandar a la víctima sola y prometerle que la alcanzará más tarde en el destino final. Sin embargo, quien finalmente recibe a la víctima es un amigo o familiar del enganchador, lo cual él suele ser un tratante o dueño de un prostíbulo donde la víctima será explotada. En otros casos, el propio enganchador puede ser el que acompañe a la víctima durante el traslado, se quede un tiempo con ella en el destino de explotación y luego regrese a reclutar nuevas víctimas.

Mientras que sea fácil el acceso a documentos falsos, esto constituye una herramienta indispensable para muchos tratantes. Con la falta de documentos puede ser un factor que aumente la vulnerabilidad de ciertos grupos poblacionales a ser víctimas de trata de personas.

Los tratantes mediante la corrupción pueden conseguir documentos falsos y también pueden comprar voluntades en distintos niveles y órdenes de gobierno. Por ejemplo, en el caso de la trata de personas con fines de explotación sexual, esta no sería posible:

- Si las autoridades locales no toleraran la operación ilícita
- Si los operadores del transporte público fueran capaces (y tuvieran la voluntad) de identificar a una posible víctima de trata de personas
- Si el personal de un hotel identificara posibles situaciones de trata de personas que se registran en sus espacios de trabajo
- Si un taxista no le diera información a su cliente sobre dónde están las zonas rojas
- Si un policía auxiliar no se hiciera “de la vista gorda” ante una posible situación de trata de personas

- Si un cliente se cuestionara si el “servicio sexual” que compra ha sido posible gracias a una larga red de complicidades que dan vida a la trata de personas y culminan con el momento mismo en que la víctima es explotada.

1.5.3. Explotación

Después que la víctima llega a su destino final, se le obliga a desarrollar actividades como son: trabajo sexual, doméstico u otros, que permitan su explotación. O bien, el explotador confisca sus documentos, o le cobra los gastos de traslado a otra ciudad o país, de esta forma crean una deuda, dado que las víctimas nunca podrán llegar a ganar lo suficiente como para pagar. Los abusos, golpes, violaciones, chantajes y amenazas contribuye a que las víctimas sufran una explotación dolorosa y prolongada.

1.6. Las ganancias económicas

Las ganancias económicas varían dependiendo del lugar específico donde se da la explotación sexual, del perfil de la víctima, del perfil del cliente, los precios y del servicio ofrecido.

Algunas personas que participan y son beneficiadas en este mercado, son:

- Falsificadores de documentos
- Autoridades municipales
- Policías ministeriales
- Abogados
- Empresarios hoteleros

1.7. Rutas

Las rutas que utilizan los tratantes para trasladar a sus víctimas de sus lugares de origen hacia el destino de su explotación, están relacionadas con factores sociales, económicos, culturales, históricos y geográficos. Por ejemplo, es citado en UNODC (2013) que hay municipios y ciudades en Baja California, Chiapas, Guanajuato y Tlaxcala, donde se registran actividades de origen, tránsito y destino.

La Trata de personas en México

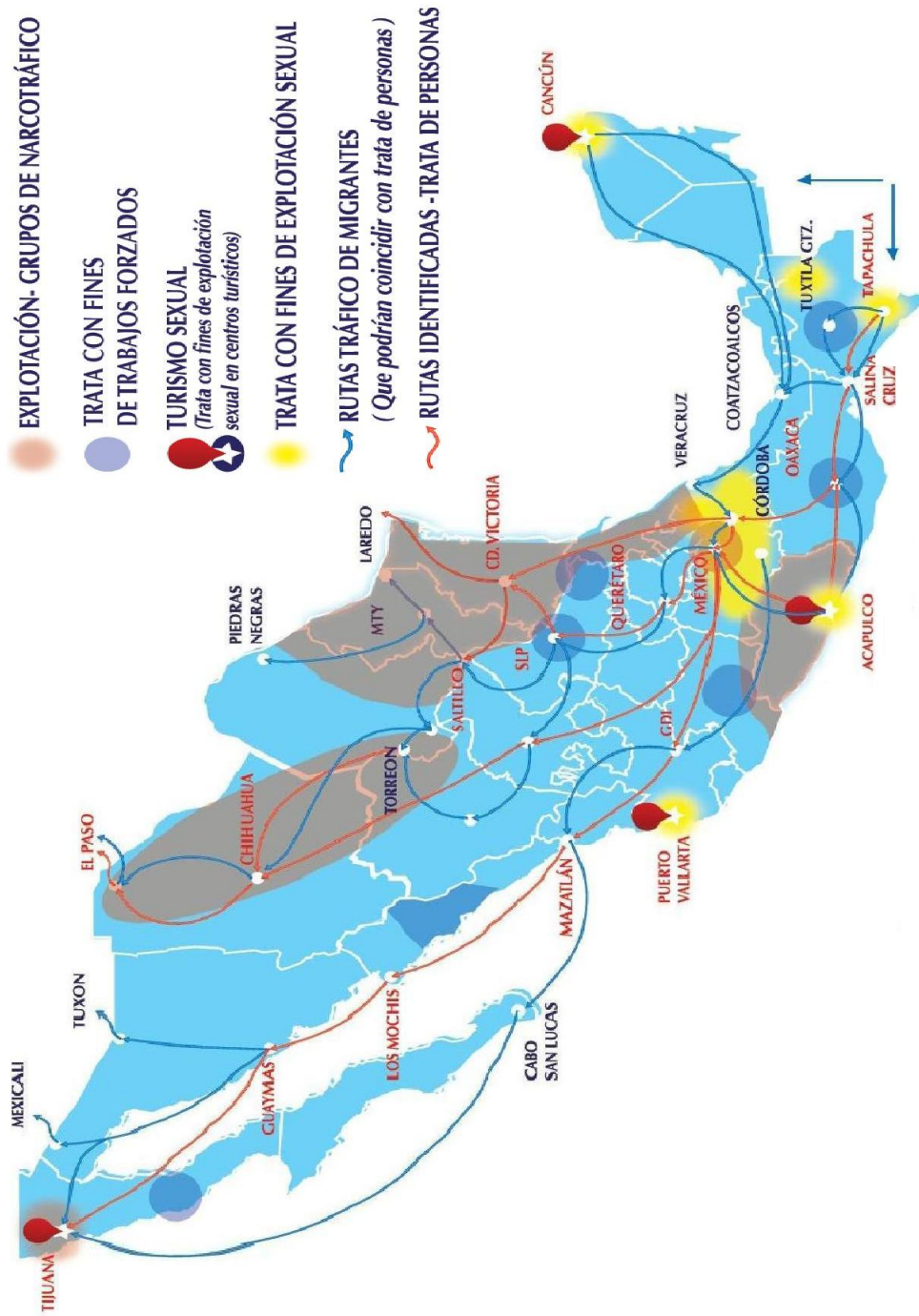
Por otro lado, una parte fundamental del modus operandi de los tratantes es la necesidad de desligar a la víctima de su familia y de sus redes sociales.

La gran extensión de México, su orografía y sus múltiples puertos de entrada terrestre, marítimos y aeroportuarios, hacen difícil establecer con precisión las rutas que utilizan los tratantes.

De acuerdo a la CNDH (2013), los estados considerados como “alta incidencia de trata de personas” son:

- Aguascalientes
- Baja California
- Baja California Sur
- Chiapas
- Chihuahua
- Colima
- Ciudad de México (Distrito Federal)
- Estado de México
- Guanajuato
- Guerrero
- Hidalgo
- Jalisco
- Michoacán de Ocampo
- Morelos
- Nayarit
- Oaxaca
- Puebla
- Quintana Roo
- Sonora
- Tabasco
- Tlaxcala
- Veracruz de Ignacio de la Llave

Se identificó rutas de trata de personas, estas rutas se muestra a continuación en el siguiente mapa (UNODC, 2013):



1.8. Estadísticas de incidencia

Se indica en Aztecanoticias (2014) que, México ocupa el segundo lugar en el mundo, en casos de trata de personas. De acuerdo con datos del Instituto Nacional de Estadística y Geografía (INEGI) cada año son explotados sexualmente entre 16 mil y 20 mil niños y niñas en México (Universal, 2010).

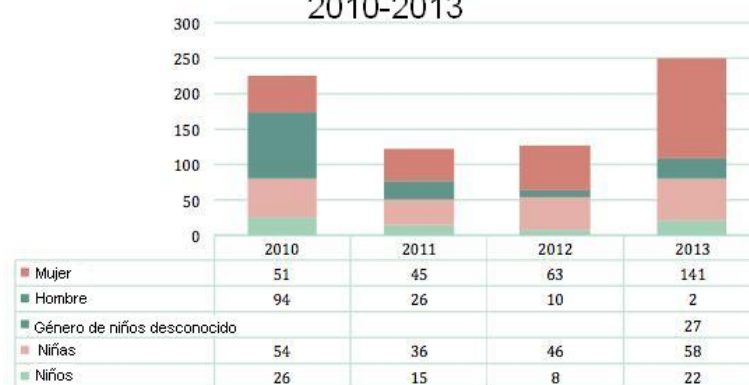
Las víctimas de trata de personas en el país, se duplicaron de 2012 a 2013 al pasar de 127 víctimas identificadas a 250, informó UNODC (2013).

Conforme a los datos oficiales que entregaron a la UNODC para su informe Mundial sobre la Trata de Personas 2014, el Instituto Nacional de Migración, la Fiscalía Especializada para la Violencia contra la Mujer y el Tráfico de Personas, la Subprocuraduría especializada en investigación de delincuencia organizada y la Unidad Especializada en Investigación de tráfico de menores, personas y órganos, que el número de niños y niñas víctimas de trata en el país aumentó 56.8 % desde 2011.

Se menciona en Víctimas de trata en México (2014) que, los datos que esa agencia de la ONU reunió sobre México indican que en 2011 se conoció el caso de 122 víctimas de trata de personas, 127 un año después y 250 en 2013; por lo que la cifra aumentó 104 %.

Según México unido contra la delincuencia (2014) cita que, la mayoría de las víctimas son mujeres (56 %), seguido de niñas (23.2 %), pero hay un aumento importante sobre el número de niños identificados como víctimas que ha pasado de 8 a 22.

Número de víctimas de la trata, detectados por edad
2010-2013



Fuente : Instituto Nacional de Migración (INM)/ Oficina de la Fiscalía Especializada para la Violencia contra la Mujer y el Tráfico de Personas (Fevimitra)/Subprocuraduría especializada en investigación de delincuencia organizada (SEIDO PGR)/Unidad Especializada en Investigación de tráfico de menores, personas y órganos.

El incremento de las denuncias se han mantenido año con año. El 2008 se registró 25 casos para el 2010 sumaron 97 averiguaciones, esto es, el triple. En 2014 el número de casos ascendió hasta los 178 lo que significa un crecimiento del 602 % en la incidencia de este delito.

La UNODC (2013) también señala que entre 2011 y 2012 sólo siete personas fueron condenadas por el delito de trata de personas; y aunque para el año siguiente la cifra aumentó a 56, sigue siendo muy inferior al número de investigaciones en curso.

La PGR le reportó a la UNODC (2013) que, el año pasado inició un proceso legal contra 215 personas por el delito de trata de personas, vinculadas a 136 casos, pero sólo el 26 % recibieron condena.

La mayoría de las víctimas de este delito fueron encontradas entre 2010 y 2013 en México, de las cuales, 354 mexicanas, 204 guatemaltecas, 76 hondureñas, 20 colombianas y 20 haitianas. En este tiempo, se reportaron 402 casos de víctimas para fines de explotación sexual, 259 casos de trabajo forzado y 33 casos de explotación laboral.

Número de víctimas de la trata, detectados por forma de explotación, 2010-2013



Fuente: Instituto Nacional de Migración (INM)/ Oficina de la Fiscalía Especializada para la Violencia contra la Mujer y el Tráfico de Personas (Fevimitra)/Subprocuraduría especializada en investigación de delincuencia organizada (SEIDO PGR)/Unidad Especializada en Investigación de tráfico de menores, personas y órganos.

1.9. Dificultades del estudio de la trata de personas

Se menciona en Requena et al (2012) que, la trata de persona presenta características que dificultan su estudio. Las víctimas son trasladadas desde su origen hasta su destino. Es claro, que este hecho complica a las investigaciones de esta actividad debido a que estos delincuentes se encuentran distribuidos por todo el país y las víctimas se encuentran desorientadas con respecto a su localización.

La trata de persona es un delito que se efectúa clandestinamente, es decir, es difícil obtener evidencias sobre su existencia, dimensión, personas que ejercen este delito e incluso víctimas, debido al entorno donde se elabora, por ejemplo, clubes nocturnos donde las mujeres ejercen la prostitución o cuartos en los que se encuentran alojadas decenas de personas que trabajan en diversos sectores laborales, como la agricultura o el sector textil (en el caso del terremoto de 1985, se evidenció este fenómeno (La Izquierda Diario, 2015)). Este fenómeno delictivo, está evolucionando hacia una mayor complejidad debido al crecimiento de las nuevas tecnologías a nivel mundial. En la actualidad, cualquier persona puede comunicarse con otras personas que se encuentra en otra parte del mundo. En este sentido, resulta cada vez menor el uso de la violencia, lo cual da menos evidencias de la actividad delictiva.

Otras dificultades en su estudio, son las víctimas, pues las personas que sufren o han sufrido no pueden probar este hecho y podrían ser sospechosas en participar, también les dificulta su situación psicológica para denunciar este fenómeno delictivo.

1.10. La Experiencia de la investigación documental

En nuestra investigación documental obtuvimos información de casos reales del delito de trata de personas en México, esta información fue recabada de las notas periodísticas y noticieros del país, y dio como resultado una interpretación del riesgo en los Estados. Así mismo, una interpretación de los bienes que tenían en su poder los presuntos tratantes, de su sexo, de su actividad económica. Igualmente, sobre las víctimas, se obtuvo información de su nacionalidad y la forma de su captación. (Véase Tabla A.2)

Recabar esta información, no fue tarea fácil debido a que la información pública en las notas periodísticas esta incompleta, por ejemplo, algunos casos reales no se da información sobre la nacionalidad de las víctimas, la edad de los tratantes, la colonia de operación, el lugar donde fueron captadas las víctimas, edad de las víctimas, etc.

Aunque la información está incompleta, nos dio ideas sobre las características del fenómeno y ayudó en la conformación de las variables que podrían ser de interés para tratar de identificar sujetos relacionados con el delito de trata de personas o sujetos que obtienen ganancias de manera marginal del mismo.

Capítulo 2

Aprendizaje no supervisado

En este capítulo revisaremos los modelos de regla de asociación, modelo de la canasta de mercado y el análisis de conceptos formales (FCA, por sus siglas en inglés), con el fin de abordar el problema de trata de personas en México.

Introducción

El aprendizaje no supervisado es un proceso automático donde un modelo se ajusta a los datos sin un conocimiento apriori del modelo adecuado. Así el aprendizaje no supervisado trata a los objetos de entrada como un conjunto de variables aleatorias para las cuales se construye un modelo de densidad, esto es, se tiene un conjunto de N observaciones (x_1, x_2, \dots, x_n) de un p vector X aleatorio con densidad conjunta $Pr(X)$. Por ejemplo, suponga que $p = 4$ y $X = (\text{sexo}, \text{estado civil}, \text{escolaridad}, \text{ingreso})$, x_1, x_2, \dots, x_n son los vectores X asociados a n individuos en una base de datos. El objetivo es determinar directamente con los datos de la base algunas propiedades de la función de densidad conjunta $Pr(X)$, sin que podamos medir el grado posible de error. La dimensión de X es a veces mucho mayor que en el aprendizaje supervisado, y las propiedades de interés son a menudo más complicadas de estimar; sin embargo se tiene la ventaja de que no hay otras variables, además de las contenidas en X que pudieran ser necesarias en el estudio. Uno de los principales problemas de los modelos no supervisados es la dimensión de X , ya que el número de posibles valores que puede tomar las coordenadas del vector X puede crecer de manera exponencial. Por ejemplo, en el caso de los vectores $X = (\text{sexo}, \text{estado civil}, \text{escolaridad}, \text{ingreso})$. El sexo tiene 2 posibles valores: femenino y masculino, el estado civil tiene 4 valores: soltero,

casado, divorciado y viudo, la escolaridad tiene 6 valores: sin estudios, primaria, secundaria, bachillerato, licenciatura y posgrado; y el ingreso tiene 4 valores: hasta 3 salarios mínimo, de 4 a 10 salarios mínimo, 10 a 20 salarios mínimo y más de 20 salarios mínimo; el número de posibles combinaciones es $2 \times 4 \times 6 \times 4 = 192$. En este sentido, en el caso de menor dimensión como $p \leq 4$, se puede estimar $Pr(X)$ para todas las posibles combinaciones, sin embargo si p es grande, el número posible de valores de X puede ser muy grande y entonces el caso de estimación puede ser complicado y lo más práctico es considerar la probabilidad para un conjunto de posibles valores de X usando estadísticas descriptivas simples.

2.1. Regla de asociación

Supongamos $X = (x_1, \dots, x_p)$ son los vectores asociados a un conjunto \mathcal{I} de individuos en una base de datos, el objetivo de la regla de asociación es encontrar el conjunto de coordenadas del vector X que en la base de datos tiene una alta frecuencia de ocurrir. El primer paso en la regla de asociación es formar una nueva base con los datos de la base de origen. Por ejemplo, en el vector que ya se vio de 4 coordenadas, $X = (sexo, estado\ civil, escolaridad, ingreso)$, la base tiene 5 columnas, la primera es para el nombre del individuo, la segunda es para el sexo, la tercera es para el estado civil, la cuarta es para la escolaridad y la quinta es para el ingreso, de esta base se genera una nueva base de 17 columnas, esto debido a que cada atributo de las variables x_i se les asocia una columna, de esta manera quedan las 17 columnas: la primera columna es para el nombre del individuo; la segunda y la tercera columna para los posibles valores de la variable “sexo”: masculino y femenino; la cuarta, quinta, sexta y séptima columna para los posibles valores de la variable “estado civil”: soltero, casado, divorciado y viudo; y así para las otras dos variables, quedando la nueva base de datos con 17 columnas.

De manera general, del vector $X = (x_1, \dots, x_p)$ se tiene, que para la coordenada j se define el conjunto S_j , cuyos elementos son los valores que puede tomar la variable x_j , para aclarar considere en el ejemplo ya visto, los conjuntos S_j son:

$x_1 = sexo$ y $S_1 = \{femenino, masculino\}$,

$x_2 = estado\ civil$ y $S_2 = \{soltero, casado, divorciado, viudo\}$,

$x_3 = escolaridad$ y $S_3 = \{sin\ estudios, primaria, secundaria, \dots, posgrado\}$,

$x_4 = ingreso$ y $S_4 = \{hasta\ 3\ salarios\ minimo, \dots, 10\ a\ 20\ salarios\ minimo\}$.

El total de atributos es igual a

$$K = \sum_{j=1}^p |S_j|,$$

donde $|S_j|$ es igual a la cardinalidad del conjunto S_j . Así, el número de columnas en la nueva base de datos es $K + 1$.

En este punto, definimos K variables ficticias para cada individuo en la base de datos; Z_1, Z_2, \dots, Z_K , tal que si el individuo tiene el atributo i se le asigna el valor $Z_i = 1$ y si no lo tiene entonces se le asigna el valor $Z_i = 0$. Con lo que un individuo de la base de datos tendrá asociado un vector $Z = (Z_1, Z_2, \dots, Z_K)$ de ceros y unos. La siguiente tabla ilustra lo explicado arriba.

Individuo	sexo		estado civil				escolaridad	...
	femenino	masculino	soltero	casado	divorciado	viudo	sin estudio	...
1	1	0	0	1	0	0	0	...
2	0	1	0	0	0	1	0	...
3	1	0	0	1	0	0	0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...

Esta forma de manejar la información es útil para el diseño de catálogos, la organización de almacenes, el marketing cruzado con correo, la segmentación de individuos de una base de datos respecto a algunas características comunes entre ellos como, por ejemplo, cuantas personas en la base de datos cumplen que sean solteros y al menos con bachillerato.

En forma general, el objetivo básico del análisis de la regla de asociación es encontrar una colección de valores del conjunto X , tal que cada miembro de esta colección (v_1, v_2, \dots, v_l) sea relativamente de frecuencia alta, con $v_i \in S_i$.

2.1.1. Análisis de la canasta de mercado

El análisis de la canasta de mercado se originó en las transacciones de un supermercado en la caja registradora para determinar que artículos compraban las diferentes personas que eran clientes. En nuestro caso, vamos a considerar las transacciones como los individuos en la base de datos y los artículos u objetos comprados como los atributos asociados a estos individuos y aplicaremos la misma metodología.

En el análisis de la canasta de mercado, consideremos que la base ya está modificada para tener los atributos Z_i tales que $Z_i = 1$ si el producto (atributo) v_i está en la canasta y $Z_i = 0$ si el producto no está en la canasta porque no se compró. El objetivo del análisis de la canasta de mercado es encontrar un subconjunto de números enteros $\mathcal{K} \subset \{1, \dots, K\}$ tal que:

$$Pr \left[\bigcap_{k \in \mathcal{K}} (Z_k = 1) \right] = Pr \left[\prod_{k \in \mathcal{K}} Z_k = 1 \right], \quad (2.1)$$

sea grande. Es decir, estamos interesados en buscar las combinaciones (v_1, \dots, v_p) , $v_i \in S_i$ con alta probabilidad.

Esta es la formulación estandar del problema de la canasta de mercado. El conjunto \mathcal{K} se llama “el conjunto de elementos (conjunto de atributos)” y a su cardinalidad se llama “tamaño” (tenga en cuenta que el tamaño no es mayor al tamaño de la base de datos). El valor estimado de (2.1) es igual a la fracción de individuos que tiene todos los atributos del conjunto \mathcal{K} .

$$\hat{Pr} \left[\prod_{k \in \mathcal{K}} (Z_k = 1) \right] = \frac{1}{N} \sum_{i=1}^N \prod_{k \in \mathcal{K}} z_{ik}, \quad (2.2)$$

donde z_{ik} es el valor de Z_k para el i -ésimo individuo de la base de datos. A la expresión (2.2) se llama “soporte” o “prevalencia” de \mathcal{K} , $T(\mathcal{K})$; esto es, $T(\mathcal{K})$ es la fracción de individuos de la base de datos que tienen los atributos del conjunto \mathcal{K} . Un individuo i para la cual se tiene $\prod_{k \in \mathcal{K}} z_{ik} = 1$ significa que tiene todos los

atributos del conjunto \mathcal{K} .

En la regla de asociación se especifica una cota inferior t y uno busca todos los conjuntos de elementos $\mathcal{K}_l \subset \{1, 2, \dots, K\}$ que se pueden formar y que tengan el soporte mayor que el número t , esto forma el conjunto \mathbf{Q} :

$$\mathbf{Q} = \{\mathcal{K}_l | T(\mathcal{K}_l) > t\} \quad (2.3)$$

Encontrar todos los elementos de \mathbf{Q} es una tarea bastante difícil. Muchos algoritmos han sido desarrollados para encontrarlos en bases grandes de datos, tales como el algoritmo Apriori, (Pang-Ning et al, 2006).

2.1.2. Algoritmo Apriori

El algoritmo Apriori se usa en minería de datos para encontrar reglas de asociación en un conjunto de datos. Este algoritmo se basa en el conocimiento previo o “a priori” de los conjuntos con alta frecuencia, esto sirve para reducir el espacio de búsqueda y aumentar la eficiencia.

Para bases de datos grandes es difícil tener un algoritmo simple para encontrar a los elementos de \mathbf{Q} . El algoritmo Apriori nos permite considerar solo una pequeña fracción de los elementos de \mathbf{Q} a la vez.

El algoritmo Apriori utiliza las propiedades para el *umbral* t , siguientes (Friedman et al, 2008):

- La cardinalidad $|\{\mathcal{K}|T(\mathcal{K}) > t\}|$ es relativamente pequeño.
- Cualquier conjunto de elementos \mathcal{L} que consiste de un subconjunto de los elementos en \mathcal{K} debe tener soporte mayor o igual a la de \mathcal{K} , $\mathcal{L} \subseteq \mathcal{K} \Rightarrow T(\mathcal{L}) \geq T(\mathcal{K})$.

El algoritmo Apriori, se basa en la búsqueda de los elementos de $2^{\{1, \dots, K\}} - \{\emptyset\}$ que satisfacen el *umbral* t , donde 2^A indica el conjunto potencia de A cuya cardinalidad es $2^{|A|}$. Aunque el tamaño del conjunto $2^{\{1, \dots, K\}}$ crece exponencialmente con la cardinalidad de $\{1, \dots, K\}$, es posible hacer una búsqueda eficiente utilizando el principio Apriori, (Pang-Ning et al, 2006), el principio Apriori utiliza el hecho de que si el soporte de un conjunto esta por arriba del *umbral* t , entonces el soporte de todos sus subconjuntos también estan por arriba del *umbral* t .

En general, el algoritmo Apriori es un método recursivo donde en el paso n se calcula el soporte de un subconjunto $\mathcal{K}_l \subset \{1, 2, \dots, K\}$ de n elementos para determinar si sobrepasa el *umbral* t ($T(\mathcal{K}_l) > t$) con la condición de que todos los subconjuntos de $n - 1$ elementos de \mathcal{K}_l sobrepasaron el *umbral* t .

Se establece una relación entre A y B como $A \Rightarrow B$, si cada vez que se tiene los atributos de A con alta probabilidad también se tienen los atributos de B . En esta relación al conjunto A se le llama “antecedente” y el conjunto B se llama “consecuente”. Por ejemplo, si una persona tiene bachillerato y no gana nada, su salario es cero es más probable que sea soltero; esto es, $A = (\text{bachillerato}, \text{sin salario})$ y $B = (\text{soltero})$, y la relación es $A \Rightarrow B$, es decir, $(\text{bachillerato}, \text{sin salario}) \Rightarrow (\text{soltero})$.

El soporte de la relación $A \Rightarrow B$, $T(A \Rightarrow B)$, es la fracción de individuos que tienen los atributos de A y de B ; en otras palabras, $T(A \Rightarrow B) = T(A \cup B)$ (Rodríguez et al, 2009). Para aclarar esté concepto se introduce la notación E_W para el conjunto $W \subset \{1, 2, \dots, K\}$ donde

$$E_W = \{x \in \mathcal{I} | x \text{ tiene los atributos de } W\}$$

$Pr(E_W)$ es la probabilidad de que un individuo elegido aleatoriamente de la base de datos tenga los atributos de W y entonces $\hat{Pr}(E_W) = T(W)$. Consecuentemente, se tiene que

$$T(A \Rightarrow B) = T(A \cup B) = \hat{Pr}(E_A \cap E_B)$$

Capítulo 2. Aprendizaje no supervisado

La confianza de la relación $A \Rightarrow B$ se define como el soporte de la relación $A \Rightarrow B$ entre el soporte de A .

$$C(A \Rightarrow B) = \frac{T(A \Rightarrow B)}{T(A)} \quad (2.4)$$

y que podemos escribir como

$$C(A \Rightarrow B) = \frac{\hat{Pr}(E_A \cap E_B)}{\hat{Pr}(E_A)} \quad (2.5)$$

Finalmente, el "levantamiento" de la relación $A \Rightarrow B$ se define como la confianza de la relación $A \Rightarrow B$ dividido por el soporte de B , es decir,

$$L(A \Rightarrow B) = \frac{C(A \Rightarrow B)}{T(B)} \quad (2.6)$$

Esta es una estimación de la asociación medida $Pr(E_A \cap E_B)/Pr(E_A)Pr(E_B)$. A manera de ejemplo, sea el conjunto de elementos $\mathcal{K} = \{\text{soltero}, \text{bachillerato}, \text{sin salario}\}$ y la relación $(\text{bachillerato}, \text{sin salario}) \Rightarrow (\text{soltero})$. El valor del soporte de está relación es 0.03, esto significa, que el 3% de la base de datos tienen los 3 atributos juntos: bachillerato, sin salario y soltero. Una confianza de 0.82 para está relación implica que el 82% de individuos de la base de datos tienen los atributos: bachillerato y sin salario mínimo, también tienen el atributo de ser soltero; el 18% de individuos en la base de datos tienen los atributos: bachillerato y sin salario pero no tiene el atributo de ser soltero. Si el atributo de ser soltero aparece con un 43% en la base de datos entonces la relación $(\text{bachillerato}, \text{sin salario}) \Rightarrow (\text{soltero})$ tendrá un levantamiento de 1.90.

El objetivo de esté análisis es encontrar reglas de asociación $A \Rightarrow B$ con altos valores en su soporte y en su confianza (2.4), y para lograr esté objetivo se utiliza el algoritmo Apriori. Se establece un *umbral* c de confianza y se encuentra todas las relaciones $A \Rightarrow B$ cuya confianza esta por arriba de este *umbral* c :

$$\{A \Rightarrow B | C(A \Rightarrow B) > c\} \quad (2.7)$$

Para cada conjunto de elementos \mathcal{K} existen $2^{|\mathcal{K}|-1} - 2$ relaciones del tipo $A \Rightarrow (\mathcal{K} - A)$, con $A \subset \mathcal{K}$.

Se considera que una relación es interesante si su soporte y su confianza son mayores que ciertos *umbrales* t y c especificados:

$$T(A \Rightarrow B) > t \text{ y } C(A \Rightarrow B) > c$$

Las reglas de asociación se han convertido en una herramienta popular para el análisis de grandes bases de datos. Esto es cuando los datos pueden ser emitidos en forma de una tabla de contingencia multidimensional. La salida es en la forma de reglas conjuntivas (2.1) que sean fácilmente comprensibles e interpretables. El algoritmo Apriori permite este análisis que debe aplicarse a bases de datos enorme, que no son susceptibles de otros tipos de análisis. Las reglas de asociación se encuentran entre los mayores éxitos de minería de datos.

2.2. Análisis de Conceptos Formales

El Análisis de Conceptos Formales (FCA, por sus siglas en inglés) es un método de análisis de datos que se ha hecho muy popular recientemente en varios dominios. FCA analiza los datos que describen la relación entre un determinado conjunto de objetos y un conjunto determinado de atributos. Este término fue introducido por Rudolf Wille en 1984, empleando tanto la Teoría de Reticulos como la Teoría del Orden, que fue desarrollada por Birkhoff y otros en la década de los treinta. Se basa en el entendimiento de que un concepto filosófico está constituido por dos partes: su **extensión** que consta de todos los objetos (entidades o instancia) pertenecientes al concepto, y su **intensión** que comprende todos los atributos (propiedades o características) compartidos por esos objetos. Estos tipos de datos aparecen comúnmente en muchas áreas de la actividad humana.

FCA produce dos tipos de salida de los datos de entrada. La primera salida del FCA es un concepto de reticulo. Un concepto de reticulo es una colección de conceptos formales en los datos que son jerárquicamente ordenados por un orden parcial. Los conceptos formales son grupos particulares que representan conceptos humanos, como el “coche con todos los sistemas de transmisión de la rueda”, “organismo vivo en el agua”, “número divisible por 3 y 4”, etc. La segunda salida del FCA es una colección de los llamados atributos de implicación. Un atributo de implicación describe una dependencia particular que es válida en los datos como “todo número divisible por 3 y 4 es divisible por 6”, “se retiró cada encuestado con edad mayor de 60”, etc, (Belohlávek, 2008).

2.2.1. Contexto y concepto formal, concepto de reticula

FCA fue introducido como una teoría matemática para modelar, el concepto de “concepto” en términos de la teoría reticular. Para permitir una descripción matemática de extensión e intención, se comienza con la definición de un contexto formal (Ganter, 2002)

Capítulo 2. Aprendizaje no supervisado

Definición 2.1. (*Contexto Formal*). Un contexto formal, es una tripleta (X, Y, I) donde X es un conjunto cuyos elementos son llamados objetos, Y es un conjunto cuyos elementos se denominan atributos e I es una relación binaria entre X e Y , es decir, $I \subseteq X \times Y$. La relación $(x, y) \in I$ afirma “el objeto x tiene el atributo y ”.

Entonces, un contexto formal correspondiente (X, Y, I) consiste en un conjunto $X = \{x_1, \dots, x_n\}$, un conjunto $Y = \{y_1, \dots, y_m\}$, y una relación definida por: $(x_i, y_j) \in I$. El concepto formal se puede representar mediante una tabla, donde cada objetos x_i tiene asignado un renglón y cada atributo y_j tiene asignada una columna, y si el vector $(x_i, y_j) \in I$ entonces en la tabla tiene una \times en la celda correspondiente a la fila i y columna j (Ganter y Wille ,1999).

I	y_1	y_2	y_3	y_4
x_1	\times	\times	\times	\times
x_2	\times		\times	\times
x_3		\times	\times	\times
x_4		\times	\times	\times
x_5	\times			

Por ejemplo, en la tabla de arriba, se tiene un FCA con 5 objetos: $\{x_1, x_2, x_3, x_4, x_5\}$, y con 4 atributos: $\{y_1, y_2, y_3, y_4\}$, tal que, el objeto $\{x_1\}$ tiene los 4 atributos, el objeto $\{x_2\}$ tiene los atributos $\{y_1, y_3, y_4\}$, los objetos $\{x_3, x_4\}$ tiene los atributos $\{y_2, y_3, y_4\}$ y por último el objeto $\{x_5\}$ sólo tiene el atributo $\{y_1\}$.

Definición 2.2. Siendo (X, Y, I) un contexto formal, definimos la operación de derivación ($'$) para $A \subseteq X$ y $B \subseteq Y$:

$$A' = \{y \in Y | (x, y) \in I \text{ para cada } x \in A\},$$
$$B' = \{x \in X | (x, y) \in I \text{ para cada } y \in B\}.$$

La definición anterior significa que A' es el conjunto de todos los atributos del conjunto Y que se aplican sobre todos y cada uno de los objetos en A . De la misma manera, dado un subconjunto B del conjunto de atributo Y . B' es el conjunto de objetos pertenecientes a X sobre los que se aplican todos los atributos de B , (Ganter et al, 2005, Petko, 2011 y Yao y Y., 2004).

Ejemplo 2.1. Para la tabla.

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

Tenemos:

- $\{x_2\}' = \{y_1, y_3, y_4\}$
- $\{x_2, x_3\}' = \{y_3, y_4\}$
- $\{x_1, x_4, x_5\}' = \emptyset$
- $X' = \emptyset$
- $\{y_1\}' = \{x_1, x_3, x_5\}$
- $\{y_1, y_2\}' = \{x_1\}$
- $\{y_2, y_3\}' = \{x_1, x_3, x_4\}$
- $\{y_2, y_3, y_4\}' = \{x_1, x_3, x_4\}$
- $Y' = \{x_1\}$

Definición 2.3. (Concepto Formal). Un concepto formal en (X, Y, I) , es una pareja (A, B) con $A \subseteq X$ y $B \subseteq Y$ tal que $A' = B$ y $B' = A$, los conjuntos A y B se denominan la **extensión** y la **intención** de un concepto formal (A, B) , respectivamente.

La noción de un concepto formal es fundamental en la FCA.

La siguiente proposición nos indica algunas propiedades importantes del operador $(\cdot)'$.

Proposición 2.1. Para todos subconjuntos $A, A_1, A_2 \subseteq X$ y $B, B_1, B_2 \subseteq Y$ se satisfacen las siguientes propiedades:

- $A'''' = A''$,
- Si $A_1 \subseteq A_2 \implies A_2' \subseteq A_1'$,
- $A \subseteq A''$,
- $B'''' = B''$,
- Si $B_1 \subseteq B_2 \implies B_2' \subseteq B_1'$,
- $B \subseteq B''$.

Definición 2.4. (Cerrado) Sea $A \subseteq X$, un subconjunto de objetos, A es cerrado si cumple que $A'' = A$. Análogamente, para B , un subconjunto de atributos, sea $B \subseteq Y$, B es cerrado si cumple que $B'' = B$.

Capítulo 2. Aprendizaje no supervisado

Ejemplo 2.2. (Concepto Formal). Para la tabla.

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

El rectángulo sombreado representa un concepto formal.

$$(A_1, B_1) = (\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\})$$

Debido a

$$\begin{aligned} \{x_1, x_2, x_3, x_4\}' &= \{y_3, y_4\} \\ \{y_3, y_4\}' &= \{x_1, x_2, x_3, x_4\} \end{aligned}$$

En esta tabla podemos encontrar otros conceptos formales, los cuales están representados por los rectángulos sombreados:

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

$$(a) \quad (A_2, B_2) = (\{x_1, x_3, x_4\}, \{y_2, y_3, y_4\}) \quad (b) \quad (A_3, B_3) = (\{x_1, x_2\}, \{y_1, y_3, y_4\})$$

I	y_1	y_2	y_3	y_4
x_1	×	×	×	×
x_2	×		×	×
x_3		×	×	×
x_4		×	×	×
x_5	×			

$$(c) \quad (A_4, B_4) = (\{x_1, x_2, x_5\}, \{y_1\})$$

Definición 2.5. (*Orden Parcial*) Sea (A_1, B_1) y (A_2, B_2) conceptos formales, y \leq un orden parcial, tal que $(A_1, B_1) \leq (A_2, B_2)$ si y sólo si $A_1 \subseteq A_2$ o $B_2 \subseteq B_1$.

Observación 2.1. $(A_1, B_1) \leq (A_2, B_2)$ significa que (A_1, B_1) es más específico que (A_2, B_2) ((A_2, B_2) es más general). Por ejemplos, \leq captura la intuición detrás de $PERRO \leq MAMÍFEROS$ (el concepto de un perro es más específico que el concepto de un mamífero).

Ejemplo 2.3. Considere los siguientes conceptos formales del **Ejemplo1.2**:

$$(A_1, B_1) = (\{x_1, x_2, x_3, x_4\}, \{y_3, y_4\})$$

$$(A_2, B_2) = (\{x_1, x_3, x_4\}, \{y_2, y_3, y_4\})$$

$$(A_3, B_3) = (\{x_1, x_2\}, \{y_1, y_3, y_4\})$$

$$(A_4, B_4) = (\{x_1, x_2, x_5\}, \{y_1\})$$

Entonces

$$(A_3, B_3) \leq (A_1, B_1)$$

$$(A_3, B_3) \leq (A_2, B_2)$$

$$(A_3, B_3) \leq (A_4, B_4)$$

$$(A_2, B_2) \leq (A_1, B_1)$$

La colección de todos los conceptos formales de un contexto formal (X, Y, I) , se llama un concepto de reticulo, otra noción fundamental en FCA (Hongliang y Dexue (2008)).

Definición 2.6. (*Concepto de Reticulo*) El conjunto de todos los conceptos formales obtenidos a partir de un contexto (X, Y, I) , denotado por $\mathcal{B}(X, Y, I)$, es decir,

$$\mathcal{B}(X, Y, I) = \{(A, B) | A' = B, B = A'\},$$

junto con el orden parcial \leq , se llama concepto de reticulo de (X, Y, I) .

A partir del conjunto de reticulo se conforma el conjunto de extensión y el conjunto de intensión, dado por:

$$Ext(X, Y, I) = \{A | (A, B) \in \mathcal{B}(X, Y, I) \text{ para alguna } B\} \quad (2.8)$$

$$Int(X, Y, I) = \{B | (A, B) \in \mathcal{B}(X, Y, I) \text{ para alguna } A\} \quad (2.9)$$

Capítulo 2. Aprendizaje no supervisado

Ejemplo 2.4. (Belohlávek, 2008, Ganter y Wille, 1999 y Pfaltz et al, 2002) Considere la siguiente tabla cruzada:

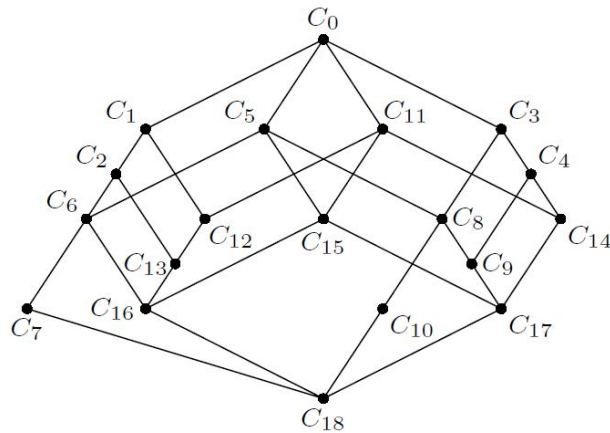
		a	b	c	d	e	f	g	h	i
Sanguijuela	1	×	×					×		
Mojara	2	×	×					×	×	
Rana	3	×	×	×				×	×	
Perro	4	×		×				×	×	×
Elodea	5	×	×		×		×			
Junco	6	×	×	×	×		×			
Frijol	7	×		×	×	×				
Maiz	8	×		×	×		×			

Donde **a**: necesita el agua para vivir, **b**: vive en el agua, **c**: vive en la tierra, **d**: necesita clorofila para producir comida, **e**: dicotiledóneas, **f**: monocotiledóneas, **h**: tiene extremidades e **i**: amamantan a sus crías.

El correspondiente contexto formal (X, Y, I) contiene los siguientes conceptos formales:

$C_0 = (\{1, 2, 3, 4, 5, 6, 7, 8\}, \{a\})$, $C_1 = (\{1, 2, 3, 4\}, \{a, g\})$, $C_2 = (\{2, 3, 4\}, \{a, g, h\})$,
 $C_3 = (\{5, 6, 7, 8\}, \{a, d\})$, $C_4 = (\{5, 6, 8\}, \{a, d, f\})$, $C_5 = (\{3, 4, 6, 7, 8\}, \{a, c, d\})$,
 $C_6 = (\{3, 4\}, \{a, c, g, h\})$, $C_7 = (\{4\}, \{a, c, g, h, i\})$, $C_8 = (\{6, 7, 8\}, \{a, c, d\})$,
 $C_9 = (\{6, 8\}, \{a, c, d, f\})$, $C_{10} = (\{7\}, \{a, c, d, e\})$, $C_{11} = (\{1, 2, 3, 5, 6\}, \{a, b\})$,
 $C_{12} = (\{1, 2, 3\}, \{a, b, g\})$, $C_{13} = (\{2, 3\}, \{a, b, g, h\})$, $C_{14} = (\{5, 6\}, \{a, b, d, f\})$,
 $C_{15} = (\{3, 6\}, \{a, b, c\})$, $C_{16} = (\{3\}, \{a, b, c, g, h\})$, $C_{17} = (\{6\}, \{a, b, c, d, f\})$ y
 $C_{18} = (\{\}, \{a, b, c, d, e, f, g, h, i\})$.

El correspondiente contexto de reticula $\mathcal{B}(X, Y, I)$ es representado en la siguiente figura:



Capítulo 3

Uso de modelos no supervisados para identificar sospechosos de trata de personas

El tráfico de personas es un fenómeno complejo que no es posible identificar usando métodos automáticos sobre algunas bases de datos, por el contrario, para la identificación de probables sospechosos se requiere hacer un análisis sobre los datos usando la experiencia e inteligencia humana. Por la amplitud y complejidad del problema, en este trabajo, solo se considerará la trata de personas para prostitución forzada.

3.1. Identificación de sospechosos de trata en los países bajos

En esta sección se va a estudiar lo expuesto en el artículo *Semi-automated Knowledge discovery: identifying and profiling human trafficking* escrito por *Poelmans Jonas, Elzinga Paul, Ignato Dmitry I. y Kuznetsov Sergei O* en 2012.

En los países bajos la prostitución está legalizada, sin embargo es un crimen que una persona obtenga beneficios económicos obligando a otra, u otras personas a ejercer la prostitución. Por lo tanto, uno de los objetivos de la justicia Holandesa es perseguir a este tipo de delincuentes.

La policía holandesa ha identificado que la mayoría de los tratantes en Holanda son de origen marroquíes, antillanos o turcos y que su principal estrategia es controlar a las víctimas a través del enamoramiento. Se han identificado que las víctimas

Capítulo 3. Uso de modelos no supervisados

proviene de países como Albania, Moldavia, Rumania, Hungría, Bulgaria, Rusia, Bielorrusia y Ucrania, y estas mujeres permanecen en su mayoría en condiciones de esclavitud.

Los datos que utilizaron Poelmans et al (2012) para su análisis fueron los reportes policíacos de la ciudad de Amsterdam-Amsterland, los cuales son llenados por los policías al hacer su recorrido por las calles de la ciudad. Para obtener información de los reportes policíacos, éstos deben ser analizados manualmente, uno por uno, buscando indicios de que se está cometiendo el delito de trata de personas.

Una vez que los investigadores encuentran indicios de que alguien es obligado a ejercer la prostitución, se hace una investigación más a fondo para determinar si existe realmente el delito.

Sin embargo, la sobrecarga de información en la base de datos de la policía y la falta de herramientas adecuadas para que estos datos sean analizados dificulta la posible identificación de sospechosos.

Enseguida se enlista algunas de las características que son indicios de que una mujer esté siendo víctima de trata:

1. Dependencia de un explotador
 - La mujer tiene un pasaporte falso.
 - La mujer no conoce la dirección donde trabaja.
2. Privación de libertad
 - La mujer no recibe asistencia médica adecuada.
 - La mujer no lleva sus documentos de identificación.
3. Condiciones malas de trabajo
 - La mujer recibe un salario por debajo del mercado laboral.
 - La mujer tiene largas jornadas de trabajo.
4. Violación de la integridad física
 - Las mujeres son amenazadas y se enfrentan a la violencia física.
 - Las mujeres pueden tener tatuajes o portar amuletos o adornos vudú.
5. Signos incidentales de abuso
 - Las mujeres trabajan en sitios diferentes.

- Las mujeres pueden tener una persona que está cerca de ella y le indica que hacer.

Igualmente se listan algunas características de un posible explotador y tratante.

1. Reclutamiento de las víctimas

- Controla el lugar de alojamiento de las víctimas.
- Análisis de la víctima en su primera entrevista para determinar los puntos vulnerables de la misma, esto es, si el control se tendrá por enamoramiento, por regalos, o por agresión física.

2. Iniciación de las víctimas en la prostitución

- Obligan o convencen a las víctimas a tener relaciones sexuales con ellos, esto inicia el control, en particular ocurre con las mujeres islámicas, las cuales no pueden regresar a su casa porque temen ser rechazadas y abandonadas por sus familiares.
- Por chantaje o amenazas, obligan a la víctima a ejercer la prostitución por primera vez.

3. Mantenimiento de la víctima en la prostitución

- Por dependencia emocional, les hacen creer que las quieren y que solo ellos pueden apoyarlas, o el explotador es el padre de sus hijos.
- Por el aislamiento social, las víctimas solo tienen relación con miembros del círculo de la prostitución.

4. Protección de su negocio.

- Se asegura de la vigilancia a las víctimas y mantiene el control sobre ellas con amenazas y violencia física.
- Interroga a las víctimas que han estado en contacto con la policía.

Para obtener información de los reportes policíacos, Poelmans et al (2012) proponen conceptos formales, en particular, conceptos de retícula y con ello crearon los perfiles de las personas que aparecían en los reportes, con las retículas formadas se encontró, para algunos individuos, evidencias suficientes para iniciar una investigación profunda, usando FCA se lograron efectuar arrestos reales y el cierre de locales donde se practicaban la prostitución forzada.

Capítulo 3. Uso de modelos no supervisados

Un ejemplo de un FCA con 5 objetos (reportes) y 6 atributos, se representa en la tabla 3.1.

		y_1	y_2	y_3	y_4	y_5	y_6
Reporte 1	x_1	×	×				×
Reporte 2	x_2			×	×	×	
Reporte 3	x_3	×	×	×	×	×	
Reporte 4	x_4						×
Reporte 5	x_5				×	×	

Tabla 3.1

Donde los atributos en la tabla son: y_1 : Prostitución, y_2 : Tratante, y_3 : Violencia, y_4 : Autos caros, y_5 : Gran cantidad de dinero y y_6 : Bulgaros.

En esta tabla podemos encontrar los siguientes conceptos formales:

$C_1 = (\{x_1, x_3\}, \{y_1, y_2\})$, $C_2 = (\{x_2, x_3, x_5\}, \{y_4, y_5\})$, $C_3 = (\{x_2, x_3, \}, \{y_3, y_4, y_5\})$, $C_4 = (\{x_1, x_4\}, \{y_6\})$ y $C_5 = (\{x_3\}, \{y_1, y_2, y_3, y_4, y_5\})$.

La relación de orden entre estos 5 conceptos formales esta dada por $C_5 \leq C_3 \leq C_2$.

Además, esta tabla es una pequeña fracción que se utilizó en la investigación. De los conceptos formales se puede elaborar la siguiente reticula:

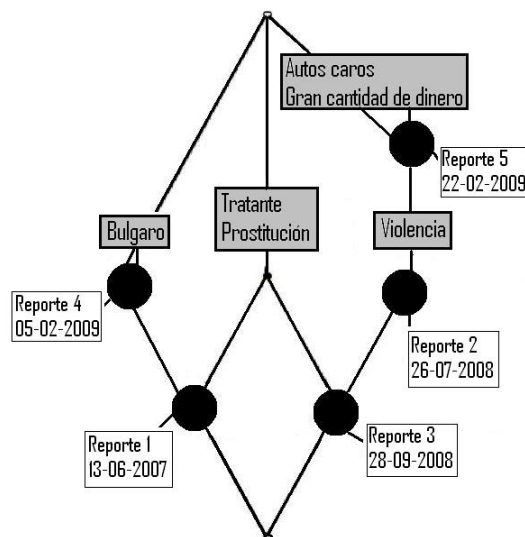


Figura 3.1: Diagrama de concepto de reticula.

Los nodos de la reticula en la figura 3.1 representan a los informes policíacos y su etiqueta se encuentra en las casillas no sombreadas, y las casillas sombreadas representa a los atributos. Los reportes 3 y 2 estan asociados con los atributos ‘violencia’, ‘autos caros’ y ‘gran cantidad de dinero’, el reporte 5 esta asociado con los atributos ‘autos caros’ y ‘gran cantidad de dinero’, los reportes 1 y 3 estan relacionados con los atributos ‘tratante’ y ‘prostitución’, el reporte 4 esta asociado con el atributo ‘Bulgaros’.

El concepto de reticula les permite a los investigadores:

- Definir términos de los reportes policíacos con precisión: Los investigadores encontraron que los oficiales de policía al llenar sus informes utilizaban diferentes palabras para los mismos conceptos, esto dificultaba determinar cuando hay indicios de trata de personas. Para ser más accesible esta información, los investigadores mejoraron la definición de violencia domestica.
- La identificación de sospechosos potenciales: Los conceptos de reticula contiene toda la información posible de una sola persona, esto se representa mediante una red de información, esta imagen facilita la toma de decisión para determinar si una persona es sospechosa o no. También puede describir redes criminales. En los reticulos formados se tiene que las personas que estan en los niveles más bajos tienen más probabilidad de estar cometiendo el delito de obligar a una tercera persona a ejercer la prostitución.
- La exploración de la estructura social: El reticulo ayuda a exponer:
 - 1) a los individuos sospechoso,
 - 2) a las redes criminales y
 - 3) el rol que tiene estas personas en estas redes.

El FCA fue capaz de ofrecer a la policía un enfoque ideal para una búsqueda más accesible de los sospechosos. Para la detección de sospechosos se elaboraron múltiples reticulos, estos reticulos contenian más de 200 características que se basan en diferentes conjuntos de atributos. Estos atributos estan basados en los indicadores de tráfico de persona anteriormente mencionados.

Casos de reticula elaborados por Poelmans et al (2012).

La figura 3.2 presenta un reticulo que describe a 4895 personas extraidas de los informes policíacos. Este reticulo esta relacionados con los sospechosos de los casos 1, 2 que revisaremos en más adelante.

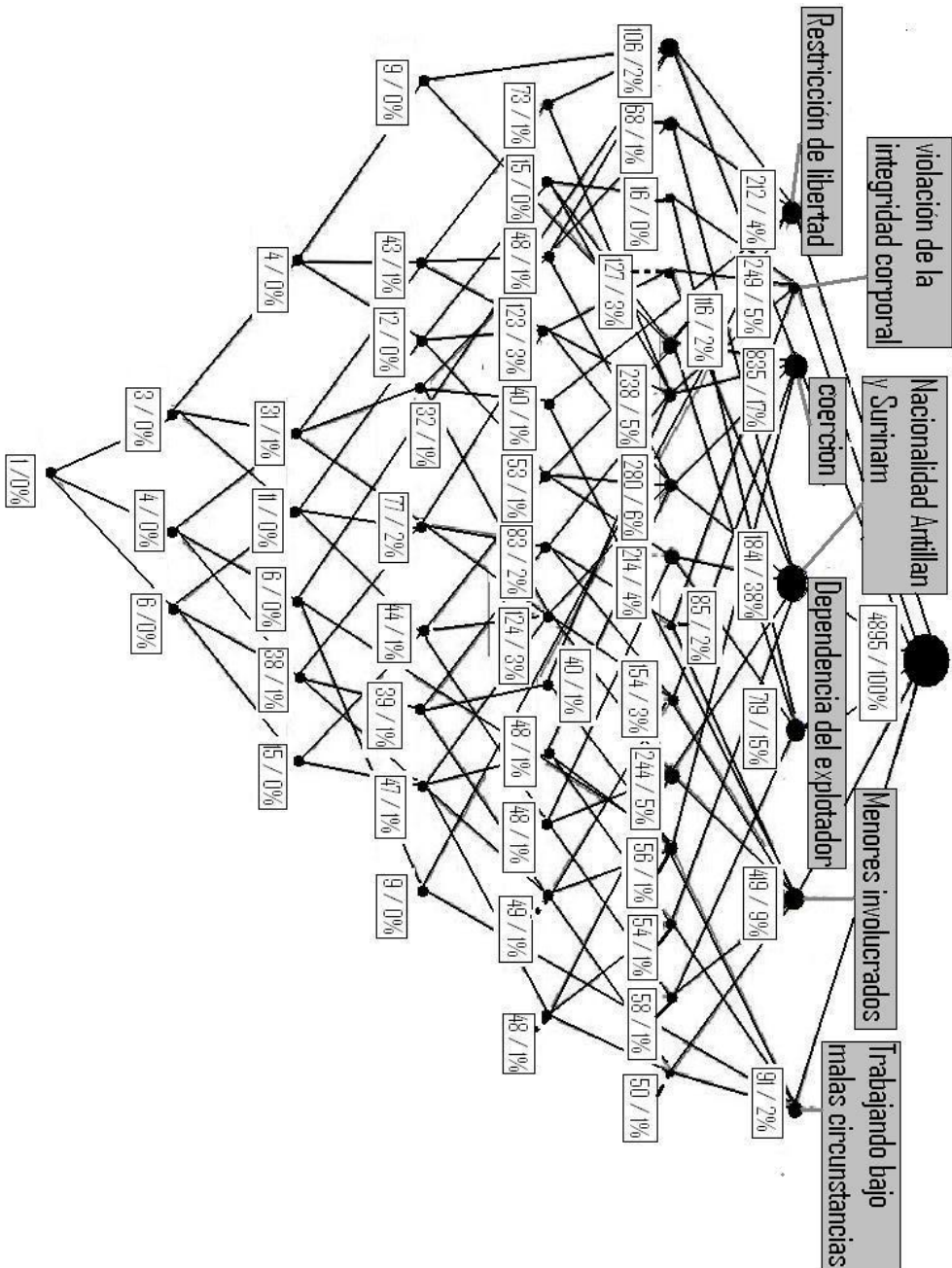


Figura 3.2: Diagrama de retículo de detección de sospechoso de tráfico humano

El reticulado de la figura 3.3 describe a 1255 personas de origen búlgaro, húngaro, y rumano, estas personas son un subconjunto de las 4895 personas del diagrama de la figura 3.2. Este reticulado esta relacionados con los sospechoso del caso 3, que revisaremos más adelante.

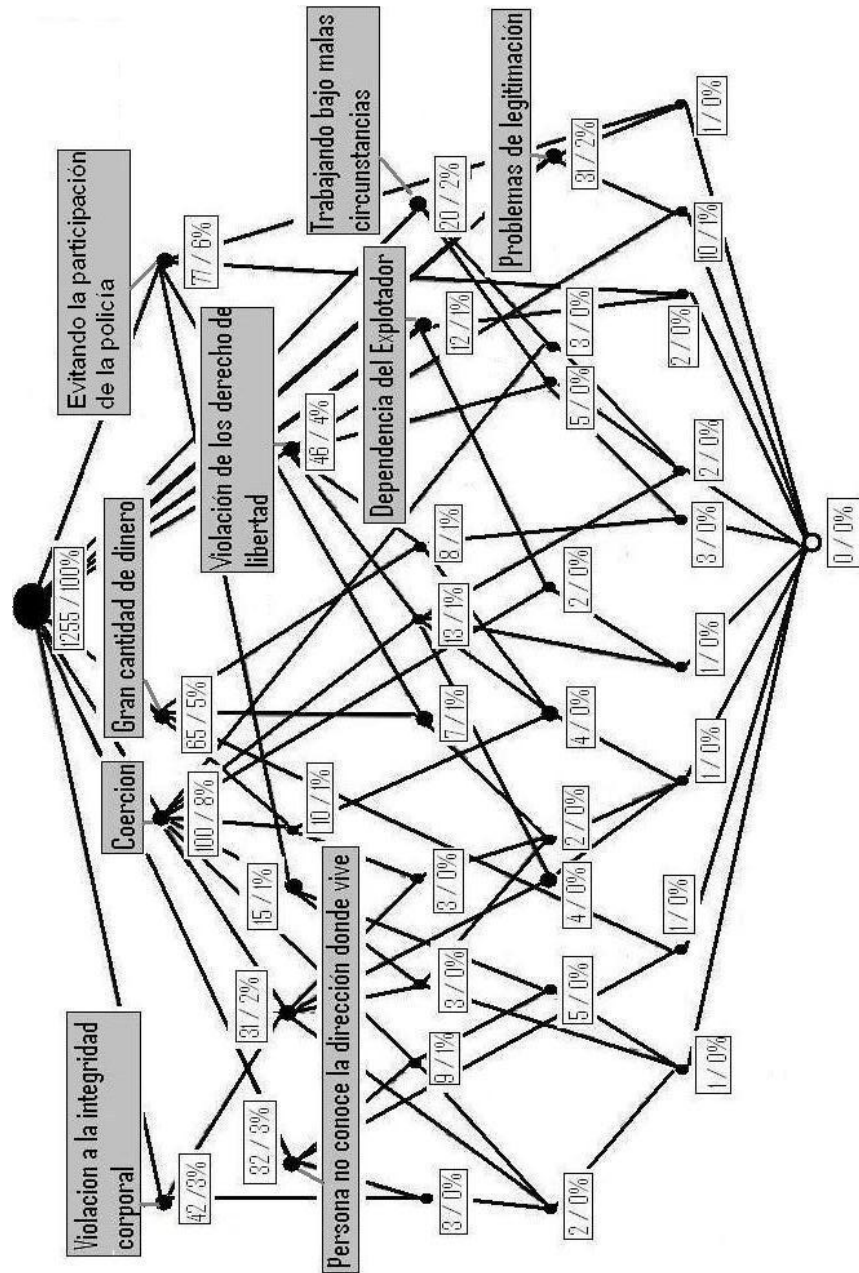


Figura 3.3: Diagrama de reticulado de detección de sospechoso de tráfico humano

Caso 1: Sospechoso B

Usando la figura 3.2 se descubrió una chica de menor edad y con el diagrama se pudo determinar quien era su explotador.

Con la ayuda de 3 informes, los investigadores encontraron información sobre una muchacha H. Estos informes los llevaron al descubrimiento del presunto tratante B. En el primer informe se indica que la muchacha H tiene tatuado el nombre 'B' en su muñeca. Resulta que el nombre de su novio es 'B', el novio es de origen surinamia y tiene 30 años. En el segundo informe se señala que la chica le dijo al agente policiaco que ella se dedicaba a la prostitución porque había apostado con una amiga que podia dedicarse a esto. También en este informe señala que la chica tiene tatuado el nombre 'B' en su muñeca y en su vientre, además se indica que tiene heridas en su brazo y quemaduras en su pierna, y que tiene largas jornadas de trabajo. En el tercer informe se describio una posible relación entre la muchacha H y un hombre, el oficial de policía que realizó este informe conoce a este hombre, él sabe que este hombre esta relacionado con el mundo de la prostitución ilegal, el nombre de pila de este individuo es 'B', el mismo nombre que tiene la muchacha H en la muñeca. Esta información les indicó que este hombre es el tratante de la víctima H.

Así se tuvo información suficiente para indicar que B era un tratante y la muchacha H era su víctima. Estos indicios fueron suficientes para empezar una investigación contra el tratante B. Los investigadores encontraron otros 12 informes que indica que el tratante B esta involucrado en la prostitución, el resultado este análisis se muestra en el diagrama 3.4.

Además, se encontraron indicios que el tratante B tenia una relación con otro tratante (Caso 2), también se descubrio a otra mujer R que es víctima del tratante B. Toda esta información en conjunto dio resultado a una investigación más profunda.

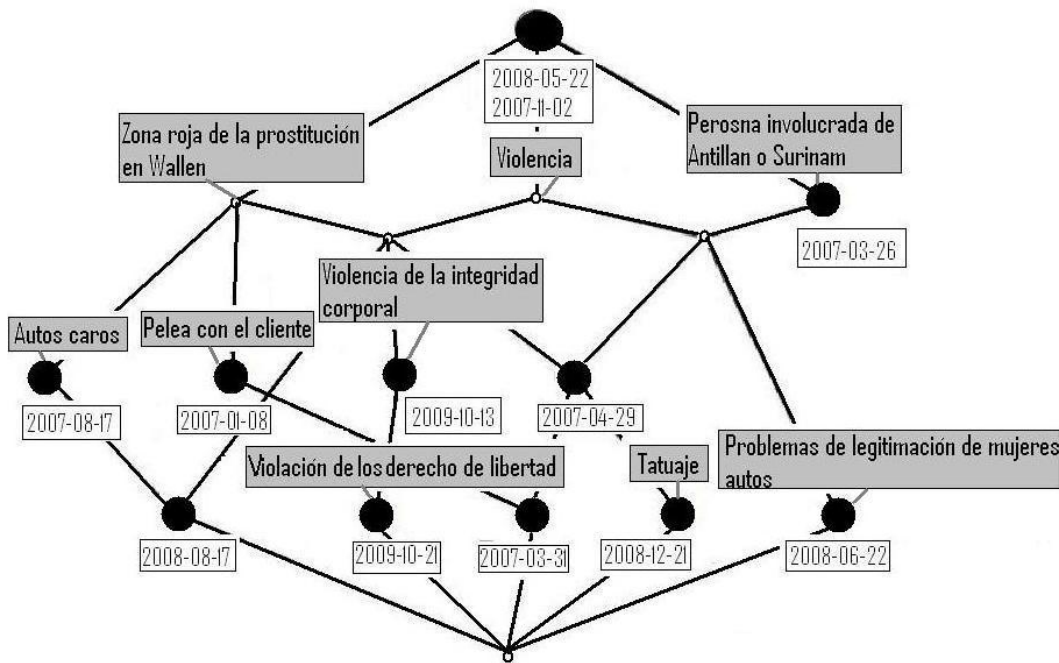


Figura 3.4: Diagrama de retículo del sospechoso B.

Caso 2: Sospechoso A

Se describe el caso del tratante A, que los investigadores descubrieron en el análisis del diagrama de retículo 3.2.

El tratante A es de nacionalidad antillanas, la mujer V es de nacionalidad holandesa. Esta mujer V es víctima del tratante A. Los informes policíacos indican que una llamada telefónica anónima advierte a la policía que un grupo de hombres jóvenes antillanos ejercen vigilancia sobre algunas mujeres y las mantienen en aislamiento, de esta manera ubican por primera vez al sospechoso A. Posteriormente en un patrullaje rutinario se encuentra que el sospechoso A conduce un auto, se detiene y sale la mujer V, corriendo y entra a un domicilio. La misma mujer V en otro informe indica que el sospechoso A siempre la recoge y la regresa a este domicilio. En otro informe se indica que el sospechoso A estuvo en la cárcel por posesión de armas. En otro informe se indica de una segunda víctima, la mujer B, quien pidió informes sobre si podría trabajar con una identificación muy dañada y ella mencionó que el sospechoso A era su exnovio y que la extorcionaba a ella y a la mujer V, también ella mencionó que tenía miedo para denunciarlo. En otro informe los oficiales de policía comprobaron que el domicilio donde recogían a la mujer V, vivía también la víctima B, así mismo, descubrieron que la víctima V

Capítulo 3. Uso de modelos no supervisados

tiene un tatuaje en su hombro derecho y otro en su brazo en la parte superior. En otro reporte se indica que el sospechoso A estuvo involucrado en una pelea que él la inicio porque un hombre queria el servicio de la víctima V. Finalmente, en otro reporte se indica que en todos los videos de vigilancia se observa al sospechoso A y la víctima V juntos. Con toda la información obtenida en los informes policiaos dio resultado a la elaboraron del reticulo 3.5 del sospechoso A .

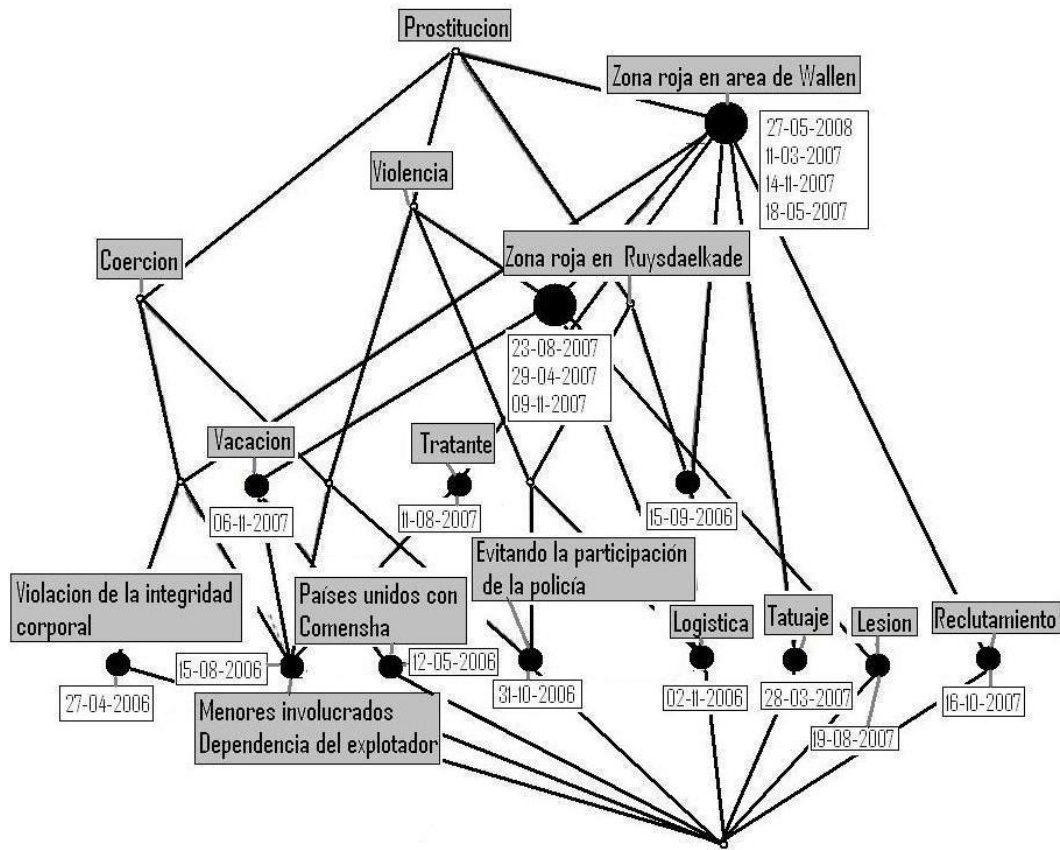


Figura 3.5: Perfil de sospechoso tratante.

Caso 3: Red de tráfico de personas turcas

Los investigadores descubrieron en el análisis del diagrama de reticulo 3.3 a una red criminal que operaba en Amsterdam, este grupo estaba involucrado en la prostitución forzada.

El reticulo 3.6 contiene 61 personas e indicios que se encontraron en los informes policiaos, estos informes mencionan actividades alrededor de un bar, este bar se

ubica en Amsterdam. El bar era lugar donde operaba este grupo criminal y fue cerrado. También los investigadores encontraron que este bar estaba controlado por hombres en su mayoría turcos, estos hombres reclutaban a muchachas en el mundo de la prostitución. Así mismo, los investigadores encontraron al menos 2 explotadores que tenían a varias muchachas que trabajaban para ellos.

En uno de los informes se menciona por primera vez a la víctima H. La víctima H declaró a los oficiales de policía que ella fue forzada a trabajar como una prostituta y no recibía ningún salario por esto, también menciona que ella estaba en aislamiento en un departamento. En otro informe indica que el sospechoso A salía del bar con una muchacha con vestimenta exótica posiblemente trabaja como prostituta en el bar. En otro informe se indica que los oficiales detuvieron un auto, la persona que conducía el auto era el sospechoso B y su acompañante era la víctima E. También se menciona que la víctima E trabaja en el bar y además el auto siempre está estacionado cerca del bar. Asimismo, el sospechoso A tenía en su bolsillo del pantalón una gran cantidad de dinero en efectivo. En otro informe se indica que los oficiales de policía fueron a inspeccionar el bar, los oficiales notaron a una muchacha nueva y le preguntaron cómo había llegado a este lugar, ella les respondió que llegó por tren, y además no tenía dinero y no conocía la dirección de su residencia. Igualmente, los oficiales notaron al sospechoso B, quien les comentó que era comerciante de autos y que viajaba mucho entre los países bajos. Así mismo, la víctima F y la víctima G estaban en el bar. En otro informe se indica que en un patrullaje rutinario notaron al sospechoso A hablar con el conductor de un auto con matrícula búlgara. Después el sospechoso A forzó a una muchacha a seguirlo, los oficiales le preguntaron al sospechoso A por esta acción, él respondió que solo son amigos. La muchacha no tenía sus documentos en su poder. Posteriormente los oficiales investigaron la dirección de esta muchacha y fueron a su domicilio. En la casa encontraron varios colchones, además encontraron a otra chica. Esta chica le dijo a la policía que no tenía trabajo. Finalmente en otro informe se indicó que los oficiales observaron el bar, ellos notaron a varias muchachas que salían del bar con clientes a un hotel y después las traían de nuevo al bar. Con toda la información obtenida en los informes policíacos dio resultado al cierre de este bar.

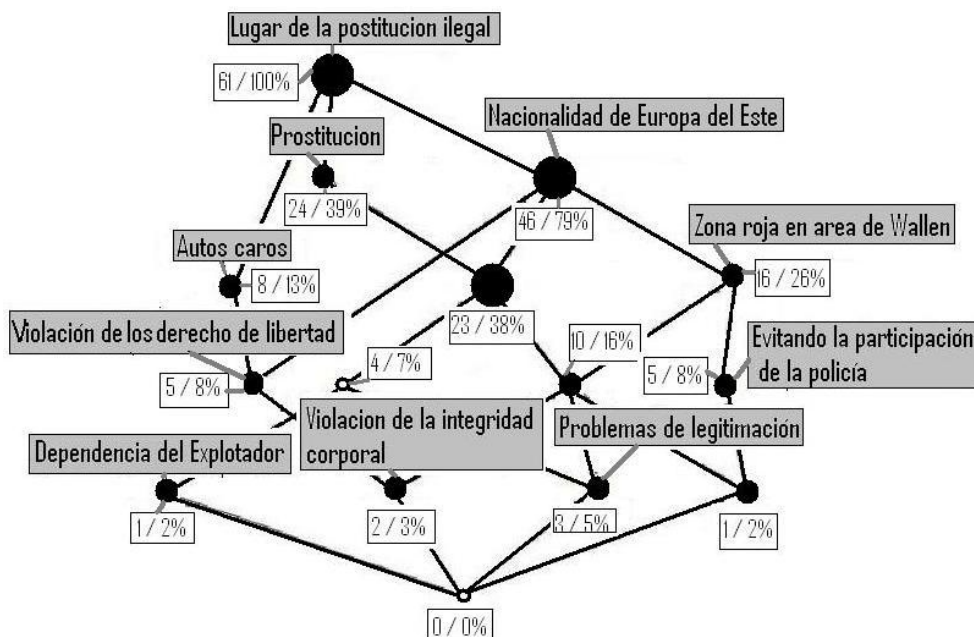


Figura 3.6: Concepto de reticula de una red de tráfico de persona.

Caso 4: Sospechoso de origen Bulgaro

En este caso, se describió un perfil de un sospechoso de origen Bulgaro, quien operaba en Amsterdam. El reticulo 3.7 describe las actividades que realizaba el sospechoso A y su grupo criminal, este reticulo fue resultado de los informes de policía.

En uno de los informes policiaco se indica que el sospechoso A fue visto por primera vez en un patrullaje rutinario. Los oficiales detuvieron un auto con matriculas de origen bulgaro, en el auto habia tres hombres y una mujer. El conductor del auto era el sospechoso B y su copiloto era el sospechoso A, en la parte de atras estaban sentados la mujer F y el hombre K. Estos ultimos mencionaron que hace tres días llegaron y son una pareja. El sospechoso A y el sospechoso B fueron llevados a la oficina de la policía, mientras que el hombre K y la mujer F huyeron del lugar. Los oficiales los persiguieron y observaron que el hombre K forzó a la mujer F a entrar a una casa. En la oficina de la policía, le encontraron al sospechoso A una gran cantidad de dinero en el bolsillo de su pantalón. En otro informe se indica que una mujer J pidio ayuda a los oficiales para supervisar un acuerdo de arrendamiento de un departamento. Ella le dijo a la policía que el sospechoso A la intimidaba y que queria destinar el departamento para fines de prostitución. Los oficiales notaron que

Capítulo 3. Uso de modelos no supervisados

la mujer J hablaba tímidamente y dedujeron que posiblemente habia sido forzada a prostituirse por el sospechoso A. En otro informe se indica que en un patrullaje rutinario notaron a dos hombres con motocicletas que estaban en una calle. Era la segunda vez que notaron a estos hombres, uno de ellos es el sospechoso A y su motocicleta esta registrada al nombre de la mujer C quien esta involucrada en actividades de tráfico de personas como una víctima. En otro informe se indica que recibieron un aviso donde se señala que un criminal de origen colombiano podria estar viviendo en un domicilio. Los oficiales de policía fueron a insperccionar el lugar, encontraron a dos hombre y a dos mujeres de origen bulgaros. El hombre X y la mujer C le dijeron a los oficiales que estaban de vacaciones y regresarian a Bulgaria. El hombre Y le dijo a los oficiales que se dedicaba a exporta autos caros a Bulgaria y por su trabajo viajaba entre los países bajos. La mujer Z le dijo a la policía que trabajaba en la prostitución. Posteriormente, los oficiales salieron del domicilio, encontraron una motocicleta registrada con el nombre del sosopechoso A. Con toda la informaci recaudada en los informes policiaco se dedujo que los sospechosos A y B son los que tienen mayor probabilidad de ser culpables de tráfico de personas, se hizo una investigación más profunda y finalmente fueron detenidos. Con la ayuda del reticulo 3.7 se encontraron a otra peronas que estan involucradas con la misma banda y seran investigados de la misma manera.

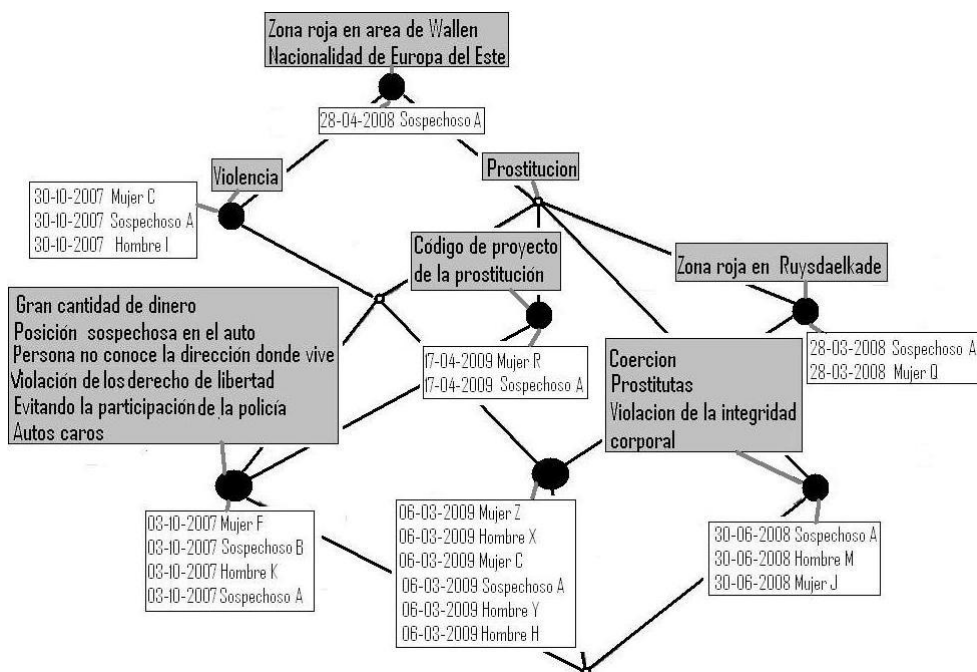


Figura 3.7: El reticulo del perfil del sospechoso y su red criminal.

Caso 5: Mujer de origen hungaro tanto víctima como sospechosa

Se describe a una mujer quien fue primero víctima y después se convirtió en una sospechosa de prostitución forzada. En el retículo 3.8 describe indicios de que la mujer SV1 fue víctima en actividades de prostitución, de igual manera, había señales que esta mujer participaba en actividades delictivas.

En uno de los informes se indica que la mujer SV1 fue vista por primera vez en un patrullaje de policía, esta mujer no habla el holandés, el inglés u otra lengua hablada por los oficiales de policía. Los oficiales notaron que la mujer SV1 podría estar en la prostitución y llevada a esta actividad por una banda criminal. En otro informe se indica que la mujer SV1 y la muchacha I tenían identificaciones falsas. En otro informe se indica que la mujer Q quien es una prostituta, le dijo a los oficiales que tenía que dar todo su dinero a un tratante de origen húngaro, este tratante trabaja para una red criminal, también la mujer SV1 le dijo a los oficiales que ella trabaja como prostituta para uno de los tratantes de esta red. En otro informe se indica que la mujer SV1 tiene un tatuaje. Los oficiales de policía notaron que la mujer SV1 se estaba convirtiendo en una explotadora por que en un informe policiaco se indicó que la mujer SV1 siempre pagaba el alquiler de un cuarto para una mujer L de origen húngaro que es una prostituta. En otro informe se indica que una mujer desconocida vive con la mujer SV1, esta mujer desconocida le dijo a la policía que desconoce su dirección y además siempre la recoge y la deja en la casa la mujer SV1. En otro informe se indica que los oficiales de policía fueron a inspeccionar una casa, esta casa a veces es arrendada por una mujer D de origen holandés quien trabaja como prostituta. La mujer D le dijo a los oficiales que es amenazada por un grupo de tratantes y además les dijo que este grupo los conoció por la mujer SV1. En otro informe se indica que la mujer SV1 realizaba llamadas telefónicas a hombres de origen húngaros, probablemente la mujer SV1 trabaja para estos hombres. Con la ayuda de todos los informes policiacos dio resultado que la mujer SV1 es una víctima de la prostitución y además no tenía otra opción mejor que convertirse en una explotadora. Así mismo, encontraron pruebas que la mujer SV1 era parte de esta red criminal. Al finalizar el análisis de los informes, reunieron pruebas suficientes contra la mujer SV1, lo cual dio resultado su arresto en una prisión holandesa.

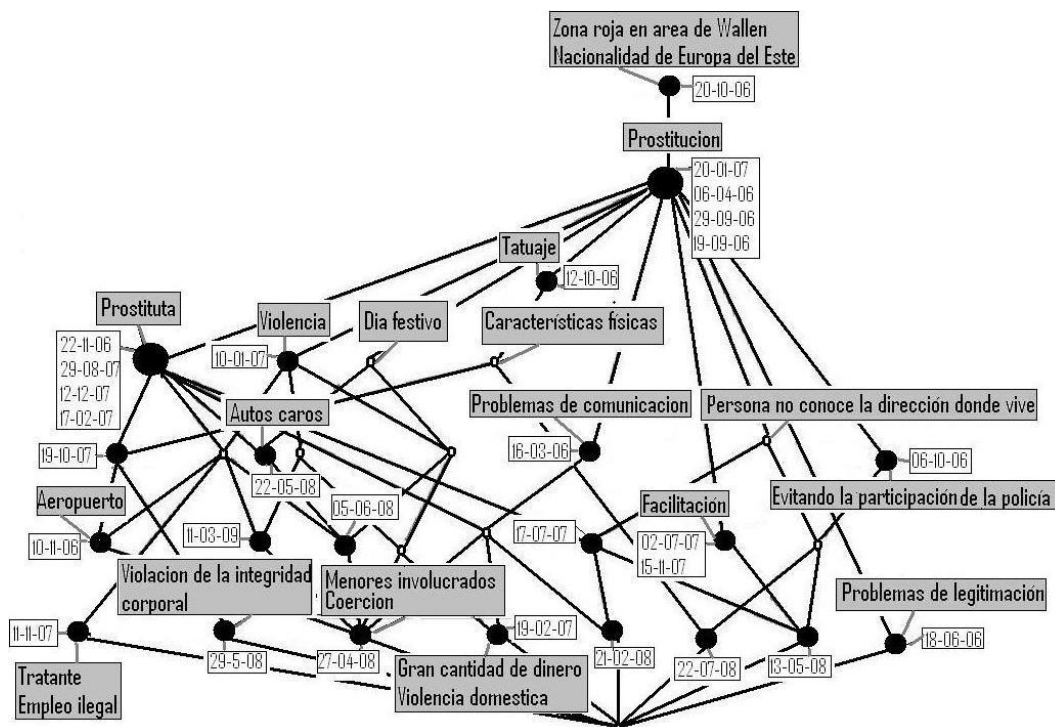


Figura 3.8: El reticula del perfil de la mujeres quien primero fue víctima y después sospechosa.

Con los ejemplos presentados se puede ver como los autores Poelmans et al (2012) del artículo *Semi-automated Knowledge discovery: identifying and profiling human trafficking* que estamos revisando pueden acomodar la información de los reportes policíacos de ciudad de Amsterdam para obtener:

- El resumen de la estructura conceptual de datos en una sola imagen: los reticulos de los casos fueron usados para mostrar este aspecto atractivo del FCA. La sobrecarga de los informes fue convertido en un artefacto intuitivamente analizable.
- Un medio eficaz para la búsqueda de los datos: a partir de los reticulos de los casos, varias personas fueron elegidas y analizadas en detalle en los casos mencionados.
- La visualización intuitiva con un orden parcial de las personas basadas en los indicadores observados: Los agentes de policía fueron guiados por el orden parcial de los conceptos cuando los analizaron los reticulos. A partir de este

Capítulo 3. Uso de modelos no supervisados

análisis la policía pudo darse cuenta que aquellos individuos cuyo nombre estaba en una posición más baja en el retículo tenía mayores indicios de ser un tratante.

- Relaciones conceptuales entre los documentos individuales, personas, fechas, etc. hicieron visibles a personas que a menudo permanecen ocultos: se pudo observar que en el retículo del caso 3 se expuso una red criminal que operaba en Amsterdam. Múltiples observaciones independientes contenían indicios de la actividad de la red, pero sólo fue identificado cuando se analizó el retículo.
- La visualización de la evolución temporal de una persona: los retículos en los casos 1, 3 y 4 mostraron evidencia contra varios sospechosos de tráfico humano. En el caso 5 mostraron cómo una mujer primero fue víctima y más adelante se convirtió en una sospechosa de trata de personas.

Los autores analizaron otros tipos de métodos automatizados desafortunadamente estos métodos tiene serios problemas para algunos dominios, tales como:

- * La clasificación de caja negra no es aceptable: los policías necesitan justificar los motivos detrás de una decisión, detrás de una etiqueta asignada, etc. Cada decisión de etiquetar a un sospechoso debe ser basada en pruebas y ser acompañado de un informe detallado de los indicios observados.
- * Los textos son cortos y escrito por autores con diferentes estilos de escritura: esto lo hace imposible e inútil aplicar las técnicas de extracción de términos como el análisis de frecuencia. Los términos que ellos obtuvieron a través de paquetes de software, como el detective de datos y Clementine tampoco no satisfacía. Las relaciones entre las personas, los informes y los redes juegan un papel esencial, pero son difíciles o imposibles de extraer de forma automática.
- * Los contextos de palabras y frases son esenciales en la interpretación de los datos: la interpretación de las palabras, frases, etc., a menudo depende fuertemente del contexto en el cual son utilizados. Por ejemplo, durante un rondin de la policía, un oficial observa a una nueva prostituta y le pregunta sobre las cicatrices de sus piernas. Él escribió “ella dijo que durante su niñez fue sexualmente abusada y golpeada” pero de repente su conversación fue interrumpida por el explotador quien trae sus alimentos. Los atributos “sexualmente abusada”, “el explotador”, “traer el alimento” y “cicatrices” puede conducir a un falso positivo, aunque este documento solo esta lejos

de ser suficiente para comenzar un caso de prostitución forzada. Además, múltiples personas son mencionadas en muchos informes y sus papeles como el sospechoso, la víctima, o ambos son difíciles de filtrar de estos informes, incluso con instrumentos avanzados como el programa NLP. También algunos atributos únicamente deberían ser atribuidos a una persona pero a menudo es imposible automáticamente deducir a cual. La toma de decisiones humana es necesaria.

- * Poca información está disponible por persona: la información que tenemos es naturalmente incompleta ya que los informes escritos por oficiales describen sólo una parte de la realidad, a saber la parte observada por ellos durante su trabajo. La policía realmente sólo tiene la información sobre los fragmentos de las vidas de estas personas basadas de ellos quienes deciden si esta persona podría ser interesante. Dada la falta de información, uno debería tomar con precaución la toma de decisiones totalmente automatizadas y se debe asignar esta tarea difícil a especialistas o policía entrenados quienes podran determinar eficazmente cuando los indicios constituyen una prueba. El centro del estudio radica en el desarrollo de un sistema temprano de advertencia que reduzca el número de sospechosos potenciales, reunir toda la información disponible en un cuadro visual apoya a los oficiales en la toma de decisiones eficiente sobre que casos deberían recibir la atención especial.
- * No se clasifican a los individuos o a los informes: Los datos no etiquetan a ningún individuo. No existe la capacitación de un clasificador automático. Para identificar frases que se refieren a la prostitución forzada durante la construcción de un algoritmo debe ser confiada al conocimiento experto.
- * Los conceptos subyacentes del dominio son confusos: las relaciones conceptuales entre personas, informes, ubicaciones, etc, eran de importancia significativa y tuvieron que hacerse visible para los oficiales ya que ellos son esenciales a la toma de decisiones. Muchas técnicas de visualización como la autoorganización de mapas sólo dan una distribución de las personas, documentos, etc. pero las relaciones entre ellos no se muestra explícitamente.

Los autores señala que una limitante de su metodología es que los reticulos con muchos atributos son difíciles e interpretar sin embargo esto no representa un problema porque los puntos de interés se conforman de un subconjunto pequeño de todos lo reportes.

3.2. Método de Monte-Carlo para analizar la Trata en México

En esta sección veremos la implementación computacional simulada para detectar las actividades de posibles sospechosos de trata de personas en México, mediante los métodos anteriormente revisados. En esta simulación se obtienen los valores de la base mediante el uso de números aleatorios debido a que los datos reales son confidenciales.

3.2.1. Variables

México es uno de los países con mayor riesgo de incidencia de operaciones financieras con recursos de procedencia ilícita (Fortuna, 2012). Por esta razón, las instituciones financieras están obligadas a emitir reportes de operaciones para prevenir o detectar posibles actos u operaciones de lavado de dinero o financiamiento al terrorismo según sus propias leyes que la regulan (Código Penal Federal, 2016).

Los usuarios de estas instituciones realizan diferentes tipos de operaciones, las cuales son catalogadas en:

- Operación inusual: es la operación, actividad o comportamiento de una persona que no concuerde con su patrón habitual de realización de transacciones.
- Operación preocupante: es la operación, actividad o comportamiento de los directivos, funcionarios, empleados y apoderados de la persona obligado, sin justificar razonablemente su nivel de vida superior al de sus posibilidades.
- Operación relevante: es la operación que se efectúa por un monto mayor o igual 10,000 dólares, ya sea con billetes y las monedas metálicas de curso legal en los Estados Unidos Mexicanos o en cualquier otro país (LFPIORPI, 2016).

Dependiendo el tipo de operación se llena un tipo de reporte (inusual, preocupante, relevante), este tipo de reporte contiene datos de identificación del sujeto que está obligado, del sujeto reportado y de la operación (SHCP, 2016), a continuación se enlistará las instituciones financieras que están obligadas a emitir estos reportes:

- Instituciones de Banca Múltiple
- Instituciones de Banca de Desarrollo

- Sociedades Financieras de Objeto Limitado
- Casas de Bolsa
- Sociedades Operadoras de Sociedades de Inversión
- Sociedades Distribuidoras de Acciones de Sociedades de Inversión
- Arrendadoras Financieras
- Empresas de Factoraje Financiero
- Almacenes Generales De Depósito
- Uniones de Crédito
- Sociedades de Ahorro y Préstamo
- Casas de Cambio
- Sociedad Financieras de Objeto Múltiple
- Entidades de Ahorro y Crédito Popular
- Centros Cambiarios
- Transmisores de Dinero
- Instituciones de Seguros
- Instituciones de Fianzas

Por otro lado, hay actividades económicas que dada a su naturaleza y características puede ser utilizada por la delincuencia para ocultar o transformar recursos ilícitos en lícitos sin que las autoridades pudieran darse cuenta, estas actividades se conocen como actividades vulnerables. La Ley Federal para la prevención e identificación de operaciones con recursos de procedencia ilícita (LFPIORPI, 2016), identifica ciertas actividades vulnerables que están obligadas a emitir avisos de operaciones para prevenir o detectar posibles actos u operaciones de lavado de dinero a la Secretaría de Hacienda, a continuación, se enlistará las actividades vulnerables (Portal de Prevención de Lavado de Dinero, 2016):

Capítulo 3. Uso de modelos no supervisados

- Juegos y sorteos
- Cheques de viajero
- Inmuebles
- Traslado o custodio de valores
- Comercio exterior
- Tarjeta de servicio y crédito
- Mutuo, préstamo o crédito
- Metales y joyas
- Servicios profesionales
- Fe Pública
(Notarios y Corredores Públicos)
- Tarjeta de prepago y cupones
- Blindaje
- Obras de Arte
- Donativos
- Vehículos

De la base de operaciones podemos construir variables en término de monto en moneda nacional dependiendo el tipo de reporte (inusual, relevante, preocupante), los montos promedios, montos en otra divisas, así como el número de operaciones por mes o por año, las localidades, los estados, los países en que se ha realizado operaciones. Además, podemos obtener variables respecto a las diferentes instituciones financieras, a las diferentes cuentas aperturadas. Igualmente, podemos obtener variables respecto de los montos por compra o venta de bienes tales como: inmuebles, vehículos, joyas, juegos o apuestas y obras de artes. Podemos construir una diversidad de variables que pueden obtenerse de la información financiera o de actividades vulnerables que puedan ayudar a identificar alguna elemento del ocultamiento o utilización de recursos procedentes de la trata de personas.

3.2.2. Simulación

En la base de datos simulada se incluye un subconjunto de variables (atributos) del total de variables que nosotros construimos, las cuales nos parece pudieran ser las más relevantes en nuestro problema haciendo la aclaración que no se pretende duplicar los datos reales.

Las variables en nuestra base de datos simulada son:

- `Riesgo_domicilio` (Estado): riesgo relacionado con la trata de personas por estados. Véase tabla A.1
- `Monto_dep_inu`: monto total en moneda nacional de reporte de operaciones por depósitos inusuales
- `Num_dep_inu`: número de reportes de operaciones por depósito inusuales

Capítulo 3. Uso de modelos no supervisados

- Num_ret_inu: número de reportes de operaciones por retiro inusuales
- Num_vehiculos: número de vehículos lujosos comprados
- Num_inmuebles: número de inmuebles adquiridos

Suponemos que para los individuos reales, las diferentes variables asociadas a ellos guardan una dependencia estadística natural, esto es, alguien que tiene ingresos altos, seguramente tendrá compras de objetos caros, de la misma manera para personas de ingresos bajos, sus compras deberán ser bajas. Por esta razón, los valores de las diferentes variables se harán con probabilidades diferenciales. En la simulación consideramos un total de 10,000 sujetos con las variables (atributos) simuladas de la siguiente forma:

- * La primera variable es “Estado”. La generación de esta variable toma en cuenta la densidad poblacional de los distintos estados que conforma la República Mexicana. Se sabe (INEGI, 2016) que el total de habitantes en México es aproximadamente 120,000,000 y que por estados el número de habitantes es:

Entidad Federativa	Densidad Poblacional	Densidad Acumulada
Aguascalientes	1 312 544	1 312 544
Baja California	3 315 766	4 628 310
Baja California Sur	712 029	5 340 339
Campeche	899 931	6 240270
Chiapas	5 217 908	11 458 178
Chihuahua	3 556 574	15 014 752
Coahuila de Zaragoza	2 954 915	17 969 667
Colima	711 235	18 680 902
Cuidad de México (Distrito Federal)	8 918 653	27 599 555
Durango	1 754 754	29 354 309
Estado de México	16 187 608	45 541 917
Guanajuato	5 853 677	51 395 594
Guerrero	3 533 251	54 928 845
Hidalgo	2 858 359	57 787 204
Jalisco	7 844 830	65 632 034
Michoacán de Ocampo	4 584 471	70 216 505
Morelos	1 903 811	72 120 316
Nayarit	1 181 050	73 301 366
Nuevo León	5 119 504	78 420 870
Oaxaca	3 967 889	82 388 759
Puebla	6 168 883	88 557 642
Querétaro	2 038 372	90 596 014
Quintana Roo	1 501 562	92 097 576
San Luis Potosí	2 717 820	94 815 396
Sinaloa	2 966 321	97 781 717
Sonora	2 850 330	100 632 047
Tabasco	2 395 272	103 027 319
Tamaulipas	3 441 698	106 469 017
Tlaxcala	1 272 847	107 741 864
Veracruz de Ignacio de la Llave	8 112 505	115 854 369
Yucatán	2 097 175	117 951 544
Zacatecas	1 579 209	119 530 753
TOTAL	≈120 000 000	≈ 120 000 000

Capítulo 3. Uso de modelos no supervisados

Nótese que, la densidad acumulada de Baja California se obtuvo sumando el número de habitantes de Aguascalientes más la densidad poblacional de Baja California, de la misma forma, la densidad acumulada de Baja California Sur fue obtenida de la suma del número de personas que residen de Aguascalientes más la densidad poblacional de Baja California más la densidad poblacional de Baja California Sur y de la misma forma con los estados restantes.

Para dar valores a esta variable, para cada sujeto de la base de datos se generó un número aleatorio $k_1 \in \mathbb{Z}$ entre $1 - 120,000,000$, si tal número aleatorio k_1 resulta estar, por ejemplo, entre la densidad acumulada de Aguascalientes entonces el sujeto reside en ese estado, si el número aleatorio k_1 resulta estar entre la densidad acumulada de Guerrero más uno y entre la densidad acumulada de Hidalgo entonces el sujeto reside en el estado de Hidalgo y de igual manera con los estados restantes.

Se seleccionó la variable “Monto_dep_inu” como la variable de mayor peso y que incide en los valores de las variables restantes. Puesto que los valores más pequeños de esta variable “Monto_dep_inu” son más probables que los valores grandes, se considero reflejar estas probabilidades al generar sus valores.

- * Para dar los valores de la segunda variable, para cada individuo de la base de datos se generó un número aleatorio $k_2 \in \mathbb{Z}$ entre $0 - 19$, si este número está entre $0-11$ al individuo se le asigna un número aleatorio en $cat_{21} = [4000, 12000)$ que corresponde al monto total por depósitos que realizó el sujeto; si el número aleatorio k_2 está entre $12 - 15$ al individuo se le asigna un número aleatorio en $cat_{22} = [12000, 20000)$ como el monto total por depósitos que realizó; si el número aleatorio k_2 está entre $16 - 18$ al individuo se le asigna un número aleatorio entre $cat_{23} = [20000, 85000)$ como el monto total por depósitos que realizó y finalmente, si el número aleatorio k_2 es el número 19 al individuo se le asigna un número aleatorio en $cat_{24} = [85000, 500000)$ como el monto total por depósitos que realizó.

Donde cat_{2j} se refiere a la categoría en la cual cayó el valor de la segunda variable.

- * La tercera variable “Num_dep_inu”, se condicionó a la segunda variable “Monto_dep_inu”, es decir, el número de reportes por depósito depende del monto total. Para dar valores a esta variable, para cada individuo de la base de datos se generó un número aleatorio $k_3 \in \mathbb{Z}$ entre $0 - 19$ y de acuerdo a

la categoría que tiene el individuo de la segunda variable se asignó el valor a la tercera variable como es mostrado en la tabla 3.2:

Categoría de la segunda variable	Categoría de la tercera variable		
	$cat_{31} = \{1, 2, 3\}$	$cat_{32} = \{5, 6, 7\}$	$cat_{33} = \{8, \dots, 15\}$
$cat_{21} = [4000, 12000)$	0-13	14-16	17-19
$cat_{22} = [12000, 20000)$	0-9	10-13	14-19
$cat_{23} = [20000, 85000)$	0-9	10-15	16-19
$cat_{24} = [85000, 500000)$	0-1	2-6	7-19

Tabla 3.2: Categoría que se asigna a la tercera variable de acuerdo al valor de la categoría cat_{2j} de la segunda variable y el número aleatorio k_2

Por ejemplo, si el individuo tiene la categoría cat_{22} de la segunda variable y k_3 está entre 0 – 9 al individuo se le asigna un número aleatorio en $cat_{31} = \{1, 2, 3\}$ como el número de depósitos que realizó, si k_3 está entre 10 – 13 al individuo se le asigna un número aleatorio en $cat_{32} = \{5, 6, 7\}$ como el número de depósitos que realizó, y finalmente, si k_3 está entre 14 – 19 al individuo se le asigna un número aleatorio en $cat_{33} = \{8, \dots, 15\}$ como el número de depósitos que realizó.

Nota: Como vimos la tercera variable “número de depósito” depende de las categorías obtenidas de la segunda variable “monto total” y esto es porque, si un individuo realiza un monto total alto es más probable que realice una gran cantidad de depósitos y viceversa, esta información es la que está contenida en la tabla 3.2.

- * La cuarta variable “Num_ret_inu”, también se condicionó a la segunda variable “Monto_dep_inu”, es decir, el número de reportes por retiro depende del monto total. Para dar valores a esta variable, para cada individuo de la base de datos se generó un número aleatorio $k_4 \in \mathbb{Z}$ entre 0 – 19 y de acuerdo a la categoría que tiene el individuo de la segunda variable se asignó el valor a la cuarta variable como es mostrado en la tabla 3.3:

Categoría de la segunda variable	Categoría de la cuarta variable			
	$cat_{41} = \{0, 1, 2\}$	$cat_{42} = \{3, 4, 5\}$	$cat_{43} = \{6, 7, 8\}$	$cat_{44} = \{10, \dots, 17\}$
$cat_{21} = [4000, 12000)$	0-15	16-17	18	19
$cat_{22} = [12000, 20000)$	0-7	8-14	15-17	18-19
$cat_{23} = [20000, 85000)$	0-4	5-10	11-17	18-19
$cat_{24} = [85000, 500000)$	0-1	2-4	5-7	8-19

Tabla 3.3: Categoría que se asigna a la cuarta variable de acuerdo al valor de la categoría cat_{2j} de la segunda variable y el número aleatorio k_2

Capítulo 3. Uso de modelos no supervisados

Por ejemplo, si el individuo tiene la categoría cat_{23} de la segunda variable y k_4 está entre 0 – 4 al individuo se le asigna un número aleatorio en $cat_{41} = \{0, 1, 2\}$ como el número de retiros que realizó, si k_4 está entre 5 – 10 al individuo se le asigna un número aleatorio en $cat_{42} = \{3, 4, 5\}$ como el número de retiros que realizó, si k_4 está entre 11 – 17 al individuo se le asigna un número aleatorio en $cat_{43} = \{6, 7, 8\}$ como el número de retiros que realizó, y finalmente, si k_4 está entre 18 – 19 al individuo se le asigna un número aleatorio en $cat_{44} = \{10, \dots, 17\}$ como el número de retiros que realizó.

Nota: Como vimos la cuarta variable “número de retiro” depende de las categorías obtenidas de la segunda variable “monto total” y esto es porque, si un individuo realiza un monto total alto es más probable que realice una gran cantidad de retiro y viceversa, esta información es la que está contenida en la tabla 3.3.

- * La quinta variable “Num_vehiculos”, está condicionada a la variable anterior, es decir “Num_ret_inu”, en otras palabras, el número de vehículos comprados que son lujosos depende del número de retiros. Para dar valores a esta variable, para cada individuo de la base de datos se generó un número aleatorio $k_5 \in \mathbb{Z}$ entre 0 – 19 y de acuerdo a la categoría que tiene el individuo de la cuarta variable se asignó el valor a la quinta variable como es mostrado en la tabla 3.4:

		Categoría de la quinta variable	
		$cat_{51} = \{0\}$	$cat_{52} = \{1, 2\}$
Categoría de la cuarta variable	$cat_{41} = \{0, 1, 2\}$	0-18	19
	$cat_{42} = \{3, 4, 5\}$	0-16	17-19
	$cat_{43} = \{6, 7, 8\}$	0-14	15-19
	$cat_{44} = \{0, \dots, 17\}$	0-7	8-19

Tabla 3.4: Categoría que se asigna a la quinta variable de acuerdo al valor de la categoría cat_{4j} de la cuarta variable y el número aleatorio k_4

Por ejemplo, si el individuo tiene la categoría cat_{41} de la cuarta variable y k_5 está entre 0 – 18 al individuo se le asigna el número en $cat_{51} = \{0\}$ como el número de vehículo que compró, si k_5 es 19 al individuo se le asigna un número aleatorio en $cat_{52} = \{1, 2\}$ como el número de vehículo que compró.

Nota: Como vimos la quinta variable “número de vehículos lujosos comprados” depende de las categorías obtenidas de la cuarta variable “número de retiro” e implícitamente de la segunda variable “Monto_dep_inu”, esto es

porque, si un individuo realiza una gran cantidad de retiro es más probable que realice alguna compra de vehículos lujosos y viceversa, esta información es la que está contenida en la tabla 3.4.

- * Finalmente, la sexta variable “Num_inmuebles”, de igual manera depende de la cuarta variable “Num_ret_inu”, en otras palabras, número de inmuebles adquiridos depende del número de retiros. Para dar valores a esta variable, para cada individuo de la base de datos se generó un número aleatorio $k_6 \in \mathbb{Z}$ entre 0 – 19 y de acuerdo a la categoría que tiene el individuo de la cuarta variable se asignó el valor a la sexta variable como es mostrado en la tabla 3.5:

		Categoría de la sexta variable		
		$cat_{61} = \{0\}$	$cat_{62} = \{1\}$	$cat_{63} = \{2\}$
Categoría de la cuarta variable	$cat_{41} = \{0, 1, 2\}$	0-17	18	19
	$cat_{42} = \{3, 4, 5\}$	0-16	17-18	19
	$cat_{43} = \{6, 7, 8\}$	0-11	12-17	18-19
	$cat_{44} = \{10, \dots, 17\}$	0-2	3-7	8-19

Tabla 3.5: Categoría que se asigna a la sexta variable de acuerdo al valor de la categoría cat_{4j} de la cuarta variable y el número aleatorio k_4

Por ejemplo, si el individuo tiene la categoría cat_{44} de la cuarta variable y k_6 está entre 0 – 2 al individuo se le asigna el número en $cat_{61} = \{0\}$ como el número de inmueble que adquirió, si k_6 está entre 3 – 7 al individuo se le asigna el número en $cat_{62} = \{1\}$ como el número de inmueble que adquirió y finalmente, si k_6 está entre 8 – 19 al individuo se le asigna el número en $cat_{62} = \{2\}$ como el número de inmueble que adquirió.

Nota: Como vimos la sexta variable “número de inmueble adquiridos” depende de las categorías obtenidas de la cuarta variable “número de reportes por retiro” e implícitamente de la segunda variable “Monto_dep_inu”, esto es porque, si un individuo realiza una gran cantidad de retiro es más probable que realice alguna compra de inmueble y viceversa, esta información es la que está contenida en la tabla 3.5.

Al final, se tienen un vector de 6 variables $x_i = (x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i})$ para cada uno de los 10,000 individuos en la base de datos, es decir, cada uno de ellos tiene asociado un vector aleatorio con las siguientes coordenadas:

$$x_{1i} = \text{Riesgo_domicilio}(\text{Estado})$$

$$x_{2i} = \text{Monto_dep_inu}$$

Capítulo 3. Uso de modelos no supervisados

$$\begin{aligned}x_{3i} &= \text{Num_dep_inu} \\x_{4i} &= \text{Num_ret_inu} \\x_{5i} &= \text{Num_vehiculos} \\x_{6i} &= \text{Num_inmuebles}\end{aligned}$$

Una vez que tenemos esta base de datos usaremos el método de variables ficticias para obtener la base de datos acondicionada para la aplicación de los métodos no supervisados, esto es, a cada categoría de las distintas variables se le asigna una columna en la nueva base de datos.

La variable x_1 toma 32 valores posibles, la agrupamos en dos categorías correspondientes al estado en riesgo o no de que se realice la trata en el mismo (los estados rojos son riesgosos y los estados blancos son no riesgosos), véase tabla A.1, en consecuencia se crearon 2 variables ficticias: z_{11}, z_{12} , para esta variable toma los siguientes valores

- La variable $z_{11} = 1$, si x_1 toma los valores de los estados del país que están en zona roja y $z_{11} = 0$ en otro caso
- La variable $z_{12} = 1$, si x_1 toma los valores de los estados del país que no están en zona roja y $z_{12} = 0$ en otro caso

La variable x_2 toma 4 categorías, por consiguiente se crearon 4 variables ficticias: $z_{21}, z_{22}, z_{23}, z_{24}$, que serán llenadas de la siguiente forma:

- si $x_2 \in [4000, 12000)$ entonces $z_{21} = 1, z_{22} = 0, z_{23} = 0$ y $z_{24} = 0$
- si $x_2 \in [12000, 20000)$ entonces $z_{21} = 0, z_{22} = 1, z_{23} = 0$ y $z_{24} = 0$
- si $x_2 \in [2000, 85000)$ entonces $z_{21} = 0, z_{22} = 0, z_{23} = 1$ y $z_{24} = 0$
- si $x_2 \in [85000, 500000)$ entonces $z_{21} = 0, z_{22} = 0, z_{23} = 0$ y $z_{24} = 1$

De la misma manera, para la variable x_3 , se crearon 3 variables ficticias; una por cada categoría de la variable x_3 : z_{31}, z_{32}, z_{33} y son llenadas de la siguiente manera:

- si $x_3 \in \{0, 1, 2\}$ entonces $z_{31} = 1, z_{32} = 0$ y $z_{33} = 0$
- si $x_3 \in \{5, 6, 7\}$ entonces $z_{31} = 0, z_{32} = 1$ y $z_{33} = 0$
- si $x_3 \in \{8, \dots, 15\}$ entonces $z_{31} = 0, z_{32} = 0$ y $z_{33} = 1$

Igualmente, para la variable x_4 , se elaboraron un total de 4 variables ficticias: $z_{41}, z_{42}, z_{43}, z_{44}$ y cuyo llenado es el siguiente:

Capítulo 3. Uso de modelos no supervisados

- si $x_4 \in \{0, 1, 2\}$ entonces $z_{41} = 1, z_{42} = 0, z_{43} = 0$ y $z_{44} = 0$
- si $x_4 \in \{3, 4, 5\}$ entonces $z_{41} = 0, z_{42} = 1, z_{43} = 0$ y $z_{44} = 0$
- si $x_4 \in \{6, 7, 8\}$ entonces $z_{41} = 0, z_{42} = 0, z_{43} = 1$ y $z_{44} = 0$
- si $x_4 \in \{10, \dots, 17\}$ entonces $z_{41} = 0, z_{42} = 0, z_{43} = 0$ y $z_{44} = 1$

Similarmente, para la variable x_5 , se elaboraron un total de 2 variables ficticias: z_{51}, z_{52} que toma los siguientes valores:

- si $x_5 \in \{0\}$ entonces $z_{51} = 1$ y $z_{52} = 0$
- si $x_5 \in \{1, 2\}$ entonces $z_{51} = 0$ y $z_{52} = 1$

Finalmente, para la variable x_6 , se crearon un total de 3 variables ficticias: z_{61}, z_{62}, z_{63} que tomaron los siguientes valores:

- si $x_6 \in \{0\}$ entonces $z_{61} = 1, z_{62} = 0$ y $z_{63} = 0$
- si $x_6 \in \{1\}$ entonces $z_{61} = 0, z_{62} = 1$ y $z_{63} = 0$
- si $x_6 \in \{2\}$ entonces $z_{61} = 0, z_{62} = 0$ y $z_{63} = 1$

La siguiente tabla ilustra lo explicado anteriormente. Esta tabla es una pequeña fracción que se utilizó en la simulación.

Sujeto	Base de datos						Categorías de la base de datos						Base de datos con variables ficticias																	
	x_1	x_2	x_3	x_4	x_5	x_6	cat_2	cat_3	cat_4	cat_5	cat_6		x_1			x_2			x_3			x_4			x_5		x_6			
	z_{11}	z_{12}	z_{21}	z_{22}	z_{23}	z_{24}	z_{31}	z_{32}	z_{33}	z_{41}	z_{42}		z_{43}	z_{44}	z_{51}	z_{52}	z_{61}	z_{62}	z_{63}											
1	1	7093	1	0	0	0	1	1	1	1	1	→	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	
2	11	7373	12	4	0	0	1	3	2	1	1	→	1	0	1	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0
3	15	12527	11	11	2	2	2	3	4	2	3	→	1	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	1
4	9	7794	1	1	2	0	1	1	1	2	1	→	1	0	1	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0
5	15	23145	6	4	0	0	3	2	2	1	1	→	1	0	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0
6	30	5395	1	1	0	0	1	1	1	1	1	→	1	0	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0
7	27	6691	6	3	0	0	1	2	2	1	1	→	1	0	1	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0
8	7	13840	1	0	0	0	2	1	1	1	1	→	0	1	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0	0
9	9	14925	14	4	0	1	2	3	2	1	2	→	1	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0
10	21	77838	1	4	0	0	3	1	2	1	1	→	1	0	0	0	1	0	1	0	0	0	1	0	0	1	0	1	0	0

Tabla 3.6

En la tabla 3.6, los valores que toma la variable x_1 son los números de las Entidades Federativas de acuerdo al orden dado en el Catálogo de Riesgo (tabla A.1).

Capítulo 3. Uso de modelos no supervisados

Del conjunto de los 10,000 individuos que se generó en la base de datos, suponemos que 50 de ellos son tratantes identificados. Estos 50 individuos los seleccionamos por su lugar de origen (todos son de un lugar catalogado como de alta incidencia de trata de personas), por el manejo de sumas altas de dinero (todos tienen una alta actividad bancaria que es un número alto de depósitos, un número alto de retiros y monto altos en estos manejos).

Algunas de estas características estuvieron presentes en un conjunto de 54 personas de casos reales que obtuvimos en una investigación documental en periódicos, revistas y noticieros (tabla A.2). De esta información se pudo extraer los estados de alta incidencia de trata de personas, también, el número de inmuebles y el número de vehículos.

Para generar las características restantes, se tomó en cuenta criterios que nos parecieron razonables, por ejemplo, si un individuo deposita un monto total alto es más probable que realice un alto número de retiros y también, es más probable que realice alguna compra de inmueble, así con las características restantes.

Entrando en materia, sea V el conjunto de sujetos que generamos en la base de datos, W el conjunto de atributos que poseen los sujetos, es decir:

$$V = \{\text{individuo } 1, \dots, \text{individuo } 10000\}$$

$$W = \{z_{11}, z_{12}, \dots, z_{62}, z_{63}\}$$

Llamemos A el conjunto de sujetos que son tratantes identificados, una vez definido nuestros conjuntos de interés, observamos que estamos interesados en aquellos sujetos del conjunto $V - A$ que satisfagan los atributos que tengan en común los individuos en A , o que poseen en común la mayoría de los atributos en A , en otra palabras, supongamos que el conjunto A tiene en común 3 atributos, es decir, cada uno de los sujetos del conjunto A posee estos 3 atributos, entonces estamos interesados en aquellos sujetos del conjunto $V - A$ que tengan estos 3 atributos conjuntamente y también estamos interesados en aquellos que tengan un cierto porcentaje de estos 3 atributos, por ejemplo, todos aquellos sujetos que tengan cualesquiera 2 atributos mencionados con anterioridad.

Aplicamos el operador “” al conjunto A , lo que nos dio como resultado fue el conjunto de atributos que poseen los individuos del conjunto A ; como sabemos este conjunto de atributos $A' \subset W$. Aplicando de nuevo el operador “” pero en este caso al conjunto A' , obtenemos un conjunto $A'' \subset V$ tal que $A \subset A''$.

Es posible que los tratantes no todos tengan los atributos en A' , es decir, hay un porcentaje que no tengan todos los atributos pero pueden ser tratantes. Nos gustaría también seleccionar a los conjuntos que tenga esas características. Para estos conjuntos usaremos la Canasta de Mercado, es decir, analizaremos los atributos que tengan un soporte alto $T(\mathcal{K})$, por ejemplo, los 50 sujetos del conjunto A tiene en común los atributos de A' , pero que pasa, si, digamos 45 sujetos de la base de datos simulada satisfacen un atributo $w \notin A'$, el FCA lo descartaría de manera automática ya que no lo tienen en común todos los sujetos, sin embargo hay que tener en consideración estos atributos ya que tienen un “alto” soporte. Entonces aplicamos un FCA modificado considerando soportes altos, con $T(\mathcal{K}) = 0.85$.

Capítulo 4

Resultados y conclusiones

Al estudiar fenómenos como la trata de personas, en las cuales, tanto las víctimas como los agresores intentan mantener sus actividades clandestinamente y en secreto, resulta difícil poder identificar situaciones y personas cuando existe esta práctica. Aunado a esta situación en las dependencias donde están en custodia datos que podrían ser utilizados para determinar posibles sospechosos, no tienen autorizado su exhibición para posibles estudios. En este sentido, se ha utilizado exitosamente el análisis de conceptos formales con datos reales de los informes policíacos en la ciudad de Amsterdam. Esto puede dar esperanza de que al ser aplicados a bases de datos existente en diversas dependencias que tienen como objetivo perseguir estos tipos de delitos.

En la literatura no pudimos encontrar metodología tendientes a modelar matemáticamente el fenómeno de trata de personas, a pesar que desde el punto de vista de las Ciencias Sociales, el problema está bastante estudiado y se tienen reportes descriptivos de las características, tanto de víctimas como de los tratantes, en los casos que han sido identificados. El único artículo encontrado por nosotros donde se da un modelo matemático y se presenta una metodología para intentar obtener la ubicación de este tipo de actividades es el escrito por Poelmans et al (2012).

Ante la imposibilidad de utilizar bases de datos reales, tomamos dos caminos: uno fue, hacer una investigación en periódicos, noticieros, etc, de las personas que en los últimos años han sido arrestados o buscados por cometer el delito de trata y con los datos recolectados se hizo un análisis descriptivo para identificar posibles características comunes. Si bien no se tiene información sobre todas las caracte-

Resultados y conclusiones

terísticas de los tratantes en los medios noticieros revisados, si se pudo realizar el análisis estadístico para ver sus características comunes.

El otro camino fue, generar una base de datos con 10000 registros, en los cuales se tiene 6 características, estas características son un subconjunto de las características que pudimos construir, para conocer que características utilizar en nuestra base simulada, se revisó la página de la Secretaría de Hacienda y Crédito Público, esta información es pública y disponible para la consulta de cualquier interesado.

Para la elaboración de nuestra base de datos supusimos relaciones de dependencia entre las distintas características que contiene, con esta base de datos probamos la metodología para seleccionar a un posible conjunto de sospechosos, dada la información de individuos ya procesados. El método seleccionó a 135 individuos de nuestra base de datos como posibles sospechosos. Sin embargo, no podemos evaluar la exactitud de nuestros resultados, una vez que estos 135 individuos deberían ser investigados para determinar cuántos de ellos son realmente tratantes.

Generamos una base de 10000 individuos mediante una simulación, teniendo la base de datos se aplicaron los conceptos formales tomando como base un conjunto de 50 individuos que consideramos que eran tratantes bien identificados, esto fue nuestro conjunto inicial A , al aplicar la operación de A' obtuvimos 135 posibles tratantes. Realizamos otro estudio en el que consideramos menos características comunes, el número de resultado fue 172, esto es razonable porque estamos quitando restricciones al conjunto seleccionado, de cualquier manera, el conjunto de posibles sospechosos tendrían que ser investigados.

Consideramos que el algoritmo funciona y que podría ser aplicado a datos de la Secretaría de Hacienda, de los Servicios de Aduana, de la Procuraduría General de la República (PGR), etc. Después de identificar a los sospechosos, se debe hacer una investigación profunda para determinar si son casos reales de trata como lo hicieron los autores del artículo *Semi-automated Knowledge discovery: identifying and profiling human trafficking* escrito por *Poelmans Jonas, Elzinga Paul, Ignato Dmitry I. y Kuznetsov Sergei O* en 2012, en el que se dieron arrestos reales.

En este trabajo se puede apreciar la importancia que tiene los modelos de aprendizaje no supervisado para el manejo de bases grandes de datos, en particular el FCA. El FCA es una técnica que se puede utilizar para exponer, investigar o relacionar grandes cantidades de información. Si el número de objetos es muy grande, los

reticulos obtenidos se pueden volver complejos, sin embargo, para simplificar los análisis se pueden considerar sólo algunos subconjuntos del reticulo para obtener conclusiones.

Finalmente, consideramos que se alcanzaron los objetivos propuestos a pesar de tener las limitaciones de no poder utilizar la metodología con los datos reales y por lo tanto no pudimos evaluar el modelo ante la imposibilidad de hacer un seguimiento para poder hacer una evaluación de los resultados.

Esperemos que este trabajo permita a otros investigadores a retomar este camino y aplicarlo a datos reales, y que puedan hacer la evaluación de este modelo.

Las organizaciones que se beneficiarán con la metodología propuesta en este trabajo, sin duda alguna serían: la Secretaría de Hacienda, los Servicios de Aduana, la Procuraduría General de la República (PGR), entre otras.

Apéndice **A**

Apéndice

En esta sección presentamos la construcción del Catálogo de Riesgo por estado relacionado con la trata de personas en México que utilizamos en la simulación en el capítulo 3.

Así mismo, presentamos una investigación de campo del delito de trata de personas en México.

A.1. Catálogo de Riesgo por estado relacionado con la trata de personas en México

El contenido de la tabla siguiente ya se mencionó con anterioridad y representa las entidades federativas con mayor riesgo o con grandes posibilidades de que están haciendo trata en ellos. En esta tabla se presenta los estados de la república, organizados en esta:

Núm	Entidad Federativa	Con o sin Riesgo
1	Aguascalientes	Con riesgo
2	Baja California	Con riesgo
3	Baja California Sur	Con riesgo
4	Campeche	Sin riesgo
5	Chiapas	Con riesgo
6	Chihuahua	Con riesgo
7	Coahuila de Zaragoza	Sin riesgo
8	Colima	Con riesgo
9	Cuidad de México (Distrito Federal)	Con riesgo
10	Durango	Sin riesgo
11	Estado de México	Con riesgo
12	Guanajuato	Con riesgo
13	Guerrero	Con riesgo
14	Hidalgo	Con riesgo
15	Jalisco	Con riesgo
16	Michoacán de Ocampo	Con riesgo
17	Morelos	Con riesgo
18	Nayarit	Con riesgo
19	Nuevo León	Sin riesgo
20	Oaxaca	Con riesgo
21	Puebla	Con riesgo
22	Querétaro	Sin riesgo
23	Quintana Roo	Con riesgo
24	San Luis Potosí	Sin riesgo
25	Sinaloa	Sin riesgo
26	Sonora	Con riesgo
27	Tabasco	Con riesgo
28	Tamaulipas	Sin riesgo
29	Tlaxcala	Con riesgo
30	Veracruz de Ignacio de la Llave	Con riesgo
31	Yucatán	Sin riesgo
32	Zacatecas	Sin riesgo

Tabla A.1

A.2. Investigación Documental

Realizamos una investigación en los diferentes medios periodísticos (La Jornada, El Universal, Excelsior, etc.), esta información es pública, no pretendemos hacer esta investigación exhaustiva pero estos datos dan una pauta del tipo de personas que comenten el delito de trata con fines de explotación sexual.

Los datos obtenidos se presentan en la siguiente tabla. Por la extensión de los encabezados de las columnas, se consideró escribir una letra con el siguiente significado:

a: Numero_Averiguacion	h: Localidad_Operacion
b: Nombre del Tratante	i: Estado_Operacion
c: Sexo	j: Pais_Nac_Tratantes
d: Edad del Tratante	k: Numero_Victimas
e: Actividad Economica	l: Pais_Nac_Victimas
f: Inmueble donde operan	m: Localidad_Origen_victimas
g: Colonia_Operacion	n: Descripcion_Captacion_victimas

Para la variable l: Pais_Nac_Victimas, se tiene diferentes valores optamos por usar su abreviatura, esto es:

Mex: México,

Ven: Venezuela,

Dom: República Dominicana

Hon: Honduras

Col: Colombia

Esl: Eslovenia

a	b	c	d	e	f	g	h	i	j	k	l	m	n
1	Sujeto 1	H	-	Fotógrafo y Empresario	Casa de huéspedes	-	-	Chihuahua	-	-	-	Chihuahua	Engaño con propuesta de trabajo como modelos y retratos
2	Sujeto 2	H	44	-	-	-	Nueva York	USA	México	25	México	México	Engaño a sus víctimas (Los retaban en México y después las llevaban al USA)
	Sujeto 3	H	28	-	-	-	Nueva York	USA	México	25	México	-	-
3	Sujeto 4	H	-	-	Bar supuestamente con karaoke	-	Chalatenunc	Ciudad de México	México	6	México, Corea	Corea del Sur	Engaño con propuesta de trabajo bien remunerado en la República Mexicana
	Sujeto 5	M	-	-	Bar supuestamente con karaoke	-	Chalatenunc	Ciudad de México	Asia	6	México, Corea	Corea del Sur	-
4	Sujeto 6	H	-	Hoteles y Servicios de masaje	Su casa (hotel)	-	Chalatenunc	Ciudad de México	-	-	-	-	Engaño con propuesta de trabajo de masajistas
5	Sujeto 7	H	-	Dueño del bar "Desexxx3"	Bar	-	Chalatenunc	Ciudad de México	-	17	-	-	Engaño con propuesta de trabajo de bailarinas
	Sujeto 8	H	-	Empleados del bar "Desexxx3"	Bar	-	Chalatenunc	Ciudad de México	-	17	-	-	-
	Sujeto 9	H	-	Empleados del bar "Desexxx3"	Bar	-	Chalatenunc	Ciudad de México	-	17	-	-	-
	Sujeto 10	M	-	Empleados del bar "Desexxx3"	Bar	-	Chalatenunc	Ciudad de México	-	17	-	-	-
6	Sujeto 11	H	-	Dueño de los bares 1 y 2	2 bares	Barrio San Antonio	Xochimilco	Ciudad de México	-	16	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
	Sujeto 12	H	-	Sotro de los bares 1 y 2	2 bares	Barrio San Antonio	Xochimilco	Ciudad de México	-	16	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
	Sujeto 13	H	-	Encargado y mesero del bar 1	2 bares	Barrio San Antonio	Xochimilco	Ciudad de México	-	16	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
	Sujeto 14	M	-	Cajera del bar 1	2 bares	Barrio San Antonio	Xochimilco	Ciudad de México	-	16	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
	Sujeto 15	H	-	Encargado del bar 2	2 bares	Barrio Yaltocan	Xochimilco	Ciudad de México	-	6	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
	Sujeto 16	M	-	Cajera del bar 2	2 bares	Barrio Yaltocan	Xochimilco	Ciudad de México	-	6	México	D. F., Guerrero, Nuevo Leon, Veracruz	-
7	Sujeto 17	H	-	-	-	La jira	Tlalpau	Ciudad de México	-	9	-	-	Engaño con propuesta de trabajo de bailarinas
	Sujeto 18	H	-	-	-	La jira	Tlalpau	Ciudad de México	-	9	-	-	-
	Sujeto 19	H	-	-	-	La jira	Tlalpau	Ciudad de México	-	9	-	-	-
8	Sujeto 20	M	-	Dueña del bar "Los Amigos"	Bar	Panajé San Juan	Izapaapa	Ciudad de México	-	7	-	-	-
	Sujeto 21	H	-	Empleados del bar "Los Amigos"	Bar	Panajé San Juan	Izapaapa	Ciudad de México	-	7	-	-	-
	Sujeto 22	M	-	Empleados del bar "Los Amigos"	Bar	Panajé San Juan	Izapaapa	Ciudad de México	-	7	-	-	-
9	Sujeto 23	H	-	Encargada del bar 2	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 24	M	-	Encargada del bar 2	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 25	H	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 26	M	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 27	M	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 28	M	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 29	M	-	Encargado del bar 3	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 30	H	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 31	M	-	Encargado del bar 4	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-
	Sujeto 32	M	-	-	4 Bares	Colonia Tacuba	Miguel Hidalgo	Ciudad de México	-	33	-	-	-

Apéndice

a	b	c	d	e	f	g	h	i	j	k	l	m	n
10	Sujeto 33	H	-	-	Vehículo blanco, Laptop	-	Tapachula	Chiapas	-	-	-	-	Engaño con propuesta de trabajo
	Sujeto 34	H	-	-	Vehículo blanco, Laptop	-	Tapachula	Chiapas	-	-	-	-	-
11	Sujeto 35	H	-	-	Bar	Colonia San Pedro de Los Pinos	Ávaro Obregón	Ciudad de México	-	47	Mex, Ven, Dom, Hon, Cuba, Col y Esl	-	Engaño con propuesta de trabajo de bailarinas
	Sujeto 36	H	-	-	Bar	Colonia San Pedro de Los Pinos	Ávaro Obregón	Ciudad de México	-	47	Mex, Ven, Dom, Hon, Cuba, Col y Esl	-	-
	Sujeto 37	H	-	-	Bar	Colonia San Pedro de Los Pinos	Ávaro Obregón	Ciudad de México	-	47	Mex, Ven, Dom, Hon, Cuba, Col y Esl	-	-
	Sujeto 38	M	-	-	Bar	Colonia San Pedro de Los Pinos	Ávaro Obregón	Ciudad de México	-	47	Mex, Ven, Dom, Hon, Cuba, Col y Esl	-	-
	Sujeto 39	H	-	-	Bar	Colonia San Pedro de Los Pinos	Ávaro Obregón	Ciudad de México	-	47	Mex, Ven, Dom, Hon, Cuba, Col y Esl	-	-
12	Sujeto 40	H	-	Encargado de uno de los bares	5 bares	Colonia Centro	Cuahtémoc	Ciudad de México	-	26	-	-	Engaño con propuesta de trabajo como meseras
	Sujeto 41	H	-	Dueño del otro de los negocios	5 bares	Colonia Centro	Cuahtémoc	Ciudad de México	-	26	-	-	-
	Sujeto 42	H	-	Dueño de dos locales	5 bares	Colonia Centro	Cuahtémoc	Ciudad de México	-	26	-	-	-
	Sujeto 43	H	-	Dueño de uno más	5 bares	Colonia Centro	Cuahtémoc	Ciudad de México	-	26	-	-	-
13	Sujeto 44	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 45	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 46	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 47	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 48	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 49	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 50	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 51	H	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 52	M	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 53	M	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-
	Sujeto 54	M	-	-	4 bares	-	Venustiano Carranza	Ciudad de México	-	55	-	-	-

Tabla A.2: Investigación documental

De esta tabla podemos hacer un resumen descriptivo de los datos, lo que nos genera lo siguiente:

Número de sospechosos de este delito: 54, de los cuales, 37 son del sexo masculino y 17 del sexo femenino. Aproximadamente el 5 % pertenece la nacionalidad mexicana, el 1 % a extranjera y el 96 % desconocida. Se confiscaron 22 bares, 1 Laptop y 1 vehículo. Su actividad económica está relacionada a los bares confiscados: 6 eran dueños, 7 eran encargados, 2 eran cajeros, 5 eran empleados y 1 era socio; de profesión libre: 1 se hacía pasar como fotógrafo y empresario, 1 se hacía pasar como masajista; y 31 se desconoce su actividad económica.

Respecto a los lugares donde ejercía su actividad de explotación, aproximadamente el 96 % en México (91 % Ciudad de México, 4 % Chiapas y 1 % Chihuahua) y el 4 % en Estados Unidos de América (USA).

Respecto a las víctimas, se identificó a 247 víctimas de este delito, de las cuales, 2 son menores de edad. Aproximadamente el 19 % pertenecía a la nacionalidad mexicana, el 13 % son extranjeras (2 % corea, 11 % se distribuye en Venezuela, República Dominicana, Honduras, Cuba, Colombia y Eslovenia) y 68 % desconocida.

Respecto a la captación de las víctimas, se identificó que aproximadamente 60 % de los 13 casos encontrados, era mediante el engaño de una oferta de trabajo, 10 % de los casos era mediante el enamoramiento y 30 % se desconoce.

Bibliografía

- [1] Aztecanoticias, (2014, septiembre 28). *México, segundo país con mas casos de trata de personas*. Recuperado el 24 de octubre del 2015, de <http://www.aztecanoticias.com.mx/notas/seguridad/202166/mexico-segundo-pais-con-mas-casos-de-trata-de-personas>
- [2] Belohlávek Radim, (2008). *Introduction to Formal Concept Analysis*, Department of computer Science, Faculty of Science Palacký University, Olpmouc.
- [3] Código Penal Federal, (2016, julio 18). *Nuevo Código Publicado en el Diario Oficial de la Federación el 14 de agosto de 1931*. Recuperado el 10 de noviembre del 2016, de http://www.diputados.gob.mx/LeyesBiblio/pdf/9_180716.pdf
- [4] Comisión Nacional Bancaria y de Valores (CNBV), (2016). *Documentos*. Recuperado el 10 de noviembre del 2016, de <http://www.gob.mx/cnbv/documentos>
- [5] Comisión Nacional de los Derechos Humanos (CNDH), (2013). *Diagnóstico sobre la Situación de la Trata de Personas en México*, pp. 15-18 y 90-94.
- [6] El Universal, (2010). *Se dispara en México la trata de personas*. Recuperado el 13 de febrero del 2016, de <http://archivo.eluniversal.com.mx/primer/36044.html>
- [7] Fondo de las Naciones Unidas para la Infancia (UNICEF, por sus siglas en inglés), (2010). *Trata de personas. Una forma de esclavitud moderna. Un fenómeno mundial que afecta principalmente a niños, niñas y adolescentes*, Argentina Primera edición, mayo.

- [8] Fortuna (2012). *México, entre los países más propensos al lavado de dinero*. Recuperado el 1 de marzo del 2017, de <http://revistafortuna.com.mx/contenido/2012/07/19/ley-contra-lavado-de-dinero-en-suspenso/>
- [9] Friedman Jerome, Hastie Trevor, Tibshirani Robert, (2008). *The Elements of Statistical Learning*, Second Edition, Springer Editorial, August, pp. 485-579.
- [10] Ganter Bernhard, (2002). *Formal Concept Analysis: Methods and Applications in Computer Science*. Universität Magdeburg.
- [11] Ganter Bernhard, Stumme G., Wille R., (2005). *Formal Concept Analysis. Foundations and Applications*. Springer.
- [12] Ganter Bernhard, Wille R., (1999). *Formal Concept Analysis. Mathematical Foundations*. Springer, pp. 1-15
- [13] Hongliang Lai, Dexue Zhang, (2008). *Concept lattices of fuzzy contexts: Formal concept analysis vs. rough set theory*. Department of Mathematics, Sichuan University, 24, South Section 1, Yihuan Road, Chengdu 610064, China.
- [14] Instituto Nacional de Estadística y Geografía (INEGI), (2016). *Número de habitantes*. Recuperado el 4 de diciembre del 2016, de <http://cuentame.inegi.org.mx/poblacion/habitantes.aspx?tema=P>
- [15] Instituto Nacional de Migración (INM), la Organización Internacional para las Migraciones (OIM) y la Comisión Interamericana de Mujeres de la Organización de Estados Americanos (CIM-OEA), (2006). *Trata de personas: Aspectos Básicos*, 1a. Edición.
- [16] La Izquierda Diario (2015). *Terribles condiciones laborales para las trabajadoras de la industria textil y manufacturera*. Recuperado el 9 de marzo del 2017, de <http://www.laizquierdadiario.mx/Terribles-condiciones-laborales-para-las-trabajadoras-de-la-industria-textil-y-manufacturera>
- [17] La Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC, por sus siglas en inglés), (2013). *Diagnóstico sobre la Situación de la Trata de Personas en México*, 1a. Edición.
- [18] La Secretaría de Hacienda y Crédito Público (SHCP), (2016, noviembre). *Documentos*. Recuperado el 10 de noviembre del 2016, de <http://www.hacienda.gob.mx>

Bibliografía

- [19] México Unido Contra la Delincuencia (2014). *Víctimas de trata en México se duplican en el último año*. Recuperado el 7 de marzo del 2017. de <http://www.mucd.org.mx/V%C3%ADctimas-de-trata-en-México-se-duplican-en-el-último-año-n3392.html>
- [20] Trata de Blancas: Internet y Redes Sociales, (2014). Disponible en: <http://redesdetratadeblancas.blogspot.mx/2014/11/internet-y-redes-sociales.html>. Consultado el: 2 de noviembre del 2015
- [21] Organización de las Naciones Unidas (ONU), (2002). *Principios y Directrices recomendados sobre los derechos humanos y la trata de personas*, presentado por el Consejo Económico y Social como adición al Informe del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (E/2002/68/Add.1).
- [22] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, (2006). *Introduction to Data Mining*, Pearson Addison-Wesley, pp. 327-363.
- [23] Petko Valtchev Robert Jäschke, (2011). *Formal Concept Analysis*. 9th International Conference, ICFCA 2011 Nicosia, Cyprus, May 2-6, Proceedings, pp. 107-118.
- [24] Pfaltz John L, Taylor, C. M, (2002). *Scientific knowledge discovery through iterative transformation of concept lattices*. In SIAM Workshop on Discrete Math. and Data Mining, Arlington, VA, USA.
- [25] Poelmans Jonas, Elzinga Paul, Ignato Dmitry I., Kuznetsov Sergei O., (2012). *Semi-automated Knowledge discovery: identifying and profiling human trafficking*, Int. J. Gen. Syst. 4, no. 8, 774-804.06B23, Nombriembre.
- [26] Portal de Prevención de Lavado de Dinero (LFPIORPI), (2016, noviembre). *Actividades Vulnerables*. Recuperado el 10 de noviembre del 2016, de <https://sppld.sat.gob.mx/pld/index.html>
- [27] Requena Espada Laura, Giménez-Salinas Framis Andrea, Espinosa Manuel de Juan, (2012). *ESTUDIAR LA TRATA DE PERSONAS. Problemas metodológicos y propuestas para su resolución*. Recuperado el 10 de octubre del 2015, de <http://criminet.ugr.es/recpc/14/recpc14-13.pdf>
- [28] Rodríguez González Ansel Yoan, Martínez Trinidad José Francisco, Carrasco Ochoa Jesús Ariel, Shulcloper José Ruiz, (2009). *Minería de Reglas de Asociación sobre Datos Mezclados*. Coordinación de Ciencias Computacionales, INAOE, pp.1-11.

- [29] Sioen Mark, (2012). *Formal Concept Analysis*. Faculteit Wetenschappen en Bio-ingenieurswetenschappen, Departement Wiskunde.
- [30] Trata de Blancas: Internet y Redes Sociales, (2014). Recuperado el 2 de noviembre del 2015, de <http://redesdetratadeblancas.blogspot.mx/2014/11/internet-y-redes-sociales.html>.
- [31] Víctimas de trata en México se duplican en el último año: la mayoría son mujeres y niñas, (2014). Recuperado el 25 de diciembre del 2015, de <http://www.animalpolitico.com/2014/11/victimas-de-trata-en-mexico-se-duplican-en-el-ultimo-ano-la-mayoria-son-mujeres-y-ninas/>
- [32] Women's Link Worldwide, (2010). *Strategies for Gender Justice*.
- [33] Yao, Y.Y.,(2004). *Concept Lattices in Rough Set Theory*. Proceedings of Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS, 2004).