Instituto Politécnico Nacional

Centro de Investigación en Computación
Secretaría de Investigación y Posgrado

Caracterización y predicción espacio-temporal de patrones delictivos mediante modelos lógico-combinatorios.

T E S I S
QUE PARA OBTENER EL GRADO DE

MAESTRO EN
CIENCIAS DE LA COMPUTACIÓN

P R E S E N T A

ING. MIL. VÍCTOR MANUEL MARTÍNEZ HERNÁNDEZ

DIRECTORES DE TESIS:

DR. SALVADOR GODOY CALDERÓN
DR. FRANCISCO HIRAM CALVO CASTRO

MÉXICO, D.F., DICIEMBRE, 2009.

SIP-14

INSTITUTO POLITECNICO NACIONAL
SECRETARIA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 18:00 horas del día 27 del mes de octubre de 2009 se reunieron los miembros de la Comisión Revisora de Tesis designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis de grado titulada:

"CARACTERIZACIÓN Y PREDICCIÓN ESPACIO-TEMPORAL DE PATRONES DELICTIVOS MEDIANTE MODELOS LÓGICO COMBINATORIOS"

MARTÍNEZ	HERNÁNDEZ	VÍCTOR MANUEL
Apellido paterno	materno	nombre(s)

Con registro: B 0 7 1 4 5 5

aspirante al grado de: ***MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN***

Después de intercambiar opiniones los miembros de la Comisión manifestaron ***SU APROBACIÓN DE LA TESIS***, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Presidente

Dr. Serguei Pavlovich Levachkine

Secretario

Dr. Marco Antonio Moreno Armendáriz

Primer vocal (Director)

Dr. Salvador Godoy Calderón

Segundo vocal

Dr. José Matías Alvarado Mentado

Tercer Vocal

Dr. Ricardo Barrón Fernández

EL PRESIDENTE DEL COLEGIO

Dr. Jaime Álvarez Gallegos

INSTITUTO POLITECNICO NACIONAL
CENTRO DE INVESTIGACION
EN COMPUTACION
DIRECCION

INSTITUTO POLITECNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA CESION DE DERECHOS

En la Ciudad de México D.F. el día 03 del mes Diciembre del año 2008, el (la) que suscribe Víctor Manuel Martínez Hernández alumno (a) del Programa de Maestría en Ciencias de la Computación con número de registro B071455, adscrito a Centro de Investigación en Computación, manifiesta que es autor (a) intelectual del presente trabajo de Tesis bajo la dirección de Dr. Salvador Godoy Calderón y Dr. Francisco Hiram Calvo Castro y cede los derechos del trabajo intitulado Caracterización y predicción espacio-temporal de patrones delictivos mediante modelos lógico-combinatorios, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección Av. Juan de Dios de Batiz s/n casi esq. Miguel Othón de Mendizábal; Unidad Profesional "Adolfo López Mateos", Edificio CIC. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Víctor Manuel Martínez Hernández

Nombre y firma

Resumen

En este trabajo es presentada una nueva metodología para Análisis y Predicción Delictiva, basada en técnicas de aprendizaje inductivo empleado comúnmente en el área de Reconocimiento de Patrones. La metodología para Análisis Delictivo se basa en la construcción de definiciones inductivas para cada familia delictiva registrada dentro de un conjunto de datos delictivos; con base en esas definiciones, es posible determinar el comportamiento de la actividad delictiva dentro de un espacio-tiempo seleccionado. Posteriormente, se predice la cantidad de hechos delictivos que se espera ocurran**,** para cada una de estas familias delictivas**,** dentro de todas las áreas de una región en estudio. Los resultados así obtenidos prestan un valioso apoyo para la toma de decisiones de seguridad pública, permitiendo mejorar la planificación de rutas de patrullaje y la optimización en la asignación de recursos humanos y / o recursos financieros en esas zonas (y horarios) de mayor incidencia criminal.

Abstract

A new methodology, for crime-analysis and criminal activity forecast, based on inductive learning techniques commonly used in pattern recognition is presented. First, the proposed crime-analysis methodology is based upon the construction of inductive-definitions for each type of criminal family previously registered in a crime dataset. Based on these definitions, it is possible to determine the crime behavior within selected space and time frames; afterwards, a specific criminal activity forecast is constructed for all areas within the region under study. The results thus obtained provide valuable support for the public-security decision making process by allowing decision makers to better plan surveillance patrols and to optimize the use of human and /or financial resources on those areas (and times) of highest criminal incidence.

Agradecimientos

"Lo poco que he aprendido carece de valor, comparado con lo que ignoro y no desespero en aprender."

René Descartes

Agradezco a Dios por bendecirme con la vida y permitirme llegar hasta este momento rodeado de gente maravillosa.

Agradezco enormemente a mi familia, fuente constante de inspiración, amor y cariño. Todo lo que soy se lo debo a ustedes.

Agradezco de manera muy especial a mis directores de tesis, Dr. Salvador Godoy Calderón y Dr. Francisco Hiram Calvo Castro, así como, al Dr. Marco A. Moreno Armendáriz, por todo el apoyo y confianza depositada en mí durante la realización de este trabajo; pero sobre todo por su amistad, que va mas allá del conocimiento y que se lleva de por vida.

Agradezco a los miembros de mi comité tutorial Dr. Sergei Pavlovich Levachkine, Dr. Ricardo Barrón Fernández y al Dr. José Matías Alvarado Mentado, por sus valiosas aportaciones, las cuales me permitieron ver el problema desde diferentes ángulos y lograr con ello el desarrollo de este proyecto.

Agradezco a todo el personal administrativo de este centro de estudios y de manera muy especial al Departamento de Tecnologías Educativas, quienes están constantemente en mis oraciones.

Aquellas personas con lo que compartí momentos especiales y que me han brindaron su amistad incondicional: Adriel, Benja, Bryan, Ceci, Chirbuso, Cynthia, Fer, Jesús, Jeanine, Jenny, Jimmy, Juan, Lily, Mayra, Miguel, Sam, Gaby, Stefi, Gil, Richard, U.V.A., así como, a los demás compañeros de este centro de investigación.

A mis amigos de toda la vida: Tania, Miriam, Tocayo y Claudia.

Agradezco al Instituto Politécnico Nacional (IPN), al Centro de Investigación en Computación (CIC), al Consejo Nacional de Ciencia y Tecnología (CONACyT) y a la Secretaria de Investigación y Posgrado (SIP-IPN), por todo el apoyo otorgado durante la realización de este trabajo.

Agradezco a la Secretaria de la Defensa Nacional (SEDENA), por la beca de estudios otorgada para la realización de estudios de posgrado en esta importante institución.

Contenido

Glosario de Términos

Análisis Delictivo. El conjunto de procesos y análisis sistemáticos dirigidos a proveer información oportuna y pertinente relativa a hechos y correlaciones de tendencia delictiva a el personal operativo y administrativo durante la planeación de acciones tendientes a prevenir y evitar actividades delictivas, ayudar en los procesos de investigación, incrementar la aprensiones de delincuentes, asignación de recursos y esclarecer los casos.

Criminalística. La Criminalística es una ciencia penal auxiliar, que mediante la aplicación de sus conocimientos, metodología y tecnología, descubre y verifica científicamente la existencia de un hecho presuntamente delictuoso y a los responsables, aportando las pruebas a los órganos que procuran y administran justicia.

Criminología. Ciencia social que estudia la naturaleza, extensión y causas del crimen; características de los criminales y de las organizaciones criminales; problemas de detención y castigo de los delincuentes; operatividad de las prisiones y de otras instituciones carcelarias; rehabilitación de los convictos, tanto dentro como fuera de prisión, y la prevención del delito.

Sistema de Información Geográfica (SIG o GIS, en su acrónimo en inglés). es una integración organizada de hardware, software y datos geográficos diseñado para capturar, almacenar, manipular, analizar y desplegar en todas sus formas la información geográficamente referenciada con el fin de resolver problemas complejos de planificación y gestión.

Sistema soporte para la toma de decisiones (SSD o DSS en su acrónimo en inglés). Es un sistema de información basado en un computador interactivo, flexible y adaptable, especialmente desarrollado para apoyar la solución de un problema de gestión no estructurado que mejorar la toma de decisiones. Utiliza datos, proporciona una interfaz amigable y permite la toma de decisiones en el propio análisis de la situación.

Sitio Web. Conjunto de páginas Web, típicamente comunes a un dominio de Internet o subdominio en la World Wide Web en Internet.

Lista de Figuras y Tablas

CAPÍTULO UNO

Introducción

Se describe la problemática de inseguridad pública que actualmente vive México, así como la aportación del presente trabajo de tesis en esta materia.

1.1 Contexto y Antecedentes

El índice de criminalidad en México ha presentado un crecimiento preocupante durante los últimos años. El aumento de la delincuencia en el país ha sido de tal magnitud, que ha ocasionado un fuerte sentimiento de inseguridad entre la población, particularmente en la del Distrito Federal y en otras ciudades como Ciudad Juárez, Tijuana, Monterrey, Michoacán, etc. Este fenómeno, no sólo afecta el patrimonio y la integridad física de las personas, sino también conlleva un deterioro en el estado de derecho, desalienta la inversión**,** principalmente la de largo plazo, debilita la base competitiva del sector productivo nacional y, por ende, disminuye el potencial de crecimiento económico del país **[Coparmex, 2002]**.

Dado que entre las características inherentes a la existencia de un estado de derecho está el adecuado control de las actividades delictivas, es fundamental que las autoridades lleven a cabo un combate eficaz a la delincuencia**,** con el fin de mantener un ambiente institucional que propicie la competitividad y el progreso de México.

Actualmente la seguridad pública y el combate a la delincuencia constituyen una de las mayores preocupaciones sociales no sólo en México sino en todas las grandes ciudades del mundo. A pesar de la enorme cantidad de recursos, tanto humanos como materiales, que los gobiernos destinan a este rubro, aún resulta evidente la necesidad de mecanismos alternativos que permitan volver más eficaz y más eficiente el trabajo de las fuerzas policiacas. Una de las principales variables que limitan la eficacia y eficiencia de dicho trabajo es el tiempo de respuesta a los acontecimientos delictivos. Inclusive los mayores esfuerzos en la creación y administración de cuerpos especiales denominados de "reacción inmediata" evidencian que la ganancia marginal en ese aspecto siempre resulta insuficiente, tanto para disminuir el índice delictivo general de la zona, como para modificar sustancialmente la sensación de inseguridad entre los ciudadanos.

Durante muchos años, las fuerzas policíacas se han concentrado en atender los delitos una vez que éstos se han cometido. Los delitos son tratados como episodios a ser detectados y**,** si de ahí resulta una aprehensión**,** se dice que el asunto está resuelto. Esta forma de combatir a la delincuencia es completamente errónea, que sorprende que no existan reclamos por parte de la sociedad. Es necesario que en la actualidad el combate a la delincuencia se realice con un enfoque distinto tendiente a la prevención en vez de la reacción **[Arango, 2006]**.

Cuando los ciudadanos consideran las causas que ocasionan los delito, se tiende a hablar de ellas como si fueran temas o asuntos distantes que no pueden ser cambiados rápidamente y, por ello, se espera a que las autoridades vayan actuando, poco a poco, para que dicha situación se vaya corrigiendo. Bajo este enfoque, se descuidan las causas más inmediatas,

aquellas que a menudo se pueden modificar con relativa facilidad y sobre las cuales es posible influir de forma rápida y oportuna, así, no es una exageración señalar que las acciones efectivas en el combate a la delincuencia, están alrededor de nosotros, la mayoría de las veces al alcance de la mano, y que sólo esperan a ser descubiertas.

Una mejor perspectiva de la situación se logra si se transporta el problema al ámbito de la prevención en lugar del de la reacción. Si las fuerzas del orden público fueran capaces de anticipar cuándo y dónde se verá incrementada la actividad delictiva de algún tipo específico, se lograría un beneficio doble. Por una parte, sería posible concentrar los recursos y actividad logística necesarios para combatir ese tipo específico de actividad delictiva en la zona y tiempo anticipados. Por otra parte, se podría establecer, de manera dinámica pero bien fundamentada, varios de los parámetros comunes del trabajo cotidiano en seguridad pública como el diseño específico de rondas de vigilancia, la distribución de fuerzas en espacio y tiempo y por supuesto, la realización de operativos de seguridad e inclusive de campañas de información y prevención por medios masivos de comunicación.

Como beneficio adicional para los funcionarios responsables de la seguridad pública, la capacidad de conocer ese comportamiento (denominado en términos técnicos: tendencia espacio-temporal) de la actividad delictiva proporcionaría un fundamento más adecuado para evaluar su desempeño en la toma de decisiones. La simple comparación entre las tendencias anticipadas, las medidas de prevención por ellos adoptadas y el resultado de las acciones emprendidas proporcionaría el mencionado fundamento. Todo lo anterior permitiría una mejor planeación del presupuesto que anualmente se destina a seguridad pública, así como una mejor distribución del mismo en los diferentes rubros que componen tan compleja actividad.

1.2 Planteamiento del Problema

Con base en los antecedentes expuestos anteriormente, es necesario señalar la importancia de que todos funcionarios responsables de la seguridad pública, cuenten con instrumentos de Análisis Delictivo, que les permita determinar las tendencias espacio-temporales que presentan las familias delictivas dentro de la zona de su responsabilidad, a fin de realizar una planeación que tenga como base la prevención en vez de la reacción.

El hecho de poder anticipar la tendencia de cada familia delictiva dentro de una zona y en un tiempo determinado, no es el único problema al que se enfrentan los mandos policiacos a la hora de concebir estos planes, la falta de recursos humanos, económicos y materiales es un escenario común en los países en vías de desarrollo como México. Esta falta de recursos hace necesario implantar un mecanismo que permita asignar prioridades de atención entre diferentes familias delictivas, con el objetivo de hacer un uso eficiente de los recursos para

atender primeramente a aquellas que tienen un mayor impacto en la población, para posteriormente atender aquellas que presentan un impacto menor y con ello ir reduciendo de manera gradual el índice delictivo de una región de estudio.

El hecho de que se asignen prioridades de atención a las familias delictivas no debe significar el total desentendimiento de fuerzas de orden público para aquellas familias que no presentan un problema mayor. Es necesario dividir las fuerzas del orden público en dos grupos: el primero encargado de establecer una vigilancia continua y, el segundo, encargado de atender aquellas actividades delictivas que, de acuerdo con el análisis realizado, requieran de atención inmediata.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar una metodología para el Análisis y Predicción Delictiva que permita determinar:

- Las familias delictivas que presentan actividad notable dentro de un espacio-tiempo de referencia. Se entenderá por espacio-tiempo en este trabajo a una región de estudio a la que se le realiza Análisis y Predicción Delictiva en un periodo de tiempo específico (Año, Mes, Semana, etc.)

- Los periodos del año y horarios en que estas actividades presentan mayor actividad.

- Los espacio-tiempos en los cuales es posible disminuir la asignación de recursos humanos y materiales para atender a las familias delictivas notables de otros espacio-tiempos.

- Los espacio-tiempos en donde la mayoría de las familias delictivas se presentan.

- Finalmente, la cantidad de delitos que, según la predicción basada en el historial delictivo, ocurrirán en un espacio-tiempo determinado.

1.3.2 Objetivos Específicos

- Desarrollar una metodología para el Análisis Delictivo que estudie el comportamiento espacio-temporal de las familias delictivas y sus patrones dentro de una región de estudio propuesta.

- Desarrollar una metodología para Predicción Delictiva que, a partir de la información obtenida de una región de estudio propuesta, anticipen por cada una de las familias delictivas, el número de hechos delictivos que se espera ocurran en una zona y tiempo determinado.

- Con base en los resultados obtenidos de los dos análisis anteriores, generar recomendaciones que permitan a las fuerzas del orden público mejorar sustancialmente la asignación de recursos necesarios, tanto materiales como humanos, para hacer frente a las familias delictivas que afectan un área de manera característica y con problemas de incrementos delictivos identificados.

- Implementar un sistema de información que incorpore la metodología para el Análisis y Predicción Delictiva y que trabaje dentro de una arquitectura de red cliente-servidor, a fin de ser utilizada por los responsables de la de seguridad pública en diferentes planes preventivos, desde varias locaciones y en cualquier momento.

1.4 Contribuciones

Dentro de los resultados que se obtuvieron del desarrollo de este trabajo de tesis se tienen las siguientes contribuciones:

- Se desarrolló una nueva metodología para Análisis Delictivo basada en la caracterización espacio-temporal de patrones delictivos.

- Se desarrolló una metodología de Predicción Delictiva basada en el análisis del comportamiento de una actividad delictiva durante sus últimos 12 meses y su comportamiento a través de los años en la muestra disponible, que permite calcular la cantidad de hechos delictivos se espera ocurran en un espacio-tiempo con un error mínimo.

- Se desarrolló un sistema de información que permite incorporar las dos contribuciones anteriormente planteadas dentro de una arquitectura cliente-servidor, con el objetivo de que dicho sistema esté disponible para todos los elementos involucrados en el proceso de planeación y aplicación de la actividades tendientes a reducir los índices delictivos, desde cualquier locación y en cualquier momento.

- Finalmente, en el capítulo de conclusiones, existe un apartado (7.2) en donde se explica de manera detallada las aportaciones técnicas que el presente trabajo

desarrolló sobre diferentes algoritmos de clasificación supervisada, empleados en el enfoque Lógico-Combinatorio en Reconocimiento de Patrones.

1.5 Trabajos publicados

1. Congresos Internacionales

   - *"Automatic Spatio-Temporal Characterization of Criminal Activity. A New Algorithm and a Public Security Decision-Support System"*, Research in Computing Science, Advances in Computer Science and Engineering, vol. 42, Mayo 2009.

   - *"The Kora-Ω Algorith for Spatio-Temporal of Criminal Activity",* 4to. Congreso Internacional "Tendencias Tecnológicas en Computación 2008", CIDETEC, 2008.

2. Revista Internacional Arbitrada

   "The CR-Ω+ Classification Algorithm for Spatio-Temporal Prediction of Criminal Activity", Journal of Applied Research and Technology, vol. 7, 3 de diciembre 2009.

3. Publicaciones en Proceso.

   Al momento de imprimir este trabajo se encuentra en elaboración dos artículos, el primero de ellos, relacionado con la metodología de Predicción Delictiva propuesta y su aplicación en el área de Análisis Delictivo; el segundo enfocado a presentar las modificaciones realizadas a los algoritmos de clasificación KORA-Ω y CR+.

1.6 Organización de la Tesis

Con base en los objetivos expresados, esta TESIS se estructura de la manera siguiente:

En el capítulo de introducción se describe la problemática que el presente trabajo de tesis aborda dentro del contexto de seguridad pública.

En el capítulo 2 se presenta el estado del arte sobre criminología y criminalística, así como, lo relativo a los diferentes tipos de Análisis Delictivos que se desarrollan tanto en México como en diferentes partes del mundo.

En el capítulo 3 se presentan los fundamentos teóricos (conceptos y definiciones) que se emplearon durante el desarrollo de la metodología para el Análisis y Predicción Delictiva propuestos.

En el capítulo 4 se describe de manera detallada la metodología desarrollada para el Análisis y Predicción Delictiva, así como su interpretación semántica.

En el capítulo 5 se exponen los resultados obtenidos de la aplicación de las metodologías propuestas para Análisis y Predicción Delictiva sobre una base datos de hechos delictivos de la ciudad de Sacramento, en el estado de California, E.U.A.

En el capítulo 6 se describe el diseño de un sistema de información que incorpora las metodologías propuesta dentro de una arquitectura de red cliente-servidor, así como los requerimientos mínimos necesarios para su implementación.

Las conclusiones derivadas de los resultados obtenidos de la experimentación realizada anteriormente se presentan en el capítulo 7.

Por último, en la parte final, se presenta:

- La bibliografía consultada durante la realización de esta tesis.

CAPÍTULO DOS

Estado del Arte

Se presenta el estado del arte sobre Criminología**,** Criminalística y lo relativo a los diferentes tipos de Análisis Delictivos que se desarrollan tanto en México como en diferentes partes del mundo.

La realidad actual exige un cambio en la forma de percibir, analizar y juzgar los programas y operaciones de prevención y combate a la delincuencia. La situación económica mundial evidencia la necesidad de detener, o incluso revertir, el constante incremento en los presupuestos destinados a la seguridad pública. Cada vez resulta más complicado justificar la asignación de fondos para corporaciones policiacas que, en términos generales, ofrecen pobres resultados y un mínimo impacto de su actividad cotidiana en la calidad de vida de los ciudadanos.

Al mismo tiempo, también resulta claro que el papel de las organizaciones y cuerpos policiacos es decisivo en la prevención del delito. Por ello, en los últimos años se han puesto de manifiesto fuertes preocupaciones sociales por la eficacia y eficiencia de las policías, tanto a nivel de la integridad y capacitación de los agentes que las componen, como a nivel de las políticas y estrategias de las organizaciones mismas.

> *"...Estas fuerzas brindan un nuevo ímpetu para examinar medidas de largo plazo y necesidades descuidadas, tales como generar en la policía la capacidad institucional de evaluar el producto de su trabajo, de analizar constantemente su marcha antes de hacer más de lo mismo, determinar lo que se espera cumpla la policía y encontrar la manera más fácil de lograrlo"* **[Arango, 2006]**.

En términos generales, esta preocupación por la eficacia y eficiencia de la policía, aunque presenta sólidos argumentos, dispone también de suficientes y variados elementos para proponer soluciones. Colectivamente se conoce mucho sobre el amplio rango de problemas de comportamiento que constituyen el trabajo de la policía, así como las mejores formas de prevenirlos **[Gottlieb-Arenberg-Singh, 1994]**. Es posible, hoy en día, encontrar gran cantidad de información sobre prácticas exitosas de las agencias de seguridad y sobre propuestas específicas de agentes experimentados en el tema. Dicha información puede encontrarse en la literatura básica sobre delincuencia y prevención del crimen, especialmente en los textos sobre prevención situacional del delito. Para que la disponibilidad de toda esa información realmente redunde en mejoras palpables y de fuerte impacto, tanto social como económico, esa información debe ser rigurosamente analizada para determinar qué actividades resultan óptimas según la problemática específica a enfrentar en cada demarcación **[Arango-Lara, 2008(2)]**.

No existe actualmente un consenso, nacional o internacional, sobre el concepto de Prevención Delictiva. Son múltiples las definiciones existentes y cada una de ellas enfatiza aspectos diferentes. Como ejemplo de esto basta con mencionar dos de las consideradas más representativas. En primer lugar, el investigador catalán ***Soria Verde***, en su tratado sobre Victimología, define la Prevención Delictiva como:

> *"...la acción social dirigida a la mejora de la calidad de vida, mediante acciones tendientes a que un determinado problema no aparezca o atenúe sus efectos..."* **[Verde, 1993]**.

Por su parte, ***Rodríguez Manzanera***, en su obra *Criminología*, expresa un sentido más específico del término:

> *"...en materia Criminológica, prevenir es conocer con anticipación la probabilidad de una conducta criminal, disponiendo los medios para evitarla..."* **[Manzanera, 1993].**

2.1 Criminalística y Criminología

El enorme progreso tecnológico experimentado por nuestro siglo ha acarreado, por una parte, el nacimiento de nuevas ciencias y, por otra parte, el desarrollo de ciencias cuyos orígenes no se remontan más allá del siglo pasado. En este último caso se encuentran la Criminalística y la Criminología que sirven de auxiliares de la justicia para la determinación de las posibles causales de los delitos ocurridos.

2.1.1 Criminalística

La criminalística es una disciplina que aplica fundamentalmente los conocimientos, métodos y técnicas de investigación**,** de las ciencias naturales**,** en el examen de material sensible, significativo y relacionado con un presunto hecho delictivo, con el fin de determinar, en auxilio de los órganos de administrar justicia, su existencia, reconstruirlo o bien señalar y precisar la intervención de uno o varios sujetos en el mismo **[Montiel, 1997]**

Una definición más formal de la criminalística es la siguiente:

> *"La Criminalística es una ciencia penal auxiliar, que mediante la aplicación de sus conocimientos, metodología y tecnología descubre y verifica científicamente la existencia de un hecho presuntamente delictuoso y al o los responsables aportando las pruebas a los órganos que procuran y administran justicia"* **[Montiel, 1997]**.

2.1.2 Criminología

Esta ciencia social estudia la naturaleza, extensión y causas del delito; características de los delincuentes y de las organizaciones delictivas; problemas de detención y castigo de los

delincuentes; operatividad de las prisiones y de otras instituciones carcelarias; rehabilitación de los convictos, tanto dentro como fuera de prisión, y la prevención del delito.

La ciencia de la Criminología tiene dos objetivos básicos: la determinación de causas, tanto personales como sociales, del comportamiento delictivo y el desarrollo de principios válidos para el control social del delito. Para la consecución de estos objetivos, la Criminología investiga a partir de los descubrimientos de otras disciplinas interrelacionadas con ella, tales como la Biología, Psicología, Psiquiatría, Sociología, y Antropología **[1].**

El presente trabajo de tesis, aunque evidentemente se trata de una investigación dentro del marco de las ciencias de la computación y, en particular, en el área de Inteligencia Artificial y Reconocimiento de Patrones, puede también ser ubicado dentro de esta ciencia auxiliar de la justicia, ya que determina el comportamiento de la actividad delictiva que se presenta dentro de un área de estudio.

2.2 Análisis Delictivo

En la actualidad, no existe una definición universalmente aceptada de lo que es Análisis Delictivo. En algunos departamentos de policía es considerado como el estudio de reportes policiacos y la extracción de información que permita la captura de delincuentes, en particular, asesinos seriales **[Osborne-Wernicke, 2006]**. En otras agencias, el análisis delictivo consiste en extraer datos estadísticos de las bases de hechos delictivos que ocurren dentro un área y dividirlos en familias delictivas y épocas del año **[Osborne-Wernicke, 2006]**.

Sea cual sea la definición de Análisis Delictivo, éste tiene por objetivo, encontrar información importante dentro de los datos contenidos en cada uno de los hechos delictivos y diseminarla entre los oficiales e investigadores para ayudarlos en la captura de posibles delincuentes, así como frenar la actividad delictiva **[Osborne-Wernicke, 2006]**.

La definición formal de Análisis Delictivo empleada en esta tesis es la siguiente:

> *"Análisis Delictivo es como el conjunto de procesos y técnicas de análisis dirigidos a proveer información oportuna y pertinente relativa a hechos y correlaciones de tendencia delictiva a el personal operativo y administrativo durante la planeación de acciones tendientes a prevenir y evitar actividades delictivas, ayudar en los procesos de investigación, incrementar la aprensiones de delincuentes, asignación de recursos y esclarecer los casos"***[Osborne-Wernicke, 2006]**.

Cualquier estudio u observación por sencillo que sea, que se lleve a cabo sobre la actividad delictiva ocurrida dentro de una región de estudio y un periodo de tiempo puede ser considerado como Análisis Delictivo. Esto le permite a cualquier departamento de policía alrededor del mundo, proponer la forma en que realiza este análisis, lo que ha permitido que éstos incursionen en diferentes campos y ciencias auxiliares de esta disciplina, a fin de dar respuesta a las necesidades de información de los cuerpos de policía operativos que se encargan del trabajo de campo.

Por otra parte esa sencillez del concepto, ha provocado que en la actualidad el simple hecho de determinar el incremento o decremento de una cierta actividad delictiva dentro de una región de estudio y plasmar esta información dentro de un informe sea considerado como Análisis Delictivo en ciertos departamentos de policía.

2.2.1 Tipos de Análisis Delictivo

A pesar de no existir una definición unificada de Análisis Delictivo en posible agrupar los diferentes tipos de análisis que se desarrollan en los departamentos de policía de la siguiente manera **[Osborne-Wernicke, 2006]**:

1. **Análisis Delictivo Táctico**

Se basa en determinar dónde, cuándo y cómo se están cometiendo los delitos para que a los oficiales de policía e investigadores. Sin embargo, otro rol de este tipo de análisis es el de encontrar patrones comunes a cierto tipo de delito, que prediga nuevos hechos delictivos dentro de la misma área y modo de actuar de los delincuentes.

2. **Análisis Delictivo Estratégico**

Desarrollado por los mandos superiores encargados de la de seguridad pública para resolver problemas delictivos de gran impacto y planear proyectos a largo plazo. El resultado obtenido de este análisis es proporcionado como información a niveles menores para que estos a su vez lo consideren dentro de la planeación de sus respectivos sectores.

3. **Análisis Delictivo Administrativo**

Se centra en proveer a los diferentes organismos policiacos datos, estadísticas e información referente a la tendencia delictiva presentada dentro de una zona.

4. **Análisis Delictivo de Investigación**

Se encarga del estudio de la conducta criminal para la creación de perfiles policiacos (socio-económicos, culturales y psicológicos), de sospechosos, delincuentes y víctimas, a efecto de detectar posibles delincuentes**,** asesinos en serie y fundamentar los estudios de victimología.

5. **Análisis Delictivo de Inteligencia**

Se desarrolla en los niveles federales y estatales, centrándose principalmente en el estudio del crimen organizado, terrorismo, etc.

6. **Análisis Delictivo de Operaciones**

Tiene como objetivo evaluar el aprovechamiento de los recursos dentro de una organización policiaca (presupuestos, recursos humanos, inventarios de armas, estados de fuerza, aprovechamiento tecnológico, etc.).

2.2.2 Análisis Delictivo en México

En la actualidad, las organizaciones policiacas en México presenta muchas deficiencias en la metodología para la realización de Análisis Delictivo, tradicionalmente el trabajo en este campo es realizado por oficinas denominadas de "Análisis Delictivo", las cuales generalmente no cuentan con personal capacitando en y únicamente se limitan al conteo y clasificación de los delitos que ocurren en su región de responsabilidad para posteriormente ser mostrados mediantes graficas en informes que son enviados a sus jefes inmediatos o distribuidos a las áreas operativas en forma de información que en teoría debería de ayuda a estos en la prevención del delito.

Independientemente del tipo de enfoque que se emplee para la realización del Análisis Delictivo, es vital contar con información verídica, completa y oportuna de los hechos delictivos ocurridos dentro de una región de estudio. Sin embargo, el proceso de generación de datos es el proceso más complicado de realizar. El Instituto Ciudadano de Estudios sobre Inseguridad (*I.C.E.S.I.*) publicó en junio del 2008 un artículo titulado ***"Generación de Información en Seguridad Pública"*** donde plantea**,** de manera clara**,** los principales problemas que actualmente tiene México para recolectar información adecuada para la realización de Análisis Delictivo. A continuación se describen estos problemas **[I.C.E.S.I., 2008]**:

1. Maquillaje de origen

Actualmente los sistemas de información, al menos los conocidos, captan únicamente lo relacionado a la averiguación previa (AP) y la hacen equivalente a un delito. Es decir que aunque en un solo incidente concurran dos o más delitos (por ejemplo en un robo puede haber alguna lesión o agresión física de la víctima) en la AP que inicia el agente del ministerio público (MP) registra estadísticamente un solo delito, por lo general el que considera más grave.

Existen otros casos en ciertas entidades federativas en donde hay registro de un mayor número de AP que delitos ocurridos, lo cual es imposible.

2. Problemas relacionados con la información

Cantidad de la información: No existe una norma que señale a las entidades cuánta y qué tipo de información de debe generarse de cada uno de los delitos denunciados.

Tipo de la información: Aunque existe el formato CIEISP (cuyas siglas significan Comité Interinstitucional de Estadística e Información en Seguridad Pública), que podría servir como base para la generación de información estadística, las entidades terminan por reportar con diferentes niveles de agregación y de conceptos, que incluso internamente cambian de un año a otro. Hay casos en los que ni en una misma entidad es posible hacer comparaciones, por estas diferencias de información.

Homogenización: El problema anterior está ligado con el de la estandarización de la información. Algunas entidades o dependencias reportan por mes, otros por trimestre, otros por año, algunas por día, aunque de un momento a otro cambian conceptos y agregan o desagregan información sin que medie alguna guía o documento que permita entender los motivos, y menos aún que información estaba contemplada en el rubro anterior y donde se contempla bajo los nuevos esquemas de clasificación.

Falta de confiabilidad: En muchas ocasiones la información se maneja para alcanzar ciertos propósitos. Así por ejemplo si es para dar a conocer a la población los resultados en materia de seguridad o para presentar en informes de gobierno se reportan datos a la baja; si es para contar con más recursos por parte del Sistema Nacional de Seguridad Pública los datos son más altos. Es así que muchas entidades cuentan con dos o más sistemas de información, que se mueven de manera distinta, de acuerdo con los objetivos que persigan.

Disponibilidad: En el mejor de los casos la información se hace pública hasta con dos años de retraso y, en el peor, la información nunca llega a hacerse pública. Por ejemplo, un caso típico donde no se cuenta con mayor información (y que debería ser objeto de rendición de

cuentas), corresponde a aquel en donde las AP se "resuelven" por prescripción, es decir extinción de la acción penal por el paso del tiempo que hace que por ley las AP pasen al archivo de asuntos concluidos sin que se haya resuelto nada al respecto. A veces este tipo de "asuntos concluidos" llegan a representar un porcentaje mayor al 50% del total.

Asimismo, otro de los puntos que tiene que ver con la disponibilidad se refiere a que en muchas ocasiones la información se mantiene sin darse a conocer en períodos muy largos y que para efectos de análisis deja de ser útil. Para que la información pueda tener un uso eficiente especialmente para delinear políticas de prevención debiera estar disponible máximo con un mes de atraso.

Temporalidad. A nivel nacional, con la información por entidad federativa, no se sabe nada del día y la hora en que se cometen los delitos, si estos son mayores en quincenas o días de pago, los fines de semana, si ocurren por la mañana o por la tarde, si se dan en días de fiesta patronales, etc.

Lo anterior resulta relevante ya que**,** de acuerdo con la temporalidad**,** se pueden definir políticas diferentes de atención incluso sobre un mismo delito.

3. Contextualización de los delitos

No es posible elaborar políticas de prevención si los delitos no se contextualizan. Por ejemplo, en el caso de un homicidio interesa, entre otras cosas, conocer el contexto, es decir, si había o no relación previa entre víctima y victimario, si el homicidio ocurrió en el interior de una vivienda o negocio, si se dio en la calle, si se utilizó o no un arma y qué tipo de arma, si se dio bajo el influjo del alcohol o alguna otra droga, si participaron una o varias personas, si fue en riña.

De cada una de estas preguntas se desprenden medidas de intervención diferentes.

Por ejemplo, es sabido que no se pueden hacer políticas de prevención del homicidio cuando existe relación previa entre víctima y victimario. Es decir, cuando una persona quiere matar a otra por cuestiones afectivas no hay poder que lo impida. Así puede señalarse el caso donde dos conocidos compadres, uno del otro, empiezan tomando en una reunión de amigos, continúan tomando en la casa de uno de ellos y, de repente, uno de ellos ve o cree ver que las miradas que está realizando el compadre sobre la comadre no son las adecuadas, vienen los reclamos, éstos suben de tono y uno de ellos acaba matando al otro. Ante esta situación no hay nada que hacer.

Lo mismo sucede con los homicidios por ajustes de cuentas entre delincuentes participantes en crimen organizado. No existe poder humano, ni política de seguridad pública, que vaya a impedir este hecho.

Otro delito es el de la violación, donde la contextualización es importante. No sirve de nada tener a 50 policías fuera de la casa de un padre violador, ya que la violación se va a dar en el interior del hogar.

4. Análisis longitudinal

En este momento no existe posibilidad de realizar un análisis, a lo largo del tiempo, de lo que sucede con cada asunto que lleva el aparato de seguridad pública en su conjunto (prevención, procuración y administración de justicia, rehabilitación social), debido a que en cada etapa del proceso se asignan números distintos o a que cada parte del proceso se considera como independiente una de la otra.

Así no se sabe para un evento en particular, si hubo o no intervención policíaca, por qué delitos se inicia una averiguación previa, cuántos y cuáles fueron los delitos por los que se inició, cuántos, cuáles, en su caso, y de qué sexo eran los delincuentes, cuántos, cuáles y de qué sexo eran las víctimas, por qué delitos se consigna una averiguación previa, por qué delitos se procesa, por qué delitos se sentencia, por qué delitos están los delincuentes presuntos o sentenciados en prisión, etc.

5. Información faltante

La mayoría de los eventos en que intervienen las fuerzas policiales quedan en el olvido, lo que impide evaluar dónde son eficientes, dónde no lo son y dónde se necesitan pequeños actos de administración para lograr esa eficiencia.

No se conoce el personal policiaco de los tres niveles de gobierno: cuántos son, dónde están, cómo se distribuyen, cuántos y cuáles están capacitados, qué turnos tienen, las edades, sexos, etc.

Falta información sobre el equipamiento: en qué estado se encuentra, cuánto es, cuántas armas existen, etc.

La información de los penales es completamente deficiente, no se conocen flujos de entrada- salida, por delito, tiempo en prisión, sexo, ocupación, reincidencia, etc.

6. Georeferencia

No es posible ubicar cada hecho delictivo mediante coordenadas geográficas. Asimismo, no se describen las características del espacio donde se cometió el crimen (iluminación, abundancia de árboles, cuadrantes de calles, etc.).

7. Información de los victimarios

Aunque parezca increíble, de los delincuentes se tiene poca información: no se conoce cuántos son, el sexo, su edad, su apariencia, si actúan o no en grupo, las horas y días preferidos para cometer delitos, las armas que utilizan, el tipo de víctimas que escogen, los lugares donde atacan principalmente, el modus operandi, etc.

8. Información de la víctima

Por supuesto que, si la información anterior presenta deficiencias importantes, en la estadística oficial la víctima no existe en lo absoluto.

Así, con esa información, es imposible conocer cuántos hombres y cuántas mujeres son víctimas del delito, los lugares donde han sido victimizados, los horarios más frecuentes, el delito o delitos que padecen, las edades que tienen, etc.

En conclusión, podemos determinar que mientras en México no exista un sistema unificado de información delictiva que contemple a los tres niveles de Gobierno (Federal, Estatal y Municipal) y no se establezca un canal de comunicación entre los organismos encargados de la seguridad pública y el sector educativo, principalmente con universidades, quienes podrían aportar las herramientas científicas y tecnológicas, será casi imposible realizar un Análisis Delictivo que realmente permita hacer frente a los problemas de inseguridad que actualmente se viven en el país, limitándose como hasta el momento a atacar los delitos una vez que estoy ya ocurrieron.

2.2.3 Sistemas de Análisis Delictivo en el Mundo

Los principales sistema de Análisis Delictivo que se pueden encontrar en la actualidad han sido desarrollados por empresas privadas o los departamentos de policía de diferentes estados en países como: Estados Unidos de América, Inglaterra, España, etc. A continuación se describen algunos de ellos.

2.2.3.1 STAC (Spatial and Temporal Analysis of Crime, 1988)

STAC es un paquete de software que se compone de dos programas: un analizador temporal y un analizador espacial. El paquete de STAC se completó en 1988, con el apoyo del Departamento de Justicia de E.U.A., Oficina de Estadísticas de Justicia.

Este paquete fue creado en respuesta a las solicitudes de los organismos policiales locales para mejorar sus capacidades de Análisis Delictivo. Su objetivo principal es localizar agrupaciones de actividad delictiva dentro de una comunidad mediante la automatización de funciones como el análisis de tiempo en los datos de los delitos cometidos y la colocación manual de puntos de interés en mapas (Map Pin).

El analizador temporal ayuda a determinar el momento más probable del día y el día de la semana que un determinado tipo de delito se produce, incluso cuando la información de que dispone la policía es imprecisa.

El analizador espacial ayuda a encontrar las áreas en una comunidad donde hay focos rojos delictivos (*hot Spots*). Este analizador no es un paquete de mapas, sino que ayuda a los analistas que ya tienen la capacidad de mapeo en la representación de resultados.

STAC primero encuentra las agrupaciones más densas de puntos en el mapa, y luego ajusta una "elipse de desviación estándar" a cada una de estas agrupaciones. El analista puede entonces hacer comparaciones entre diferentes agrupaciones.

Esta aplicación y su información técnica esta está disponible de manera gratuita en:

http://www.icjia.state.il.us/public/index.cfm?metaSection=Data&metaPage=STACfacts

La figura 1 muestra la forma en que este sistema muestra algunos de los resultados obtenidos de su análisis.

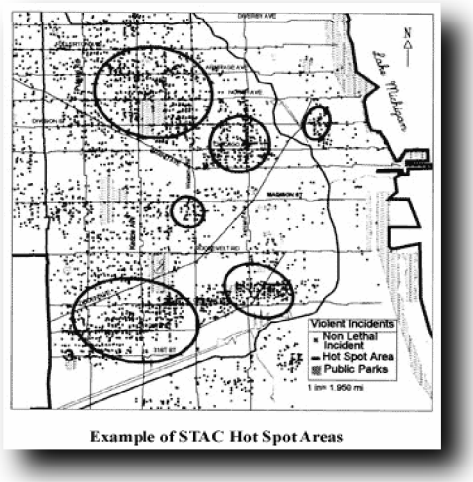

Figura 1. Sistema STAC

2.2.3.2 CrimeLink (PCI Precision Computing Intelligence 1993)

CrimeLink es una herramienta visual de Análisis Delictivo diseñado para ayudar a los analistas e investigadores en la conversión de grandes cantidades de datos aparentemente independientes y sin relación en productos gráficos de análisis ordenados, comprensibles y comprensibles.
Automatizando las probadas técnicas de análisis como: Matrices de Asociación, Gráficas de evento-hora, rueda de análisis de patrones y Análisis de llamadas. CrimeLink es la solución perfecta para comprender y resolver problemas complejos que involucran lucha contra el terrorismo, la delincuencia organizada, operaciones militares, las drogas, el fraude financiero, investigaciones importantes, al por menor, y cualquier otro conjunto de actividades complejas, de conspiración o de redes sociales.

La matriz de asociación visualiza en una sola matriz, las personas y las entidades involucradas en los hechos delictivos, lo que permite encontrar posibles relaciones entre todas ellas. Cada celda de la matriz se asemeja a la tabla de distancias en un atlas de carreteras y representa una posible relación. La matriz de asociación es un registro de quién conoce a quién, quién ha hecho cada quien y quien ha estado donde. La figura 2 muestra esta matriz.

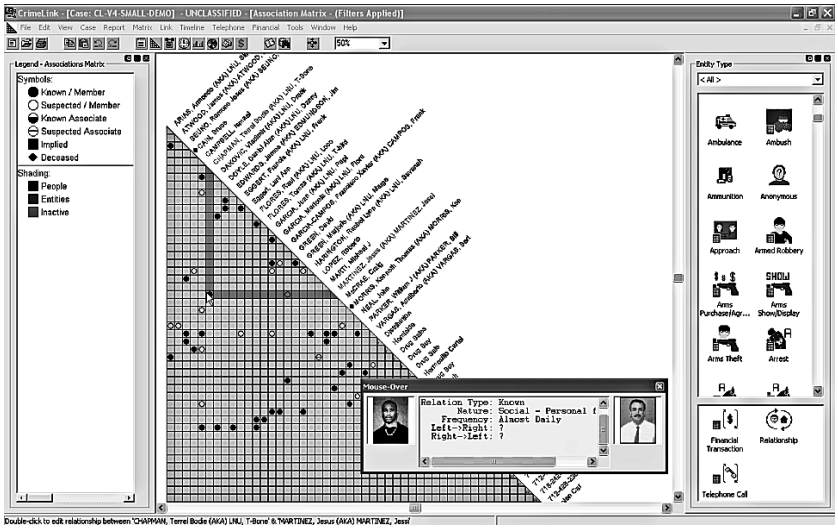

Figura 2. Matriz de asociación de CrimeLink

El Gráfico Eventos-Hora (GEH) y la Rueda de análisis de patrones son un grupo de registros cronológicos diseñados para almacenar y mostrar grandes cantidades de información en el menor espacio posible. Cada registro cronológico es único para cada grupo de estudio y muestra las actividades delictivas relacionadas con el grupo en una variedad de formatos y pantallas.

El GEH es fácil de preparar, entender y utilizar. Al registrar la cronología, es posible analizar las actividades del grupo, las transiciones, las tendencias, y, en particular, los patrones de funcionamiento, es decir, los patrones en el tiempo y actividad. Si se desea, el evento puede ser un código de color para indicar un evento en particular o tipo de evento que ayude en el reconocimiento de patrones.

El GEH es una herramienta de análisis de patrones que permite predecir los acontecimientos futuros de cierta actividad delictiva. La figura 3 muestra la interfaz de usuario del **GEH.**

El análisis de llamadas está basado en información disponible de una variedad de fuentes que muestran el tipo y ubicación de los teléfonos o las comunicaciones relacionadas con el universo de estudio; el momento, la dirección y duración de las llamadas, los patrones de tiempo, orden de frecuencia de llamadas, facturación, y lo más importante, muestra toda una red de telecomunicaciones, así como la identificaciones individuales y corporativas. La figura 4 muestra la interfaz de usuario del analizador de llamadas.

CrimeLink y su información técnica está disponible en: http://www.pciusa.us/.

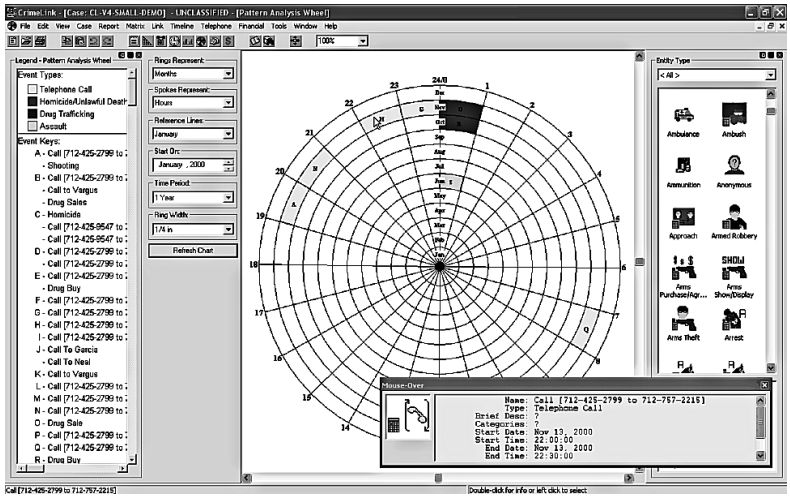

Figura 3. Rueda de análisis de patrones

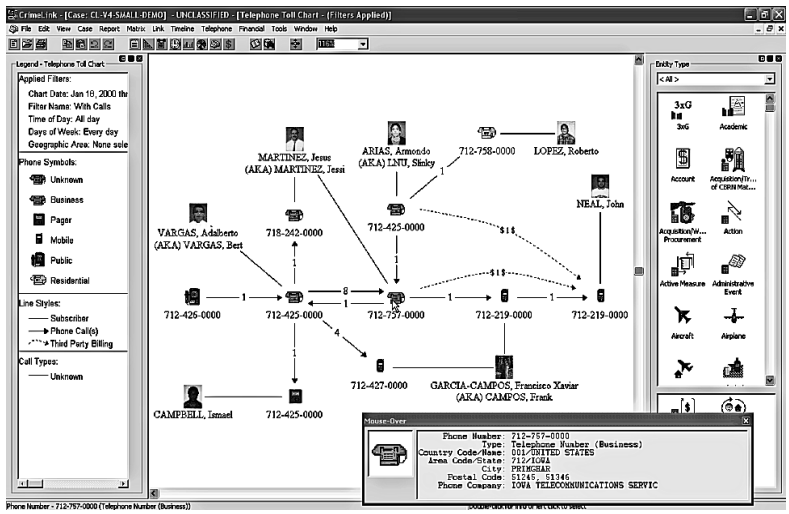

Figura 4. Análisis de llamadas de CrimeLink

2.2.3.3 CrimeView (Omega Group 1992)

CrimeView el software de Análisis Delictivo y mapeo más utilizado en América del Norte. Cientos de agencias utilizar este software para las investigaciones, el despliegue de recursos y la gestión de emergencias. Permite a los usuarios finales importar automáticamente datos delictivos hacia una plataforma de mapas que visualiza geográficamente la actividad delictiva.

CrimeView ofrece a los analistas delictivos una interfaz sencilla para encontrar focos rojos delictivos (*hot spots*) y reporte de llamadas. Con funciones automatizadas, como umbrales de alertas y la generación de informes.

Características Principales de CrimeView

1. **Filtra:** Permite la búsqueda de hechos delictivos por categorías (tipo de delito, fecha y hora).

2. **Fijar:** Encuentra áreas con problemas graves de delincuencia (*hot spots*) y crea mapas de llamadas que ayudan a aislar las áreas problemáticas. La figura 5 muestra la forma en que son presentados los focos rojos delictivos (*hot spots)* encontrados.

3. **Automatiza:** Las tareas repetitivas como la generación de informes periódicos.

4. **Alerta:** hace notificaciones cuando la delincuencia supera un nivel específico.

5. **Informa:** Realiza reportes detallados y resumidos de cualquier capa de la delincuencia y las clasifica por numerosas categorías.

6. **Gráfica:** Crea tablas y gráficos por día de la semana, hora del día y mes del año.

Esta herramienta y su información técnica se encuentra disponible en: http://www.theomegagroup.com/police/crime_mapping_solutions.html.

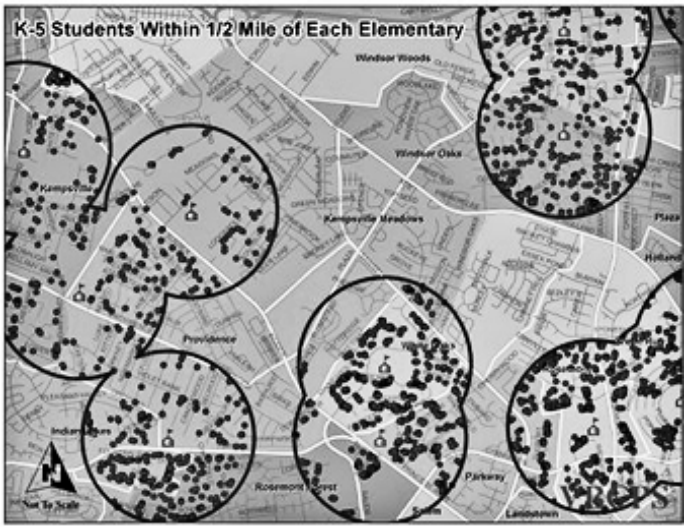

Figura 5. Sistema CrimeView

2.2.3.4 ArcGis (Extensión para Análisis Delictivo 2000)

En la actualidad, los departamentos de policía están utilizando análisis de datos espaciales para evaluar patrones delictivos, optimizar la asignación de recursos, y mejorar la respuesta

de llamada de emergencia. ArcGis es un sistema de información geográfica que actualmente cuenta con un módulo para Análisis Delictivo que emplea herramientas estadísticas diseñadas específicamente para su uso con los datos espaciales.

Las herramientas con las que cuenta este software son:

1. **Análisis de focos rojos delictivos (*hot spots*) en ArcGIS 9.** Permite encontrar aquellas áreas con una densidad delictiva alta. La figura 6 muestra la forma en que ArcGis presenta los resultados obtenidos de este análisis.

2. **Cálculo del centro principal.** Permite calcular el centro principal de una actividad delictiva mediante el promedio de las coordenadas en x y y de cada una de los hechos delictivos ocurridos en una región de estudio.

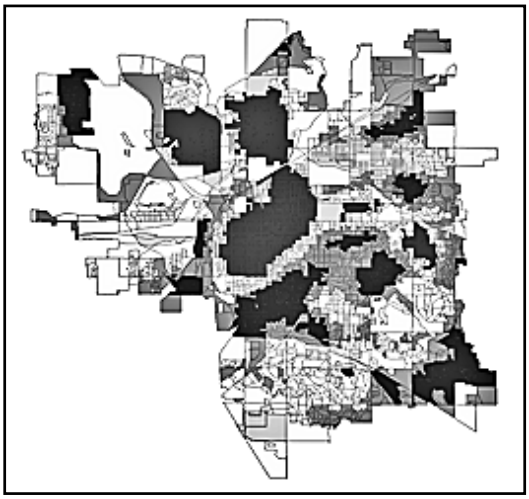

Figura 6. Análisis de focos rojos delictivos (hot spots) de ArcGis

3. **Calculo de la dirección del delito**. Determina la posible dirección que seguirá una familia delictiva. La manera más común para determinar la tendencia de puntos o áreas es calculando la desviación estándar de sus direcciones en x y y. Estas dos medidas definen los ejes de una elipse que muestra si la distribución de puntos se alarga, lo que indica que tiene una orientación particular. La figura 7 muestra la forma en que ArcGis presenta los resultados obtenidos de este análisis.

Figura 7. Cálculo de la dirección del delito

2.2.3.5 A.T.A.C (Automated tactical analysis of Crime 2004)

Es un paquete de estadísticas espaciales que puede analizar datos sobre la ubicación de la incidencia delictiva. Su objetivo es ofrecer una variedad de herramientas para el análisis espacial de la delincuencia. Está diseñado para operar grandes conjuntos de datos provenientes de hechos delictivos. Sin embargo, puede ser utilizado para otros tipos de aplicaciones que impliquen localización de puntos, como la ubicación de las detenciones, accidentes de vehículos, servicios médicos de urgencia, o instalaciones (por ejemplo, estaciones de policía).

La aplicación permite realizar investigación y Análisis Delictivo con un amplio rango de procedimientos espaciales estadísticos que pueden ser incorporados dentro de un sistema de Información Geográfica.

Características Principales de A.T.A.C:

1. **Identificación de patrones:** este software identifica patrones delictivos mediante la ordenación de datos y su consulta.

2. **Análisis de Series de Tiempo:** Este módulo tiene incorporado la mayoría de los métodos de análisis de series de tiempo existentes, los cuales pueden ser aplicados a la totalidad o sólo una parte de los datos con los que se cuente. Cuando un módulo ha terminado su análisis, se puede elegir un rango de tiempo específico para presentar los resultados. El módulo grafica los resultados de las distribuciones temporales utilizando un área, una línea o un gráfico de barras. La figura 8 muestra la forma en que son presentados los resultados de este módulo.

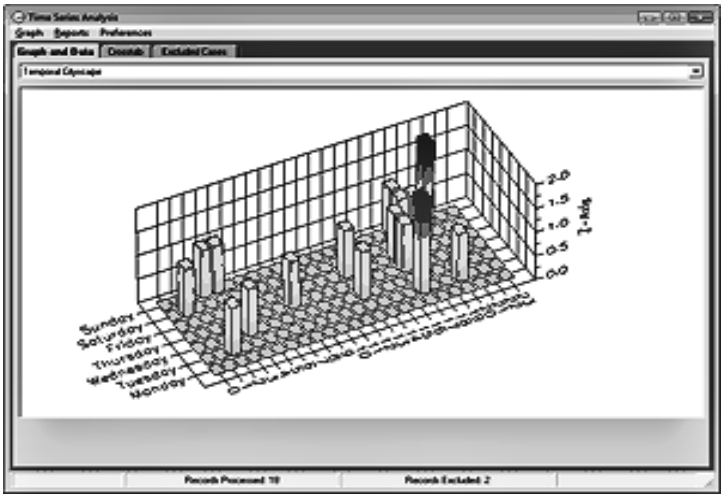

Figura 8. Análisis de series de tiempo en el sistema A.T.A.C.

3. **Predicción Temporal:** El Módulo de Análisis Temporal Dinámico proporciona los medios para analizar y descubrir cualquier patrón temporal de los datos.

   Este análisis permite encontrar lo siguiente:

   - Predicción del próximo evento usando más de una docena de técnicas probadas de predicción.

   - Encontrar las relaciones entre las dos variables para descubrir motivaciones ocultas y pronosticar el próximo evento basándose en estas relaciones

4. **Integración de Google Earth:** Los resultados del análisis espacial pueden ser exportados a un formato de archivo KML de Google Earth. La figura 9 muestra la forma en que presentan los resultados en esta aplicación.

5. **Mapeo:** A.T.A.C., brinda un medio fácil para visualizar datos que contengan coordenadas X y Y en forma tabular, además puede leer archivos Shape ESRI para una comparación rápida entre los datos tabulares y la información geográfica existente.

   Este módulo permite calcular lo siguiente:

   • Áreas de densidad delictiva importantes (*hot spots*). La figura 10 muestra la forma en que son presentados los resultados de este análisis.

   • Cálculo de Distancias Euclidianas y de Manhattan.

- Usar la tabulación X/Y para visualizar inmediatamente las posiciones en Google Earth.

Figura 9. Integración del sistema A.T.A.C. con Google Earth.

6. **Análisis de densidad:** La función de hot spots está optimizado para identificar grupos estadísticamente significativos en los datos, fácilmente determinar el tamaño de la célula y la distancia de búsqueda adecuada. Independientemente del nivel de zoom, ATAC configurará la distancia óptima de búsqueda basado en la perspectiva actual del mapa en donde se trabaje. La figura 11 muestra la forma en que son presentados los resultados de este análisis

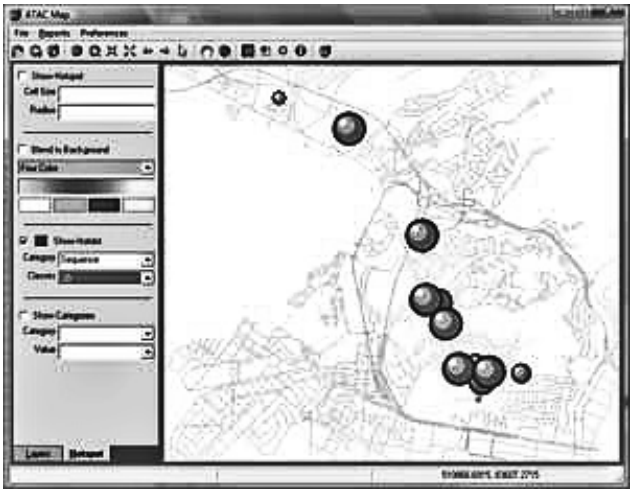

Figura 10. Análisis de hot spots de A.T.A.C.

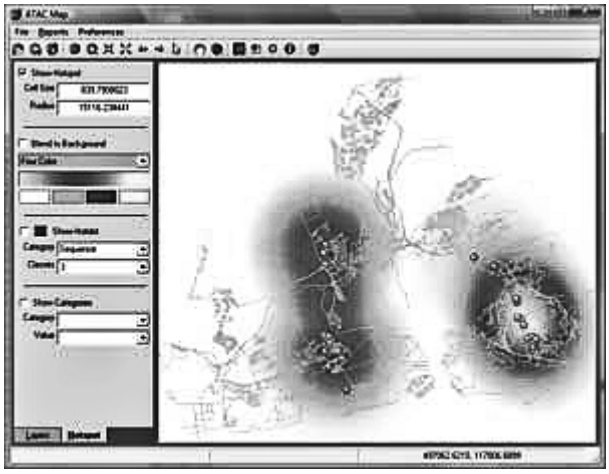

Figura 11. Análisis densidad de A.T.A.C.

Esta herramienta y su información técnica se encuentran disponibles en: http://www.bairsoftware.com/atac/features/identify.html

2.3 Conclusión

El Análisis Delictivo es una herramienta indispensable en todas las dependencias encargadas de la seguridad pública, tanto en México como en cualquier parte del mundo. Sin embargo, en la actualidad, países como el nuestro están apenas en la etapa de creación de una estructura organizacional y tecnológica capaz de registrar de manera eficiente cada uno de los hechos delictivos que se presente en el país. Sin esta estructura, es imposible aplicar cualquier modelo matemático que sirva para hacer Análisis Delictivo, limitándose únicamente a la creación de informes que no aportan ninguna utilidad a los cuerpos operativos de la policía. En otros países, esta estructura ya existe y actualmente ponen a prueba muchos modelos matemáticos que están auxiliando a los cuerpos policiacos de todos los niveles en la toma de decisiones de seguridad pública.

CAPÍTULO TRES

Marco Teórico

Se detallan los conceptos teóricos empleados durante el desarrollo del presente trabajo de tesis.

3.1 Soporte a la toma de decisiones

En las últimas décadas se han desarrollado diferentes sistemas de información que auxilian a los usuarios en la toma de decisiones, estos sistemas han recibido el nombre general de **Sistemas de Apoyo a las Decisiones** (**DSS** por sus siglas en inglés *Decision Support Systems*). Existen diferentes definiciones acerca de lo que es un DSS, sin embargo, para efectos de esta tesis consideraremos la siguiente:

> *"Un DSS es un sistema de información basado en un computador interactivo, flexible y adaptable, especialmente desarrollado para apoyar la solución de un problema de gestión no estructurado que mejorar la toma de decisiones. Utiliza datos, proporciona una interfaz amigable y permite la toma de decisiones en el propio análisis de la situación"* **[Turban, 1995]**

Es evidente que los DSS pertenecen a un entorno con fundamentos multidisciplinarios, incluyendo (pero no exclusivamente) la investigación en base de datos, inteligencia artificial, Interacción hombre-máquina, métodos de simulación, ingeniería de software y telecomunicaciones.

Arquitectura de un Sistema de Apoyo a las Decisiones

Para diferentes autores existe una gran cantidad de componentes que conforman un DSS. Sin embargo, la mayoría de ellos cuentan**,** cuando menos**,** con los siguientes componentes: **[Haag- Cummings- Phillips, 2000]**

1. **Sistema de Gestión de Base de Datos**.- Almacena información de diversos orígenes, puede proceder de los repositorios de datos de una organización tradicional, de fuentes externas (como Internet), o del personal (de ideas y experiencias de los usuarios individuales).

2. **Sistema Gestor de Modelos**.- Se ocupa de las representaciones de los acontecimientos, hechos o situaciones utilizando diversos lenguajes simbólicos.

3. **Sistema Gestor y Generador de Diálogos**.- Se trata de la interfaz ~~de~~ con el usuario; es por supuesto, el componente que permite a un usuario interactuar con el sistema.

3.2 Reconocimiento de Patrones

Se entiende por Reconocimiento de Patrones, a la rama del conocimiento científico, de carácter multidisciplinario, cuyo objeto de estudio son los procesos de identificación, caracterización, clasificación y reconstrucción**,** sobre conjuntos de fenómenos o entes, tanto físicos como abstractos, así como el desarrollo de teorías, metodologías y tecnologías relacionadas a dichos procesos.

En la definición anterior se entiende por ***identificación*** al proceso por el cual se llega al conocimiento o localización precisa de un resultado u objeto buscado en un conjunto. Por ***Caracterización*** se entiende la determinación de los atributos particulares de un objeto, de forma tal que sea posible distinguirle claramente de entre otros semejantes. ***Clasificación*** es el ordenamiento y disposición de los objetos en diferentes categorías o clases según criterios establecidos de semejanza. Por último, ***reconstrucción*** se entiende como el proceso de reunir o evocar las diferentes partes constitutivas de un objeto o fenómeno para completar el concepto o conocimiento del mismo.

Tomando como punto de partida la definición propuesta, este trabajo se ubica como parte del desarrollo de metodologías y tecnologías relacionadas con los procesos propios del reconocimiento de patrones.

3.2.1 Enfoques en reconocimiento de Patrones

En la actualidad, no existe hoy en día una teoría unificada en reconocimiento de patrones (RP) **[Duda-Hart, 73]**, menos aún una visión uniforme sobre los supuestos fundamentales, las metodologías o las herramientas que han de ser usadas pata tal actividad. Por ello, en el trascurso de la corta historia de la disciplina, han surgido diferentes enfoques para abordar el reconocimiento de patrones. Los enfoques más comúnmente estudiados y con mayor impacto en términos de sus resultados son el enfoque *Estadístico*, el *Neuro-reticular*, el *Sintáctico-estructural*, y el *Lógico-Combinatorio*. Se presenta**,** a continuación**,** una breve caracterización de cada uno de ellos en términos de sus fundamentos teóricos y herramientas distintivas.

1. Enfoque Estadístico

Su fundamento es la *Teoría de probabilidad* y la *Teoría bayesiana* de la decisión. Los procedimientos que comúnmente usa son los de la Estadística inferencial y descriptiva, en especial el análisis discriminante. Opera bajo el supuesto fundamental de que los patrones son un conjunto de medidas numéricas, definidas sobre un espacio métrico o normado y con distribuciones de probabilidad conocidas **[Fukugama, 72]**, **[Devroye-Lugosi, 96]**.

2. Enfoque Neuro-Reticular

Se fundamenta en la Teoría *conexionista* surgida en el ámbito de la neuropsicología. Para clasificar patrones utiliza una estructura de neuronas artificiales interconectadas que se estimulan unas a otras, llamadas *Red Neuronal* y en la que cada neurona puede ser "entrenada" para entregar una respuesta específica ante ciertos estímulos predeterminados **[Marques de Sá, 01]**, **[Haykin, 99]**.

3. Sintáctico-Estructural

Se basa en encontrar las relaciones estructurales que guardan entre sí los patrones en estudio, utilizando la teoría de lenguajes formales y funciones recursivas para construir gramáticas que describan la estructura del universo de patrones. Emplea la *Teoría de autómatas* para construir algoritmos que puedan reconocer los patrones **[Fu, 82]**, **[González-Thomason, 78]**.

4. Enfoque Lógico-Combinatorio

El objetivo de este enfoque es modelar **situaciones** sin los supuestos restrictivos de otros enfoques, es decir, en donde los patrones puedan estar formados por cualquier combinación de características tanto cualitativas como cuantitativas. Los problemas de clasificación se modelan según la *Teoría clásica de conjuntos* o la *Teoría de subconjuntos difusos.* Típicamente se aplica la *Teoría de Testores* para determinar las características esenciales de cada patrón. Es común**,** en este enfoque**,** estudiar y modificar algoritmos generados originalmente en otros enfoques para remover algunas de sus limitaciones y extender su capacidad de aplicación. **[Martínez-Trinidad-Guzmán-Arenas, 01]**, **[Shulcloper-Guzmán Arenas-Martínez Trinidad, 99]**, **[Shulcloper-Lazo, 90]**, **[Shulcloper-Abidi, 02]**.

Resulta importante ubicar el presente trabajo en el contexto de un enfoque específico, el lógico-combinatorio, para permitir la comprensión cabal de los problemas abordados, las soluciones propuestas y el impacto de las mismas en la actividad cotidiana de investigadores, docentes y estudiantes de RP. Todos los planteamientos, propuestas y soluciones contenidas en este trabajo deben de entenderse primordialmente desde la perspectiva de este enfoque.

3.2.2 Enfoque Lógico Combinatorio

Los principios que fundamentan el enfoque lógico-combinatorio surgen a mediados de la década de 1960, en la ex Unión de Repúblicas Socialistas Soviéticas (URSS), a partir de los

trabajos de **[Dimitriev-Zhuravlev-Kendrelev, 66]** sobre la formalización matemática de los problemas de clasificación. Estos trabajos retoman las investigaciones sobre Teoría de testores publicados casi una década antes en **[Chenguis-Yabloskii, 55]** y **[Chenguis-Yabloskii, 58]** para darle un nuevo contexto más general e integrarlos como parte del enfoque que en esos momentos estaba surgiendo. A pesar de que la mayoría de los primeros trabajos en esta rama fueron de origen soviético, nunca se concibió en la URSS la integración de estos trabajos como un enfoque diferente en reconocimiento de patrones y, aunque el naciente enfoque se extendió en Europa del este y los países del bloque socialista, no es común, incluso hoy en día, que se le considere como un enfoque en RP, sino como trabajos en modelado matemático.

La característica más trascendente de este enfoque y seguramente la que atrae la atención de cada vez mas número e investigadores **[Martínez-Trinidad-Guzmán-Arenas, 01]**, es que aborda problemas de clasificación y selección de variables cuya representación no es forzosamente numérica, sino que se extiende a casos que involucra descripciones de objetos y fenómenos en términos cualitativos y, sobre todo, espacios de representación que combinan tanto variables cuantitativas como cualitativas e incluso condiciones de información incompleta, imprecisa y ausencia de información **[Goldfarb, 85]**, **[Pons-Shulcloper-Martínez Trinidad, 02]**, **[Shulcloper-Sánchez Díaz-Abidi, 02]**. Esta característica se logra al desasociar el concepto de similitud entre patrones de las condiciones requeridas para una métrica, de esa forma se puede modelar problemas cuyo espacio de representación resulta en una complicada combinación de rasgos cualitativos y cuantitativos. Todas las características anteriores amplían considerablemente el espectro de áreas de estudio y brindan al enfoque lógico-combinatorio un enorme potencial de aplicación, no solo en las ciencias exactas y la ingeniería, sino particularmente en las humanidades, disciplinas sociales y ciencias poco formalizadas, donde la contribución de otros enfoques restringidos a espacios numéricos y métricos, ha sido mínima **[López Reyes, 81]**, **[Cheremesina-Shulcloper, 92]**.

Sin embargo, contrastando con su amplio espectro de investigación y enorme potencial de aplicación, este enfoque presenta también la característica de que los algoritmos tienen, por lo general complejidad alta y baja eficiencia computacional. Los resultados de estos algoritmos y sus interpretaciones son altamente sensibles a los detalles del modelado matemático, en particular, a las características del conjunto de atributos que describen los patrones, así como a las propiedades de las funciones planteadas de semejanza entre patrones **[Shulcloper-Pico, 92]**.

3.2.3 Enfoque Estadístico

Históricamente, una de las principales herramientas que se han empleado en el reconocimiento de patrones ha sido la Estadística, la cual utiliza principalmente: Análisis

discriminante, Teoría bayesiana de la decisión, Teoría de las Probabilidades y Análisis de Agrupamientos **[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]**.

El Enfoque estadístico tiene las siguientes características fundamentales:

1. Se basa en variables numéricas.

2. A estas variables se le presupone propiedades sobre un espacio métrico o normado.

3. Es muy frecuente el uso de probabilidades cuando se considera la presencia de elementos de incertidumbre.

4. Estas herramientas han sido aplicadas con muy buenos resultados a problemas principalmente de imágenes y señales. Sin embargo su uso se ha extendido indebidamente a problemas para los cuales estas herramientas no fueron concebidas **[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]**.

Clasificación Bayesiana

El **Teorema de Bayes**, enunciado por Thomas Bayes en el siglo XVIII, proporciona la distribución de probabilidad condicional de un evento "A" dado otro evento "B" (probabilidad posteriori), en función de la distribución de probabilidad condicional del evento "B" dado "A" y de la distribución de probabilidad marginal del evento "A" (probabilidad simple o apriori).

Sea $\{A_1,A_2,...,A_i,...,A_n\}$ un conjunto de sucesos mutuamente excluyentes, tales que la probabilidad de cada uno de ellos es distinta de cero. Sea "*B"* un suceso cualquiera del que se conoce la probabilidad condicional $P(B \mid A_i)$. Entonces, la probabilidad $P(A_i \mid B)$ está dada por la siguiente expresión: **[Lind ,Marchal, Wathen, 06]**

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)} = \frac{P(B \mid A_i)P(A_i)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)} \qquad (3.1)$$

Donde:

$P(A_i)$ Es la probabilidad de que ocurra el evento A_i.

$P(B)$ Es la probabilidad de que ocurra el evento B.

$P(B \mid A_i)$ Es la probabilidad condicional de que ocurra el evento B dado que ya ocurrió el evento A_i.

$P(A_i \mid B)$ Es la probabilidad condicional de que ocurra el evento A_i dado el evento B.

3.3 Aprendizaje Inductivo

El proceso de aprendizaje inductivo consiste en la generación de nuevo conocimiento después de realizar inferencia inductiva (inducción) sobre los datos proporcionados por un entorno o área de estudios **[Moreno, Armengol, Béjar, 94]**.

Se pueden distinguir dos grandes tipos de aprendizaje inductivo:

1. **Adquisición de conceptos**. También conocido como aprendizaje a partir de ejemplos. Se caracteriza por que existe la descripción de algunos objetos, ya clasificados en una o más clases (Conceptos). La hipótesis que se induce puede ser vista como una regla del **reconocimiento del concepto**, lo anterior significa que si un objeto nuevo satisface las condiciones de la regla entonces representa al concepto dado (clases).

   Algunos de los problemas tratados con este tipo de aprendizaje son:

   - **Obtención de la descripción característica** de una clase de objetos, que especifica las propiedades comunes todos los objetos conocidos de la clase.

   - **Obtención de la descripción discriminante** de una clase de objetos, que la distinguen de un número limitado de clases diferentes.

2. **Aprendizaje a partir de la observación.** También conocida como generalización descriptiva. Su objetivo es determinar una descripción general que caracterice un conjunto de observaciones.

   Algunos de los problemas tratados en este tipo de aprendizaje son:

   - Formular una teoría que caracterice un conjunto de elementos.

   - Descubrir regularidades en los datos.

A continuación se describen dos algoritmos de clasificación que emplean el aprendizaje inductivo de tipo de **adquisición de datos** para la construcción de ***definiciones inductivas.*** Estos algoritmos constituyen la base de la metodología de Análisis Delictivo propuesta en el presente trabajo.

3.3.1 Algoritmos tipo KORA

Los algoritmos de esta familia son extensiones del original algoritmo KORA-3 desarrollado por M. Bongard para solucionar problemas de Geología y Geofísica en los años 60 **[Bongard, 66]**. Este algoritmo ocupa subconjuntos de rasgos descriptivos, de cardinalidad tres, para caracterizar cada una de las clases en la muestra de supervisión de un problema de clasificación supervisada. En 1998, este algoritmo es modificado y adopta el nuevo nombre de *Fuzzy Kora-*Ω *Algorithm* **[De la vega, Shulcloper, 98],** siendo su característica principal la de poder usar subconjuntos de rasgos de cualquier cardinalidad para la caracterización de las clases; asimismo, este algoritmo introduce nuevos umbrales que flexibilizan el tipo de caracterizaciones hechas.

El algoritmo KORA-Ω, en particular su fase de aprendizaje, desempeña un papel fundamental es este trabajo de investigación, por lo que a continuación se presentan los elementos teóricos previos para su comprensión y después se presenta el algoritmo mismo.

Conceptos Previos:

Un conjunto $O = \{o_1,...,o_n\}$ de objetos en estudio, definidos por un conjunto de atributos o rasgos descriptivos $\Re = \{x_1,...,x_r\}$, de manera tal que cada objeto o_i se representa mediante una tupla $(x_1(o_i), x_2(o_i),..., x_r(o_i))$ llamada Patrón. Estos patrones son colocados en una muestra de supervisión donde son clasificados en un conjunto $C = \{C_1,...,C_k\}$ de clases o categorías. El problema de clasificación supervisada consiste en que, dado un conjunto $P = \{P_1,...P_s\}$ de patrones nuevos sin clasificar, cada patrón debe ser asignado a una de las k clases conservando las propiedades de la muestra de referencia.

Los algoritmos de la familia KORA, para resolver este problema, construyen a partir de la muestra de supervisión un conjunto de definiciones inductivas que caracterizan a cada una de las clases contenidas en dicha muestra. Las definiciones inductivas se componen de rasgos complejos, que no son otra cosa más que subconjuntos $\tau \subseteq \Re$ de rasgos descriptivos y sus valores asociados. En el caso del algoritmo original KORA-3 estos rasgos descriptivos (también llamados Propiedades) son exclusivamente de cardinalidad tres; en

posteriores extensiones al algoritmo se consideran subconjuntos de cualquier cardinalidad. La familia de subconjuntos que sirven de referencia para la construcción de definiciones inductivas se denomina *Familia de Conjuntos de Apoyo*.

Una familia de conjuntos de apoyo es un conjunto $A = \{\omega_1, \omega_2, ..., \omega_g\}$ donde $\omega_i \subseteq \Re$, $i = 1...g$.

Un ***Rasgo Complejo*** o ***Propiedad*** de una clase tiene la forma:

$$P_i = \left\{ \begin{matrix} rasgo_{i_1} & & rasgo_{i_m} \\ < valor_{i_1} > & ,..., & < valor_{i_m} > \end{matrix} \right\}$$

Una ***Propiedad es Característica de una clase*** C_i si ocurre β_1 veces o más en esa clase y β_1' veces o menos en las otras clases. La constate β_1 se denomina *Umbral Positivo de Caracterización Positiva* y β_1' se denomina *Umbral Negativo de Caracterización Positiva,* donde $0 < \beta_1 < \min_i |C_i|$, $0 < \beta_1' < \min_i |C_i|$ y, generalmente $\beta_1 > \beta_1'$.

Una ***Propiedad es Complementaria de una clase*** C_i si ocurre β_2 veces o más en esa clase y β_2' veces o menos en las otras clases. La constate β_2 se denomina *Umbral Positivo de Complementariedad* y β_2' se denomina *Umbral Negativo de Complementariedad,* donde $0 < \beta_2 < \min_i |C_i|$, $0 < \beta_2' < \min_i |C_i|$ y, generalmente $\beta_2 > \beta_2'$.

Una ***Definición Inductiva Característica de una clase*** C_i es la unión de todas las propiedades que son características en dicha clase:

$$D(C_i) = \{P_{i_1}\} \cup \{P_{i_2}\} \cup ... \cup \{P_{i_m}\} \qquad (3.2)$$

De igual manera, una *Definición Inductiva Complementaria*, tiene la misma estructura pero está conformada por propiedades complementarias en la clase.

***Algoritmo Kora-Ω**:*

Etapa de Aprendizaje

1. Determinar la familia de conjuntos de apoyo
2. Determinar los valores de los umbrales β_1, β_1'
3. Calcular todas las propiedades características de cada clase.

Etapa de re-aprendizaje

4. Determinar los valores de los umbrales β_2, β_2'
5. Calcular todos los rasgos complementarios de cada clase.

Etapa de clasificación

6. Esquema de Votación: para cada patrón nuevo a clasificar, contabilizar el número de propiedades que ocurren en cada una de las definiciones inductivas características y complementarias.

7. Regla de Solución: con base en el resultado anterior determinar a qué clase se asigna el nuevo patrón y con qué grado de pertenencia.

Se puede encontrar más información en **[De la vega, Shulcloper, 98]**, **[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]**.

3.3.2 Algoritmo CR (Conjuntos Representantes)

Este algoritmo es una idea desarrollada por Baskakova y Zhuravlëv, en la cual se trata de evaluar la evidencia disponible, a favor y en contra**,** de que un patrón pertenezca a una determinada clase; para ello se construyen definiciones inductivas con propiedades representantes Positivas y Negativas para cada una de las clases**,** así como una definición inductiva Neutra para toda la muestra de supervisión. **[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]**.

Las definiciones para trabajar con este algoritmo son muy similares a las expuestas en el algoritmo Kora-Ω. Los rasgos complejos característicos de una clase C_i forman una definición inductiva positiva $D^+(C_i)$.

Una ***Propiedad es Característica Negativa de una clase*** C_i si aparece α_2 veces o menos en esa clase y α_2' veces o más en las otras clases. La constate α_2 se denomina *Umbral Positivo de Caracterización Negativa* y α_2' se denomina *Umbral Negativo de Caracterización Negativa*, donde $0 < \alpha_2 < \min_i |C_i|$, $0 < \alpha_2' < \min_i |C_i|$ y, generalmente $\alpha_2 < \alpha_2'$.

Una ***Definición Inductiva Negativa de una clase*** $D^-(C_i)$ es la unión de todas las propiedades que son características negativas en dicha clase, su estructura es idéntica a la descrita en la fórmula 3.2.

Una ***Propiedad es Neutra en la Muestra de Supervisión*** MS si aparece γ veces en todas las clases, La constate γ se denomina *Umbral de Neutralidad*, donde $0 < \gamma < \min_i |C_i|$.

Una ***Definición Inductiva Neutra de la Muestra de Supervisión*** $D^0(MS)$ es la unión de todas las propiedades neutras de la muestra de supervisión, su estructura es idéntica a la descrita en la fórmula 3.2.

Algoritmo CR

Etapa de Aprendizaje

1. Determinar la familia de conjuntos de apoyo.
2. Determinar los valores de los umbrales β_1, β_1'
3. Construir una definición inductiva positiva para cada clase $D^+(C_i)$.
4. Determinar los valores de los umbrales α_2, α_2'
5. Construir una definición inductiva negativa para cada clase $D^-(C_i)$.
6. Determinar los valores del umbral γ.
7. Construir una definición inductiva Neutra para toda la muestra de supervisión $D^0(MS)$.

Etapa de clasificación

8. Esquema de Votación: para cada patrón nuevo a clasificar, contabilizar el número de propiedades que ocurren en cada una de las definiciones inductivas características y complementarias.

9. Regla de Solución: con base en el resultado anterior determinar a qué clase se asigna el nuevo patrón y con qué grado de pertenencia.

Para mayor información revisar: **[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]**, **[Shulcloper-Abidi, 02]**.

3.4 Predicción

3.4.1 Series de Tiempo

Una serie de tiempo es un conjunto de observaciones tomadas y medidas a lo largo del tiempo. Estas observaciones pueden ser hechas de manera continua o de manera discreta. **[Chatfield, 00].**

Un ejemplo de serie de tiempo discreta se observa en la Figura 12.

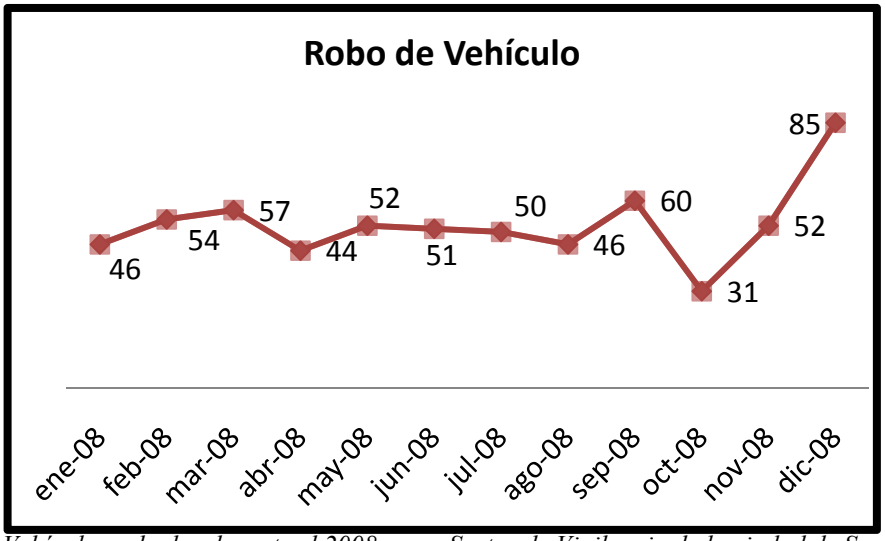

Figura 12. Vehículos robados durante el 2008 en un Sector de Vigilancia de la ciudad de Sacramento, California, E.U.A.

Los principales objetivos del área de estudio conocida como Análisis de Series de Tiempo, según **[Chatfield, 00]** son:

a) Descripción. Busca explicar los datos o hecho registrados a través del tiempo mediante estadística descriptiva.

b) Modelación. Determina modelos estadísticos que describan la forma en que se generan los datos a través del tiempo.

c) Predicción. Estima los valores futuros de la serie con base en la información existente.

d) Control. Un buena predicción de un serie de tiempo permite a los analistas tomas acciones de control sobre un proceso que está siendo analizado.

Los principales componentes de una serie de tiempo según **[Lind, Marchal, Wathen, 08]** son:

a) Tendencia secular. Es la dirección uniforme de una serie de tiempo a largo plazo.

b) Variación estacional. representa la variabilidad en los datos debida a influencias de las estaciones. Esta variación corresponde a los movimientos de la serie que recurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad.

c) Variación cíclica: Con frecuencia las series de tiempo presentan secuencias alternas de puntos abajo y arriba de la línea de tendencia que duran más de un año, esta variación se mantiene después de que se han eliminado las variaciones o tendencias estacional e irregular.

d) Variación Irregular: Esta se debe a factores a corto plazo, imprevisibles y no recurrentes que afectan a la serie de tiempo. Como este componente explica la variabilidad aleatoria de la serie, es impredecible, es decir, no se puede esperar predecir su impacto sobre la serie de tiempo.

Para el propósito de esta tesis**,** se describen a continuación**,** diversos métodos de predicción para series de tiempo que se emplearon para comparar la metodología de predicción propuesta.

1. Promedio Móvil

El modelo de promedio móvil adopta la siguiente forma **[Hanke, 08]**:

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \ldots \omega_q \varepsilon_{t-q}$$

Donde:

Y_t = Es la variable de respuesta en un tiempo *t.*

μ = Valor promedio que permanece constante durante todo el proceso.

$\omega_1, \omega_2, \ldots, \omega_q$ = Coeficientes a ser estimados durante el proceso.

ε_t = Termino que representa en el error en un tiempo *t*.

$\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$ = Errores anteriores al tiempo *t,* que se incorporan a la repuesta de Y_t.

Este modelo proporciona predicción para la variable dependiente de *Y,* con base en una combinación lineal de un número finito de errores.

2. Suavizamiento Exponencial Simple.

La característica principal de este método es dar (exponencialmente) mayor peso a las observaciones más recientes en la predicción que a las más antiguas. Para este método se toma en cuenta una constante de suavizamiento (α) que será determinada o estimada para asignar el peso a las observaciones. La predicción estará formada por la suma de la predicción anterior más una porción del error de la predicción. La ecuación y la definición de variables para este método son las siguientes **[Janacek, 2001]**:

$$F_{t+1} = \alpha Y_t + (1-\alpha)F$$

Donde

F_{t+1} = Predicción del valor de la serie para el periodo *t*+1.

Y_t = Valor real de la serie de tiempo en el periodo *t*

F_t = Predicción de la serie de tiempo para periodo anterior

α = Constate de Suavizamiento con valores entre 0 y 1.

3. Suavizamiento Exponencial Doble

El suavizamiento exponencial doble también es conocido como método de Holt y su característica es que incorpora el componente de tendencia. Se llama suavizamiento exponencial doble ya que tanto la estimación del promedio de los errores como la estimación de la tendencia son suavizadas. La constante de suavizamiento α seguirá siendo para el promedio, mientras que β, será la constante de suavizamiento para la tendencia. Las fórmulas y la definición de variables para el desarrollo de este método se describen en seguida **[Janacek, 2001]**:

$$F_{t+1} = \alpha Y_t + (1-\alpha)(Y_{t-1} + T_{t-1})$$
$$T_t = \beta(Y_t - Y_{t-1}) + (1-\beta)T_{t-1}$$

Donde:

F_{t+1} = Predicción del valor de la serie para el periodo *t*+1.

Y_t = Valor real de la serie de tiempo en el periodo *t*

F_t = Predicción de la serie de tiempo para periodo anterior

$\alpha \ y \ \beta$ = Constate de Suavizamiento con valores entre 0 y 1.

4. A.R.I.M.A.

Es un modelo autorregresivo integrado de media móvil o A.R.I.M.A. (por su nombre en inglés, AutoRegressive Integrated Moving Average) es un modelo estadístico que utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para una predicción hacia el futuro. Este modelo fue desarrollado a finales de los 60's. Box y Jenkins (1976) lo sistematizaron.

ARIMA es un modelo con tres parámetros $ARIMA(p,d,q)$ en donde **[Janacek, 2001], [Chatfield, 00], [Hanke, 08]**:

p: Autoregresión
d: Integración o Diferenciación
q: Media Móvil

Para más información acerca de los métodos de predicción mencionados anteriormente se recomienda **[Chatfield, 00], [Hanke, 08]** y **[Lind, Marchal, Wathen, 08].**

CAPÍTULO CUATRO

Análisis y Predicción Delictiva.

Se plantea la metodología a seguir para el desarrollo del Análisis Delictivo, así como el algoritmo de predicción propuesto.

En éste capítulo se presentan, los fundamentos teóricos de la metodología propuesta para análisis y predicción delictiva. En la primera parte, se muestra de manera detallada el uso de técnicas de aprendizaje inductivo descritas en el capítulo anterior, que permiten la construcción de definiciones inductivas; estas definiciones permiten determinar aquellos espacios-tiempos que son afectados de manera característica por alguna actividad delictiva. Asimismo, sienta las bases para la elaboración de recomendaciones que impactan directamente en la reasignación de recursos humanos y materiales que atiendan aquellas zonas con una incidencia delictiva mayor. Es importante señalar que la metodología propuesta tiene como base la idea de impactar primeramente en aquellas familias delictivas con presencia característica en un espacio-tiempo para posteriormente impactar en aquellas actividades que no los son y de esta manera ir reduciendo el índice delictivo de manera progresiva. Esta forma en que se pretende reducir el índice delictivo, surge de la gran escasez de recursos que actualmente tienen todos los organismos policiacos en el país, por lo que resulta casi imposible atender todas las actividades delictivas que se presentan en una región de estudio de manera simultánea.

En la segunda parte se presenta la metodología de predicción empleada, con la cual se determina la cantidad de delitos que**,** en cada familia delictiva**,** se espera ocurran en un lugar determinado y en una unidad de tiempo posterior a la cubierta por el universo de estudio.

Todas las recomendaciones obtenidas de la metodología para el Análisis Delictivo, así como el número de delitos que se obtiene de la predicción, establecen una base importante en la planeación y toma de decisiones de seguridad pública.

La metodología propuesta permite la implementación de acciones de seguridad pública preventivas y no de reacción, lo anterior con fundamento en lo que se menciona en:

> "*Las Políticas (o planes) de Seguridad deberían estar basadas en la prevención de las conductas antisociales, delictivas y no delictivas, más que en soluciones de tipo represivo. Es más, el problema de la seguridad debería ser abordado desde políticas sociales que contemplen en su diseño la participación activa de toda la sociedad. Los planes de prevención que no cumplen con los requisitos básicos de tener en cuenta la realidad, la dimensión social del problema, la articulación con otros planes y el monitoreo periódico, están destinados al fracaso. Pero sobre todo fracasarán si no están formulados sobre la base de un diagnóstico que esté dentro de un marco teórico adecuado*" **[Arango-Lara (2), 2008].**

4.1 Modelo de Análisis Delictivo

4.1.1 Universo de Estudio

Está integrado por diferentes hechos delictivos registrados dentro de una región de estudio **y** en un periodo de tiempo determinado. Para el nivel de análisis que se propone en la presente tesis es necesario que cada uno de estos hechos contenga información relativa a su ubicación espacial y temporal.

4.1.2 Proceso de Aprendizaje

Antes de comenzar con el proceso de aprendizaje, es necesario introducir un nuevo concepto denominado *Índice de Ocurrencia*. Este índice nos permite observar qué tanto se presenta una familia delictiva dentro de un espacio-tiempo seleccionado con respecto a las otras familias. El Índice de Ocurrencia (I) de cada propiedad espacio-temporal (P) en cada clase (C_i) o familia delictiva está dado por

$$I_{C_i}(P) = \frac{nP_{C_i}}{nP} \qquad 0 \le I_{C_i}(P) \le 1 \qquad (4.1)$$

Donde nP_{C_i} es el número de ocurrencias de la propiedad (P) en la familia delictiva (C_i), y nP es el número total de ocurrencia de la propiedad (P) en toda la clasificación de referencia.

El proceso de Análisis Delictivo comienza a partir de un universo de estudio generalmente integrado en una gran base de datos delictiva, donde cada uno de los hechos delictivos registrados en ella deben de ser representados como un patrón delictivo que consta de tres componentes:

$$Patron\ Delictivo = \left(\langle Hecho\ Delictivo\rangle, \langle Localización\ Espacial\rangle, \langle Localización\ Temporal\rangle\right)$$

- Hecho delictivo.- específica el tipo de delito ocurrido (robo, violación, homicidio, etc.), así como muchas de sus características, como pueden ser el

grado de violencia, número de sospechosos implicados, tipo de armamento empleado, *modus operandi*[1], etc.

- Localización Espacial.- especifica en lugar donde se presentó el fenómeno delictivo (Sector de Vigilancia, colonia, calle, etc.)

- Ubicación Temporal.- indica en el momento en que ocurrió el hecho delictivo (año, mes, día, hora, etc.)

Una vez que se tienen los patrones delictivos, el proceso de aprendizaje comienza con el desarrollo de las siguientes tres fases: construir una clasificación de referencia, construir dos definiciones inductivas, una **positiva** y una **negativa** para cada familia delictiva y, finalmente construir una definición inductiva neutra para todas las familias delictivas. A continuación cada una de estas etapas es descrita:

4.1.2.1 Construcción de una Clasificación de Referencia.

Una vez que cada hecho delictivo es transformado en un patrón delictivo, el siguiente paso es construir una clasificación de referencia. Para construir esa clasificación, todos los patrones son agrupados en familias delictivas de acuerdo al tipo de medidas preventivas, materiales y recurso logísticos necesarios para prevenirlos. La selección del número y naturaleza de estas familias (clases) depende de las políticas y estrategias locales. Es importante señalar, que muchas ocasiones la clasificación jurídica de los delitos puede no ser útil para efecto de lo que se busca pues resulta más relevante crear familias delictivas que agrupen hechos delictivos que tengan en común ciertos recursos necesarios para su prevención y combate.

4.1.2.2 Construcción de definiciones inductivas.

En esta etapa es necesario determinar el subconjunto de rasgos que cada patrón deberá contener para formar las propiedades que serán buscadas dentro de la clasificación de referencia. Una propiedad se define como un subconjunto de rasgos que cada patrón tiene con su respectivo valor asociado. Las propiedades deben incluir componentes espaciales y temporales:

$$P_i = \left\{ \begin{matrix} rasgo_{i_1} & & rasgo_{i_q} \\ < valor_{i_1} > & \ldots & < valor_{i_q} > \end{matrix} \right\}$$

[1] Expresión empleada, fundamentalmente en criminalística, para hacer referencia al modo característico de actuar de un delincuente. También en criminalística se puede interpretar como el análisis de los esquemas intelectuales, estrategias y diversas formas utilizadas comúnmente en la comisión de secuestros, robos y delincuencia organizada.

Una vez que las propiedades son definidas, una consulta sobre la clasificación de referencia es realizada. Todas aquellas propiedades que aparezcan más comúnmente en una clase y menos en otras son consideradas como **Características positivas** de la clase. La unión de propiedades características positivas de una clase (C_i) forma una ***Definición Inductiva positiva*** de dicha clase con la siguiente estructura:

$$C_i^+ : \left\{ Propiedad_1^{C_i} \right\} \cup \left\{ Propiedad_2^{C_i} \right\} \cup \ldots \cup \left\{ Propiedad_{m_i}^{C_i} \right\}$$

De igual forma todas aquellas propiedades que aparezcan poco en una clase y mucho en las otras son consideradas como **Características negativas** de la clase. La unió de propiedades características negativas de una clase (C_i) forma una ***Definición Inductiva negativa*** de dicha clase con la siguiente estructura:

$$C_i^- : \left\{ Propiedad_1^{C_i} \right\} \cup \left\{ Propiedad_2^{C_i} \right\} \cup \ldots \cup \left\{ Propiedad_{m_i'}^{C_i} \right\}$$

La presente metodología propone la construcción de ambas definiciones inductivas para cada clase de estudio. Estas definiciones se construyen de la siguiente manera:

a) Definición Inductiva Positiva $D^+(C_i)$

i. Inicializar el umbral β_1 *(Umbral Positivo de Caracterización Positiva)* entre 0 y 1.

ii. Determinar el Índice de Ocurrencia (I) de cada propiedad espacio-temporal (P) en cada clase (C_i).

iii. Todas las propiedades espacio-temporales con un Índice de Ocurrencia mayor o igual a β_1 son consideradas para ser incluidas en la definición inductiva positiva. Cuando el umbral β_1 se establece con valor mayor o igual a 0.51, las propiedades que conforman una definición inductiva forzosamente deben ser propiedades características, pues su Índice de Ocurrencia debió superar al umbral, garantizando con ello, que más del 50 % de ocurrencia de esa propiedad se observan en la familia delictiva caracterizada. Sin embargo, también es posible construir definiciones inductivas con propiedades no-características $(I_{C_i}(P) <= \beta_1)$, en cuyo caso, la

propiedad ocurre en la clase caracterizada menos del 50% del total de sus ocurrencias y, por tanto, puede ser parte de la definición inductivas en más de una familia delictiva.

b) Definición Inductiva Negativa $D^{-}(C_i)$

i. Inicializar dos umbrales β_2 *y* β_2' *(Umbral Positivo de Caracterización Negativa* y *Umbral de Presencia para Caracterización Negativa* respectivamente*)* entre 0 y 1.

ii. Todas las propiedades espacio-temporales con un Índice de Ocurrencia menor a un umbral β_2 y con una aparición en el β_2' porciento de las otras clases son incluidos en la definición inductiva de la clase.

iii. El umbral β_2 indica el valor máximo del Índice de Ocurrencia, que una propiedad debe tener para ser considerada como característica negativa. El umbral β_2' define el mínimo porcentaje de otras clases en que la propiedad debe aparecer para ser considerada como característica negativa.

4.1.2.3 Construcción de una definición neutra $D^0(MS)$

Todas aquellas propiedades que se presenten con un número de ocurrencias igual o similar **en la mayoría de las familias delictivas** son consideradas como **propiedades neutras**. La unión de propiedades neutras forma una ***Definición neutra*** de la clasificación de referencia (con la siguiente estructura:

$$D^0 : \{Propiedad_1\} \cup \{Propiedad_2\} \cup \ldots \cup \{Propiedad_{m0}\}$$

a) Inicializar dos umbrales β_3 y β_3' *(Umbral de Holgura entre Clases* y *Umbral de Porcentaje de Presencia para Caracterización Neutra* respectivamente*)* entre 0 y 1.

b) Determinar el Índice de Ocurrencia (I_{C_i}) de cada propiedad (P) en todas las clases mediante (4.1).

c) Determinar el valor máximo del Índice de Ocurrencia de cada propiedad en cualquier clase $\underset{C_i \in MS}{Max}\{I_{C_i}(P)\}$.

d) Para que una propiedad (P) sea incluida en la definición inductiva neutra, esta debe de satisfacer la siguiente condición:

$$\frac{I_{C_i}(P)}{\underset{C_i \in MS}{Max}\{I_{C_1}(P)\}} \geq \beta_3$$

e) Finalmente, el número de propiedades que deben cumplir la condición anterior deben de ser cuando menos β_3' por ciento de todas las clases.

Una vez que la etapa de análisis es terminada, las definiciones inductivas positivas, negativas y neutra son integradas de la siguiente manera:

Tabla 1 Definiciones Inductivas

$C_1^+ = \left\{\text{Propiedad}_1^{C_1}\right\} \cup ... \cup \left\{\text{Propiedad}_{m_1}^{C_1}\right\}$ $C_1^- = \left\{\text{Propiedad}_1^{C_1}\right\} \cup ... \cup \left\{\text{Propiedad}_{m_1'}^{C_1}\right\}$	*Familia Delictiva* #1
. . .	. . .
$C_k^+ = \left\{\text{Propiedad}_1^{C_k}\right\} \cup ... \cup \left\{\text{Propiedad}_{m_k}^{C_k}\right\}$ $C_k^- = \left\{\text{Propiedad}_1^{C_k}\right\} \cup ... \cup \left\{\text{Propiedad}_{m_k'}^{C_k}\right\}$	*Familia Delictiva #k*
$D^0 = \left\{\text{Propiedad}_1\right\} \cup ... \cup \left\{\text{Propiedad}_{m^0}\right\}$	

4.1.2.4 Interpretación

Con base en las definiciones inductivas obtenidas durante el proceso de aprendizaje se establecen los siguientes criterios que deben ser tomados en cuenta durante la creación de planes preventivos de seguridad pública:

a) Las propiedades características contenidas en una definición inductiva positiva constituyen aquellos espacio-tiempos que son afectados por la ocurrencia, de forma característica**,** de alguna familia delictiva y**,** por lo tanto**,** sugieren cuándo y dónde se deben incrementar las actividades de patrullaje y vigilancia en concordancia con el tipo de actividad delictiva de que se trate. Es importante recordar que la estrategia que se plantea en la presente metodología para el Análisis Delictivo, parte de la necesidad de aprovechar de manera eficaz y eficiente los recursos económicos, humanos y materiales de que dispone la policía, por lo que primeramente se propone enfocar todos los esfuerzos policiacos a aquellos espacio-tempos característicos de cada una de las actividades delictivas y con esto logar disminuir de manera escalonada el índice delictivo dentro de la región de estudio. Esta información puede ser empleada en la planeación de acciones de seguridad pública a nivel de sectores de vigilancia.

b) Las definiciones inductivas positivas, construidas con propiedades no-características, tienen la misma semántica que las descritas en el párrafo anterior (señalan las familias delictivas que deben ser atendidas de forma inmediata). Sin embargo, no expresan una necesidad tan apremiante de atención como lo haces las definiciones formadas por propiedades características. La decisión de atender más de una actividad delictiva, depende única y exclusivamente de la cantidad de recursos económicos, humanos y materiales de que dispone la policía. En tal caso, el umbral de caracterización positiva β_1 *(Umbral Positivo de Caracterización Positiva)*, debe ser establecido con un valor menor a 0.51 para lograr el efecto deseado.

c) Las propiedades características negativas constituyen aquellos espacio-tiempos donde se sugiere traer recursos para atender otros que requieran una mayor atención, cabe señalar que una propiedad de este tipo solo implica que en un espacio-tiempo no ocurrirá de manera característica una actividad delictiva, pero no significa que no ocurrirán otras, por lo que la reasignación de recursos debe de ser solo aquellos destinados a atender aquellas actividades no características, sin dejar elementos necesarios de vigilancia que atiendan a otras actividades.

d) Las propiedades neutras revisten una gran importancia dentro del Análisis Delictivo que se propone, ya que estas propiedades son aquellos espacio-tiempos en los cuales se presentan la mayoría de actividades delictivas y, en consecuencia, presentan

grandes problemas de inseguridad. Estas propiedades proponen la aplicación de campañas más intensas de vigilancia, así como la aplicación de campañas preventivas e informativas.

4.2. Predicción

La metodología propuesta en esta etapa, permite predecir el número de delitos que ocurrirá**,** de cada una de las familias delictivas consideradas**,** dentro de una unidad espacial determinada (Sectores de Vigilancia, Colonias, Calles, etc.) y en la próxima unidad de tiempo contenida en la base de datos de referencia (Mes, Semana, Día), esta unidad de tiempo generalmente es el mes del año. Sin embargo es factible reducir esta unidad hasta el nivel de día, lo anterior dependerá de la frecuencia con que sean elaborados los planes preventivos de seguridad pública dentro de la región de estudios y en los cuales se incluyen las recomendaciones obtenidas de la interpretación planteadas en la sección anterior.

La metodología de predicción consta de cuatro etapas: Selección de una unidad espacio-temporal a ser predicha, análisis del comportamiento histórico de la actividad delictiva en la unidad espacio-temporal durante todo el período de la base datos de referencia, análisis del comportamiento de la actividad delictiva en la unidad espacio-temporal seleccionada, pero sólo durante los últimos doce meses y, por último, la construcción de la predicción resultante.

4.2.1. Selección de la unidad espacio-temporal a predecir.

Debido a la naturaleza continua de los diferentes tipos de análisis a realizar, la unidad de tiempo a ser predicha debe de ser siempre la subsecuente a la contenida en la base de datos de referencia. Por lo que únicamente se debe seleccionar la unida espacial a pronosticar.

4.2.2. Análisis Histórico

Este análisis plantea que en ciertos periodos del año algunas familias delictivas son más comunes que otras y que además tiende a comportarse casi de manera periódica a través de los años.

Para realizar el análisis, todos los delitos de cada familia delictiva son seleccionados y contados dentro de la unidad espacio-temporal seleccionada y en cada uno de los años de toda la base de datos de referencia.

4.2.3. Análisis del los 12 últimos meses

Este análisis se desarrolla de manera similar al descrito en el párrafo anterior, con la diferencia de que los delitos de cada familia delictiva son seleccionados y contados dentro de la unidad espacial seleccionada durante los últimos 12 últimos meses de la base de datos de referencia. Lo anterior con el objeto de determinar la tendencia que actualmente tiene cada familia delictiva en el espacio seleccionado.

4.2.4. Predicción Final

El número de delitos predichos (F), pertenecientes a la familia delictiva (c), observados dentro de la unidad espacial (s) y durante la unidad de tiempo (t), está determinado por:

$$F(c,s,t) = \alpha\left[f(c,s,t)\right] + \gamma\left[g(c,s,t)\right] \tag{4.2}$$

Donde:

$f(c,s,t)$ Predicción parcial obtenida del análisis histórico de delitos, de cada familia delictiva, observados dentro de la unidad espacio-temporal seleccionada, pero durante la misma unidad de tiempo (usualmente el mes de la año) de cada año disponible en la base de datos de referencia. La unidad en que se expresa esta predicción es número de delitos.

$g(c,s,t)$ Predicción parcial obtenida del análisis de delitos (misma familia delictiva y misma unidad espacial) observados durante los 12 últimos meses en la base de datos de referencia. La unidad en que se expresa esta predicción es número de delitos.

α,γ Valores de ponderación asignados a cada una de las predicciones anteriores.

El propósito de las dos predicciones parciales es aproximar la tendencia observada en el espacio-tiempo seleccionado. Cuando la tendencia es conocida es posible predecir el siguiente valor únicamente con observar el último valor en la base de datos de referencia. Para alcanzar esta predicción la herramienta básica a ser usada es simplemente una consulta a la base de datos para contar el número de delitos, de una familia delictiva especifica (c) y observada dentro un espacio y tiempo especifico (s,t).

Sea $\theta(t) =$ número de delitos, de una familia delictiva especifica (c) y observada dentro un espacio y tiempo especifico (s, t).

Se tiene que ambas predicciones parciales pueden ser expresadas como:

$$f(c,s,t) = \theta(t-1) + \frac{1}{K_1} \sum_{p=primer\ año}^{último\ año} \left[\theta(t-p) - \theta(t-p+1)\right] \qquad (4.3)$$

$$g(c,s,t) = \theta(t-1) + \frac{1}{K_1} \sum_{p=1}^{11} \left[\theta(t-p) - \theta(t-p+1)\right] \qquad (4.4)$$

Donde K_1 y K_2 son igual a la suma de cambios en el incremento y decremento en el número de delitos a través del tiempo.

CAPÍTULO CINCO

Resultados experimentales.

Se detallan los resultados obtenidos de diferentes experimentos realizados siguiendo la metodología para el Análisis y Predicción Delictiva propuesta.

5.1 Universo de Estudio

Para los experimentos realizados en esta sección, se empleó como universo de estudio una base de datos de hechos delictivos registrados en la ciudad de Sacramento, en el estado de California, E.U.A. Esta información se encuentra disponible para el público en general a través de la página de su departamento de policía en *http://www.sacpd.org/crime/stats/reports/*.

5.1.1 Sacramento, California, E.U.A.

Sacramento es la capital del estado de California, en los Estados Unidos. Situada en el interior, al Norte del Área de la Bahía de San Francisco, es de menor tamaño e importancia económica que Los Ángeles, San Francisco, San Diego y San José las otras cuatro principales ciudades californianas. Según el censo de 2000 tenía una población de 407.018 habitantes, en 2005 contaba con 456.4411 y en 2007 de 467.343, en la actualidad cuenta con una población estimada de 486,990. Es la séptima ciudad más poblada de California.

Figura 13. Mapa de Sacramento, California E.U.A.

5.1.2 Consideraciones importantes sobre la base de datos

La base de datos contiene información delictiva desde enero del 1997 hasta el mes mayo del 2009.

Los hechos delictivos contenidos en la base de datos empleada contiene la siguiente información:

- Hecho delictivo. Contiene información relativa al tipo de delito que aconteció (Robo, Violación, Homicidio, etc.).

- Ubicación Temporal. Cada uno de los hechos delictivos están ubicados en tiempo con suficiente nivel de detalle (año, mes, día y hora).

- Ubicación Espacial. Cada uno de los hechos delictivos están ubicados en espacio con suficiente nivel de detalle (Distritos, Beats, Grids y coordenadas en "x" y "y").

Sobre este aspecto, el departamento de policía de Sacramento, para ubicar en espacio cada hecho delictivo, divide el mapa de la ciudad en seis (6) distritos. Los distritos a su vez son divididos en sectores de vigilancia denominados "Beats", a los cuales son asignados las patrullas de vigilancia (19 Beats en total); cada sector o Beat es a su vez dividido en subsectores más pequeños denominados "Grids". La Figura 14 muestra esta división territorial.

Con respecto al párrafo anterior, es conveniente señalar que en México un problema con el que se enfrentas las oficinas de Análisis Delictivo, es que en cada nueva administración política tiende a cambiar sus sectores de vigilancia dentro de una determinada región, esto provoca que un delito que fue ubicado en un Sector de Vigilancia en una administración anterior, no corresponda con la actual división sectorial. Esto provoca que sea muy difícil establecer una tendencia correcta del comportamiento de una familia delictiva dentro de una región de estudio. Esta situación es manejada de una forma muy sencilla en el departamento de policía del Sacramento al dividir todo el territorio en "Grids", esta última división es fija y nunca cambia, lo que permite que aunque existan múltiples reasignaciones de sectores de vigilancia, cada hecho delictivo se puede ubicar mediante su "Grid".

El Departamento de policía de Sacramento, establece, que el margen de error existente en su base de datos es aproximadamente del 10 %, debido a que los datos son producidos a través de un conjunto de procesos muy complejos. Sin embargo, se desconoce el método específico para calcular dicho margen de error.

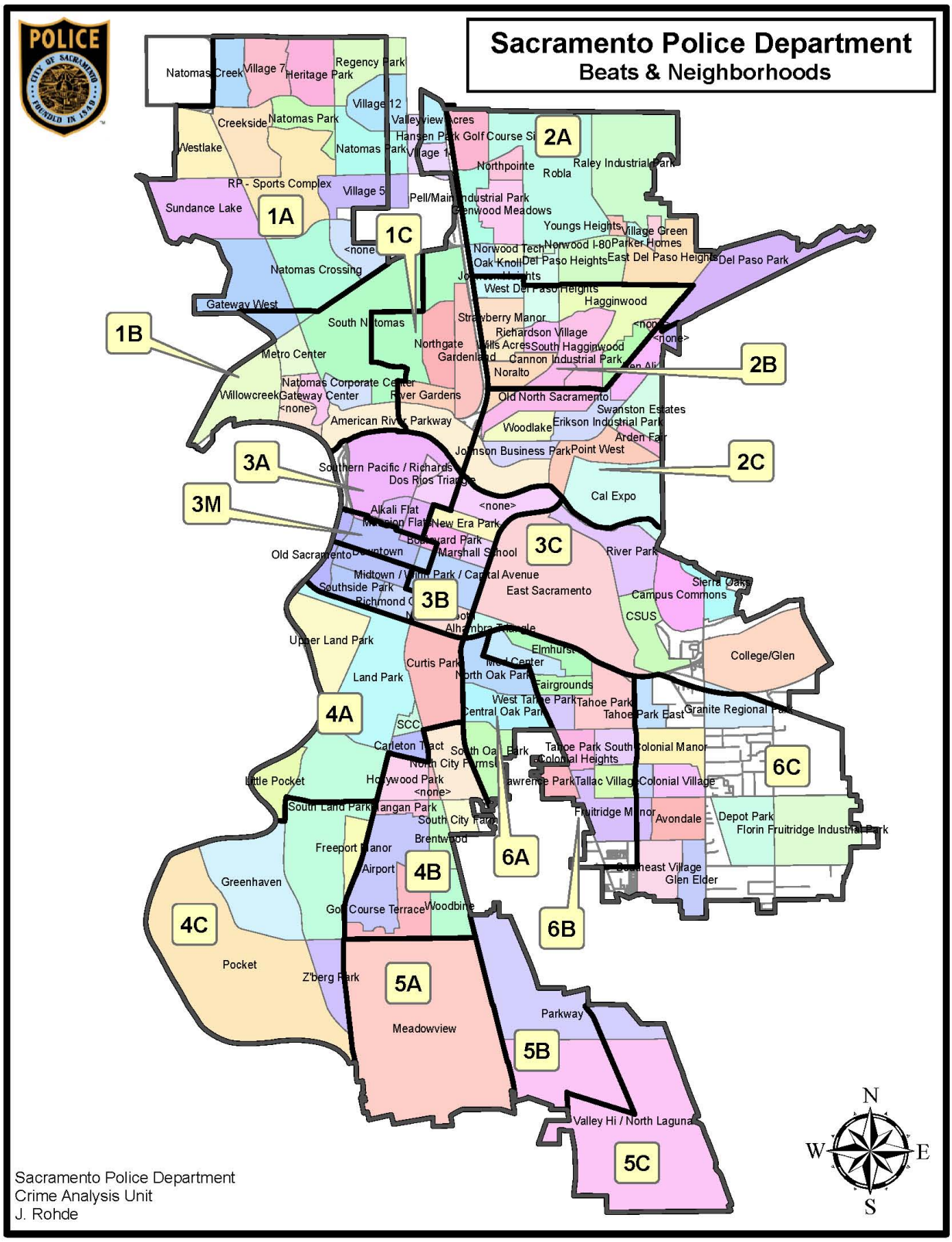

Figura 14. División Territorial por sectores de vigilancia de la ciudad de Sacramento, california, E.U.A.

A continuación se describen diversos experimentos realizados, con el propósito de probar la metodología de predicción delictiva propuesta.

5.2 Análisis Delictivo (Experimento 1)

5.2.1 Objetivo del experimento 1

Aplicar la metodología propuesta de Análisis Delictivo para determinar lo siguiente:

1. Cuáles son las familias delictivas que están ocurriendo de manera notable (característica o no-característica) dentro de la región de estudio.

2. Cuáles son las épocas del año y horarios en que estas actividades delictivas son más propensas a cometerse.

3. Los espacio-tiempos de los cuales es posible extraer recursos humanos y materiales para reasignar a la prevención de familias delictivas que están ocurriendo de manera notables (característica o no-característica).

4. Finalmente, cuáles son los espacio-tiempos en donde se presentan la mayoría de las familias delictivas.

Para probar la metodología, se emplearon 152,802 hechos delictivos registrados en la base de datos de referencia durante el periodo de Enero del 2004 a diciembre de 2009.

5.2.2 Construcción de una clasificación de referencia

Se seleccionaron las 16 familias delictivas con más cantidad de hechos delictivos e impacto en la sociedad para ser consideradas como **clases delictivas** dentro del experimento.

Cada uno de los hechos delictivos fue convertido en patrones delictivos y agrupados dentro de una clase delictiva.

a. La unidad espacial empleada para cada patrón delictivo fue el Sector de Vigilancia.

b. Las unidades de tiempo empleadas para cada patrón delictivo fueron el mes del año y el octante horario. Este último se determina dividiendo las 24 horas del día en ocho octantes de tres horas cada uno y sustituyendo la hora del hecho delictivo por el Octante Horario que le corresponda. En la Tabla 2 se muestran los Octantes Horarios resultantes.

Cada patrón delictivo quedo formado de la siguiente forma:

$$P_{C_i} = \begin{Bmatrix} \text{Sector de Vigilancia, Mes, Octante Horario} \\ 1A, \quad \textit{Enero}, \quad 8 \end{Bmatrix}$$

Tabla 2 Octantes horarios

Octante Horario	Horas Cubiertas
1	(00:01-03:00 hrs)
2	(03:01-06:00 hrs)
3	(06:01-09:00 hrs)
4	(09:01-12:00 hrs)
5	(12:01-15:00 hrs)
6	(15:01-18:00 hrs)
7	(18:01-21:00 hrs)
8	(21:01-00:00 hrs)

En la Tabla 3 se muestran las familias delictivas empleadas y el número de patrones delictivos agrupados en cada una de ellas.

Tabla 3 Familias delictivas y sus patrones

No. de Clase	Clase Delictiva	No de Patrones
1	Robo de Vehículo	28,816
2	Robo	451
3	Robo a casa habitación	18,735
4	Robo a negocio	7,827
5	Vandalismo de poco Impacto < 400 dólares	10,269
6	Vandalismo de Gran Impacto > 400 dólares	5,229
7	Asalto de poco Impacto (Robo Hormiga)	27,934
8	Asalto de gran Impacto	6,583
9	Accidentes de Trafico	10,610
10	Delitos Sexuales	3,439
11	Violencia Domestica	11,614
12	Actividades de Pandilla	3,005
13	Delitos relacionados con armas	7,233
14	Homicidio	348
15	Riñas	5,695
16	Delitos Contra la Salud	5,014
	Total	152,802

5.2.4 Construcción de definiciones inductivas

En la Tabla 4 se muestran los umbrales empleados en la construcción de las definiciones inductivas:

Tabla 4 Umbrales empleados en la construcción de definiciones inductivas.

Umbral	Valor
β_1 *(Umbral Positivo de Caracterización Positiva)*	**0.40**
β_2 *(Umbral Positivo de Caracterización Negativa)*	**0**
β_2' *(Umbral de Presencia para Caracterización Negativa)*	**0.80**
β_3 *((Umbral de Holgura entre Clases)*	**0.65**
β_3' *(Umbral de Porcentaje de Presencia para Caracterización Neutra)*	**0.45**

Usando la metodología propuesta en el capitulo anterior, fueron encontradas todas las propiedades espacio-temporales características positivas y negativas de cada una de las familias delictivas, así como el número de sectores de vigilancia en que estas familias son características. La Tabla 5 muestra estos resultados.

Finalmente, las propiedades espacio-temporales neutras de todas las familias delictivas fueron encontradas. La Tabla 6 muestra estos resultados.

5.2.5 Conclusiones del Experimento 1

Algunas de las conclusiones que se pueden obtener de la experimentación realizada son las siguientes:

a. De la Tabla 5 se puede determinar que las familias delictivas que tienen mayor impacto en la ciudad de Sacramento, California, E.U.A. son**: el robo de vehículo y el robo a casa habitación**. El robo de vehículo afecta de manera característica a 8 Sectores de Vigilancia mientras que el robo a casa habitación a 10 de ellos. Estas actividades suceden de manera característica en todo el año. Los octantes horarios más comunes para el robo de vehículo son: 1, 2, 7 y 8, es decir en un horario de entre de las 6 de la tarde y las 6 de la mañana.

b. Por otra parte, en la misma tabla se aprecia que otras actividades delictivas importantes son: el robo a negocio y el robo de poco impacto, este último afecta fuertemente sólo a dos sectores de vigilancia, durante todo el año y en un horario de entre las 9 de la mañana y 9 de la noche.

c. Igualmente, es posible observar que existen 27 propiedades características negativas de la clase de robo a casa habitación y 878 de la clase de homicidio, que nos indican espacio-tiempos en donde no suceden de manera característica estas actividades delictivas y por lo tanto es posible reasignar recursos humanos y materiales a otros espacio-tiempos donde se requieran.

d. Finalmente, es importante señalar que las conclusiones obtenidas se obtienen de manera directa del resumen de propiedades características positivas y negativas encontradas. Sin embargo, es posible conseguir conclusiones más específicas como el mes del año en que es más común el robo de vehículo. Esto último se logra analizando de manera detallada cada una las propiedades espacio-temporales obtenidas en la experimentación.

Tabla 5 Número de propiedades espacio-temporales características positivas y negativas encontradas.

Familia Delictiva	# de Propiedades Características Positivas	# de Propiedades Características Negativas	Sectores de Vigilancia en que la familia delictiva es característica.	Meses en que la familia delictiva es característica.	Octantes Horarios en que la familia delictiva es característica
Robo de Vehículo	41	5	1A, 1B, 2C, 3B, 3C, 3M, 4A, 4C	Todos los meses del año	1, 2, 7 Y 8
Robo	0	802	0	0	0
Robo a casa habitación	27	27	1C, 2B, 4A, 4B, 4C, 5A, 5B, 5C 6A, 6B	Todos los meses del año	3,4 Y 5
Robo a negocio	2	68	3M, 6C	Agosto y Febrero	2
Vandalismo de poco Impacto < 400 dólares	0	8	0	0	0
Vandalismo de Gran Impacto > 400 dólares	0	63	0	0	0
Asalto de poco Impacto (Robo Hormiga)	43	0	2C, 3M	Todos los meses del año	4,5,6,7
Asalto de gran Impacto	0	32	0	0	0
Accidentes de Trafico	0	30	0	0	0
Delitos Sexuales	0	108	0	0	0
Violencia Domestica	0	5	0	0	0
Actividades de Pandilla	0	288	0	0	0
Delitos relacionados con armas	0	13	0	0	0
Homicidio	0	878	0	0	0
Riñas	0	37	0	0	0
Delitos Contra la Salud	0	130	0	0	0

Tabla 6 Espacio-tiempos afectados por la mayoría de las familias delictivas

Sector de Vigilancia (Beat)	Mes	Octante Horario	# de Familias Delictivas que afectan al espacio-tiempo
3A	Febrero	2	7
2B	Marzo	7	10
2B	Marzo	8	6
1C	Mayo	8	9
6A	Junio	8	8
2B	Julio	1	7
6C	Agosto	6	7
3A	Abril	7	10
1C	Mayo	2	6
6B	Junio	2	10
2B	Agosto	2	8
2A	Noviembre	2	6
2B	Diciembre	1	9
3A	Diciembre	2	6
6A	Noviembre	1	9
5A	Enero	2	9
3M	Abril	2	7

5.3 Predicción delictiva (Experimento 2, 3 y 4)

5.3.1 Experimento 2

5.3.1.1 Objetivo del experimento 2

El objetivo principal del presente experimento consistió en probar y evaluar sobre los meses de enero, febrero, marzo y abril del 2009 de la base de datos de Sacramento, la metodología de predicción propuesta.

5.3.1.2 Selección de la unidad espacio-temporal a predecir

En este experimento se seleccionaron como unidades espaciales los 19 sectores de vigilancia (*beats*) que conforman al estado de Sacramento y como unidad temporal el mes del año.

5.3.1.3 Análisis Histórico

Para este análisis se ocuparon 5 años de hechos delictivos obtenidos de la base de datos de referencia. Para predecir enero del 2009, se emplearon datos de enero del 2004 a diciembre del 2008; para predecir febrero del 2009 los datos empleados fueron de febrero del 2004 a enero del 2009, y así sucesivamente para los 4 meses en que se realizó la predicción.

5.3.1.4 Análisis del los 12 últimos meses

Para este análisis únicamente se ocuparon los 12 meses anteriores a cada uno de los meses a predecir. Para enero del 2009, los hechos delictivos seleccionados fueron tomados de enero del 2008 a diciembre del 2008, Para febrero del 2009, los hechos fueron de febrero del 2008 a enero del 2009, y así sucesivamente.

5.3.1.5 Predicción Final

La predicción se realizó para todas las familias delictivas y en cada una de las propiedades espacio-temporales seleccionadas.

Los valores de los pesos asignados a cada uno de los análisis anteriores fue: para el análisis histórico, $\alpha = 0.2$ y para el análisis de los 12 últimos meses $\gamma = 0.8$. Estos valores se obtuvieron mediante experimentación.

5.3.1.6 Evaluación del experimento 2

Para evaluar los resultados de la predicción realizada a cada propiedad espacio-temporal se empleó el **Error Cuadrático Medio** definido por:

$$ECM = \frac{\sum_{p=1}^{N} e_p^2}{N} \qquad (5.1)$$

Donde:

$e_p = Y_p - P_p$, Y_p es la cantidad de delitos que ocurrieron de manera real en el espacio-tiempo seleccionado y P_p es la cantidad de delitos predichos por la metodología.

En las Tablas 7, 8, 9 y 10 se muestran de manera general los resultados obtenidos en la predicción realizada con la metodología propuesta sobre todas las familias delictivas y en todos los espacios en los mese de enero, febrero, marzo y abril del 2009.

Finalmente, en la Tabla 11 Se muestra el resumen general de Errores Cuadráticos Medios obtenidos por la metodología para cada uno de los meses en los que fue empleado.

Tabla 7 Predicción realizada sobre enero del 2009

Familia delictiva	# real de delitos cometidos (Ene-2009)	# de delitos predichos por la Metodología propuesta	Error Medio	Predicción propuesta. (ECM)
Robo de Vehículo	417	384	10.04	100.89
Robo	11	6	1.15	1.32
Robo a casa habitación	328	343	6.48	42.05
Robo a negocio	135	118	2.93	8.58
Vandalismo de poco Impacto < 400 dólares	108	129	3.61	13.00
Vandalismo de Gran Impacto > 400 dólares	80	63	3	9.00
Asalto de poco Impacto (Robo Hormiga)	449	393	8.2	67.26
Asalto de gran Impacto	93	107	2.1	4.42
Accidentes de Trafico	170	194	3.43	11.79
Delitos Sexuales	51	48	1.85	3.42
Violencia Domestica	152	185	3.27	10.68
Actividades de Pandilla	51	31	2.08	4.32
Delitos relacionados con armas	105	82	3.05	9.32
Homicidio	2	6	0.73	0.53
Riñas	67	52	2.42	5.84
Delitos Contra la Salud	57	50	2.61	6.79
Total	**2,276**	**2,191**	**56.95***	**299.21***

* Para obtener el **ECM** y el **Error Medio**, es necesario que el valor mostrado sea dividido entre las 16 familias delictivas. En la Tabla 5.11 se muestran estos resultados.

Tabla 8 Predicción realizada sobre febrero del 2009

Familia delictiva	# real de delitos cometidos (Feb-2009)	# de delitos predichos por la Metodología propuesta	Error Medio	Predicción propuesta. (ECM)
Robo de Vehículo	415	402	6.31	39.84
Robo	5	11	1.12	1.26
Robo a casa habitación	294	325	5.44	29.63
Robo a negocio	89	127	2.88	8.32
Vandalismo de poco Impacto < 400 dólares	152	114	4.12	16.95
Vandalismo de Gran Impacto > 400 dólares	61	79	2.45	6.00
Asalto de poco Impacto (Robo Hormiga)	403	432	7.83	61.32
Asalto de gran Impacto	86	95	3.44	11.84
Accidentes de Trafico	159	174	3	9.00
Delitos Sexuales	39	46	1.99	3.95
Violencia Domestica	159	159	2.6	6.74
Actividades de Pandilla	60	49	1.91	3.63
Delitos relacionados con armas	88	100	2.38	5.68
Homicidio	1	2	0.4	0.16
Riñas	57	65	2.49	6.21
Delitos Contra la Salud	51	59	1.97	3.89
Total	**2,119**	**2,239**	**50.33***	**214.42***

* Para obtener el **ECM** y el **Error Medio**, es necesario que el valor mostrado sea dividido entre las 16 familias delictivas. En la Tabla 5.11 se muestran estos resultados.

Tabla 9 Predicción realizada sobre marzo del 2009

Familia delictiva	# real de delitos cometidos (Mar-2009)	# de delitos predichos por la Metodología propuesta	Error Medio	Predicción propuesta. (ECM)
Robo de Vehículo	385	409	8.89	79.05
Robo	7	6	1	1.00
Robo a casa habitación	385	301	8.71	75.89
Robo a negocio	97	92	3.73	13.95
Vandalismo de poco Impacto < 400 dólares	144	155	4.08	16.68
Vandalismo de Gran Impacto > 400 dólares	62	65	2.82	7.95
Asalto de poco Impacto (Robo Hormiga)	450	407	5.47	29.95
Asalto de gran Impacto	99	94	2.72	7.42
Accidentes de Trafico	175	158	3.3	10.89
Delitos Sexuales	63	45	2.6	6.74
Violencia Domestica	162	165	3.49	12.16
Actividades de Pandilla	57	57	2.1	4.42
Delitos relacionados con armas	98	87	2.87	8.26
Homicidio	1	2	0.4	0.16
Riñas	74	59	2.58	6.68
Delitos Contra la Salud	55	55	2.15	4.63
Total	**2,314**	**2,157**	**56.91***	**285.84***

* Para obtener el **ECM** y el **Error Medio**, es necesario que el valor mostrado sea dividido entre las 16 familias delictivas. En la Tabla 5.11 se muestran estos resultados.

Tabla 10 Predicción realizada sobre abril del 2009

Familia delictiva	# real de delitos cometidos (Abr-2009)	# de delitos predichos por la Metodología propuesta	Error Medio	Predicción propuesta. (ECM)
Robo de Vehículo	362	378	7.74	59.89
Robo	9	7	0.97	0.95
Robo a casa habitación	314	374	7.33	53.79
Robo a negocio	76	97	3.2	10.26
Vandalismo de poco Impacto < 400 dólares	158	148	4.58	20.95
Vandalismo de Gran Impacto > 400 dólares	90	67	3.1	9.63
Asalto de poco Impacto (Robo Hormiga)	392	444	5.09	25.89
Asalto de gran Impacto	113	104	3.14	9.84
Accidentes de Trafico	186	169	5.37	28.89
Delitos Sexuales	50	55	2.65	7.00
Violencia Domestica	180	166	3.77	14.21
Actividades de Pandilla	50	61	3.2	10.26
Delitos relacionados con armas	80	97	2.93	8.58
Homicidio	4	1	0.4	0.16
Riñas	78	74	2	4.00
Delitos Contra la Salud	58	57	3.19	10.16
Total	**2,200**	**2,299**	**58.66***	**274.47***

* Para obtener el **ECM** y el **Error Medio**, es necesario que el valor mostrado sea dividido entre las 16 familias delictivas. En la Tabla 5.11 se muestran estos resultados.

Tabla 11 ECM general para cada uno de los meses predichos.

Mes	Error Medio	ECM
Enero 2009	4.32	18.70
Febrero 2009	3.66	13.40
Marzo 2009	4.23	17.86
Abril 2009	4.14	17.15

5.3.1.7 Conclusiones del experimento 2

Los resultados obtenidos en la experimentación muestran que la metodología propuesta tiene una alta efectividad en la predicción de delitos, con un ECM máximo para el mes de enero de 18.70 y un mínimo de 13.40 para el mes de febrero. Este resultado significa que, en promedio**,** la metodología falla en menos de 5 ocurrencias por cada familia delictiva en la predicción de cada mes. Esta precisión es más que aceptable para cualquier sistema automático de predicción delictiva.

5.3.2 Experimento 3

5.3.2.1 Objetivo del experimento 3

El objetivo del presente experimento consiste determinar el **grado de tolerancia** de la metodología de predicción propuesta cuando se cuenta con menos años en la base de datos de referencia. La predicción fue hecha sobre el mes de **febrero del 2009**.

Para este experimento, la selección de la unidad espacio-temporal a predecir, el análisis del los 12 últimos meses y los valores de los pesos asignados a cada uno de los análisis continúan de la misma manera a la expuesta en el experimento anterior.

5.3.2.2 Análisis Histórico

En este experimento, el análisis se realizó reduciendo la cantidad de años tomados de la base de datos de referencia. Primeramente se emplearon solo 4 años (Enero del 2005 a diciembre del 2008), después 3 años (Enero del 2006 a diciembre del 2008), y así sucesivamente hasta emplear solo el último año de la base de datos.

5.3.2.3 Evaluación del experimento 3.

Los resultados obtenidos de la experimentación se muestran en la Tabla 12.

Tabla 12 ECM obtenido al reducir la cantidad de años en la base de datos de referencia

# de años tomados de la base de datos de referencia	Error Medio	ECM (Febrero 2009)
4 años (2005-2008)	3.67	13.50
3 años (2006-2008)	3.67	13.50
2 años (2007-2008)	3.69	13.60
1 año (2008)	3.7	13.70

5.3.2.4 Conclusiones del experimento 3

Los resultados obtenidos en la experimentación demuestran que la metodología de predicción propuesta tiene una alta tolerancia a la reducción de datos al pasar de un ECM de 13.50 utilizando 4 años de la muestra a 13.70 utilizando solo uno. Lo anterior permite demostrar que la metodología arroja buenos resultados con un año de la base de datos de referencia. Sin embargo el procesamiento que se emplea es muy poco y por lo tanto es recomendable llevarlo a cabo con la mayor cantidad de datos disponibles.

5.3.3 Experimento 4

5.3.3.1 Objetivo del experimento 4

Con base en los resultados obtenidos en el experimento anterior, podría suponerse que el análisis histórico que se propone en la metodología de predicción podría descartarse, ya que el ECM aumenta muy poco cuando se emplea un año en este análisis.

El objetivo de este experimento consiste determinar que tanto influye en los resultados de la metodología de predicción propuesta si no se considera el análisis histórico. La predicción fue hecha sobre el mes de **febrero del 2009**.

Para este experimento, la selección de la unidad espacio-temporal a predecir, el análisis del los 12 últimos meses y los valores de los pesos asignados a cada uno de los análisis continúan de la misma manera a la expuesta en el experimento 2.

5.3.3.2 Evaluación del experimento 4

Los resultados obtenidos de la experimentación se muestran en la tabla 13.

Tabla 13 ECM obtenido al no considerarse el análisis histórico

Error Medio	ECM (Febrero 2009)
4.12	16.99

5.3.3.3 Conclusiones del experimento 4

Los resultados obtenidos en la experimentación demuestran que la metodología de predicción propuesta incrementa de manera considerable su ECM de 13.50 a 16.99 cuando no se utiliza el análisis histórico. Lo que permite concluir que este análisis es necesario. Además representa un costo computacional muy bajo para el sistema al requerir trabajar solo con un dato por cada año de la base de datos de referencia.

5.3.4 Experimento 5

5.3.4.1 Objetivo del experimento 5

Este experimento tuvo como objetivo comparar la metodología de predicción propuesta con los siguientes algoritmos empleados en el análisis y predicción de series de tiempo:

1. Probabilidad Bayesiana
2. Promedio Móvil (Calculado con MINITAB versión 15).
3. Suavizamiento Exponencial Doble (Calculado con MINITAB versión 15).
4. ARIMA (Calculado con MINITAB versión 15).

5.3.4.2 Consideraciones especiales

A fin de poder comparar de manera equitativa varios métodos de análisis y predicción de Series de tiempo fue necesario tomar en cuenta lo siguiente:

- Cada algoritmo trabajo con 60 datos. Cada dato corresponde al número de hechos delictivos de cada familia, presentados en un espacio seleccionado (Sector de Vigilancia) en cada uno de los meses de todo el universo de estudio (de enero del 2004 a Diciembre del 2008).

- La predicción fue hecha sobre enero y febrero del 2009.

5.3.4.3 Análisis del los 12 últimos meses

La metodología de predicción propuesta únicamente ocupa los 12 meses anteriores a cada uno de los meses a predecir. Sin embargo, a fin de trabajar con la misma cantidad de datos que los otros algoritmos propuestos se realizaron los siguientes cambios sobre la metodología propuesta:

1. El universo de estudio empleado se divide en años.
2. En cada año se realiza el análisis de los 12 últimos meses a fin de predecir el mes propuesto en cada año.

3. Se determina la ***predicción ponderada parcial*** mediante:

$$P_p = \frac{\sum_{i=\text{Primer año}}^{\text{último año}} \eta_i . g(c,s,t_i)}{\eta_i} \qquad (5.2)$$

Donde:

P_p = Al promedio ponderado de la predicción realizada en el mes seleccionado en todos los años del universo de estudio.

η_i = A una ponderación que se le da a la predicción realizada cada año del universo de estudio. Entre más reciente sea el año recibe ponderaciones más altas.

4. Finalmente, se determina el error obtenido en cada una de las predicciones realizadas a través de todo el universo de estudio mediante:

$$e_p = \sum_{i=\text{Primer año}}^{\text{último año}} (Y_i - g(c,s,t_i)) \qquad (5.3)$$

Donde:

e_p = Al error obtenido en cada una de la predicciones parciales realizadas en todo el universo de estudio.

Y= a la cantidad de hecho delictivos ocurridos realmente en el mes y año seleccionado.

5.3.4.4 Predicción Final

La predicción final está dada por la siguiente fórmula:

$$P_F = P_p + (\varphi . e_p) \qquad (5.4)$$

Donde:

φ = a un factor que permite determinar el porcentaje de error que se le suma a la predicción ponderada parcial a fin de obtener un mejor resultado.

φ Toma valores entre 0 y 1.

5.3.4.5 Predicción bayesiana

La predicción bayesiana se realiza de la siguiente forma:

1) Haciendo uso del Teorema de Bayes, se calcula la probabilidad condicional de que, delitos de una familia delictiva específica ocurran en un espacio-tiempo dado.

2) La probabilidad obtenida en el paso anterior se multiplica por el número promedio de delitos, de cualquier familia delictiva, que ocurren en el espacio de referencia y durante todos los meses contenidos en la muestra.

3) El producto así obtenido se interpreta como el número de delitos, de la familia delictiva estudiada, que se predice ocurrirán, en el espacio y durante el tiempo referidos.

5.3.4.6 Evaluación del experimento 5

La metodología de predicción propuesta emplea los siguientes parámetros:

- $\varphi = 0.5$
- $\eta_{2005} = 0.10,\ \eta_{2006} = 0.15,\ \eta_{2007} = 0.20,\ \eta_{2008} = 0.25,\ \eta_{2009} = 0.30$

Los resultados obtenidos de la experimentación se muestran en la tabla 14.

Tabla 14 ECM obtenido de cada uno de los algoritmos de predicción propuestos

Algoritmo de predicción	Error Medio (Enero 2009)	ECM (Enero 2009)	Error Medio (Febrero 2009)	ECM (Febrero 2009)
Promedio Móvil MA(3)	3.99	15.90	3.99	15.94
Suavizamiento Exponencial Doble	4.33	18.77	3.92	15.34
ARIMA (2,1,2)	3.91	15.25	3.99	15.91
Bayes	4.24	17.96	4.42	19.50
Metodología Propuesta SIAPD	**3.66**	**13.40**	**3.59**	**12.91**

5.3.4.7 Conclusiones del experimento 5

Los resultados obtenidos en la experimentación muestran que la metodología de predicción propuesta presenta un buen desempeño en comparación con otros algoritmos de predicción empleados de manera clásica en este tipo de problemas. Por otra parte se muestra que la metodología mejora de manera considerable al utilizar diferentes predicciones parciales y la inclusión del error en la predicción final, disminuyendo el ECM que se tenía de **18.70** a **13.40** para el mes de enero del 2009 y de **13.50** a **12.91** para el mes de febrero del 2009.

Resulta relevante observar, en los resultados de este experimento, que no necesariamente se tiene que trabajar únicamente con un solo algoritmo de predicción; el uso combinado de varios de estos algoritmos puede arrojar información relevante para el analista delictivo, tal es el caso de la predicción realizada sobre el mes de **Enero del 2009** para el **Sector de Vigilancia 3A**, en el cual, la predicción resultante en cada uno de los cinco métodos empleados, es notoriamente semejante; sin embargo, el número real de delitos ocurridos resulto muy superior a todos las predicciones realizadas. En la Tabla 15 se muestran esta situación.

Tabla 15 Número de delitos predichos por cada algoritmo para el mes de enero del 2009 en el Sector de Vigilancia 3A

Algoritmo de Predicción	# de delitos predichos por cada algoritmo	# de delitos ocurridos realmente
Promedio Móvil	17	32
Suavizamiento Exponencial Doble	18	
ARIMA	19	
Bayes	17	
SIAPD	18	

Como podemos observar en la tabla anterior, todos los algoritmos predijeron un número similar de delitos. Sin embargo, la cantidad real que se presento es casi el doble de lo predicho, esto le muestra claramente al Analista Delictivo la posibilidad de que ese espacio-tiempo haya existido un factor político, social, económico, etc. que haya tenido como consecuencia un cambio notorio en la tendencia de comportamiento de la actividad delictiva.

CAPÍTULO SEIS

Desarrollo de Aplicación.

Se detallan los aspectos relacionados al desarrollo de una aplicación informática que permita probar la metodología de Análisis y Predicción Delictiva (SIAPD) propuesta.

En el presente capitulo, se detallan los aspectos de diseño e implementación de una aplicación informática que permita incorporar y probar la metodología de Análisis y Predicción Delictiva propuesta. La construcción de la aplicación tiene como propósito permitir la realización de pruebas de funcionamiento de la metodología teórica propuesta.

La metodología fue probada e implementada primeramente utilizando el lenguaje técnico de computación "*Matlab*". Sin embargo, este lenguaje matemático generalmente no permite la visualización de resultados a través una intranet o de internet. Esta visualización es necesaria para convertir la metodología propuesta en una verdadera herramienta de Análisis y Predicción Delictiva que proporcione información de manera descentralizada a cada uno de los elementos encargados de la elaboración de planes de seguridad pública dentro de sus áreas de su responsabilidad. Estas áreas generalmente serán los sectores de vigilancia en que es dividida la región de estudio y en donde se desarrollan las actividades de patrullaje tendiente a disminuir el índice delictivo.

Con base en lo anterior, fue necesario diseñar y construir una aplicación informática que incorpora la metodología de Análisis y Predicción delictiva con las siguientes características:

1. Funciona dentro de una **arquitectura cliente-servidor**, a fin de que los resultados obtenidos de cada uno de los diferentes análisis se encuentren disponibles en todo momento, para múltiples usuarios y desde diferentes locaciones.

2. Toda la información delictiva histórica empleada en cada uno de los procesos señalados, así como aquella que se vaya generando después de haberse implementado la aplicación sea administrada de manera centralizada mediante una base de datos.

3. Cada uno de los análisis propuestos deben de estar desarrollados de manera modular para facilitar su administración, mejoramiento y facilidad de uso por parte de los usuarios.

4. Cuenta con interfaces de usuario que permitan la configuración de diferentes parámetros necesarios para los análisis.

A continuación se describe la forma en que la aplicación fue desarrollada para cumplir con las características descritas arriba.

6.1 Arquitectura cliente-servidor

La arquitectura del *SIAPD* responde al modelo general denominado cliente-servidor. Esto significa que en el mayor nivel de abstracción el sistema está compuesto por dos sub-sistemas: el primero de ellos, denominado servidor, realiza el contacto con la base de datos, lleva a cabo los procesos de Análisis y Predicción delictiva y la elaboración de reportes finales. El segundo sub-sistema, denominado cliente, tiene como función el de proporcionar una interfaz con los usuarios.

Gracias a esta división de actividades entre los dos sub-sistemas, es posible dividir el sistema de forma tal que ambos sub-sistemas se encuentren ubicados en computadoras diferentes que pueden o no encontrarse físicamente en la misma ubicación. El mecanismo de comunicación y sincronización entre los sub-sistemas es el protocolo *http* usado hoy en día para la comunicación en la *Internet.* Esto posibilita que, de acuerdo a la arquitectura de *SIAPD,* el sub-sistema servidor sea instalado en una computadora con mayor capacidad de procesamiento, almacenamiento y con conexión hacia la *Internet*, o si se prefiere en una *Intranet,* mientras que el sub-sistema cliente se puede encontrar en una o varias computadoras de menor capacidad y costo. La base de datos y los programas que analizan los datos se encontrara en el sub-sistema servidor. Cuando los datos han sido analizados, el servidor recopila los resultados en archivos de tipo *html*, es decir, en formato de páginas Web y los muestra a los clientes. Lo anterior permite, reducir la funcionalidad del sub-sistema cliente a simples labores de entrada-salida que pueden ser desempeñadas por cualquier programa navegador de Web.

De esa forma, la arquitectura general del *SIAPD* puede resumirse en el diagrama mostrado en la figura 15.

Figura 15. Arquitectura cliente-servidor empleada

La arquitectura presentada, es la indicada para soportar las necesidades de cualquier organización encargada del Análisis Delictivo que lleva a cabo la planeación y elaboración descentralizadas de planes preventivos de seguridad pública, producto de la división territorial de cada región de estudio.

El servidor empleado para el desarrollo de la aplicación informática puede visualizarse como un cerebro que desarrolla las siguientes actividades tendientes a proporcionar a los clientes diferentes tipos de resultados del análisis y predicción delictiva realizada:

- Mantiene una única copia de la base de datos sobre la que se realiza el análisis.
- Acceso restringido, mediante claves de acceso, al personal autorizado para ver los datos y realizar análisis sobre ellos.
- Capacidad de consultar el proceso de Análisis y Predicción Delictiva, desde cualquier ubicación y desde cualquier tipo de computadora.
- Reducir al mínimo los requerimientos de la(s) computadoras clientes.
- Contar con la facilidad de tener más de un cliente concurrente, no forzosamente ubicados en el mismo sitio.
- Capacidad para emitir reportes indexados por Sector de Vigilancia.

6.2 Base de Datos

La información delictiva utilizada durante los procesos de Análisis y Predicción Delictiva es almacenada y administrada mediante una base de datos diseñada bajo el modelo relacional. Esta base de datos se encuentra alojada en el lado del servidor, lo que permite que múltiples usuarios acecen a ella de manera recurrente. Asimismo permite el establecimiento de normas de seguridad que regulan el acceso a ella.

6.2.1. Diseño de la Base de Datos.

Para el diseño de la base de datos se empleo el modelo entidad-relación en donde se considera información detallada de aspectos relacionados a la división territorial de la región de estudio, hechos delictivos ocurridos en espacio y tiempo, así como sus características más importantes.

En la figura 16 se muestra el modelo entidad-relación empleado.

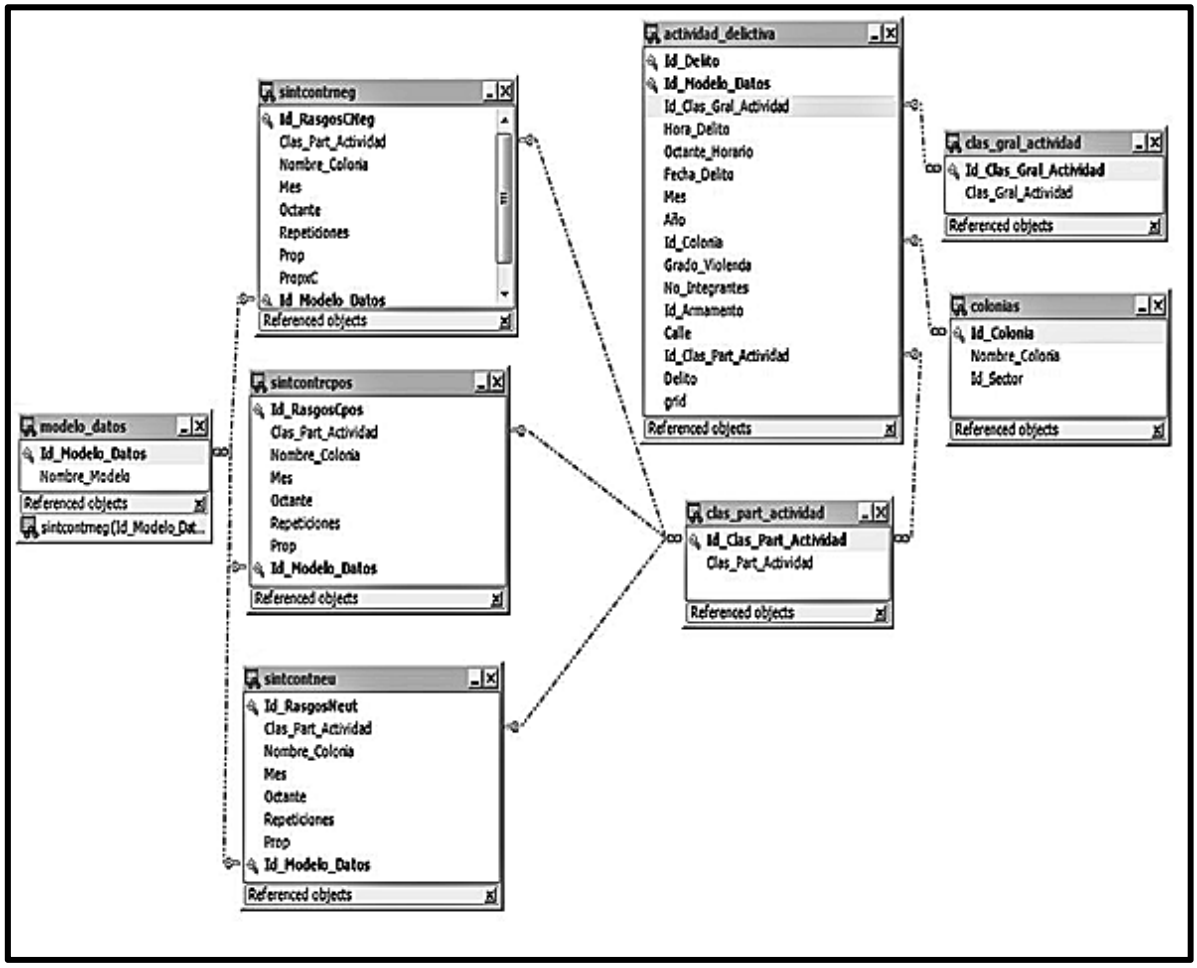

Figura 16. Modelo entidad-relación de la aplicación informática desarrollada

6.2.2. Implementación de la Base de Datos.

Para la implementación del modelo entidad relación se empleo el manejador MySQL que es un sistema de gestión de base de datos relacional, multihilo y multiusuario. Al contrario de proyectos como Apache, donde el software es desarrollado por una comunidad pública y el copyright del código está en poder del autor individual, MySQL es propietario y está patrocinado por una empresa privada (SUN Microsystems), que posee el copyright de la mayor parte del código.

MySQL es muy utilizado en aplicaciones Web, es una base de datos rápida en la lectura, pero en la modificación de datos puede provocar problemas de integridad en entornos de alta concurrencia. En aplicaciones Web hay baja concurrencia en la modificación de datos y un entorno intensivo de lectura de datos, lo que hace a MySQL ideal para este tipo de aplicaciones.

El manejador de bases de datos MySQL usado fue la versión 5.0 y fue elegido por ser el manejador de software libre más usado, documentado y uno de los más poderosos, además de que el flujo de datos en la aplicación informática propuesta no demanda un manejador más poderoso como Oracle o Informix.

6.3 Módulos de la aplicación

La aplicación informática fue desarrollada de manera modular, a fin de que cada uno de los procesos realizados durante el Análisis y Predicción Delictiva cuente con una interfaz propia que administre al módulo de manera independiente, permita la configuración de parámetros por parte de los usuarios y muestre los resultados obtenidos de cada análisis.

Cada uno de estos módulos fue implementado como una página web en el lado del servidor utilizando el lenguaje interpretado de programación PHP que es ampliamente usado para este propósito y puede ser incrustado dentro de código HTML.

Para el diseño de cada módulo se empleo la herramienta computacional CS3 de DreamWeaver, la cual es una aplicación destinada para la construcción y edición de sitios y aplicaciones basadas en estándares Web. Es uno de los programas de este tipo más utilizados en el sector del diseño y la programación Web, por sus funcionalidades, su integración con otras herramientas como Adobe Flash y, recientemente, por su soporte de los estándares del World Wide Web Consortium.

Tiene soporte tanto para edición de imágenes como para animación a través de su integración con otras herramientas.

Dreamweaver tiene alta compatibilidad entre la mayoría de los navegadores Web instalados en el ordenador cliente para previsualizar las páginas Web. También dispone de herramientas de administración de sitios dirigidas a principiantes, tales como, la habilidad de encontrar y reemplazar líneas de texto y código por cualquier tipo de parámetro especificado, hasta dentro del sitio Web completo. El panel de comportamientos también permite crear código JavaScript básico sin contar con amplios conocimientos del lenguaje.

Un aspecto de alta consideración de Dreamweaver es su arquitectura extensible. Las extensiones, tal y como se conocen, son pequeños programas, que cualquier desarrollador Web puede escribir (normalmente en HTML y Javascript) y que cualquiera puede descargar e instalar, ofreciendo así funcionalidades añadidas a la aplicación.

Los módulos implementados en la aplicación fueron:

- Módulo de inicio de aplicación.
- Módulo para el Análisis Delictivo.
- Módulo de predicción delictiva.

A continuación se describe de manera detallada cada uno de estos módulos.

6.3.1. Módulo de inicio de aplicación

Este módulo tiene como objetivo principal servir de enlace entre los módulos encargados de los diferentes análisis propuestos. Este módulo permite ejecutar la metodología para el Análisis Delictivo.

Los parámetros de entrada configurable por el usuario son los siguientes:

1. Modelos de datos. Es posible seleccionar el modelo de datos al que se va a realizar el análisis. Un modelo de datos puede ser considerado como una clasificación de referencia única, cada clasificación de referencia puede contener diferentes familias delictivas y hechos delictivos, el tipo de modelo a seleccionar dependerá del tipo de estrategia policiaca a emplear.

2. Valores de cada uno de los umbrales empleados par la caracterización positiva, negativa y neutra.

La figura 17 Muestra una esta página web.

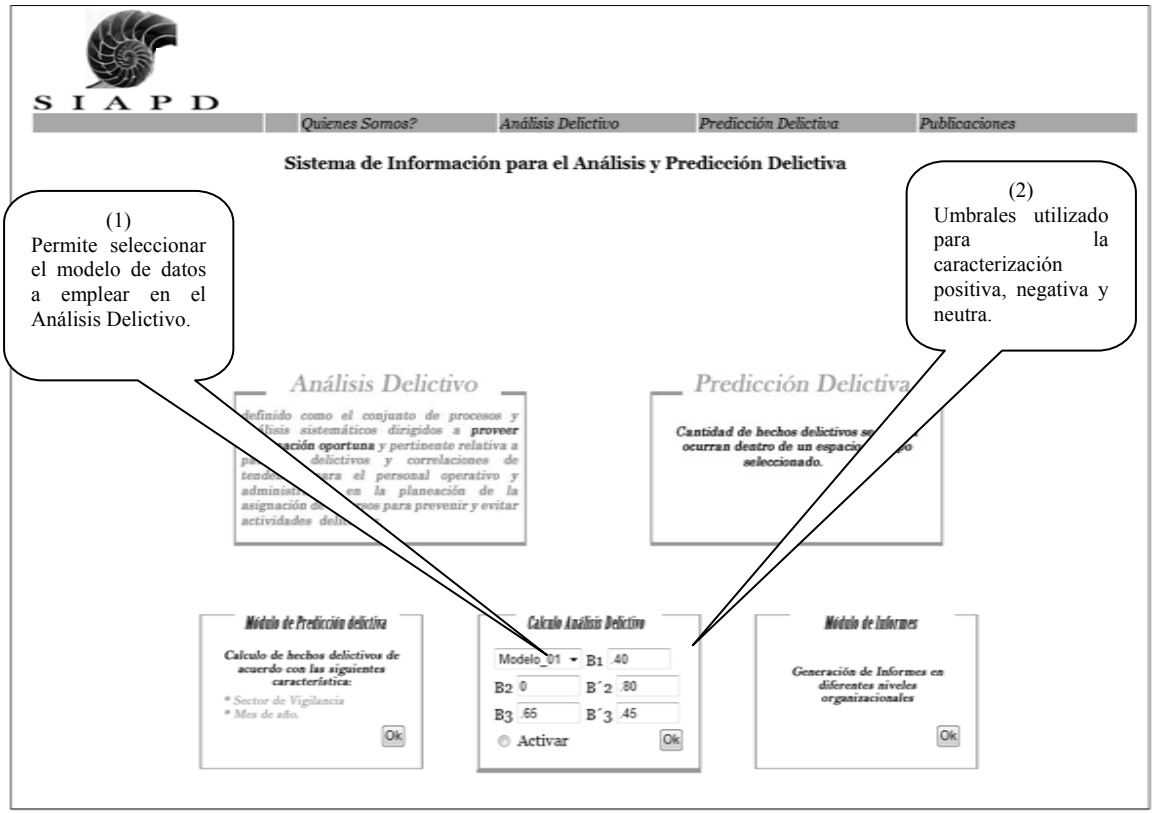

Figura 17. Módulo integrador

6.3.2. Módulo para el Análisis Delictivo

Este módulo implementa la metodología para el Análisis Delictivo. Dentro de su estructura interna cuenta con una programación hecha mediante PHP que establece una conexión con la base de datos, a fin de extraer información relacionada con los hechos delictivos suscitados dentro de cada uno de los sectores de vigilancia y poner en marcha los algoritmos que permiten determinar todas la propiedades características positivas y negativas de cada familia delictiva, así como aquellas propiedades neutras a todas la familias. Estos procesos se desarrollan de manera interna dentro del módulo y ningún usuario tiene acceso a estos. Sin embargo, el módulo cuenta con una interfaz de usuario que permite la visualización de los resultados que se obtienen del Análisis Delictivo.

Esta interface contiene un panel de opciones que permite visualizar las propiedades encontradas del análisis anterior mediante la selección de la familia delictiva, Sector de Vigilancia, mes, tipo de caracterización.

En la figura 18 se muestran los parámetros de entrada configurables y los datos de salida del módulo de Análisis Delictivo.

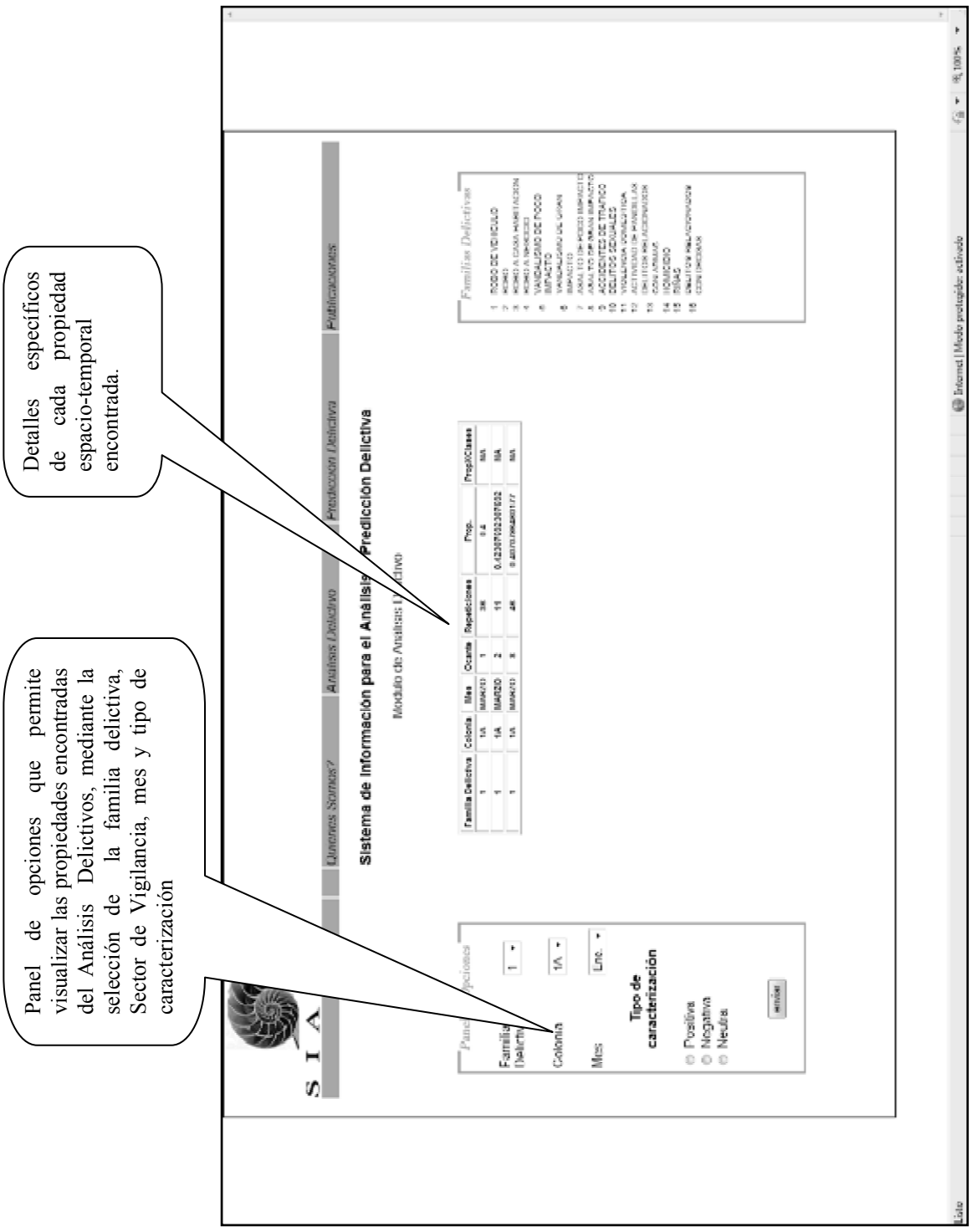

Figura 18 Parámetros de entrada y salida del módulo de Análisis Delictivo.

6.3.3. Módulo para la predicción delictiva

Este módulo implementa la metodología de Predicción Delictiva. Dentro de su estructura interna cuenta con una programación hecha mediante PHP que establece una conexión con la base de datos, a fin de extraer información relacionada con los hechos delictivos suscitados dentro de cada uno de los sectores de vigilancia y así poner en marcha la metodología que permita determinar dentro de un espacio-tiempo seleccionado la cantidad de hecho delictivos se espera que ocurran en una unidad temporal posterior a la contenida en la base de datos de referencia. Estos procesos se desarrollan de manera interna dentro del módulo y ningún usuario tiene acceso a estos. Sin embargo, el módulo cuenta con una interfaz de usuario que permite la configuración de diferentes parámetros de entrada que personaliza la forma en que se realiza la predicción. Los parámetros de entrada configurable por el usuario son los siguientes:

1. Modelos de datos. Es posible seleccionar el modelo de datos al que se va a realizar el análisis.

2. Espacio-Tiempo en donde se realizara la predicción (Sector de vigilancia y mes del año).

Los datos de salida que el módulo muestra para cada familia delictiva una vez concluida la predicción son:

3. La cantidad de delitos por cada familia delictiva se predice ocurrirán utilizando probabilidad bayesiana en el espacio-tiempo seleccionado.

4. La cantidad de delitos por cada familia delictiva se predice ocurrirán utilizando la metodología de predicción propuesta en el espacio-tiempo seleccionado.

5. El valor obtenido de cada uno de los análisis involucrados dentro de la predicción delictiva.

6. Número de delitos que ocurrieron de manera real en el espacio tiempo seleccionados. Este valor solo aparece para efectos de evaluar la eficiencia de la metodología de predicción. Sin embrago, esta información no estará disponible una vez implementada la aplicación.

7. Octantes horarios característicos de cada familia delictiva en el espacio-tiempo seleccionado.

8. Octantes horarios en donde existen actividades delictivas comunes al espacio-tiempo seleccionado.

En la figura 19 Se muestran los parámetros de entrada configurables en la interfaz y los datos de salida del módulo de predicción delictiva.

Una vez desarrollados los módulos descritos arriba, se configuro en el lado del servidor un servidor de páginas web. Este servidor permite que múltiples usuarios acecen a los módulos de manera recurrente.

El servidor de páginas web empleado fue el **servidor HTTP Apache** de código abierto para plataformas Unix (BSD, GNU/Linux, etc.), Windows, Macintosh y otras, que implementa el protocolo HTTP/1.1 y la noción de sitio virtual.

Apache presenta entre otras características altamente configurables, bases de datos de autenticación y negociado de contenido, pero fue criticado por la falta de una interfaz gráfica que ayude en su configuración.

Apache tiene amplia aceptación en la red: desde 1996, Apache, es el servidor HTTP más usado. Alcanzó su máxima cuota de mercado en 2005 siendo el servidor empleado en el 70% de los sitios web en el mundo, sin embargo ha sufrido un descenso en su cuota de mercado en los últimos años. (Estadísticas históricas y de uso diario proporcionadas por Netcraf).

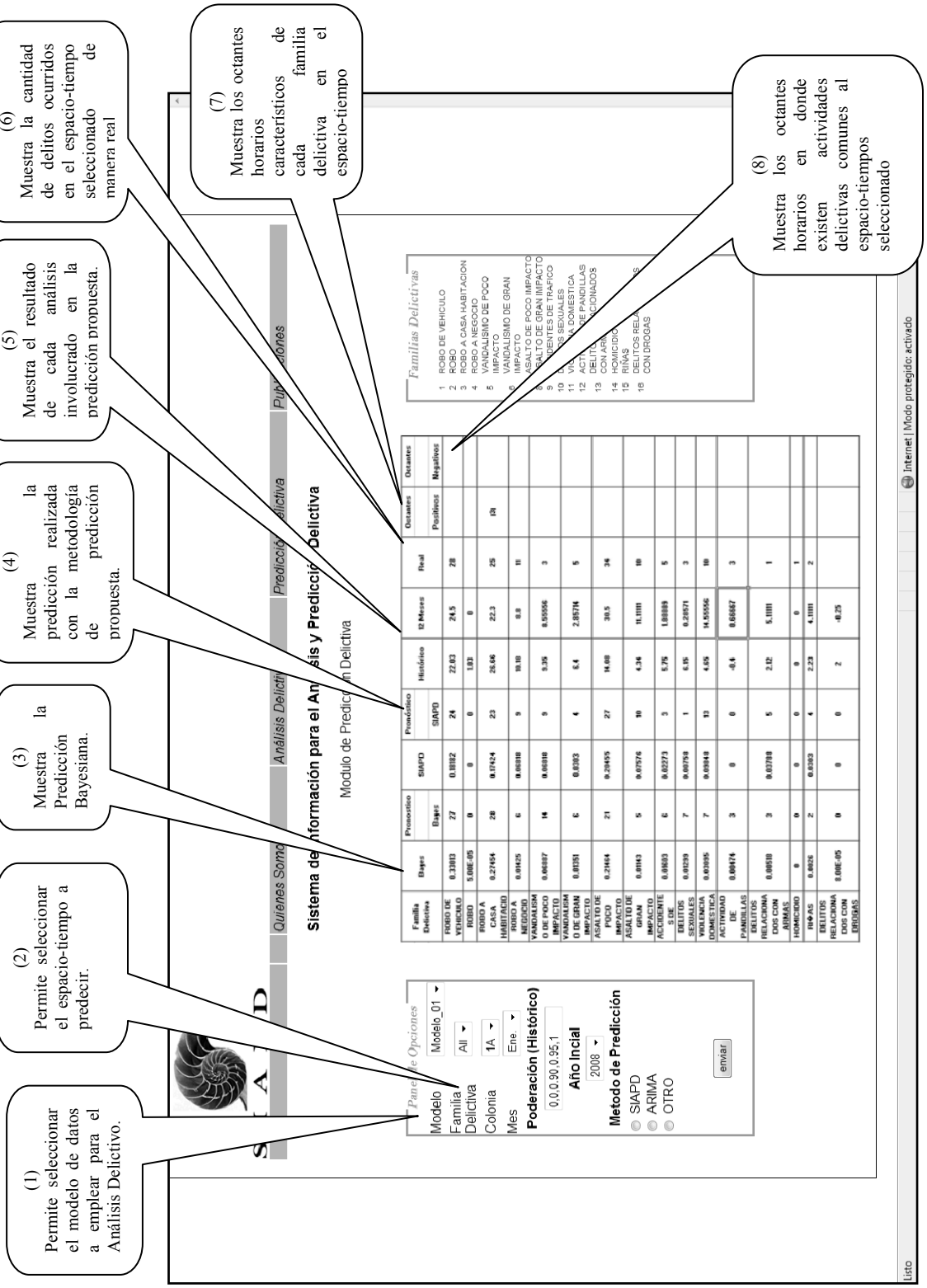

Figura 19. Parámetros de entrada y salida del módulo de predicción delictiva.

6.3.4. Requerimientos de implementación de la aplicación

El objetivo de la presente tesis no considera las cuestiones técnicas de implementación de la aplicación informática tales como: Seguridad, número de posibles usuarios recurrentes, red, etc. Sin embargo, plantea las características mínimas que deben tener de los equipos que integran los dos sub-sistemas propuestos en el presente capitulo.

Servidor

- → Conexión a Internet o en su caso una intranet de preferencia de banda ancha.
- → Un equipo de cómputo con características mínimas para soportar el manejo de Internet:
  - CPU (Procesador Pentium IV o superior)
  - Memoria RAM 1 GB
  - Disco Duro de 250 GB
  - Puertos USB
  - Monitor
  - Teclado
  - Mouse
  - Software: Windows Xp en adelante, Office 2003 en adelante.

Clientes

- → Conexión a Internet o en su caso una intranet de preferencia de banda ancha.
- → Un equipo de cómputo con características mínimas para soportar el manejo de Internet:
  - CPU Procesador Pentium IV o superior
  - Memoria RAM 256 MB
  - Disco Duro de 40 GB
  - Puertos USB
  - Mouse
  - Monitor
  - Teclado
  - Software: Windows Xp en adelante, Office 2003 en adelante.

CAPÍTULO SIETE

Conclusiones y trabajo futuro.

Se detallan las conclusiones obtenidas de los resultados observados en la aplicación de la metodología de análisis y predicción delictiva propuesta.

7.1 Conclusiones.

Difícil resulta la tarea de formular conclusiones acerca de una investigación que abarca aspectos que son estudiados desde tan diversos ángulos y perspectivas y que abre tan numerosas rutas para cuestiones de investigación y aplicación. Más aún si se reconoce que los efectos de la investigación en cuestión no resultan evidentes de manera inmediata o siquiera en el corto plazo. Es claro que el impacto de las propuestas presentadas en este trabajo sólo podrá juzgarse con el tiempo, cuando se hayan puesto a prueba en el entorno para el que fueron diseñadas. Cuando ese momento llegue, seguramente habrán todas ellas sido enriquecidas por el trabajo e investigación de diferentes miembros de la comunidad, y muchos de sus conceptos, fundamentos y especificaciones habrán sido revisados, corregidos o simplemente refinados. Por lo pronto, el único curso de acción factible es presentar las diversas reflexiones que surgen del estudio objetivo sobre cada uno de los tópicos abordados en este trabajo.

El estudio realizado sobre los patrones delictivos observados en la ciudad de Sacramento, California, en los E.U.A., y su contextualización en el campo del Análisis y Predicción Delictiva, han sido orientados, en todo momento, por el afán de aportar una guía en la toma de decisiones operativas y estratégicas sobre seguridad pública. Resulta claro que los logros de este trabajo no son, ni pretenden ser, suficientes para modificar de forma rápida y drástica la situación delictiva de nuestro país. Sin embargo, también resulta evidente que las propuestas de esta investigación pueden constituirse en importantes aportaciones que abran brecha en el largo camino de volver más eficaces y eficientes a los cuerpos de prevención delictiva. Las metodologías propuestas en este trabajo resultan logros muy modestos al ser comparados con las posibilidades de estudio e investigación que se han abierto. Ahora, para finalizar, se evaluarán las bondades y limitaciones de los conceptos planteados y se resaltarán aquellos aspectos que llaman a estudios más profundos desde nuevas perspectivas, o bien, a cambios sustanciales en la concepción y dinámica cotidiana de las fuerzas de seguridad pública.

Con base en lo resultados obtenidos de la experimentación realizada es posible llegar a las siguientes conclusiones acerca de las metodologías propuestas:

1. La metodología para el análisis y predicción delictiva posibilita la formación de una visión global de la situación delictiva. Esta investigación logró aportar una nueva forma de estudiar el comportamiento delictivo, fuera del contexto estadístico, que tradicionalmente se lleva a cabo, empleando para ello, modelos generados en un enfoque de Reconocimiento de Patrones (el enfoque Lógico-Combinatorio) que ha demostrado ser útil en el modelado de diversos problemas de tipo social para los que no se cuenta con una formalización estricta o completa.

2. Los resultados obtenidos del Análisis Delictivo propuesto permite la determinación de las familias delictivas que están ocurriendo de manera notable (característica y no-característica) en una región de estudio. Asimismo, este análisis identifica las unidades espacio-temporales donde es posible re-asignar recursos humanos y materiales para la atención (prevención o investigación) de las familias delictivas con mayor impacto en la sociedad.

3. Los experimentos realizados entregan resultados de predicción con un error mínimo comparado con los otros métodos de series de tiempo examinados. De igual forma, la metodología ha demostrado ser capaz de realizar predicciones aún cuando la cantidad de hechos delictivos ocurridos, durante un periodo de tiempo, sean escasos. De manera importante, la metodología de predicción propuesta en este trabajo presenta una alta tolerancia a la disminución en el número de datos disponibles para cada predicción, logrando obtener resultados muy cercanos a la realidad con sólo un año de la muestra de supervisión.

4. La metodología para Análisis Delictivo cuenta con una característica muy importante, permite, con base en la construcción de definiciones inductivas positivas, establecer prioridades sobre las actividades delictivas que deberán ser atendidas primero si se desea disminuir los índices delictivos de un Sector de Vigilancia. Esta característica es de suma importancia, ya que la mayoría de los planes de seguridad pública están orientados al nivel de Sector de Vigilancia. Esta jerarquización en las prioridades de atención a familias delictivas no es considerado por la mayoría de las técnicas de análisis delictivo. Estas técnicas analizan cada actividad delictiva de manera independiente, sin considerar otras familias delictivas. Por otra parte, estas técnicas tampoco consideran la división política o administrativa del espacio geográfico estudiado, situación que limita seriamente su capacidad para sugerir la asignación de tareas especificas en materia de seguridad pública a cada jefe de Sector de Vigilancia.

5. La metodología propuesta es original, desarrollada en México y aportan una visión nueva sobre análisis y predicción delictiva.

6. El trabajo conjunto con el sector académico puede cambiar significativamente el proceso de combate y prevención de la delincuencia.

7.2 Aportaciones

A continuación se muestra, en la Tabla 16, las principales modificaciones realizadas a dos algoritmos clásicos de aprendizaje empleados por las familias KORA y Conjuntos Representantes.

Tabla 16 Aportaciones realizadas a dos algoritmos clásicos de aprendizaje.

Paso	Algoritmo CR	Metodología Propuesta	Observaciones
		Se incluye el concepto de *Índice de Ocurrencia.*	Este índice permite determinar la proporción de ocurrencia, de cada propiedad, dentro de cada una de las clases en la muestra de supervisión.
Caracterización Positiva	Dos umbrales β_1 y β_1'	Un solo umbral β_1	**El Algoritmo CR** ocupa un concepto de caracterización estricta, es decir, para que una propiedad sea característica debe ocurrir β_1 veces o más en la clase caracterizada y β_1' veces en todas las demás clases. Por otra parte, la metodología propuesta en este trabajo solo requiere que el Índice de Ocurrencia de la propiedad sea mayor a β_1.
Caracterización Negativa	Dos umbrales α_2 y α_2'	Dos umbrales β_2 y β_2'	Para el Algoritmo CR α_2 es el número máximo de ocurrencia de la propiedad en una clase y α_2' es el número mínimo de ocurrencia de la propiedad en las otras clases. En la metodología propuesta β_2 es el valor máximo del Índice de Ocurrencia de una propiedad y β_2' establece el porcentaje de clases en la muestra, distintas a la que se analiza, en las cuales debe ocurrir la propiedad a fin de ser considerada negativa.
Caracterización Neutra	Un umbral γ	Dos umbrales β_3 y β_3'	Para el Algoritmo CR γ es el número mínimo de ocurrencia de la propiedad en todas las clases de la muestra de supervisión. En la metodología propuesta β_3 es el valor mínimo del cociente entre Índice de Ocurrencia de la propiedad y el máximo de dicho índice en todas las clases. β_3' es el porcentaje mínimo de clases, en las cuales debe ocurrir la propiedad a fin de ser considerada neutra.

Para el caso del Algoritmo KORA-Ω, durante la etapa de aprendizaje encuentra todos aquellos rasgos complejos característicos de cada clase, este paso es similar al que realiza el algoritmo CR durante la construcción de definiciones inductivas positivas. Es por esta razón que las modificaciones presentadas por la metodología propuesta también son

aplicables en la etapa de aprendizaje del algoritmo KORA-Ω, aunque en este algoritmo no existe la construcción de definiciones negativas ni neutras.

7.3 Trabajo futuro

El presente proyecto constituye solo un paso inicial dentro de una nueva forma de realizar Análisis y Predicción Delictiva. A pesar de ello, tiene una estructura secuencial que inicia con la determinación de aquellos espacio-tiempos que son afectados por ciertas actividades delictivas de manera notable (característica y no-característica), espacio-tiempos de donde es posible extraer recursos humanos y materiales y, finalmente aquellos espacio-tiempos donde se presentan la mayoría de las familias delictivas. Sin embargo, las recomendaciones de cómo se deben de distribuir los recursos policiacos a fin de reducir el índice delictivo es los espacio-tiempos señalados no es un problema sencillo, es necesario considerar aspectos relacionados con los estados de fuerza policiales, los cuales nos indican la cantidad de recursos de que se disponen (número de policías, armamento, patrullas, etc.) para hacerle frente a la delincuencia.

Tomando en cuenta lo anterior, se propone como trabajo futuro lo siguiente:

- Desarrollo de una metodología que con base en los resultados obtenidos del Análisis y Predicción Delictiva, haga una asignación optima de los recursos humanos y materiales de que disponga un cuerpo policiaco, con el objetivo de cubrir los espacio-tiempos de mayor incidencia delictiva.

- Esta metodología deberá además de considerar aspectos relacionados con la planeación de rutas de vigilancia que permitan que todos los espacio-tiempos determinados en el Análisis Delictivo sean cubiertos por patrullas con el mínimo de recursos y tiempo.

Referencias bibliográficas

[Arango, 2003]	Arturo Arango Durán ***Información confiable: Los problemas de la información estadística*** Instituto Nacional de Ciencias Penales, 2003.
[Arango-Lara, 2006]	Arturo Arango Durán y Cristina Lara Medina ***La Eficiencia del Sistema de Justicia*** Investigadores del Instituto Ciudadano de Estudios Sobre la Inseguridad A. C. ICESI, 2004.
[Arango-Lara, 2008(2)]	Arturo Arango Durán y Cristina Lara Medina ***Respuesta Gubernamental Frente al Delito*** Investigadores del Instituto Ciudadano de Estudios Sobre la Inseguridad A. C. ICESI. Publicado en el libro: Schutte Fernando (Coord), Una propuesta ciudadana en materia de Seguridad Pública, 2006.
[Bishop, 2006]	Christopher M. Bishop ***Pattern Recognition and Machine Learning*** Springer, 2006.
[Bongard, 66]	Bongard M. N. ***Pattern Recognition*** Spartan Books, New York-Washington, USA, 1966.
[Cheremesina-Shulcloper, 92]	Cheremesina N. E., Ruiz Shulcloper J ***Cuestiones Metodológicas de la Aplicación de Modelos Matemáticos de Reconocimiento de Patrones en Zonas del Conocimiento poco Formalizadas*** Revista Ciencias Matemáticas, XIII(2), La Habana, CUBA, 1992
[Clarke-Eck, 2002]	Ronald V. Clarke & John E. Eck, 2002. ***Análisis Delictivo para la Resolución de Problemas en 60 Pequeños Pasos*** Office of Community Oriented Policing Services, U.S. Department of Justice Traducción al español por: Arturo Arango Durán y Juan Pablo Arango Orozco.
[Chatfiel, 00]	Chatfield Chris ***Time series Forecasting*** Chapman & Hall/CRC, 2000
[Coparmex, 2002]	Centro de Estudios Económicos del Sector Privado ***La Inseguridad Pública en México*** Coparmex, 2002.
[De la vega-Shulcloper, 98]	De la Vega L., Carrasco Ochoa, Riuz-Shulcloper J. ***Fuzzy Kora-Ω Algorithm.*** 6th European Congress on Intelligent Techniques and Soft Computing, Pág. 90, 1190-1194. Vol. Vol. 2, 1998.
[Devroye-Lugosi, 96]	Devroye L., Gyöffi L., Lugosi, G. ***A Probabilistic Theory of Pattern Recognition*** Springer-Verlag, New York, USA, 1996.
[Dimitriev-Zhuravlev-Kendrelev, 66]	Dimitriev N.A., Y.I. Zhuravlev, F.P. Kendrelev ***Acerca de los Principios Matemáticos de la Clasificación de Objetos y Fenómenos*** Sbornik Diskretnii Análisis, 7, Novosibirsk, RUSIA, 1996
[Duda-Hart, 73]	Duda O.R., Hart E., Peter, Stork, G. Davis ***Pattern Classification and Scene Analysis*** Hohn Wiley and Sons, USA, 1973

[Fu, 82]	Fu K. S. ***Syntactic Pattern Recognition and Applications*** Prentice Hall, Englewood Cliffs, USA, 1972
[Fukugama, 72]	Fukugama Keinosuke ***Introduction to Statistical Pattern Recognition*** Academic Press, Inc. New York, USA, 1972
[Goldfarb, 85]	Goldfarb Lev ***A New Approach in Pattern Recognition*** Ed. L. Kanal, A. Rosenfeld. Vol. II, 1985
[González-Thomason, 78]	González C. Rafael, Thomason G. Michael. ***Syntactic Pattern Recognition: An Introduction*** Addison-Wesley, 1978
[Gottlieb-Arenberg-Singh, 1994]	Gottlieb, S.L. Arenberg, S. & Singh, R. ***Crime Analysis From First Report to Final Arrest.*** Alpha Publishing, Montclair, CA. USA, 1994.
[Haykin, 99]	Haykin Simons ***Neural Networks: A Comprehensive Foundation.*** Prentice Hall, 1999
[Hanke, 08]	Hanke E. John ***Pronóstico en los negocios*** Pearson Education, 6/a Edición, 2008
[Haag- Cummings- Phillips, 2000]	Haag Stephen, Cummings Maeve, Phillips Amy ***Management Information Systems: For The Information Age.*** McGraw-Hill, 2000.
[I.C.E.S.I.,2008]	Instituto Ciudadano de estudios sobre inseguridad pública A.C. ***Generación de información en seguridad pública.*** 2008.
[Janacek, 2001]	Janacek Gareth ***Practical Time Series*** Arnold Text in Statistics, 2001
[Lind, Marchal, Wathen, 08]	Lind Douglas, Marchal William, Wathen Samuel ***Estadística aplicada a los negocios y a la economía*** McGrawHill , México, 2008
[López Reyes, 81]	López-Reyes N. et al. ***Un Sistema Para el Pronóstico a Corto Plazo de Tormentas Ionosféricas*** Reporte de invstigación ICIMAF-ACC, 76, La Habana , CUBA, 1981
[Manzanera, 1993]	Rodríguez Manzanera, Luis. ***Criminología.*** 8ed. Porrúa, México, 1993.
[Marques de Sá, 01]	Marques de Sá J.P. ***Pattern Recognition. Concepts, Methods and application*** Springer-Verlag, New York, USA, 2001.
[Martínez-Trinidad,Guzmán-Arenas, 01]	Martínez Trinidad J.F., Guzmán Arenas A. ***The Logical Combinatorial Approach to Pattern Recognition, an overview*** Pattern Recognition, Vol. 34, 2001
[Mena, 2003]	Jesús Mena ***Investigative Data Mining for Security and Criminal Detection*** Butterworth Heinemann, Elsevier, 2003.
[Montiel, 1997]	Montiel Sosa ***Criminalística Tomo I*** Limusa Noriega Editores, México, 1997
[Moreno, Armengol, Béjar, 94]	Moreno A., Armengol E, Béjar Javier ***Aprendizaje Automático*** Ediciones Ups, 1994

[Osborne-Wernicke , 2006]	Osborne Deborah, Wernicke Susan. ***Introduction to Criminal Analysis*** 2ed. Routledge, E.U.A, 2006.
[Pons-Shulcloper-Martínez Trinidad, 02]	Pons- Porrate A., Ruiz Shulcloper J., Martínez Trinidad J. ***A New Conceptual Clustering Algorithm For Mixed Incomplete data sets*** Journal of Mathematical and Computer Modeling , 2002
[Ramírez, Chimal, 07]	Ramírez-Amaro K., Chimal-Eguía J.C. ***On a New Approach to to Times Series Tracking.*** Advance in computer Science and Engineering, eds. Torres S., López I. and Calvo H., Research in Computer Science 27, pp 205-215, México 2007
[Ronda, 1996]	Carlos C. Ronda **Métodos científicos de investigación aplicados al diseño de planes de seguridad** Subcomisario de Policía del Neuquén, Argentina, 1996.
[Shulcloper-Abidi, 02]	Ruiz Shulcloper J., Abidi Mongi A. ***Logical Combinatorial Pattern Recognition: A Review*** Recent Research Developments in Pattern Recognition, Pub. Transword Research Networks, vol.3 pp 133-176, USA, 2002
[Shulcloper, Guzmán-Arenas, Martínez Trinidad, 99]	Ruiz Shulcloper J., Guzmán-Arenas Adolfo, Martínez-Trinidad J.F. ***Enfoque Lógico Combinatorio al Reconocimiento de Patrones I: Selección de Variables y Clasificación Supervisada*** IPN. Colección de Ciencias de la Computación, 1999
[Shulcloper-Lazo, 90]	Ruiz Shulcloper J., Lazo Cortés Manuel ***Modelos Matemáticos para el Reconocimiento de Patrones*** Ed. UCLV, Cuba, 1990
[Shulcloper-Sánchez Díaz-Abidi, 02]	Ruiz Shulcloper J., Sánchez-Díaz G., Abidi M.A. ***Clustering in Mixed Incomplete Data*** Heuristic & Optimization for Knowledge, Discovery, Idea Group Publishing USA, 2002
[Turban, 1995]	Turban Efaim ***Decision support and expert systems: management support systems.*** Englewood Cliffs, N.J., Prentice Hall., 1995
[Verde, 1993]	Soria Verde, Miguel Ángel ***La Víctima: entre la justicia y la delincuencia. Aspectos psicológicos, sociales y jurídicos de la victimización.*** Promociones y Publicaciones Universitarias, Barcelona, 1993.

Páginas de Internet

1. http://www.criminalistica.com.mx/index.php/criminologia