

18
2 es.



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

CAMPUS ACATLAN

**"PLANEACION Y DISEÑO DE UN SISTEMA DE
ESTIMACION INMEDIATAS DE TENDENCIAS DE OPINION
EN CONTEOS MASIVOS"**
(Aplicación a Procesos Electorales)



TESIS CONJUNTA

QUE PARA OBTENER EL TÍTULO

LICENCIADO EN MATEMATICAS APLICADAS
Y COMPUTACION

PRESENTAN:

MIGUEL ANGEL HERNANDEZ SANDOVAL
YAYOI ROSA MA. MEGURO YUNO



ACATLAN, EDO. DE MEX.

JULIO DE 1998

**TESIS CON
FALLA DE ORIGEN**

264470



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*La observancia de los deberes convierte el mundo
en el paraíso, y la insistencia de los derechos, en el
infierno. La libertad es la hija del sacrificio.*

Benito E. Meguro.

*El hombre que cree en sus sueños
corre el riesgo de que estos se
conviertan en realidad.....*

Waltner Elias Disney

*Propónte hacer cada día algo que no desees hacer.
Tal es la regla de oro para adquirir el hábito de cumplir
con tu deber sin que te pese.*

Mark Twain.

*Ten cuidado de las cosas de la tierra:
haz algo, corta leña, labra la tierra,
planta nopales, planta magüeyes;
tendrás qué beber, qué comer, qué vestir,
con eso estarás en pié, serás verdadero,
con eso andarás.
Con eso se hablará de tí, se te alabará,
con eso te darás a conocer.*

Huchuatolli.

A mi mamá.

Orvelina Sandoval de Hernández

Porque gracias a tu coraje, perseverancia
y amor, logramos esto juntos.

A mi papá.

Emilio Hernández Guerra.

Porque siempre confiaste en mí,
y me has brindado incondicionalmente
tu apoyo, consejo y ejemplo.

A mi esposa y a mi hijo.

Yayoi y Yoneki

Por no desistir, por servir de inspiración,
motivación, y por su gran apoyo.

A mis hermanos.

Mercedalia, Emilio, Patricia y Hugo.

Porque también ustedes pusieron su granito
de arena, y principalmente por compartir el
orgullo de ser hermanos y amigos.

**A la memoria de quienes ya no están con
nosotros.**

Antonia, Delia, Josefina, Roberto, Emilio.

Porque en donde quiera que se encuentren,
siempre estarán con nosotros.

A mi familia adoptiva.

*Cristóbal, Kaoru,
Harumi, Hiroshi, Eiji y Kiyoshi.*

Porque únicamente ustedes me han aceptado
de esta forma, y por honrarme con la oportunidad
de ser parte de la familia Meguro.

A todos mis amigos y seres queridos.

Por enriquecer mi vida, y llenarla de valores.

Muy especialmente a una gran amiga.

Profa. Beatriz Clavel.

Por su paciencia, sacrificio y orientación, y porque
en estos días es muy difícil encontrar a alguien con
la camiseta tan bien puesta.

. A mi "Alma Mater", la U.N.A.M. y a todos mis maestros....

Agradezco a Dios.

Por brindarme la vida,
y darme a los padres,
hermanos, esposo, hijo,
familia y amigos que tengo.

A mis padres.

Cristóbal Meguro y Rosa Ma. Yuno.

Por su amor, cariño, comprensión y educación.
Sin ustedes no lo hubiera logrado,
sobre todo en estos últimos meses.

A mi esposo. Miguel.

Has hecho de este mundo
un lugar maravilloso,
sin ti, no hubiera logrado
llegar a donde estamos.

A mi hijo. Yoneki.

Me has hecho "super feliz"
y me has brindado las fuerzas
para el último empujón.

A mis hermanos.

Harumi, Hiroshi, Eiji y Kiyoshi.

No saben cuánto los extraño
y les agradezco todo el cariño y apoyo
que nos han brindado a Miguel y a mí,
son los mejores hermanos que se pude tener.

A mis abuelitos.

Los extraño mucho,
me gustaría que estuvieran conmigo,
y que conocieran a mi nueva familia.
Ustedes me han dado un gran ejemplo.

A mi segunda familia.

*Emilio Hernández, Orvelina Sandoval,
Mercedalia, Emilio, Patricia y Hugo.*

Les agradezco toda su aceptación y cariño
que me han demostrado.
Los tengo en un lugar muy especial
dentro de mi corazón.

A toda mi familia.

Poca gente puede decir con orgullo
que tiene una familia
tan maravillosa como la mía,
que cuenta con sus tíos, primos y sobrinos
para las cosas buenas y malas de la vida.

A mis amigos.

Soy una persona afortunada
pues tengo unos amigos estupendos,
no los cambio por nada en el mundo.

A mi maravillosa maestra. Beatriz Clavel.

Sé que un profesor da clases
en nivel superior,
sé que un maestro enseña
no sólo las materias de la escuela
sino educa, en otros aspectos de la vida.

Usted me ha asombrado
y me ha inspirado
poca gente tan activa
y tan linda como usted.
La queremos mucho

A la UNAM y a MAC.

Qué feliz fui allí.
Les agradezco mi carrera,
mis amigos y mi esposo.

INDICE

| | pág. |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| INTRODUCCION | 1 |
| I) El Sistema de Estimación Inmediata de Tendencias SEIT): Definición, composición y posibles aplicaciones. | |
| 1.1. Antecedentes y Aplicaciones de los Sistemas de Estimación Inmediata de Tendencias..... | 6 |
| 1.2. Herramientas teóricas básicas para el desarrollo y el sustento del SEIT..... | 12 |
| 1.3. El SEIT visto desde la perspectiva de la Teoría de sistemas..... | 17 |
| 1.4. La relevancia de la muestra y el manejo eficiente de los datos..... | 28 |
| II) El diseño de la Muestra como un elemento esencial de la confiabilidad del SEIT. | |
| 2.1. Las Técnicas de Muestreo y su Contribución al Diseño e Instrumentación del Sistema..... | 34 |
| 2.2. La estimación del tamaño de la muestra..... | 39 |
| 2.3. Definición del Marco Muestral y de la Población Objetivo..... | 47 |
| 2.4. Algunas observaciones para el diseño de la muestra, propiedades de las estimaciones y algunos aspectos del muestreo estratificado para proporciones..... | 51 |
| 2.5. Algunos aspectos que influyen en la precisión de la estimación y el error cuadrático medio como una herramienta apropiada para calcular la estimación..... | 58 |

| | |
|-------------------------------------------------------------------------|-----|
| V) Aplicación, evaluación y posibilidades de evolución del SEIT. | |
| 5.1. Recopilación de los resultados de las casillas en la muestra..... | 168 |
| 5.2. Organización y procesamiento de los datos obtenidos..... | 170 |
| 5.3. Presentación de los resultados..... | 176 |
| 5.4. Comparación de los resultados obtenidos..... | 184 |
| 5.5. Perspectivas y posibilidades de desarrollo del SEIT..... | 193 |
| | |
| CONCLUSIONES Y RECOMENDACIONES..... | 197 |
| | |
| ANEXO I | |
| | |
| ANEXO II | |
| | |
| APENDICE | |
| | |
| BIBLIOGRAFIA | |
| | |
| HEMEROGRAFIA | |

INTRODUCCION

En la época actual se han desarrollado una infinidad de teorías y procedimientos que por complejos y laboriosos son poco atractivos para muchas personas. El mejor ejemplo de esto son las matemáticas, ya que espantan de entrada a cualquier estudioso que desconozca la nomenclatura y significado de los modelos matemáticos, los cuales a veces son más simples de lo que se pudiera pensar. Es irónico, pero esos números y signos complicados rigen gran parte de la vida del ser humano, ya sea individual o colectivamente. Resulta difícil de creer que con una proporción mínima de una población se puedan predecir comportamientos, padecimientos y preferencias de aquella con un fundamento científico.

La Estadística Inferencial es una de las ramas de las matemáticas más interesantes y valiosas, también aplicable por difícil que parezca, que proporciona las herramientas para poder medir con la precisión deseada, debido a este tipo de características de la población en estudio, con sólo una pequeña parte de ella. Para poder desarrollar un proceso como éste es necesaria la fundamentación que proporciona la Teoría de Muestreo, ya que da las técnicas para escoger esa fracción de la población que va a hablar de las características que se buscan.

Por otro lado la tecnología está llegando a límites realmente intolerables, ¿cómo es posible entablar una conversación en segundos con una persona que se encuentre al otro lado del mundo?, ¿que una computadora pueda realizar varios millones de operaciones en fracciones de segundo, y que se esté viendo la forma de hacerlas más rápidas?, hoy día ya es posible ver una película o escuchar música en una computadora, conectar ésta a una línea telefónica y entablar comunicación con otra computadora que esté en posibilidad de recibir información, ésta bien pudiera ser una computadora portátil (LAP-TOP ó NOTEBOOK) que se encuentre en la calle en una ciudad a 3,000 kms. de

distancia y que se comunique con la primera por medio de un teléfono celular, a un costo realmente accesible.

Como se puede ver, la ciencia y la tecnología son un par de "herramientas" que ha creado el hombre y que realmente no ha faltado el ingenio suficiente para hacerlas producir en beneficios.

Por otra parte, el hombre es un ser social por naturaleza, y como tal ha tenido que desarrollar reglas para poder convivir en sociedad, aun así, resulta complicado poder organizar a una sociedad, ya que si el hombre es un ente social, también es racional y crítico, por lo cual es difícil que 90 millones de mexicanos se pongan de acuerdo, motivo por el cual debe existir algo o alguien que diga lo que hay que hacer y cómo hacerlo. Esto da paso a que se forme otro "ente" que rebasa los gustos y preferencias del hombre e impone reglas de convivencia, este "ente" es el Gobierno. En México se cambia de gobiernos constantemente, ya sea a nivel Presidente, Gobernadores, Presidentes Municipales, Senadores, Diputados, etc. Estos son elegidos en su mayoría por elecciones populares, es decir, la gente indica quién quiere que los gobierne mediante un proceso electoral que a veces tarda hasta dos meses después de las votaciones para dar a conocer los resultados oficiales, lo cual llega a causar desconfianza en la legitimidad de los comicios.

A todo esto no podría faltar la creatividad del hombre, ya que si se cuenta con una ciencia que es capaz de medir las preferencias de una población y con la tecnología capaz de manejar y enviar información a cualquier parte del mundo de una manera fácil y rápida, ¿por qué no ver la forma de aplicar todos estos conocimientos para influir en la credibilidad de la población y hacer más fácil el proceso de Gobierno?. Una metodología que ha venido tomando fuerza en México últimamente a este respecto, es los Conteos Rápidos.

El hecho de poder estimar las preferencias y proporciones de un sufragio momentos después del término de la Jornada Electoral, requiere de un operativo

muy complicado a nivel científico, técnico y organizacional, debido a la gran diversidad de tareas que hay que cumplir, es necesario dividir el sistema y delimitar responsabilidades, para ello se pueden crear varias áreas de apoyo, dichas áreas llegan a ser elementos del sistema o incluso subsistemas; un área de planeación, una de cómputo, una encargada del operativo en campo, un administrativo, etc., en esta tesis se tiene la intención de apoyarse en diferentes materias para sustentar teóricamente los planteamientos propuestos.

En el **Capítulo I** se definen algunos aspectos básicos del Sistema de Estimación Inmediata de Tendencias (SEIT), en forma muy breve se repasará su historia y se conocerán sus orígenes, se verá cuán relevante ha sido su aplicación en México, además de analizar un poco en torno a su uso y difusión. Se intentará orientar el operativo del conteo rápido desde el punto de vista de la teoría de sistemas, para poder delimitar sus elementos y analizarlos por separado. Como es lógico en esta tesis se desarrollarán las aportaciones a las áreas de cómputo y planeación.

El **Capítulo II** se dirigirá a la parte teórica para el diseño de la muestra, ya que la credibilidad de todo el trabajo posterior depende de cómo se haya seleccionado ésta. Para ello se verá cómo se calcula el tamaño de la muestra y se tendrán que analizar las diferentes técnicas de muestreo que puedan ayudar a la realización del sistema. Se comenzará a desarrollar la validación de este sistema diseñando una muestra de secciones dentro del país. Para poder realizarlo se tendrá que revisar el sistema electoral mexicano para poder plantear aspectos como son: tamaño de muestra, marco muestral, etc. Para concluir con este capítulo se verán ciertos aspectos de la realidad que llegan a chocar con la teoría y que sin embargo se deben tomar en cuenta.

En el **Capítulo III** se analizará la parte técnica, el hardware y el software del sistema. Se realizará una breve revisión de las diferentes estructuras computacionales y los distintos paquetes y lenguajes de cómputo que puedan llegar a formar parte de la estructura computacional que servirá para alcanzar los

objetivos planteados.

Se propondrán varias formas de manejar la información a nivel de software y hardware para poder evaluar cada una de las posibilidades y elegir la más viable. En este capítulo se elegirá la estructura que se considere más apropiada para continuar con el diseño del SEIT.

En el Capítulo IV se analizará la información disponible para poder diseñar la muestra, se describirá la metodología del diseño de ésta y se harán algunas observaciones al respecto. Se propondrán diferentes opciones para el manejo de la información mediante varios planteamientos, en los cuales se revisará el flujo de la información dentro de los límites del sistema. Se desarrollará de forma básica y global un modelo funcional para poder delimitar algunas funciones de los diferentes subsistemas que componen el SEIT y se mencionarán algunas medidas de seguridad que en un momento determinado pueden ser vitales para el buen funcionamiento del sistema.

Por último, en el Capítulo V se llevarán a la práctica los métodos antes sugeridos, para esto se espera ya contar con la muestra seleccionada, una vez cumplido este requisito se formará una base de datos para efectos de manipular la información las veces necesarias. Se espera tener más de una muestra diseñada para poder comparar los datos entre muestras y con los resultados oficiales dados a conocer por el I.F.E. además de los resultados obtenidos por los diferentes sistemas de conteo rápido llevados a cabo el día de las elecciones. Para concluir este capítulo se diseñará una presentación vía computadora de los resultados obtenidos, y se especulará un poco acerca de las posibilidades de evolución del SEIT.

CAPITULO I
EL SISTEMA DE ESTIMACION INMEDIATA DE TENDENCIAS (SEIT):
(Estimación, composición y posibles aplicaciones)

1.1. Antecedentes y Aplicaciones de los Sistemas de Estimación Inmediata de Tendencias.

En la actualidad se vive en un mundo en el que se toman un sinnúmero de decisiones, las cuales siempre llevan una probabilidad de no ser siempre las mejores, las personas que se ven en esta necesidad siempre buscan puntos de apoyo para determinar el camino a seguir, debido a esto, se han desarrollado una gran cantidad de modelos como apoyo, los que varían dependiendo de las necesidades de cada persona. Las decisiones son importantes en medida que afecten a más personas o que los intereses sean demasiado elevados.

Una de las preocupaciones más grandes en estos momentos a cualquier nivel es fomentar el buen gobierno, ya sea de una empresa, una escuela, y hasta de una nación. Para esto el ingenio del ser humano es ilimitado, y siempre se emplean las herramientas con las que se cuenta para hacer más fácil el trabajo. Las encuestas de opinión pública son una alternativa válida que se ha venido desarrollando últimamente, y que inclusive se han creado metodologías a la medida para apoyar procesos tan complicados como son los electorales.

Algunas de las herramientas para detectar malestares sociales, opiniones, gustos, etc., son los sondeos o encuestas de opinión pública. Dentro de este tipo de herramientas, existen algunas especialmente creadas para apoyar los procesos electorales como son las Encuestas Pos-Voto o Exit-Pool, y los Conteos Rápidos. En la Encuesta Pos-Voto se realiza la entrevista al votante cuando sale de la casilla, esto dentro de una metodología previamente desarrollada, en el caso de los Conteos Rápidos se recogen los resultados de una casilla que previamente fue seleccionada y forma parte de una muestra representativa del total de casillas. Cabe mencionar que un Conteo Rápido es un Sistema de Estimación Inmediata de Tendencias, aplicado a un proceso electoral.

Se empezará por definir lo básico, ¿Qué es un Sistema de Estimación Inmediata de Tendencias (SEIT)? Primeramente, es un sistema, es decir, es un conjunto de elemen-

tos que están interactuando entre sí buscando un fin común; ¿cuál es el "fin común" que busca este sistema?, la "estimación inmediata de tendencias", desarticulando esta frase, el término más relevante es el de la estimación, y esto lleva a abordar directamente la Estadística Inferencial, ya que cuando se realiza una inferencia estadística se está tomando como base las relaciones que existen entre una población y sus muestras para sacar conclusiones acerca de la población. Es decir, con un grupo de la población seleccionado por alguna metodología, su puede conocer cuáles son los gustos, las opiniones, molestias, etc., de toda la población. Siguiendo con el análisis de la frase, se encuentra que debe generarse la estimación en forma inmediata, lo que quiere decir, que los resultados deben presentarse en el menor tiempo posible.¹

En la época actual la forma más conocida de aplicación de un SEIT son los Conteos Rápidos y hasta este momento es la mejor aplicación que se ha dado a esta clase de sistemas, debido a que en los procesos electorales existen demasiados factores de tipo político que se ven beneficiados por los resultados que los Conteos arrojan, además de que el hecho de imaginar tener los resultados de un escrutinio en un lapso muy pequeño de tiempo al final de la jornada electoral resulta de interés para muchos sectores de las poblaciones implicadas, aparte de apoyar y afianzar los resultados finales de las elecciones, entre otros muchos aspectos.

El doctor Jorge Carpizo, siendo Secretario de Gobernación y Presidente del Consejo Técnico del IFE define un Conteo Rápido de la siguiente forma: *"El Conteo Rápido se basa en los resultados de la casilla, en resultados reales, en las cartulinas con los resultados de la votación que se fijan en las afueras de cada casilla. Se llama rápido, porque se basa en una muestra representativa de las casillas"*.²

¹ Cuando se habla de menor tiempo posible, es referente a los factores que influyen para la presentación de dicha información (tipo legal, técnico, estratégico, etc.).

² Carpizo, Jorge, *"Algunas consideraciones sobre el Proceso Federal Electoral"*, Versión Estenográfica del Mensaje pronunciado el 12 de septiembre de 1994, Secretaría de Gobernación, 1994, pág. 5.

Esto a simple vista realmente no tiene razón de ser, quizás elaborar un sistema para diseñar una muestra de una determinada población y realizar una inferencia estadística de forma rápida no sea razón suficiente para elaborar una investigación de este tipo. Si la población en estudio se encuentra delimitada geográficamente en un espacio reducido y es necesario realizar un operativo en campo, y si además los resultados pueden esperar al día siguiente para ser presentados, dados a conocer o lo que se requiera, realmente no tiene razón de ser un SEIT. Pero bajo el mismo esquema, cambiando algunas condiciones de trabajo, es decir, realizar una inferencia estadística de una población que geográficamente está demasiado distribuida y físicamente no es posible reunir el producto del trabajo en campo en un punto en común de forma casi inmediata, además de que se esperan los resultados casi instantáneamente, entonces es necesario crear alguna metodología para que este operativo resulte con éxito.

Ejemplificando lo anterior y ubicando la aplicación básica de esta tesis, que como se planteó son los "Procesos Electorales", suponiendo que en una escuela primaria se va a elegir a la reina de la primavera, la escuela tiene una población de 200 alumnos con capacidad de voto, por lo tanto el director estima que con 4 casillas de votación es más que suficiente para que todos los alumnos pasen a sufragar. Se plantea un horario de votación de las 10:00 am. a las 12:00 pm., y al final de la jornada, los encargados de las urnas realizan un conteo que, si se supone que la cantidad de votos se distribuye uniformemente, se tiene un promedio de 50 papeletas por casilla, quizá contando varias veces dichas papeletas ellos tarden 1 hora en este procedimiento, al final se reúnen dichos representantes y hacen una serie de cálculos que no les pueden llevar más de 30 minutos y en ese momento ellos ya conocen a la ganadora del sufragio, es más, ellos pueden dar los resultados finales sin tener presiones con respecto al tiempo.

Ahora poniendo el caso de una escuela pero a otro nivel, quizá una Universidad como la UACH (Universidad Autónoma de Chapingo). Esta Institución cuenta con una población de aproximadamente 8500 estudiantes que se distribuyen a lo largo de toda

la república³, y se disponen a elegir a su comité estudiantil, para dicho sufragio se requieren al menos de 110 a 130 urnas debido a la cantidad de estudiantes y a la distribución de sus instalaciones, como los votantes pueden tener diferentes turnos escolares el horario de votación es de 10:00 am. a 18:00 pm., al final de la jornada cada persona responsable de las urnas empaqueta las boletas y al día siguiente las manda a la UACH en el Edo. de México, bajo este esquema quizá después de 3 días estén reunidas todas las boletas del sufragio y después de 5 días ya se haya realizado su conteo y dado a conocer los resultados, esto es bajo un planteamiento optimista.

Suponiendo que por motivos de credibilidad la comunidad estudiantil de esta Institución desea conocer los resultados de las votaciones al menos 2 horas después de que la jornada electoral haya sido cerrada, ese es un problema. ¿Por qué?, porque para empezar no es lo mismo realizar el conteo de 200 papeletas que 8500, ni reunir los resultados de 4 casillas que de 110 casillas, y más aún, si se toma en cuenta que el 70% de las casillas se encuentran en instalaciones que pueden estar desde Campeche hasta Mexicali. ¿Cómo sería posible que después de dos horas de cerrada la jornada se pudieran tener al menos las tendencias de las votaciones?. Este es el problema al que se pretende proponer una solución.

A pesar de que la aplicación del SEIT ha sido algo familiar en algunos países europeos y en la Unión Americana, es una herramienta innovadora en nuestro país, ya que fue hasta las elecciones federales de 1994 en la República Mexicana, cuando con mucha novedad se vio por primera vez esta clase de sistemas, de manera admirable se vio como el día 21 de agosto a las 10 de la noche por los diferentes medios de comunicación se daban a conocer las "*tendencias de los resultados*" en un sufragio en el cual había un padrón electoral de 47'480,159 personas distribuidas a lo largo de casi 1,972,546 km², en más de 100,000 casillas de votación, y fue mayor la sorpresa cuando algunas semanas después se dieron los resultados oficiales y éstos encajaban de manera sorprendente con los datos a conocer el mismo día de las

³ La UACH tiene instalaciones en todo el país, desde Quintana Roo hasta Baja California, y tiene su sede en Texcoco Edo. de Méx.

elecciones por los Conteos Rápidos, con una anotación que decía: *"Con un nivel de confianza y un margen de error del 5%".*

Los resultados que arrojaron los diferentes Conteos fueron hasta cierto punto uniformes, como lo menciona el Dr. Jorge Carpizo. *"Lo novedoso en nuestro país es que hubo muchos Conteos Rápidos serios. Decidieron hacerlos: la Cámara Nacional de la Industria de la Radio y la Televisión, el Consejo Nacional de Asociaciones de Observación Electoral, la Organización Nacional de Observación Electoral del Magisterio, Alianza Cívica, COPARMEX, Cruzada Democrática por el Sufragio Efectivo, Presencia Ciudadana, Diario Reforma, Partido Verde Ecologista de México, Partido Revolucionario Institucional y Partido del Frente Cardenista de Reconstrucción Nacional. Sus resultados fueron muy parecidos entre sí".*⁴

Así, se puede observar que un SEIT en su expresión más conocida, un "Conteo Rápido", puede representar un arma demasiado poderosa para dar solución a problemas de tipo político, en especial en la estimación de resultados electorales, en donde quizá completando el término deberían llamarse "Sistemas de Estimación Inmediata de Tendencias de Resultados Electorales", porque eso es realmente lo que hacen.

Un SEIT puede ser empleado en diferentes clases de procesos, la aplicación que se le da a los procesos electorales es la más conocida, como ya se mencionó anteriormente, pero no es la única, puede servir como parte de una estrategia de mercadotecnia incluso a nivel mundial, hay empresas transnacionales que pelean el mercado de una forma tan reñida que podrían aplicar al SEIT para desarrollar alguna estrategia o política empresarial, como por ejemplo sin necesidad de levantar un censo conocer cuál es el porcentaje de personas que prefieren la Pepsi a la Coca, etc., o quizá planear de forma inmediata el número de técnicos capacitados de IBM para dar soporte a usuarios de su equipo en todo el mundo, etc.

⁴ Carpizo, Jorge, Op. cit., p.p. 5-6.

Hoy día existen empresas que comercializan con estudios como los sondeos, Exit-Pool, SEIT, etc. Entre las más grandes podemos mencionar a: Gallup, Elmer Roper, National Center Research of Public Opinion, National Institute of Public Opinion y recientemente en México se pueden mencionar a: el Centro de Estudios de Opinión (CEO), de la Universidad de Guadalajara, Cobarrubias y Asociados, Gaos y Ferrer, BIMSA, Mori de México, Indermel Louis Harris y GALLUP de México, entre otras.

Paul Perry ex-vicepresidente de Gallup, comenta que existe la posibilidad de que las encuestas de opinión fuesen un método empleado por los romanos y los griegos. Gallup es una empresa norteamericana que nace en el año de 1824, sin embargo, no es sino hasta 1916 de manera documentada que la revista Literary Digest comenzó a realizar "Sondeos" entre sus suscriptores, en aquella ocasión se trataban de predecir los resultados de la contienda electoral que llevaría acabo Franklin D. Roosevelt y su oponente Landon. (Este estudio lo realizaron Gallup y Roper & Crossley.)

En aquella ocasión los resultados del sondeo fueron tan parecidos a los resultados finales, que no solo confirmaron las posibilidades de cálculo iniciadas en elecciones anteriores, sino que sirvió a los investigadores para darse cuenta de que estaban en buen camino.

Cabe mencionar que la actitud estadística o empírica fue desarrollada principalmente por R. A. Fisher y R. Von Mises. La idea de espacio de las muestras se debe a Von Mises. Esta idea hizo posible construir una teoría estrictamente matemática basada en la teoría de la medida. Ese punto de vista emergió gradualmente durante los años 20's bajo la influencia de muchos autores.

A partir de ese momento fueron muchas las experiencias motivadas por este tipo de técnicas, aunque no todas fueron positivas, ya que en algunos casos se trató de manipular la opinión pública dejando la realización de los sondeos fuera de toda metodología válida, lo cual tuvo como consecuencia lógica el fracaso de dichos estudios. Sin embargo esto no ha sido motivo para que se detengan los adelantos que

hasta el día de hoy se han logrado en esta materia, sólo queda mencionar que si de por sí mismos los sondeos abren un sin fin de posibilidades, ¿hasta qué punto podrán ser éstos mejorados con las herramientas con las que contamos hoy en día?

1.2. Herramientas teóricas básicas para el desarrollo y el sustento del SEIT.

Es necesario hacer una revisión del flujo y el procesamiento de la información para saber cuáles son los fundamentos metodológicos de un SEIT, esto es muy importante pues no se puede asegurar que los resultados que arroje el sistema son confiables si de alguna forma no se justifican los procesos y se fundamentan todos los pasos que aquí se mencionan, realmente esto no es algo fácil, es más, no está al alcance de los autores realizar esta actividad, pero sí es posible apoyarse en las ciencias que de alguna forma están aportando algo al diseño del SEIT, de tal manera lo que se plantea en esta investigación está fundamentado y respaldado por disciplinas como son: Las Ciencias Sociales, la Estadística, la Ingeniería de Software, la Teoría de la Computación, las Telecomunicaciones, la Teoría de Sistemas, etc. Aquí se presenta el por qué de estas disciplinas:

Las Ciencias Sociales.

Son la contextualización del sistema, le dan una aplicación a la teoría, ya que la mayoría de los estudios en los que se realizan encuestas, y sobre todo en los que se necesitan los resultados de las mismas de forma rápida, es precisamente en este campo. Además, esta rama de la ciencia permite administrar el proyecto y capacitar al personal, aparte de realizar junto con el área de cómputo, el diseño de la presentación de los resultados, ya que esta ciencia es la que se encarga de estudiar, analizar, y dar un uso práctico a éstos.

La Ingeniería de Software.

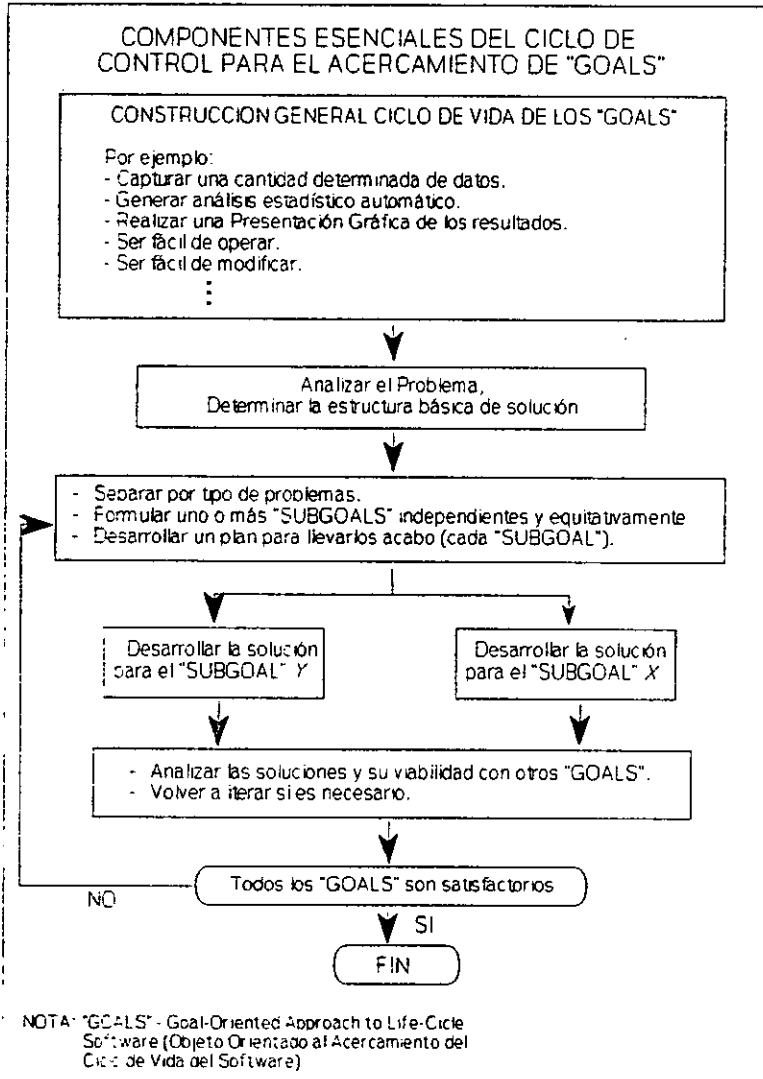
Permite, junto con la computación evaluar y verificar aspectos inmediatos del sistema, como son: el área de trabajo, los elementos necesarios y la planeación del sistema, desde el punto de vista computacional, estudiando primero las necesidades y analizando los resultados que se esperan del proyecto, además de evaluar los recursos con los que se cuenta para la realización del SEIT, esta disciplina es de tal importancia que, basándose en ella, se puede asegurar la vigencia o actualización del sistema desde el punto de vista del Software.

Esta disciplina es un punto muy delicado a tratar, empezando por la definición: *"La ingeniería de software es la aplicación de la ciencia y las matemáticas mediante la cual la capacidad del equipo de cómputo se hace útil al hombre vía programas computacionales, procedimientos; todo esto con una documentación asociada"*.⁵ Esta definición por sí sola no dice mucho. Es más, hay puntos aquí que pueden prestarse a discusión, la figura 1.2.1 puede ayudar a comprenderla, como disciplina la ingeniería de software marca ciertas reglas para desarrollar el software, inclusive técnicas que mejoran de forma sustancial el desarrollo del mismo, además da las pautas para poder actualizar el sistema cuando se considere necesario.

El término "GOALS", que se observa en la figura 1.2.1. va ligado a la ingeniería de software y significa (Goal-Oriented Approach to Life-cycle Software), Objetivo Orientado al Acercamiento del Ciclo de Vida del Software.

⁵ BEIZER, Boris, BOEHM, Barry W.. *"Software Engineering Economics"*, New Jersey, Ed. Prentice Hall, Cap.3.2, pag.16

Figura 1.2.1. Componentes esenciales del ciclo de control para el acercamiento de "Goals".



La Estadística y la Teoría de Muestreo.

La Teoría de Muestreo es un elemento vital para esta clase de estudios, ya que por las características del proyecto se explotan las cualidades y bondades que brinda dicha teoría en la selección de una subpoblación, comúnmente llamada muestra, que cuente con las características en estudio de toda la población, lo cual trae como beneficio que el costo del proyecto sea reducido y que éste se pueda llevar a cabo con la rapidez esperada. Todo esto hace posible que el estudio se lleve a cabo, es decir, hay más posibilidades de realización, sin contar que la calidad de los resultados es mayor, ya que los recursos empleados son menos, y por tanto existe mejor control sobre ellos, con esto se quiere decir que el personal que se emplee debe ser el más capacitado, además de poder someterlo a un entrenamiento intensivo, otra de las facilidades que brinda es poder llevar a cabo una supervisión cuidadosa del trabajo en campo y del procesamiento de los resultados, hay autores que aseguran que un estudio realizado con una muestra de la población es más efectivo que la realización de un censo. *"Una muestra puede producir resultados más exactos que la enumeración completa".*⁶

La Estadística apoya al proyecto en el análisis de los datos, la estimación de las tendencias, en el control de los sesgos que se puedan dar y el margen de error al que se está propenso, pues en la publicación de los resultados es un requisito dar a conocer esta información. De modo que se puede decir que esta ciencia es la que le da la validez metodológica al proyecto.

Las Ciencias Computacionales.

El SEIT no podría funcionar como tal sin la ayuda de las computadoras para la concentración de los datos y su procesamiento, ya que es la forma más rápida de presentar resultados, además de ser la más eficiente y exacta. Para hacer esto el encargado del departamento de cómputo necesita planear desde la ubicación de las

⁶ COCHRAN, William G. *"Técnicas de Muestreo"*, México, 1a. Edición, Ed. CECSA, 1980, pág.20.

computadoras, su número, forma de conexión, forma de captura de datos, el software que va a utilizar, los programas que se necesitan elaborar, la presentación de los resultados, equipo de resguardo y seguridad: no-brakes, vacunas, diskettes, software de respaldo, etc. Además de revisar la ventilación e iluminación del lugar donde se localizarán las computadoras y de dar capacitación al personal que va a hacer la captura.

Las Telecomunicaciones.

¿Se puede imaginar que cuenta con varias computadoras, que necesita que la información fluya entre ellas de una manera rápida y que esas computadoras se encuentran a una distancia de una cuadra una de otra?. Imagínese ahora que no se cuenta con redes o módems, sería realmente desastroso el hecho de tener a las computadoras conectadas entre sí por medio de cables por toda la ciudad. Se puede hacer, pero sería muy costoso y nada práctico, o también se podría intercambiar la información por medio de discos, pero hay veces que la información es tanta que se necesitaría de muchos de ellos, además de que sería bastante riesgoso pues se puede perder la información. Gracias a que se cuenta con las Telecomunicaciones, el mandar la información de una computadora a otra se hace de una manera mucho más práctica y segura.

De la enciclopedia *"Hombre, Ciencia y Tecnología"* se obtiene la siguiente definición. *"Las Telecomunicaciones son los tipos de comunicación efectuados entre una estación transmisora y otra receptora situadas a gran distancia"... "Los medios de transmisión empleados varían según la distancia entre transmisor y receptor. En las uniones por cable se utilizan líneas aéreas y cables terrestres y submarinos"*⁷. Por la naturaleza del SEIT los datos se obtienen de lugares dispersos y la concentración de éstos debe ser eficiente, esta eficiencia se obtiene en mayor grado si se cuenta con la información en algún medio magnético ya que tenemos la capacidad de comunicar dos computadoras vía módem sin importar la distancia en que éstas se encuentren, si no se cuenta

⁷ Enciclopedia *"Hombre Ciencia y Tecnología"*. Tomo B, México, Ed. Encyclopedía Británica de México, S.A de C.V., 1982.

con computadoras a grandes distancias, se pueden dictar los datos por teléfono y la persona que los recibe los puede capturar, cualquiera de estas dos formas de transmisión de datos, es por medio de vía telefónica, también se podría manejar una red, donde los datos se mandan por satélites, que también son telecomunicaciones, pero por vía aérea, las cuales también podrían servir para mandar los resultados arrojados por el sistema, al lugar donde es requerida para su observación y análisis.

Teoría de Sistemas.

Esta disciplina es quizá la más importante, ya que sirve como el pegamento para juntar todas las piezas que conforman el SEIT. Esta rama de la ciencia permitirá distinguir cuáles son las partes que conforman el Sistema y dar un acomodo a cada una de ellas. Es tan importante que se ha considerado necesario tratarlo detalladamente más adelante.

Además de las disciplinas anteriormente mencionadas, hay que tomar en cuenta la participación de otras disciplinas como pueden ser: la Administración, la Organización operativa, el Análisis social, económico y político que esto pudiese traer consigo, etc.

1.3. El SEIT visto desde la perspectiva de la Teoría de Sistemas.

El SEIT como su nombre lo dice es un sistema, es necesario hacer un especial énfasis en esto, ya que si bien es muy importante distinguir las diferentes disciplinas que aportan valiosas bases y elementos de seguridad en este "sistema" es más importante el "pegamento" que las une en una situación muy específica, sería muy poca la ayuda que pudiera aportar cada una las disciplinas mencionadas anteriormente si se trabajara

por separado, sin orden. Siempre es necesario imponer el ritmo de trabajo y trazar las directrices a seguir cuando se trata de un proyecto en el que participan más de dos personas. Para poder abordar el tema es necesario revisar algunos conceptos de dicha teoría e irlos aplicando al SEIT. Desde el punto de vista personal (de los autores de esta tesis), *la Teoría de Sistemas es la forma de conjuntar el esfuerzo de varias áreas o módulos de trabajo encaminado en un objetivo común.*

Generalmente todo lo que rodea al ser humano es un sistema, o al menos es parte de alguno, desde el lenguaje con el que se comunican millones de personas, el auto en el que se transportan, el lugar en donde laboran, la naturaleza, etc.; si se analizan bien todas estas cosas se puede observar que casi ninguna se escapa a pertenecer a alguno, ya que incluso el mismo cuerpo humano es un sistema, quizás el más perfecto de todos. Algunos sistemas son pequeños y controlables como podría ser un equipo de fútbol, la familia, etc., otros pueden crecer desmedidamente como son la política, la economía, la comunicación, etc.

El problema que se encuentra más comúnmente cuando se observa un sistema, es su complejidad, *"La complejidad es el resultado de la multiplicidad y embrollo de la interacción del hombre en los sistemas."*⁸ Una actividad común como puede ser el alimentarse, generalmente es el resultado de un largo proceso, desde la generación de esos alimentos en los campos de cultivo, su procesamiento, la transportación a los lugares de consumo, su compra-venta, etc., hasta la preparación de un guiso que se servirá a la hora de la comida. En realidad estos alimentos han sido parte de varios sistemas, han formado parte de sus insumos y también han sido productos de otros, todo esto cuando originalmente cada quien cultivaba sus propios alimentos. Lo anterior trata de ejemplificar cuan complejo se ha hecho el proceso alimenticio, que sin lugar a duda en sus orígenes era algo muy simple. El proceso de la transportación, es otro ejemplo de la complejidad moderna, inicialmente la gente se transportaba por sus propios medios, caminando, hoy día se cuenta con modernos taxis espaciales que transportan al hombre al espacio y aun no satisface.

⁸ VAN Gigch, John P. Teoría General de Sistemas, México, 2a. Edición, Ed. Trillas, 1987.

Todas estas situaciones generan problemas a grandes niveles y siempre es recomendable ver el problema desde el punto de vista más amplio, lo que en Teoría de Sistemas se llama "Holístico", es decir, desde la "Totalidad del sistema". Es obvio que para poder hacerlo es necesario tener la capacidad de poder ver y analizar la situación y los problemas desde el punto de vista del TODO, y no caer en la tentación de solucionar pequeños problemas que nunca terminan y que no representan ninguna mejoría al funcionamiento general del sistema. Para esto el hombre ha creado sus propias herramientas. Se piensa que se ha hecho un buen avance mediante la introducción y adopción del Enfoque de Sistemas, que se presenta como una filosofía práctica y una metodología del cambio.

"El enfoque de sistemas puede, muy posiblemente, ser la única forma en la que podamos volver a unir las piezas de nuestro mundo fragmentado: la única manera en que podamos crear coherencia del caos".⁹

Después de esta breve contextualización parece conveniente dar alguna definición de lo que es un sistema, pues es de vital importancia ya que el proyecto propuesto es eso, un Sistema de Estimación Inmediata de Tendencias, John P. Van Gish lo define así *"Un sistema es una reunión o conjunto de elementos relacionados"*¹⁰, otra definición la da Milton Keynes, *"Es la unión de partes o componentes, conectados en una forma organizada"*¹¹, o como lo definiera G.M. Weinberg, *"Un sistema puede existir como un agregado natural de partes componentes encontradas en la naturaleza, o ésta puede ser un agregado inventado por el hombre - una forma de ver el problema que resulta de una decisión deliberada de suponer que un conjunto de elementos están relacionados y constituyen una cosa llamada 'un sistema'"*¹². Conjuntando todo lo anterior, se definirá un sistema como *"Un conjunto de elementos diferenciados y*

⁹ C.W.Churchman, I. Auerbach y Simcha Sadan, *"Thinking for Decissions Deductive Cuanitative Methods"*. Palo Alto California: Science Research Associates, 1975, pág.421.

¹⁰ VAN GIGCH, John P. *"Teoría General de Sistemas"*. México, 2a. Edición, Ed. Trillas, 1987, pág. 17.

¹¹ KEYNES, Milton. *"The Open University, Systems Performance. Human factores and systems failures"*. Inglaterra; The Open University Press, 1976, p.p. 8-11.

¹² WEINBERG, G.M. *"An Introduction to General System Thinking"*. Nueva York, Wiley, 1975, p.p.62-63.

organizados, que tienen un propósito, objetivo o fin común".

En estas definiciones se ha hablado de "elementos", "Partes o propósitos". ¿Qué son éstos?, los elementos de un sistema pueden ser conceptos, lo cual llevaría a un sistema conceptual, como ya se mencionó un lenguaje puede ser un sistema, lo que no se dijo es ¿qué es un "Sistema conceptual"?, por otra parte los elementos también pueden ser objetos como un automóvil. O pueden ser sujetos, por ejemplo un equipo de fútbol. Y por otro lado un sistema puede estar formado por diferentes tipos de elementos: sujetos, objetos y conceptos, como puede ser una ciudad; todas las combinaciones son posibles. En la mayoría de los casos se puede pensar en sistemas más grandes a los cuales se les llama "sistema total" o "sistema integral", los cuales absorben sistemas inferiores, denominados subsistemas. Uno de los problemas más comunes se deriva de la capacidad para saber qué tanto descomponer un sistema en sistemas componentes o qué tanto componer u organizar un sistema en sistemas mayores.

Cuando se habla de *diseño de sistemas* es necesario hacer referencia a un término básico, el "*Paradigma de sistemas*", este término es sinónimo de proceso de diseño de sistemas. Un paradigma es un proceso, un procedimiento (no definido necesariamente en forma de pasos secuenciales), que se puede utilizar en forma repetida para abordar un tipo específico de problema.

El enfoque de sistemas es generalmente una metodología de diseño, debido a esto cuestiona al mismo sistema y a su desempeño dentro del campo de trabajo de un sistema mayor. Cuando se va a diseñar un sistema se hace pensando en los objetivos que se desean alcanzar y cómo va a interactuar éste con los sistemas mayores, a esta visión se le llama *extrospectiva*, debido a que se emplea una visión desde dentro hacia fuera. Con esto se quiere decir que *el enfoque de sistemas procede de lo particular a lo general, e infiere el diseño del sistema a través de un proceso de inducción.*

El enfoque de sistemas es un método de investigación, una ideología que enfatiza el sistema total, en lugar de sistemas componentes. En contraste con la "Metodología de cambio" que es la que se encarga del mejoramiento de sistemas, el enfoque de sis-

temas es una metodología de diseño que se caracteriza por:

1. Para poder plantear el problema, se tiene que tomar en cuenta a los sistemas superordinales, o los sistemas a los cuales pertenece el sistema en estudio y está relacionado con ellos por los aspectos comunes en los objetivos, por ejemplificarlo así, es necesario evaluar el impacto en la sociedad, o en el sistema político mexicano en el momento que se den a conocer los resultados obtenidos por el SEIT.
2. Los objetivos del sistema deben plantearse de acuerdo a sistemas mayores o al sistema total y no a los subsistemas, por ejemplo, los objetivos que se impongan en el SEIT, se deben pensar de acuerdo a este mismo y no al área de Cómputo o al de Administración, por decir algo.
3. Al diseñar un sistema, se tienen que tomar en cuenta los costos de oportunidad, pensar en que el diseño del sistema no sea sólo para un momento, sino tratar de planearlo para un futuro y calcular qué tanto se aleja del diseño óptimo. En el caso de los Conteos Rápidos en las elecciones de 1994, su diseño no sólo funcionó para ese momento, pues se siguen utilizando actualmente, y pueden seguir siendo utilizados para elecciones venideras como ya se vio en las pasadas elecciones del 6 de julio de 1997.
4. En la mayoría de los casos, el diseño óptimo de un sistema realizado tiempo atrás, no puede seguir siéndolo actualmente, pues cada día se cuenta con nuevas alternativas que ofrecen mejores salidas para el sistema total. En el SEIT esto es muy claro, ya que la tecnología mejora cada día y la situación económica, política y social de los diferentes países cambia continuamente, creando diferentes ambientes que influyen en la realización de un proyecto como éste.
5. Para el diseño de sistemas se emplea el paradigma de sistemas que dice que se verá el problema desde el punto de vista de TODO el sistema, yendo de lo particular a lo general y resolviendo el problema globalmente, es decir con un

pensamiento inductivo y de síntesis. En el caso de querer mejorar el sistema en vez de diseñarlo, se emplea el paradigma de la ciencia, y su pensamiento es deductivo y de reducción.

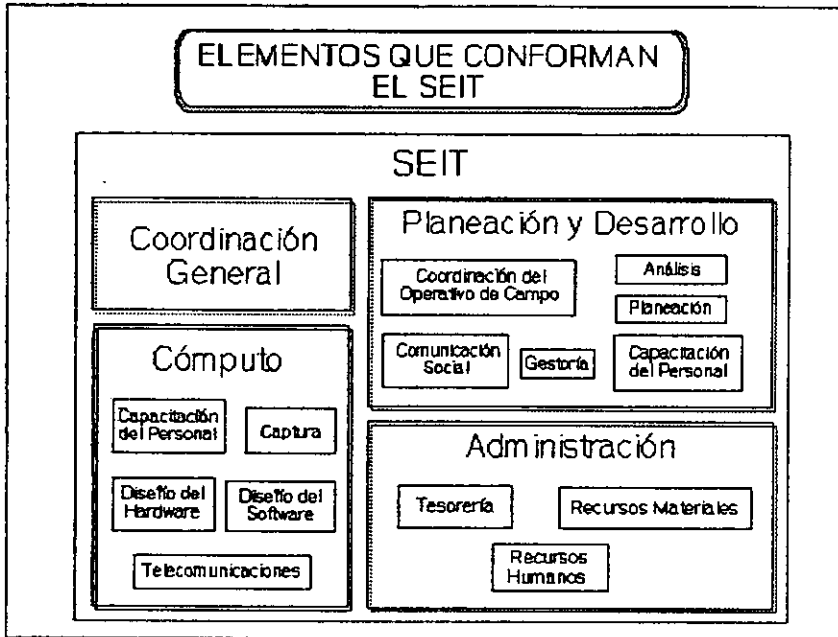
6. Para la realización del sistema, la persona que realiza el planteamiento del mismo debe ser el líder del proyecto, analizando las opciones que se tienen y tratando de ver nuevas alternativas, para la optimización del sistema, y no una persona que siga las órdenes de otra sin tener idea de lo que está haciendo o que trate de continuar con el sistema existente tal como está, sin intentar arreglar sus defectos. Para la realización del SEIT, la persona que asuma el papel de líder debe de tener idea de lo que cada área va a realizar para poder planear todo el sistema, preguntando las necesidades de cada área y tratando de resolverlas todas o por lo menos las necesidades más indispensables.

Aplicando lo anterior en el SEIT, se analizará cuáles son las partes que integran un sistema, para hacer esto se empleará un SEIT aplicado a un proceso electoral, ya que así será más fácil ubicar los componentes del Sistema y facilitará la explicación y la comprensión de cada parte.

Los Elementos.

Son los componentes de cada sistema. Cada elemento del sistema puede a su vez ser otro sistema, es decir, un subsistema, aunque también puede no serlo. Los elementos deben de carecer de un proceso decisorio que afecte significativamente al sistema. Los elementos de sistemas pueden ser vivientes o carecer de vida. La mayoría de los sistemas con los cuales se trata, tienen elementos de ambos. Los elementos que entran al sistema se llaman entradas y los que lo dejan son las salidas o resultados. En el caso del SEIT se pueden ubicar una serie de elementos como el Area Administrativa, el Area de Planeación y Desarrollo y el Area de Cómputo; éstos a su vez son subsistemas que se pueden desmenuzar obteniendo así otros elementos. Ver figura 1.3.1.

Figura 1.3.1.



Los Procesos de Conversión.

Un proceso de conversión cambia a los elementos de entrada en elementos de salida en un sistema donde los elementos están organizados, agregando, generalmente, valor o utilidad en el sistema, imponiendo costos e impedimentos. El proceso central de conversión del SEIT toma como elementos de entrada los resultados de las casillas, y mediante dicho proceso de conversión tendrá como elementos de salida las tendencias de las elecciones.

Las Entradas y los Recursos.

Por lo general se consideran como costos al momento de ser evaluada la eficacia del sistema. La diferencia entre entradas y recursos depende sólo del punto de vista. Generalmente, a los elementos de entrada se les aplican los recursos en el proceso de conversión. En el caso del SEIT, las entradas son los datos que se obtienen de las casillas y los recursos pueden ser las personas que mandan los datos y las que capturan esos datos, por mencionar algunos recursos. Al identificar las entradas y recursos de un sistema, es importante especificar si están o no bajo el control del diseñador del sistema, es decir, si pueden ser considerados como parte del sistema o parte del medio.

Las Salidas o Resultados.

Como ya se mencionó anteriormente, los resultados del proceso de conversión del sistema son las salidas, y se cuentan como resultados, éxitos o beneficios. Los resultados en el SEIT son obviamente, las tendencias estimadas para cada partido.

El Medio.

En los sistemas que interactúan con otros sistemas, es muy importante localizar los límites del sistema, esto es tal vez lo más difícil de lograr, ya que al definir los límites del sistema, se distinguen los sistemas que quedan bajo el control de quienes toman las decisiones y los sistemas que se quedan fuera de dicho control conformando el medio. No importa dónde se pongan los límites del sistema, se sigue interactuando con el medio, pues no se está aislado, a menos que las salidas del sistema no afecten al medio. El medio en el que se ubica el SEIT es un proceso electoral.

El Propósito y Función.

En el caso de sistemas inanimados, éstos no cuentan con un propósito sino hasta cuando se relacionan con otros subsistemas y crean un sistema más grande. Por lo que son muy importantes las conexiones entre subsistemas, y entre subsistemas y sistema total.

Los Atributos.

Cada sistema, subsistema o elemento, cuenta con atributos o propiedades. Estos atributos pueden ser cualitativos o cuantitativos, diferenciándose en su forma de medición. Los atributos cualitativos son más difíciles de medir que los cuantitativos, pues es más fácil una medida contable que una medida en la que tenga que ver algún punto de vista. Los atributos en el SEIT son cualitativos y cuantitativos, pues existen varios subsistemas que interactúan, los cuales están a su vez compuestos por personas, computadoras, conceptos, técnicas, etc.

Las Metas y los Objetivos.

Es muy importante para el diseño de un sistema el poder identificar cuáles son sus metas y objetivos. Cuanto mejor especificados estén los propósitos del sistema, serán mejor definidos y más operativos, y mientras más objetivos se cumplan, y lo realicen de la mejor manera, aumentarán las medidas de eficiencia del sistema. Estas medidas representan el valor de los atributos del sistema. La meta a la que se quiere llegar en el SEIT será determinar las tendencias electorales llegando a esos resultados en un mínimo de tiempo.

Los Componentes, Programas y Misiones.

Cuando se quiere que un sistema cumpla con sus objetivos, se establecen componentes, programas o misiones que forman al proceso de conversión, éste consiste de elementos compatibles para que juntos lleguen a un objetivo definido. Un aspecto muy

representativo en el enfoque de sistemas es que en la mayoría de los casos, los límites de los componentes no coinciden con los de la estructura organizacional. En el SEIT, los componentes son los equipos de trabajo, los que se dedican a la administración, los que se dedican a la captura, los que se dedican al trabajo de campo, etc.; los programas y las misiones van de acuerdo a cada equipo, pues cada uno tiene un objetivo diferente.

Administración, Agentes y Autores de Decisiones.

Tienen la responsabilidad de encaminar el sistema a la realización de sus objetivos mediante acciones y toma de decisiones. Estas personas se encuentran principalmente en sistemas con un propósito o función definido y que se esfuerzan por alcanzar esos objetivos. Los administradores del SEIT, son los líderes de cada equipo, los que están enterados de todas las tareas que su equipo debe realizar y cómo las deben de realizar, además de conocer las tareas que el resto de los equipos van a efectuar.

La Estructura.

La estructura es la forma en que se relacionan los elementos del sistema, puede ser simple o compleja, de acuerdo al número de interrelaciones entre las partes del sistema y de la manera en que estas interrelaciones se encuentren. Los sistemas complejos cuentan con jerarquías que pueden ser niveles ordenados, partes o elementos de subsistemas. La eficacia de los sistemas depende del tipo y forma de interrelaciones entre los componentes del sistema. La estructura de el SEIT se ve en la figura 1.3.1.

Los Estados y Flujos.

El estado de un sistema se define por las características de los elementos del sistema en un punto en el tiempo. Los flujos del sistema surgen con los cambios que sufre un elemento del sistema al pasar de un estado a otro. La conducta es el cambio de un estado en el sistema sobre el tiempo. Para distinguir los estados dentro del SEIT, se

tomarán a los votos que se recogen de las casillas al final de la jornada electoral como los datos de entrada, y se seguirán esos datos como los elementos que cambian de un estado a otro, así, el momento en que se introducen esos datos al sistema se cuenta como un estado, el momento en que se están procesando los datos es otro estado, el momento en que se obtienen los resultados es otro y el momento de mandar los resultados a donde serán analizados será otro estado del sistema. El flujo se verá en la forma en que los datos van cambiando de un estado a otro.

Como ya se mencionó, y como es lógico, debe de existir lo que se llamaría el cerebro de la operación, el Administrador del Sistema, este administrador es la persona que tiene el punto de vista "holístico", el que dirige. Además hay cuatro áreas importantes en la aplicación del enfoque de sistemas en organizaciones que requieren de particular atención:

1. **Definir los límites del sistema total y del medio.** El medio son todos esos sistemas sobre los cuales el que toma las decisiones no tiene control. El sistema total comprende todos esos sistemas que de alguna forma se ven afectados por el problema al que se pretende dar solución, aunque pertenezcan a diferentes organizaciones. Todo lo que no está dentro del sistema total es parte del medio.
2. **Establecer los Objetivos de los Sistemas.** El determinar cuál es el objetivo del sistema es conocer perfectamente qué es lo que se espera como salidas de éste y no aceptar alguna desviación, no se debe perder de vista que el problema de establecer un sistema total y delimitarlo correctamente está completamente relacionado con la implantación de los objetivos.
3. **Determinación de programas y relación de programas agentes.** En el momento en que los objetivos de la organización han quedado completamente claros, se puede empezar a agrupar las actividades que buscan objetivos similares o el logro de funciones relacionadas en programas o misiones.
4. **Descripción de la dirección de Sistemas.** Cuando se habla de dirección, se habla

de dirección en todo el sistema, actividades, programas, autores de decisiones, agentes involucrados en la planeación, evaluación, implantación y control del diseño del sistema, generalmente no se distingue al diseñador del director debido a que las decisiones tomadas por uno afectan directamente al otro.

Los sistemas también se dividen en sistemas abiertos y sistemas cerrados, el abierto es aquel que posee medio, es decir que existen otros sistemas con los cuales se relaciona, intercambia y comunica; y un sistema cerrado es aquel que carece de medio, es decir, no hay sistemas externos que lo alteren, se retroalimentan a sí mismos. Por lo anterior, se tiene que el SEIT, por su naturaleza, se clasifica como un sistema abierto, debido a su forma de operar y al número de sistemas que naturalmente interactúan con él.

Para terminar este subcapítulo, vale la pena señalar que al Enfoque de Sistemas puede llamársele correctamente **Teoría General de Sistemas Aplicada**.

1.4. La relevancia de la muestra y el manejo eficiente de los datos.

Es necesario resaltar cuáles son las partes más importantes del SEIT, cabe mencionar que un SEIT no es mas que una encuesta de opinión pública como de las que ya se ha hablado anteriormente, nada más que con una serie de variaciones que crean las diferencias fundamentales. En sí, el proceso es el mismo, sólo que se tienen algunos aspectos que no se pueden pasar por alto, por ejemplo, la selección adecuada de la muestra, con esto no se pretende decir que en un sondeo la muestra no sea tan relevante o que no se diseñe adecuadamente, por el contrario, sólo que existe una diferencia, en un sondeo los resultados llegan a un escritorio y si se considera por alguna razón que los resultados no son representativos de la población, se puede corregir el rumbo y/o revisar el procedimiento, ya que es posible que la técnica de muestreo no sea la adecuada o que ésta no cumpla con los requerimientos mínimos de

representatividad, lo cual no es tan grave, debido a que es posible detectar la falla y corregir el curso de la investigación, ya que regularmente el trabajo en campo de un sondeo puede durar 1 o varios días, dependiendo de los requerimientos, y de los recursos del mismo estudio.

No es el mismo caso en el SEIT, en donde el tiempo es el principal enemigo, y el operativo, por lo tanto, debe ser lo más eficiente posible ya que sólo se cuenta con algunos minutos para recolectar los datos, una vez concentrados éstos y procesados, el tiempo de revisión es mínimo y quizá en algunos casos completamente nulo. Por un lado si se dan a conocer resultados erróneos es posible imaginar todas las consecuencias que esto podría traer, desde muchas perspectivas de tipo social, económico, de confianza en el proceso estudiado, etc. Por otro lado, si es posible revisar los datos y se detecta la falla, ¿qué se podría hacer en ese momento?, es posible que se puedan hacer muchas cosas, pero el 100% de éstas requieren tiempo, si la respuesta está al alcance de la mano, esto puede no trascender de manera importante, de lo contrario se pone en riesgo el éxito del proyecto, ya que como se ha mencionado el tiempo es el principal enemigo a vencer.

Cabe señalar que la parte genial del proceso es precisamente ésta, seleccionar una muestra que sea representativa del total de las personas involucradas en el proceso, y de ahí realizar una inferencia de las preferencias de la población en estudio. Cuando se hace esto, se está empleando a fondo las facilidades y opciones que brinda la inferencia estadística.

El hecho de manejar una muestra brinda una serie de ventajas en comparación con la enumeración completa. Estas ventajas son:

- * **Costo Reducido.**- Es decir, si los datos se obtienen únicamente en una pequeña fracción del total, los gastos son menores de los que se realizarían si se llevara a cabo un censo completo. Cuando la población es muy grande los resultados son lo suficientemente exactos si se analizan muestras que representan sólo una pequeña fracción de la población. "En los Estados Unidos las investigaciones periódicas más importantes las lleva a cabo el gobierno que usa muestras de alrededor de 105 mil personas, o sea aproximadamente una persona por cada 1,240".
- * **Mayor rapidez.**- Por la misma razón los datos son recolectados y procesados con mayor facilidad y rapidez cuando se trata de una muestra y no de una enumeración completa. "Esta es una consideración vital cuando se necesita la información con urgencia".
- * **Más posibilidades.**- Si la población es muy grande, un censo completo en algunos casos puede llegar a ser impracticable, por lo tanto, siempre va a ser más factible que se realice una encuesta basada en una muestra a un censo. Para obtener la información en ciertos tipos de encuestas se utilizan los servicios de personal altamente calificado o equipo muy especializado con algunas limitantes específicas
- * **Mayor precisión.**- Debido a que al reducir el volumen de trabajo se puede emplear personal mejor capacitado y someterlo a un entrenamiento intensivo además a que en estas condiciones también será factible la supervisión cuidadosa del trabajo de campo y del procesamiento de los resultados.

Como prueba de la confiabilidad de las técnicas de muestreo se puede mencionar los resultados de los Conteos Rápidos realizados en México el 21 de agosto de 1994. El doctor Jorge Carpizo comenta: "*Si la muestra está bien diseñada y bien realizada, los Conteos Rápidos coinciden con los resultados finales de la elección en forma asombrosa*".

Es claro que es posible conocer algunas aproximaciones de estos resultados al menos en lo que a sus tendencias se trata, y depende mucho de la planeación del operativo y de la velocidad en la que se logren reunir los resultados desde los diferentes puntos que esto involucre, pero si bien, esto es muy importante, no es menos saber qué hacer con esos datos, es decir, suponiendo que en este momento se tienen todos los insumos que el sistema requiere, al menos en lo que a información se refiere. ¿Cuál es el siguiente paso?. Es ver la forma más rápida de realizar un conteo, es decir, el procesamiento de la información, y ¿cómo lograr esto?. De la forma que sea debe ser rápido, en el menor tiempo posible, desde el proceso de captura hasta el procesamiento de los datos y la presentación de la información.

Es importante recordar que *ésta es la parte novedosa en estos procesos, la velocidad en la que se pueden recolectar los datos, procesarlos y presentar los resultados*. Para ello se puede decir que se necesitará la herramienta más popular de la época, "La Computadora", ahora es necesario mencionar que ésta es la forma más eficiente, segura y rápida que existe en esta época para el procesamiento de datos. El asunto no termina ahí, ya que si la computadora puede procesar una base de datos de 5 mil registros en algunos segundos, es necesario primeramente introducir los datos a la computadora.

¿Cómo reunir la información?, es decir, si se captura en 30 computadoras, la captura sería relativamente rápida, pero después cuando se necesite procesar la información cómo se concentra?, si se trabaja en una red de área local ese problema queda resuelto porque se está accedendo una sola base de datos y se va tomando para el proceso cada vez que se necesita hacer un corte, por otro lado, siempre existe el riesgo latente de que se caiga el sistema, estos detalles parecen no ser tan importantes, pero si algo de esto sucede durante el transcurso de las 4 horas que puede durar la recolección de los datos, es posible que el proyecto fracase.

Por otro lado se puede plantear que dentro del área en estudio, sean instaladas un número determinado de centros receptores de información, los cuales realizarán la captura de algún número determinado de casillas que posteriormente podrán transmitir

por algún medio al centro general de acopio; por llamarlo de alguna manera.

Son muchas las posibilidades que se tienen con respecto a esta tarea, sin embargo es necesario conocer todos los riesgos que pueden presentar estas diferentes opciones, antes de poder plantear siquiera una metodología para el SEIT.

Como se puede apreciar, son muchos los aspectos que hay que tomar en cuenta cuando se pretende diseñar un sistema de esta especie ya que existen demasiados puntos vulnerables que por más que uno quisiera rodear no es posible y lo único que resta es buscar el camino de que todo esto se realice de la forma más segura o al menos con la menor probabilidad de falla.

En este capítulo se han definido y revisado algunos conceptos básicos que entran en función en el SEIT, además de dar un breve vistazo a los antecedentes y orígenes de las técnicas de sondeo, su difusión y aplicación en los campos de la política y las ciencias sociales, y la importancia que han venido tomando en las estrategias de campaña a últimas fechas. Por otro lado, se ha analizado cómo el SEIT entra en esta clase de técnicas de sondeo y su expresión más conocida como un "Conteo Rápido", y de forma breve se ha descrito el funcionamiento de éste. De igual manera se han identificado las disciplinas que aportan elementos teórico-prácticos para el buen desarrollo del sistema, y a pesar de que todas estas son básicas para el caso particular del SEIT, se ha identificado que la Teoría de Sistemas sobresale de las demás debido a que esta disciplina da las pautas para trabajar en equipo y conjuntar los esfuerzos realizados por las diferentes áreas de trabajo involucradas en este sistema.

Por último, se ha resaltado la importancia de la muestra en este tipo de sistemas y lo importante que es manejar adecuadamente los datos, los cuales son un par de aspectos de suma importancia para la satisfactoria obtención de los resultados finales del sistema.

CAPITULO II.
EL DISEÑO DE LA MUESTRA COMO UN ELEMENTO ESENCIAL
DE LA CONFIABILIDAD DEL SEIT.

2.1. Las Técnicas de Muestreo y su Contribución al Diseño e Instrumentación del Sistema.

Como ya se comentó en el capítulo anterior, una forma muy popular de obtener información en estos tiempos es la realización de encuestas. La gran mayoría de éstas se basan en muestras, en una fracción de la población de la que se quiere hacer la inferencia y esto trae como consecuencia una serie de facilidades entre las que se encuentran: menor costo, menor tiempo, más posibilidades de realización del estudio en sí, mejor supervisión del trabajo en campo, mayor precisión en los resultados, etc.

Esta muestra como es lógico no puede ser diseñada sin tomar en cuenta alguna técnica que respalde su representatividad, así como tampoco puede ser un aspecto sin control alguno, el calcular el número de elementos que formarán parte de dicha muestra. En este capítulo se analizarán las técnicas que prestan apoyo en este proceso.

Las muestras se pueden seleccionar, ya sea de manera intencional, es decir, la persona que selecciona la muestra escoge, a su juicio, los elementos que forman parte de ésta; o bien de manera probabilística, en donde la selección de la muestra se puede considerar un experimento aleatorio. Cabe mencionar que el hecho de que todas las unidades en la población tengan una probabilidad conocida y no nula de ser seleccionadas es lo que científicamente permite inferir los resultados de toda la población en base a la muestra.

Por otro lado, es importante tener bien identificado el marco muestral o lista de muestreo en el desarrollo de una encuesta. Esta lista consta del grupo completo de elementos que tienen la oportunidad de ser seleccionados, dada la aproximación de muestreo que se escoja. Por lo tanto, debe cuidarse que la cobertura de la población estudiada sea adecuada, ya que una muestra sólo puede ser representativa de la población incluida en el marco muestral.

Para el diseño de muestras existe una técnica especializada dentro de la estadística: La Teoría de Muestreo, que actualmente se aplica a muchas áreas de la investigación científica, social y comercial.

Para algunos autores el diseño de las muestras tiene dos aspectos:

- *"Un proceso de selección, que consiste en las reglas y operaciones mediante las cuales se incluyen en la muestra algunos miembros de la población."*
- *"Un proceso de estimación para calcular las estadísticas de la muestra, que son estimaciones muestrales de valores de la población."¹*

Tratando el primer aspecto del diseño de la muestra, existen diferentes métodos de selección, de los cuales éstos son los más solicitados:

Muestreo Aleatorio Simple.

Es un método para seleccionar n unidades de una población total N de tal modo que cada una de las combinaciones posibles tenga la misma probabilidad de ser elegida. El número posible de muestras lo da la expresión ${}_N C_n$, para ésto se enumeran las unidades de 1 a N , después de una tabla de números aleatorios se extraen n números al azar entre 1 y N , esto también se puede hacer mediante un programa en una computadora. Cabe señalar que en cada extracción la probabilidad de selección es igual para todas las unidades que no hayan salido.

Muestreo Aleatorio Estratificado.

En este caso la población total N , se subdivide en subpoblaciones N_1, N_2, \dots, N_L , estas subpoblaciones no comparten elementos y en su conjunto comprenden al total de la población de modo que,

¹ KISH, Leslie. *"Muestreo de Encuestas"*. Ed. Trillas, México, 1975, pág.24

$$N_1 + N_2 + \dots + N_L = N$$

A cada una de estas poblaciones se les pueden llamar estratos, una vez determinados los estratos se extrae una muestra de cada uno de ellos, esta extracción es independiente en cada uno de los diferentes estratos. Los tamaños de la muestra en cada estrato varían y se identifican con

n_1, n_2, \dots, n_L , respectivamente.

Cabe señalar que una vez delimitados los estratos se calculará el tamaño de la muestra para cada uno, y que la selección dentro de cada estrato se toma por medio de un muestreo aleatorio simple.

Esta técnica de muestreo es muy empleada por sus bondades, entre las cuales se tiene:

- Que en un momento determinado es posible conocer los datos que se esperan dentro de un estrato establecido, por lo que es recomendable que cada estrato se tome como si éste fuera una población.
- La administración se puede simplificar y la calidad del levantamiento de la información es superior ya que es posible tener sucursales en el campo, cada una de las cuales supervisaría el trabajo en el campo.
- Los problemas de muestreo pueden llegar a tener diferencias en diversas partes de la población, las cuales se pueden tratar con el cuidado necesario y plantearse dichas diferencias entre estratos.
- La estratificación puede dar lugar a mejorar la precisión de las estimaciones de características de la población total. Quizás sea posible dividir una población heterogénea en varias subpoblaciones intensamente homogéneas. De aquí que si cada subpoblación es lo más homogénea posible, una pequeña muestra bastará para poder inferenciar sobre ésta, ya que entre unidades

sus medidas varían ligeramente, y finalmente podrán juntarse estas estimaciones en una estimación precisa para toda la población.

Para el mismo tamaño de la muestra, este procedimiento de selección usualmente produce un estimador que posee una varianza más pequeña de la que se puede obtener por el muestreo aleatorio simple, si las mediciones en estudio son homogéneas dentro de los estratos, pero las medidas entre estratos varían entre ellas, aunque cabe mencionar que la estratificación sólo aumenta la precisión de las estimaciones de las variables que se relacionan con las variables de la estratificación.

Muestreo Sistemático.

Esta clase de muestreo consiste en que conocido el coeficiente de elevación, es decir el número de elementos de la población a que equivale cada elemento de la muestra, denotado por k , se obtiene al azar un número no superior a este coeficiente y éste es el primer elemento de la muestra m , a continuación a ese número m se le suma el coeficiente de elevación y el número obtenido es el siguiente elemento de la muestra $m + k$, así sucesivamente hasta obtener el tamaño de muestra requerido.

$$m, m + k, m + (k \cdot 2), m + (k \cdot 3), \dots$$

Este procedimiento se ocupa en el trabajo de campo para hacer más práctica la selección de los elementos en muestra, y como consecuencia se reducen los costos en la recolección de los datos.

En este tipo de muestreo es muy importante revisar el marco muestral, o lista de la población, con el objetivo de poder ver si hay o no alguna razón para pensar que la muestra resultante de un arranque aleatorio será sistemáticamente diferente que la generada con otro arranque, de tal forma que pudiera afectar los resultados de la encuesta.

Muestreo por conglomerados.

Esta clase de muestreo difiere de las anteriores ya que se parte del hecho de que si se emplean los elementos como unidades de muestreo, es necesario tener una lista confiable de los elementos de la población, por otro lado el pensar en hacer dicha lista sería demasiado costoso cuando el estudio se traslada a poblaciones grandes y dispersas. Para esto es posible ver la forma de seccionar la población en unidades de área como pudiesen ser: cuadras, colonias, delegaciones, ciudades, etc.

Con esto se quiere decir que en el muestreo por conglomerados la unidad de muestreo es una colección o conjunto de elementos llamado conglomerado, y generalmente se emplea para reducir costos, o porque no se cuenta con un marco adecuado de los elementos.

Este muestreo se puede aplicar en una o varias etapas. En el primer caso se seleccionan conglomerados y se obtiene la medición de cada uno de los elementos que pertenecen a los conglomerados seleccionados, de lo que se trata es de tener un censo en el conglomerado seleccionado. En el segundo, cuando ya se ha hecho la selección de los conglomerados, se realiza una subselección de elementos de cada uno de éstos.

Por otro lado, aunque se contará con la lista confiable de la población, es posible que por cuestiones económicas se recomiende una unidad conglomerada mayor, por algún tipo de facilidad que esto ofreciera, aunque antes hay que tomar en cuenta lo que opina W. Cochran. *"Para un tamaño de muestra dado, una unidad de muestreo pequeña es más precisa que una unidad grande."*²

El muestreo por conglomerados puede ser más preciso que el muestreo aleatorio simple o que el estratificado si las mediciones dentro de los conglomerados son heterogéneas y las medidas entre los conglomerados son aproximadamente iguales.

² COCHRAN, William G. *"Técnicas de Muestreo"*, México, 1a. Edición, Ed. CECSA, 1980.

2.2. La estimación del tamaño de la muestra.

Para conseguir estimar el tamaño de la muestra, se necesita tener en mente algunos aspectos importantes de la población en estudio y también conocer el nivel de precisión esperado para poder considerar satisfactorios los resultados.

Este es uno de los aspectos más delicados del sistema, pues si por alguna razón se determina un tamaño de muestra demasiado pequeño, es seguro que a la hora de conocer los resultados, éstos no se adecuarán a la realidad, y por otro lado, si el tamaño de la muestra es demasiado grande, específicamente para el SEIT, puede representar un retraso en tiempo innecesario, además de un desperdicio de recursos.

Para calcular el tamaño de la muestra se puede partir del hecho de que se espera un estimador p cuya diferencia con el parámetro poblacional P no sea mayor de d y existe un riesgo α que se está dispuesto a correr aunque el error real supere a d , entonces:

$$P(|p - P| \geq d) = \alpha$$

Para esto se necesita plantear que el estimador $\hat{\mu}$ de la media poblacional μ , tenga a lo más una varianza

$$V = \left(\frac{d}{t}\right)^2$$

en donde:

d , es el error máximo aceptado

t , es la abscisa de la curva normal que corta un área de α en las colas de la distribución.

Se debe verificar que:

$$V(\hat{\mu}) = \frac{S^2}{n} \left(1 - \frac{n}{N}\right) \leq V = \left(\frac{d}{t}\right)^2$$

en donde: S^2 = varianza poblacional.

n = tamaño de la muestra.

N = tamaño de la población.

Es decir:

$$\frac{S^2}{n} - \frac{S^2}{N} \leq V$$

despejando n se tiene:

$$n \geq \frac{S^2}{V + \frac{S^2}{N}} = \frac{\frac{S^2}{V}}{1 + \left(\frac{1}{N}\right) \frac{S^2}{V}}$$

si se define n_0 como

$$n_0 = \frac{S^2}{V} = \frac{S^2 t^2}{d^2}$$

se tiene

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Ahora, bajo este planteamiento se tiene la necesidad de conocer la varianza poblacional S^2 , lo cual no ayuda mucho pues éste problema es parecido al que se trata de resolver, pero mediante un artificio básico se puede plantear una alternativa, para cualquier unidad en la población se define y_i como 1 si la unidad está dentro de la población que votó por determinado partido y 0 si no fue así, para esta población, está claro que, como comenta Cochran³:

³ Op.cit. pág.80.

$$Y = \sum_1^N y_i = A$$

$$\bar{Y} = \frac{\sum_1^N y_i}{N} = \frac{A}{N} = P$$

de modo que:

$$\sum_1^N y_i^2 = A = NP$$

por lo tanto:

$$\begin{aligned} S^2 &= \frac{\sum_1^N (y_i - \bar{Y})^2}{N-1} = \frac{\sum_1^N y_i^2 - N\bar{Y}^2}{N-1} \\ &= \frac{1}{N-1} (NP - N P^2) = \frac{N}{N-1} PQ \end{aligned}$$

donde $Q = 1 - P$, entonces se puede definir a n_0 de la siguiente manera:

$$n_0 = \frac{S^2}{V} = \frac{S^2 t^2}{d^2} = \frac{PQ t^2}{d^2}$$

cuando N es muy grande o tiende a infinito, y siempre y cuando P se tome como si estuviera normalmente distribuido.

Como no se conoce el valor de P , para calcular el tamaño de muestra, se deriva la función de varianza con respecto a P y se encuentra que alcanza un valor máximo en $P = 0.5$

$$\frac{dP(1-P)}{dP} = 1 - 2P$$

$$1 - 2p = 0 \Rightarrow p = 1/2 = 0.5$$

$$P + Q = 1 \Rightarrow (1/2) + Q = 1, Q = 1/2$$

Como se puede observar n varía en función de N .

El uso de cpf (correcciones debidas a población finita ó corrección por finitud) provoca de manera indirecta que el tamaño de la muestra esté en función de N . De esta forma, cuando el efecto del factor de corrección para poblaciones finitas es significativo, la precisión de una estimación por muestreo se determina no sólo por el valor absoluto n de la muestra, sino también por la proporción de la población que se incluye en la muestra.

De lo antes visto se tiene:

$$\lim_{N \rightarrow \infty} (n) = \lim_{N \rightarrow \infty} \frac{n_0}{1 + \frac{n_0}{N}} = n_0$$

de aquí se puede deducir que en la práctica primero se calcula n_0 , si $\frac{n_0}{N}$ es despreciable, una buena aproximación de n será n_0 .

Se puede observar que, aproximadamente n es inversamente proporcional al cuadrado de la precisión, si la población es demasiado grande, lo cual indica que si se quiere reducir el error a la mitad, hace falta una muestra no doble, sino cuádruple. Como se puede observar en el cuadro 2.2.1 para $p = 0.5$ y $n = 50$, se tiene un error absoluto de 0.14, con la misma p se debe tener una muestra $n = 200$ aproximadamente para que el error sea de 0.069.

Con esto es posible deducir que existe una relación demasiado importante entre la precisión y el tamaño de la muestra, ya que al disminuir el error de muestreo aumenta el tamaño de la muestra, y que además de éste, es posible alcanzar cualquier precisión deseada con una muestra suficientemente grande.

Para el SEIT, en la aplicación ocupada, los procesos electorales se puede evaluar lo siguiente:

Error de muestreo en relación al tamaño de la muestra y a la proporción

| | | p | | | | | | | | |
|------|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | 0.1 | 0.2 | 0.3 | 0.04 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
| n | | | | | | | | | | |
| 35 | | 0.101 | 0.134 | 0.154 | 0.165 | 0.168 | 0.165 | 0.154 | 0.134 | 0.101 |
| 50 | | 0.084 | 0.112 | 0.128 | 0.137 | 0.140 | 0.137 | 0.128 | 0.112 | 0.084 |
| 75 | | 0.068 | 0.091 | 0.104 | 0.112 | 0.114 | 0.112 | 0.104 | 0.091 | 0.068 |
| 100 | | 0.059 | 0.079 | 0.090 | 0.097 | 0.098 | 0.097 | 0.090 | 0.079 | 0.059 |
| 200 | | 0.042 | 0.056 | 0.064 | 0.068 | 0.069 | 0.068 | 0.064 | 0.056 | 0.042 |
| 300 | | 0.034 | 0.045 | 0.052 | 0.056 | 0.057 | 0.056 | 0.052 | 0.045 | 0.034 |
| 500 | | 0.026 | 0.035 | 0.040 | 0.043 | 0.044 | 0.043 | 0.040 | 0.035 | 0.026 |
| 1000 | | 0.019 | 0.025 | 0.028 | 0.030 | 0.031 | 0.030 | 0.028 | 0.025 | 0.019 |
| 1500 | | 0.015 | 0.020 | 0.023 | 0.025 | 0.025 | 0.025 | 0.023 | 0.020 | 0.015 |

La variabilidad reflejada en la tabla describe el error potencial que se tendría en las estimaciones por el hecho de utilizar una muestra con un nivel del 95% de confianza. La tabla y las ecuaciones en las que se basa se aplican a muestras obtenidas con un diseño aleatorio simple.

Cuadro 2.2.1.

Si es permitido un error de estimación del 5%, con un nivel de confianza del 95%, y con una población igual a 63'596 secciones electorales en todo el país, suponiendo que estas son las unidades de muestreo, aplicando el modelo antes visto se tiene:

$$n_0 = \frac{(2)^2(0.5)(0.5)}{(0.05)^2} = 400$$

si se aplica

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = 397.5$$

Así las muestras, entre más grandes, generan intervalos de confianza más cortos para las estimaciones, ya que aumenta la precisión de la estimación. Esto lleva a lo que comenta Erwin Kreyszig⁴, "cuanto más cortos se quieran los intervalos de confianza, más grande se debe escoger el tamaño de la muestra".

⁴ KREYSZIG, Erwin. "Introducción a la Estadística Matemática", México, Ed.Limusa, 1987.

Reforzando lo antes mencionado, se verá lo que comenta Leslie Kish, "Hay dos razones por las que las muestras probabilísticas con diseño de muestreo complejo sean típicamente grandes. Primera, la organización y las operaciones necesarias para obtener muestras complejas de poblaciones grandes son tan pesadas y costosas que solamente muestras grandes pueden justificar económicamente su utilización. Segunda, la mayoría de las muestras para encuestas complejas tiene como meta medir efectos que son demasiado pequeños para poderlos detectar con muestras pequeñas."

En muchas situaciones prácticas, la distribución de muestreo de la media estimada y de la mayoría de los estimadores comúnmente utilizados para representar resultados de una encuesta es aproximadamente normal, esta tendencia a la normal de la distribución de muestreo, cuando las muestras son de tamaño grande, no se basa en la normalidad de la distribución de los elementos en la población. Por el contrario las distribuciones de las características de la encuesta en la población están generalmente muy lejos de la normal. Pero aún en muestras pequeñas, la aproximación es buena para muchas variables comunes. Es muy frecuente que las distribuciones de muestreo de las medias de muestras aunque sean de tamaño moderado, se aproximen suficientemente a la normal como para poder hacer inferencias estadísticas, a partir de estimaciones de la muestra, a valores de la población entera, y esto es consecuencia del Teorema Central del Límite.

Teorema Central del Límite.

Si se obtienen repetidas muestras de tamaño n de cualquier población (de cualquier forma) con una media μ y una varianza σ^2 , entonces tanto como n crezca, la distribución de muestreo de la media de la muestra se aproxima a la normalidad, con una media μ y una varianza σ^2/n .

La aproximación a la normalidad de la distribución de muestreo de una estadística es más rápida para unas variables que para otras. Cuando la aproximación es más rápida a la normalidad, indica que hay mejor aproximación de la distribución de

muestreo para un tamaño de muestra dado, o que la aproximación que se desea se puede alcanzar con un tamaño de muestra menor.

Por otro lado cabe mencionar que el tamaño de muestras para subpoblaciones posee algunas variaciones. Tomando como subpoblación a una parte de la muestra, a la cual se le considera como una muestra de la porción correspondiente de la población de la encuesta, y como es evidente, pueden considerarse estimaciones de la varianza y precisión diferentes en cada dominio.

Cochran propone que para determinar el tamaño de la muestra total, cuando se tienen varios dominios de estudio se debe considerar lo siguiente:

1. Si los dominios se pueden identificar con anterioridad, se calcula por separado la n necesaria en cada dominio. En caso de querer estimaciones con varianza y precisión iguales en cada una de las k subpoblaciones, el tamaño de la muestra debe ser k veces más grande que el necesario para un estimador global con la misma precisión (kn). Pero si en lugar de levantar una muestra de tamaño kn , se mantiene una muestra de tamaño n para k dominios, el error en la estimación en cada subpoblación se verá afectado por "raíz cuadrada de k " en forma directamente proporcional.
2. Si los dominios no se pueden identificar con anterioridad pero se conocen las proporciones π_i de las unidades que pertenecen a diferentes subpoblaciones, entonces en un muestreo aleatorio simple, el tamaño de la i -ésima subpoblación es m_i y la varianza promedio de la proporción en dicha subpoblación, si m_i es grande, es:

$$v(p_i) = \frac{p_i q_i}{n \pi_i}$$

Por otro lado, es conveniente resaltar que el tamaño de la muestra puede llegar a depender de otros factores como pudieran ser:

1. **Recursos disponibles.** El tamaño de la muestra debe ser necesariamente consistente con los recursos disponibles para realizar la investigación, para ello se requiere de una estimación de costo, tiempo y materiales que serán necesarios para un determinado tamaño de la muestra.
2. **Número de características que se investigan.** Dados el costo y el trabajo administrativo y de campo que se involucran en el desarrollo de una encuesta, en la mayoría de las encuestas se busca captar información de más de una característica.

Una forma de determinar el tamaño de la muestra es considerar márgenes de error para cada una de las características captadas más importantes, y escoger el valor más grande de n , claro, si es que éste no sobrepasa el presupuesto. Es necesario tener presente que cuando el número de variables es grande y se especifica un grado de precisión para cada una, los cálculos conducen a una serie de tamaños de muestra diferentes, lo cual puede llegar a convertirse en un dolor de cabeza; una opción para evitar esto es disminuir la precisión en algunas características para poder reducir el tamaño de n .

3. **Estimaciones para subpoblaciones.** Al determinar el tamaño de la muestra se debe tomar en cuenta el número de categorías o clases en las que se van a agrupar los datos. El tamaño de muestra total debe ser lo suficientemente grande para asegurar tamaños de muestra que representen bien a las categorías independientes, inclusive a las más pequeñas.
- 4.- **Homogeneidad o heterogeneidad de la población.** Entre más homogénea sea una población se requerirá una muestra más pequeña, por el contrario entre más heterogénea sea una población el tamaño de la muestra deberá ser mayor.

2.3. Definición del Marco Muestral y de la Población Objetivo.

Cuando se habla de **población** en lo que es la teoría de muestreo, se hace referencia al conjunto completo de elementos de donde se extraerá una muestra. La definición de la población puede no ser un problema si se tiene bien claro cuál es el universo de trabajo, por ejemplo, cuando se toma un grupo de automóviles con placas del D.F. para conocer cuál es la cantidad de contaminación que éstos expiden, aquí es claro que la población total son todos los automóviles con placas del D.F. Además, es muy importante que en el muestreo de estos automóviles, se deben fijar las reglas para definir lo que son éstos, es decir, es necesario diferenciar lo que es un automóvil y un camión, por trivial que parezca, en el desarrollo del proyecto puede surgir la duda cuando se trata de integrar algún elemento a la muestra. Para esto resulta práctico fijar dichas reglas, las cuales, además deben ser aplicables a la práctica, ya que el encargado de hacer el levantamiento en campo debe ser capaz de decidir sin demasiadas complicaciones, si un caso dudoso pertenece o no a la población.

La población en la que se ha realizado un muestreo, debe coincidir con la población sobre la cual se desea la información, es decir la **población objetivo**, existen casos en los que por factibilidad, facilidad, rapidez, conveniencia, etc., la población muestreada llega a poseer alguna restricción, cosa que no sucede con la población objetivo; por ejemplo, se desea hacer un estudio para conocer cuál es la opinión de los mexicanos con respecto a la calidad de los servicios que otorga el gobierno, cuando se emplea el término mexicanos, se hace referencia a todos los mexicanos, (según INEGI, 82 millones), aquí existe un pequeño problema, quizá por los factores que ya se mencionaron anteriormente sea más cómodo tomar una muestra de mexicanos que vivan en grandes ciudades, debido a que es en éstas en donde se concentra el 80% de todos ellos, pero qué pasa con las poblaciones rurales, y más aún, ¿qué pasa con las poblaciones que aún no cuentan con caminos o medios de comunicación?, ellos también son parte de la población objetivo, y probablemente tendrán un punto de vista opuesto al de sus paisanos de las grandes urbes. Si éste fuese el caso, es necesario recordarlo en las conclusiones de la muestra extraída,

estas conclusiones son aplicables a la población muestreada y será necesario recurrir a otras fuentes de información para decidir en un momento dado hasta qué grado son aplicables estas conclusiones a la población objetivo. Para esto, toda clase de información que se tenga respecto a las diferencias entre ambas poblaciones será de mucha utilidad.

Otro de los puntos que aquí se tratarán es el marco o marco muestral, antes de comenzar con la selección de la muestra, la población debe ser dividida en lo que se llamarán unidades de muestreo o unidades, éste es un trabajo que merece un trato muy preciso, ya que, si bien, identificar estas unidades de muestreo puede no ser tan complicado como en el caso de tomar una muestra de fichas bibliográficas en el que obviamente las unidades serán las fichas en sí, es posible encontrarse un caso en el que se tenga que sacar una muestra de un embarque de chocolates en donde podría existir la duda en si se toma como unidad de muestreo una pieza de chocolate o un estuche de éstos mismos. Es conveniente comentar que estas unidades de muestreo deben cubrir la totalidad de la población y no caer en el caso de que todos los elementos de la población pertenezcan a una sola unidad de muestreo. En la mayoría de los casos, la complejidad de esto radica en la decisión de definir los límites de la unidad, como ya se comentó con los chocolates.

Es común que la construcción de esta lista de unidades de muestreo llamada marco muestral sea prácticamente el problema a vencer. Ya que regularmente por el proceso de recolección da como resultado una lista que casi siempre contiene un número considerable de errores, o en otros casos la lista es ilegible, casi siempre están incompletas, seguramente contendrán duplicaciones de magnitud desconocida, etc.

Para el caso que se estudia en esta investigación, recordando que se tiene una aplicación a los procesos electorales, es necesario realizar el siguiente análisis.

El sistema electoral mexicano es uno de los más amplios y confiables con los que se cuenta actualmente, éste se compone por una serie de divisiones, se empezará por las unidades básicas, que en este caso serían los votantes.

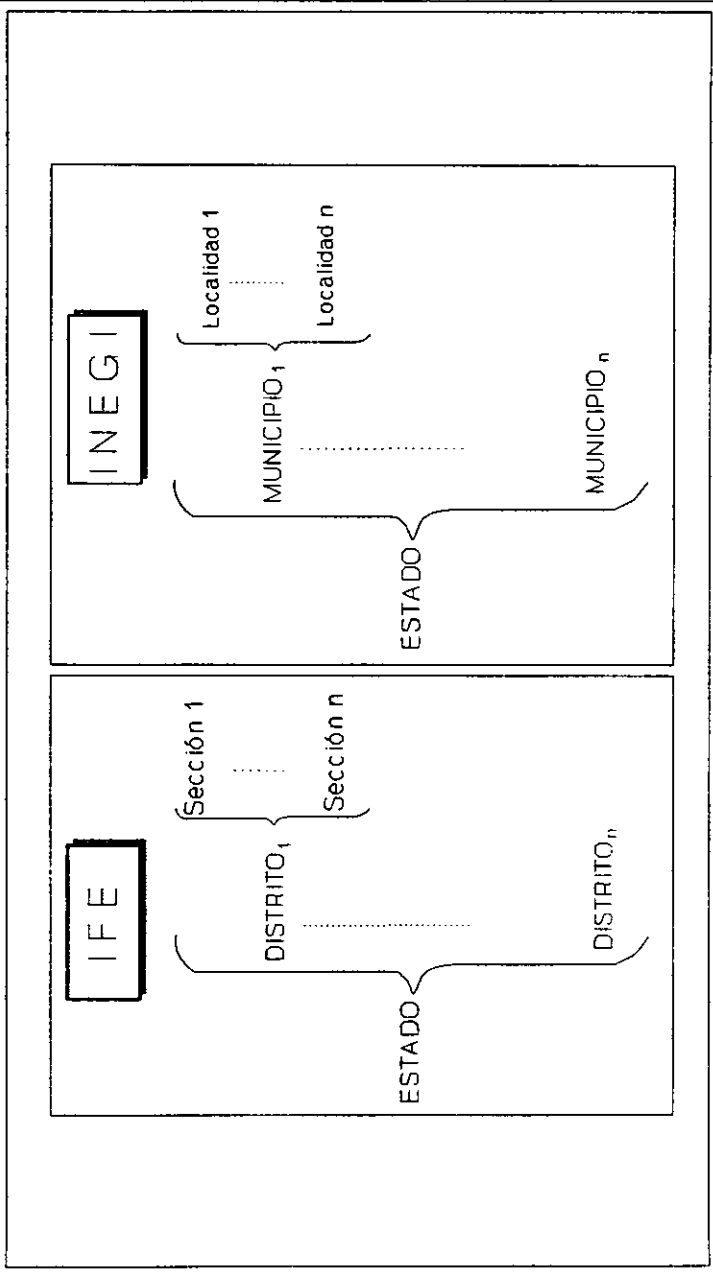
Los votantes son todas aquellas personas con capacidad de voto, esto significa que estas personas cumplen con cierto número de condiciones como son: ser mayor de 18 años, estar apuntado en el padrón electoral nacional, formar parte de las listas nominales (existe gente que está apuntada en el padrón y por alguna razón como puede ser cambio de domicilio, no actualizar su registro, inclusive algún caso de fallecimiento, etc., no está en las listas nominales), tener físicamente en su poder su propia credencial para votar, sin la cual es impensable poder sufragar, etc. En México el número de personas empadronadas a 1994 era de 47'480,159.

Se continuará con la agrupación mínima de votantes para tener alguna representatividad ante el proceso, la casilla de votación, ésta es un espacio físico a donde acuden los votantes a sufragar el día del proceso, esta casilla concentra los votos de las personas que se encuentran en la lista nominal asignada a dicha casilla, regularmente son personas cuya vivienda se localiza en la misma zona, en las casillas normalmente se encuentran representantes de diferentes partidos, ciudadanos sin filiación política, jefe de casilla, funcionarios de casilla, etc., los cuales se encargan de dar legalidad al proceso y al conteo que al final de la jornada se llevará acabo en todas y cada una de éstas. De modo que al final de dicho conteo se exhibe un anuncio mostrando los resultados finales del sufragio en cada una de las casillas. Para las elecciones federales de 1994 fueron necesarias casi 100'000 casillas de votación a lo largo de todo el país.

Continuando con esta agrupación se encuentra lo que el sistema electoral denomina como secciones político-electorales, éstas están conformadas por casillas, es decir un grupo de casillas conforma una sección, normalmente en una sección existe una casilla, esto es para la mayoría de los casos, pero puede suceder que por la densidad de la población se requiera reforzar ésta con más casillas, hay secciones que llegan a contar hasta con 3 o 4 casillas. El sistema electoral mexicano tiene un seccionamiento actual de 63,596 secciones. Ver figura 2.3.1.

La siguiente clasificación es la de distrito electoral, que es un conjunto de secciones. En México se ha seccionado en 300 distritos electorales. Ver figura 2.3.1.

Figura 2.3.1. Distribución y Organización de la población según IFE e INEGI



Y la siguiente agrupación es el estado o entidad política, el cual electoralmente hablando, consta de varios distritos electorales, el número de distritos electorales está en función del número de habitantes que tenga el estado y su densidad de población, para aclarar un poco esta estructura se recomienda ver la figura 2.3.1.

Cabe mencionar que es notorio que a nivel poblaciones quizá pueda ser otra la distribución, por ejemplo, que de votantes y casillas sea igual, pero en lugar de secciones se podrían plantear colonias y en lugar de distritos se seccionara por municipio o delegaciones políticas; en el caso de estados o entidades políticas no hay diferencia. Sin embargo, en esto se están planteando dos instituciones nacionales que difieren en sus criterio por alguna razón, estas instituciones son el Instituto Federal Electoral, "IFE" y el Instituto Nacional de Estadística Geografía e Informática, "INEGI".

2.4. Algunas observaciones para el diseño de la muestra, propiedades de las estimaciones y algunos aspectos del muestreo estratificado para proporciones.

Para el desarrollo del diseño de la muestra, así como para probar algunas de las hipótesis, se diseñarán varias muestras con el fin de poder evaluar cuál de éstas resulta mejor, cabe hacer la aclaración que el hecho de que algunas de estas muestras resulte más precisa, no quiere decir que ese método sea el ideal, sino que fue el más preciso para ese estudio. En este caso se desarrollará el diseño de la muestra en una aplicación electoral, para ello se tomarán como base las elecciones federales del 21 de agosto de 1994 en México, debido a que parece ser posible conseguir información que sirva de apoyo para desarrollar esta investigación y probar los planteamientos que se presentan.

Para la selección de la muestra es necesario revisar la estructura del aparato electoral que se encargará de organizar el escrutinio, en este caso, en el subcapítulo anterior se analizó la estructura del IFE, por lo que se partirá de este supuesto.

Como se ha mencionado con anterioridad, para diseñar la muestra es necesario contar con una lista de la población objetivo, en este caso el pretender tener una lista con 47 millones y medio de votantes resulta completamente innecesario e inoperable, además de que esa información es completamente confidencial.

Se definirán un par de conceptos básicos antes de continuar, es necesario precisar cuál es la población objetivo, y cuál el marco muestral en este estudio. Es posible definir como la población objetivo, al electorado nacional, toda la gente que participe en las elecciones, debido a que ésta será quien defina mediante el escrutinio las proporciones que tendrá cada competidor, en este caso, partido político, en el proceso, se pretende conocer cuál será el resultado de las elecciones que ellos realizarán, y por otro lado, el marco muestral, debido a las características propias del SEIT, serán el total de secciones políticas del país, para efecto de poder definir la muestra, se tomará como unidad muestral una sección, la distribución de las secciones seguramente no es información de tipo confidencial y sin embargo, tampoco está al alcance del público.

Sin embargo, con la información que el IFE tiene disponible al público se puede conocer el número de secciones que tienen los 32 estados de la república, como se muestra en la figura 2.4.1., según la organización del IFE, en cada estado las secciones son identificadas por un número que es único dentro del mismo estado, así se puede hacer referencia, por ejemplo, a la sección 487 de Veracruz, esta sección está ubicada en alguno de los distritos electorales de la entidad, y por otro lado, también físicamente se ubica en algún municipio del mismo estado, además de que en todas las entidades la numeración de las secciones comienza a partir de 1, con esto se quiere decir que existe la sección 1 de Aguascalientes y la sección 1 de Campeche.

DISTRIBUCION DE DISTRITOS, SECCIONES Y PADRON
POR ESTADOS

| | ESTADO | DISTRITOS | SECCIONES | PADRON |
|----|-----------------------|-----------|-----------|------------|
| | | FEDERALES | | 1994 |
| 1 | AGUASCALIENTES | 2 | 486 | 426,878 |
| 2 | BAJA CALIFORNIA NORTE | 6 | 1,368 | 1,087,733 |
| 3 | BAJA CALIFORNIA SUR | 2 | 352 | 190,749 |
| 4 | CAMPECHE | 2 | 489 | 308,668 |
| 5 | COAHUILA | 7 | 1,520 | 1,142,170 |
| 6 | COLIMA | 2 | 336 | 263,701 |
| 7 | CHIAPAS | 9 | 1,929 | 1,714,196 |
| 8 | CHIHUAHUA | 10 | 2,710 | 1,538,688 |
| 9 | DISTRITO FEDERAL | 40 | 5,543 | 5,535,268 |
| 10 | DURANGO | 6 | 1,390 | 733,247 |
| 11 | GUANAJUATO | 13 | 3,005 | 2,196,005 |
| 12 | GUERRERO | 10 | 2,778 | 1,244,386 |
| 13 | HIDALGO | 6 | 1,707 | 1,037,024 |
| 14 | JALISCO | 20 | 3,309 | 3,026,460 |
| 15 | MEXICO | 34 | 5,916 | 6,050,969 |
| 16 | MICHOACAN | 13 | 2,674 | 1,888,250 |
| 17 | MORELOS | 4 | 906 | 769,597 |
| 18 | NAYARIT | 3 | 878 | 475,775 |
| 19 | NUEVO LEON | 11 | 2,123 | 1,948,008 |
| 20 | OAXACA | 10 | 2,449 | 1,514,166 |
| 21 | PUEBLA | 14 | 2,548 | 2,184,055 |
| 22 | QUERETARO | 3 | 687 | 620,240 |
| 23 | QUINTANA ROO | 2 | 450 | 317,394 |
| 24 | SAN LUIS POTOSI | 7 | 1,795 | 1,076,188 |
| 25 | SINALOA | 9 | 3,789 | 1,256,285 |
| 26 | SONORA | 7 | 1,330 | 1,180,609 |
| 27 | TABASCO | 5 | 1,133 | 866,039 |
| 28 | TAMAULIPAS | 9 | 1,739 | 1,373,929 |
| 29 | TLAXCALA | 2 | 605 | 450,099 |
| 30 | VERACRUZ | 23 | 4,711 | 3,622,708 |
| 31 | YUCATAN | 4 | 1,059 | 745,587 |
| 32 | ZACATECAS | 5 | 1,882 | 695,088 |
| | TOTAL | 300 | 63,596 | 47,480,159 |

Cuadro 2.4.1.

Con esta información, es posible crear el marco muestral, lo único que se debe hacer es crear una base de datos que contenga dos campos por registro, la estructura se puede consultar en el cuadro 2.4.2.

Estructura de la base de datos que será el Marco Muestral

| Nombre del campo | Tipo de campo | Longitud |
|------------------|---------------|----------|
| Estado | Numérico | 2 |
| Sección | Numérico | 4 |

Cuadro 2.4.2.

El campo estado se utiliza para asegurar a cada registro la clave de la entidad, y el campo sección es empleado para de forma consecutiva, asignar el número de la sección, de esta forma, crear una base de datos que contenga los 63'596 registros que representa al total de secciones en el país.

Como se puede observar en la figura 2.4.1, también se dispone del padrón electoral por estado, totalizando éste, se puede tener la cifra nacional que en este caso es 47'480,159 personas empadronadas en el país, para efectos de distribución de la muestra se calculará la proporción de padrón para cada estado, con ésta se puede saber cuál será la proporción de la muestra por cada estado, por ejemplo, para el estado de Hidalgo, con un padrón de 1'037'024, del total de la muestra, tendrá una participación de 0.0219.

Con esto, de alguna forma se está estratificando la muestra, una vez que se tenga el tamaño de la muestra, es posible determinar el número de secciones por estado, de ese modo el siguiente paso será seleccionar aleatoriamente el número de secciones por entidad.

Es necesario mencionar que después de la división que se realiza por estados, existen algunos otros factores que debiesen tomarse en cuenta en la selección de casillas para la muestra, y éste es el tipo, ¿la sección es rural o urbana?, esto por trivial que pudiera parecer es un aspecto muy importante, pues como ya se mencionó, esta clasificación trasciende en una heterogeneidad de la población en estudio, y ésta debe ser representada en la muestra, de momento, no existen las herramientas para poder saber si una sección es rural o urbana, se puede suponer que en la selección aleatoria de dichas secciones, la probabilidad de que ésta sea urbana o rural está dada por su misma naturaleza.

Por otro lado, para no complicar el proceso haciendo una selección manual, se pueden elaborar algoritmos para diseñar la muestra de forma automática y rápida para todos los métodos de selección de muestra, pues no se pierda de vista que el marco muestral está definido en una base de datos.

En este momento existen varias herramientas que pueden realizar esto, como son: lenguajes de programación con rutinas orientadas a bases de datos que pueden ser de valiosa ayuda, inicialmente se trabajará en Clipper 5.1 para realizar dichos algoritmos, no porque sea la mejor opción, sino por ser una de las herramientas más prácticas.

Profundizando un poco en el muestreo aleatorio estratificado, debido a que una de las hipótesis de esta tesis involucra a esta técnica de muestreo, se analizará a grosso modo algunas de las propiedades de los estimadores, con la siguiente notación⁵.

| | |
|-------|------------------------------------------------------------------------|
| N_h | Número total de unidades del estrato h |
| N | Número total de la población $\Rightarrow N_1 + N_2 + \dots + N_L = N$ |
| n_h | Número de unidades en la muestra del estrato h |
| n | Tamaño de la muestra $\Rightarrow n_1 + n_2 + \dots + n_L = n$ |
| y_h | Valor obtenido para la i -ésima unidad del estrato h |

⁵ Notación propuesta por William Cochran, "Técnicas de Muestreo", p.126.

$$W_h = \frac{N_h}{N} \quad \text{Ponderación del estrato } h$$

$$f_h = \frac{n_h}{N_h} \quad \text{Fracción de muestreo en el estrato } h$$

$$\bar{Y}_h = \frac{\sum_{j=1}^{N_h} y_{hj}}{N_h} \quad \text{Media verdadera para el estrato } h$$

$$\bar{y}_h = \frac{\sum_{j=1}^{n_h} y_{hj}}{n_h} \quad \text{Media de la muestra del estrato } h$$

$$S_h^2 = \frac{\sum_{j=1}^{N_h} (y_{hj} - \bar{Y}_h)^2}{N_h - 1} \quad \text{Varianza verdadera del estrato } h$$

de aquí se puede deducir que la media de la población usada en el muestreo aleatorio estratificado es \bar{y}_{st} ⁶ y se describe como:

$$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{Y}_h}{N} = \sum_{h=1}^L W_h \bar{Y}_h$$

y la media muestral \bar{y}_h se describe como:

$$\bar{y} = \frac{\sum_{h=1}^L n_h \bar{y}_h}{n}$$

la diferencia entre \bar{y}_{st} y \bar{y} es que la primera se calcula tomando en cuenta los estratos y reciben su ponderación de w_h para el estrato h , y \bar{y} se calcula en base al ponderador del tamaño de la muestra, si ambos estimadores coinciden es porque:

⁶ Si, el subíndice indica que es la media "estrato".

$$\frac{n_h}{n} = \frac{N_h}{N} \text{ ó } f_h = f$$

Esto significa que la proporción del tamaño de la muestra para el estrato h es igual a la proporción del estrato con respecto a la población. A esta estratificación se le nombra estratificación con asignación proporcional y proporciona una muestra autoponderada.

Por otro lado, se tiene que si las muestras se extraen independientemente en los diferentes estratos entonces:

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 V(\bar{y}_h)$$

y si se tiene que la varianza de la media \bar{y} , de una muestra aleatoria simple es:

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{S^2 (N-n)}{n} = \frac{S^2}{n} (1-f_h)$$

entonces se puede demostrar que la varianza de la estimación \bar{y}_{st} en un muestreo aleatorio estratificado es:

$$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \frac{S_h^2}{n_h} (1-f_h)$$

pero si se toma una muestra aleatoria simple dentro de cada estrato, una estimación insesgada de S_h^2 es:

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

Aunque una mejor alternativa para la estimación de la varianza la proporciona el siguiente modelo.

$$s^2(\bar{y}_{st}) = \sum_{h=1}^L \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^L \frac{W_h s_h^2}{N}$$

Las fórmulas para calcular los límites de confianza son para la

media poblacional: $\bar{y}_{st} \pm ts(\bar{y}_{st})$

total de la población: $N\bar{y}_{st} \pm tNs(\bar{y}_{st})$

Las fórmulas antes presentadas suponen que \bar{y}_{st} está normalmente distribuida y que $s^2(\bar{y}_{st})$ está bien determinada de modo que el factor t puede encontrarse en tablas de la distribución normal.

2.5. Algunos aspectos que influyen en la precisión de la estimación y el error cuadrático medio como una herramienta apropiada para calcular la estimación.

En este subcapítulo se tratarán algunos aspectos que pueden llegar a alterar la muestra, por alguna razón completamente fuera de control.

Cabe recordar que la muestra debido al marco muestral no ha sido seleccionada por algún procedimiento razonado, es decir, no se seleccionaron las secciones que entrarán a la muestra mediante algún criterio de comodidad para el trabajo en campo. La muestra se diseñó mediante un algoritmo en la computadora que señala cuáles serán las secciones que entrarán en la muestra, de esta forma no es posible saber si la sección X se encuentra en alguna zona que por alguna razón sea inaccesible por alguna situación como las que se presentan en seguida.

Existen situaciones que complican o encarecen el trabajo de forma extraordinaria, para explicarlo se puede ejemplificar de esta manera, supóngase que en el caso que se está planteando alguna de las secciones se ubique en el poblado X, en el cual por alguna disputa entre partidos existan brotes de violencia recientemente, es posible que la presencia de una persona ajena a la casilla que esté observando, apuntando o preguntando algo, llegue a crear cierto recelo entre las partes en conflicto, lo cual puede llegar a considerarse una agresión, entonces se tendría un problema que quizá no afectase demasiado en el aspecto metodológico, si no se repite en muchas ocasiones.

Otro de los problemas que comúnmente se presentan en estos casos es el hecho de que la sección escogida por la computadora se encuentre en algún lugar distante o incomunicado, esto es un problema muy grave y de difícil solución, pues para empezar, esto encarece el proyecto, ya que para extraer los datos de esta sección se necesita más presupuesto que para las demás, de modo que se necesita más tiempo para que el encuestador la ubique, se consumen más recursos económicos, el riesgo es mayor, etc. Por otro lado si la comunidad en la que se encuentra la sección no tiene medios de comunicación, serían más las pérdidas que las ganancias, pues si la intención es realizar un conteo rápido, es seguro que la información de dicha sección podrá ser mandada a una central con algún retraso, por demás considerable, en donde lo más seguro es que dicha información no llegue a ser tomada en cuenta para el conteo final, y esto se refleja como una pérdida económica y de datos.

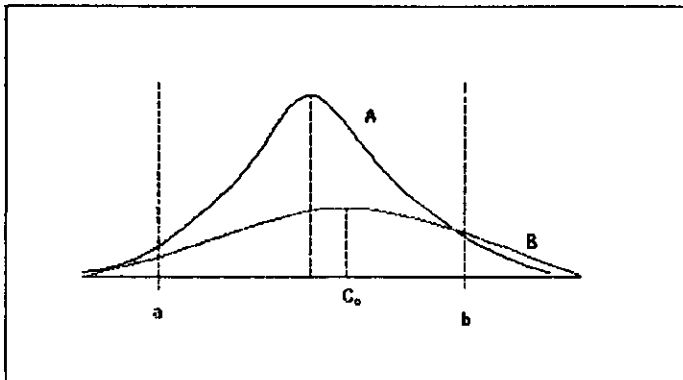
Un problema que comúnmente se encuentra en esta clase de procesos son las casillas irregulares, existen una serie de calificativos para éstas y pueden llegar a ser hasta cómicos, como el término "*Casilla Zapato*", esta clase de casillas se refiere a esas en donde el 100% de los votos son para una de las posibles opciones, es decir, todos los votos son para un sólo partido o candidato. Otro término conocido entre las personas que se dedican a esto es el de "*Mapaches*", este término se aplica a personas de conducta dudosa que se dedican a "*acarrear*" personas a una casilla, y así como estos términos existen muchos. Es evidente que por ahí existen ciertas estrategias para incrementar el número de votos movido por algún interés.

Esto de alguna forma puede llegar a alterar los datos resultantes de la casilla y por consecuencia también los del cómputo final del conteo. En apariencia, es parte del mismo proceso y no debería afectar, ya que en el conteo definitivo, estos votos ya sea correctos o no, serán contados.

De lo visto anteriormente es posible observar que, la muestra es diseñada por la computadora, en ese momento la teoría cumple su misión pero en la práctica se llega a romper esta armonía, ya que entran en el juego algunos factores prácticos de difícil solución.

Con el objetivo de medir el error que toda esta clase de factores puede producir y que se verán reflejados en las tendencias estimadas, se cuenta con una herramienta muy eficiente que es el **Error Cuadrático Medio, (ECM)**.

Se analizará el caso que se presenta en la gráfica 2.5.1.



Gráfica 2.5.1.

Si en este caso el valor que se está buscando es " C_0 ", se puede observar que de entre las dos distribuciones que se presentan, la distribución "B", es la que ofrece

una mejor estimación, pero el grado de concentración de las estimadas es mucho mayor en "A", en este caso, a pesar que el valor que se busca " C_o " prácticamente sea el centro de "B", es preferible la estimación ofrecida por "A", si es que la probabilidad de que la estimada de la muestra esté comprendida en el intervalo (a, b) , ya que como se puede observar, las probabilidades son mucho mayores en "A", que en "B". Con esto no se quiere decir que se prefiera un estimador sesgado a un insesgado, pero como lo indica Des Raj. *"En esas condiciones es preferible el estimador sesgado al insesgado."*

Desafortunadamente la varianza no es un criterio satisfactorio debido a que ésta mide las desviaciones del estimador obtenido y lógicamente el error de estimación no está considerado. Bajo esta situación, es razonable considerar las desviaciones con respecto al mismo C_o y calcular el valor esperado de sus cuadrados, a este cálculo se le llama **Error Cuadrático Medio (ECM)**, bajo este planteamiento el ECM se puede decir que para un estimador $\hat{\mu}$, está en función de μ y es:

$$\begin{aligned} ECM(\hat{\mu}) &= E(\hat{\mu} - \mu)^2 = E[\mu - E(\mu)]^2 + E[E(\mu) - \mu]^2 \\ &= V(\hat{\mu}) + [B(\hat{\mu})]^2 \\ &= V(\hat{\mu}) \left\{ 1 + \left[\frac{B(\hat{\mu})}{\sigma(\hat{\mu})} \right]^2 \right\} \end{aligned}$$

en donde: $V(\hat{\mu}) = \sigma^2(\hat{\mu})$.

$B(\hat{\mu})$ = representa el sesgo asociado con $\hat{\mu}$.

De modo que si $\hat{\mu}$ es insesgada, su varianza y el ECM coincidirán. En el caso de tener varios estimadores $\hat{\mu}_1, \hat{\mu}_2$, el que tenga un ECM menor en base al parámetro que se está estimando, es el mejor.

⁷ Des, Raj. *"Teoría de Muestreo"*, pág.43.

Debido a que es muy difícil poder asegurar que ningún sesgo se produce en las estimaciones, generalmente se habla de la precisión de un estimador más que de su exactitud, debido a que la exactitud se refiere a la magnitud de las desviaciones respecto a la media verdadera (μ), mientras que precisión se refiere a la magnitud de las desviaciones respecto a la media estimada ($\hat{\mu}$) mediante la aplicación repetida del procedimiento de muestreo.

Como se vió, a lo largo de este capítulo se han revisado las técnicas de muestreo más conocidas, describiendo brevemente cada una de ellas, se analizaron algunos modelos para calcular el tamaño de la muestra y se resaltó la importancia de este dato, pues como ya se analizó se pueden incurrir en errores que pueden traer consigo fuertes consecuencias negativas al sistema si el tamaño de la muestra no es el adecuado. Se han revisado conceptos de suma importancia como son el Marco Muestral y la Población Objetivo, aplicándolos a un caso real y práctico como fueron las elecciones del 21 de agosto de 1994. También se han hecho algunas observaciones sobre el diseño de la muestra, se han analizado las propiedades de las estimaciones, teórica y prácticamente y también se ha profundizado más en la Teoría del muestreo aleatorio estratificado para proporciones, por otro lado, de igual manera se han considerado algunos aspectos que podrían llegar a alterar la precisión en las estimaciones de las tendencias, y se ha propuesto el error cuadrático medio como herramienta para poder evaluar dichas estimaciones.

CAPITULO III

**EVALUACION DE LOS REQUERIMIENTOS COMPUTACIONALES, DESARROLLO Y
PUESTA EN MARCHA DEL SOFTWARE Y HARDWARE SELECCIONADOS.**

3.1. La aportación de la Ingeniería de Software para la selección y diseño del Sistema.

El objetivo de este capítulo es dar un panorama general de algunas de las herramientas disponibles para poder elegir la estructura computacional que será usada en el SEIT y dar los elementos con los que la persona encargada del departamento de cómputo o sistemas pueda elegir la mejor opción considerando las necesidades y restricciones con las que cuenta. Dado que día a día avanza la tecnología, al indicar una estructura o un software como el mejor, esta tesis podría quedar obsoleta en poco tiempo, por otro lado, en este momento no se sabe con qué recursos puede contar la persona interesada en realizar un trabajo como el que se plantea.

Como se mencionó en los temas anteriores, cuando se intenta diseñar un Sistema, no sólo es realizar el diseño conforme a lo que se necesita y ya, sino que hay que planearlo dependiendo de los recursos con los que se cuenta, el presupuesto, el tipo de datos que se tendrán de entrada, etc. La Ingeniería de Software aporta elementos para poder hacer un buen planteamiento, debido a que comienza por considerar aspectos de tipo real, como precisar: los recursos, el tipo de datos de entrada, la información que es requerida en la salida, los problemas a los que se puede enfrentar el diseñador, etc., todo esto tomando en cuenta la mayoría de los comentarios y sugerencias de los futuros usuarios, pues para ellos es que se está realizando el sistema.

Es necesario poder tener un panorama general de lo que es la ingeniería de software para poder evaluar la aportación de esta teoría al SEIT, para tal efecto se revisará a grosso modo algunos de los fundamentos de ésta. Se comenzará por conocer sus orígenes.

Las técnicas de programación son muy complejas y tienden a retrasar el trabajo, por lo que no es fácil realizar un proyecto de computación a gran escala. Muchas veces el software es entregado demasiado tarde y los programas no funcionan como el usuario espera, además, la mayoría de los programas tienden a tener cambios conforme pasa el tiempo, para aumentar sus funciones y/o para bajar tiempo y costos, pero en la mayoría de ellos no es fácil hacerles modificaciones, y por si esto fuera poco, la mayoría de los productos grandes de software son realizados por diferentes programadores.

Este tipo de problemas empezaron a manifestarse en los 60's, además de que los sistemas empezaron a ser muy grandes y complejos, por lo que la demanda computacional fue creciendo más rápido que la habilidad para producirlo, por lo que las necesidades de un enfoque sistemático para el desarrollo y mantenimiento del software provocaron la realización de dos conferencias una en 1968 y otra en 1969, en la que el término de *Ingeniería de Software* fue inventado, y aunque la construcción del software no es como la de edificios, puentes o casas, existe cierta relación en las bases para la realización de un proyecto en ambas ingenierías.

Al mejorar los métodos y técnicas para el desarrollo del software, puede haber grandes ahorros financieros, los sistemas se pueden ajustar mejor a las necesidades del usuario, son más confiables y tienen un mejor ambiente de trabajo, por lo que *Calidad y productividad* son los dos temas centrales en el campo de la ingeniería de software.¹

Ahora que se ha dado una idea de para qué puede ayudar la ingeniería de software, se darán algunas definiciones de la misma:

Una de las primeras definiciones fue dada en la primera conferencia NATO [Naur68]: *"La Ingeniería de Software es el establecimiento y uso de principios de*

¹ VAN VLIET, Hans. *"Software Engineering: Principles and Practice"*, Ed. John Wiley & Sons, 1993, England.

ingeniería válidas y aceptables para obtener un software económico que sea confiable y trabaje eficientemente en máquinas reales".²

La definición dada por la IEEE Standard Glossary of Software Engineering Terminology es la siguiente:

"La Ingeniería de Software es el enfoque sistemático para el desarrollo, operación, mantenimiento, y requerimientos del software",³ donde "software" es definido como: "Programas computacionales, procedimientos, reglas, y posible documentación asociada y datos concernientes para la operación de un sistema computacional".⁴

Otra definición es: *"La ingeniería de software es la disciplina tecnológica y directiva interesada en la producción sistemática y mantenimiento de productos de software que son desarrollados y modificados a tiempo y con estimación de costos".⁵*

Las metas iniciales de la ingeniería de software son para mejorar la calidad de los productos de software y para incrementar la productividad y el trabajo satisfactorio de los ingenieros de software.

Las características esenciales de la ingeniería de software son las siguientes:

² Op.cit., pág.5.

³ Op.cit.

⁴ FAIRLEY, Eichard E. *"Software Engineering concepts"*, Series in Software Engineering and Technology". Ed. McGraw-Hill, New York, México, c.1985, pág. 2.

⁵ Op.cit.

1. La Ingeniería de Software concierne a la construcción de programas grandes.

Existen dos tamaños en programas, los programas grandes y los programas pequeños. Los programas pequeños se refieren generalmente a los programas escritos por una persona en un tiempo relativamente corto. Los programas grandes se refieren a aquellos en los que trabajan muchas personas y que pueden tardar más de medio año. Los programas grandes no sólo son sostenidos por los lenguajes de programación, sino también por las herramientas (como diagramas de flujo) y métodos (como programación estructurada).

De hecho, para el desarrollo del software, el sistema cuenta con un gran número de programas interrelacionados, por lo cual un programa es una pieza del sistema.

2. El diseño del proyecto mediante módulos.

Como se vio en el subtema 1.3 (respecto a la Teoría de Sistemas), algunos problemas son tan grandes que no pueden ser observados desde su totalidad, sino que es mejor descomponer un sistema grande en sistemas pequeños, más subunidades manejables con interfaces bien definidas, así el problema total no decrece, pero se hace más manejable. En la ingeniería de software las unidades de descomposición son llamados "módulos".

3. La cooperación regular entre personas es una parte integral de un programa grande.

Cuando los proyectos son grandes, se debe administrarlo distribuyendo el trabajo en diferentes equipos, mediante métodos de comunicación, otorgando responsabilidades, etc. Se debe tener un común acuerdo con los

procedimientos y la estandarización de los mismos. Estos pueden ser soportados generalmente por herramientas.

4. Evolución del Software.

En general, los modelos de software parten de la realidad, y como ésta evoluciona, el software se hace obsoleto con bastante rapidez, es por ello que cuando se realiza un software, se tiene que pensar a futuro, para que los costos no se incrementen mucho al realizar la evolución (los cambios) del software.

5. La eficiencia del software desarrollado.

El costo total y el tiempo de desarrollo del proyecto del software y su mantenimiento es alto. La búsqueda de nuevas aplicaciones sobrepasa los recursos de la fuerza de trabajo. El hueco entre el suministro y la demanda está creciendo. Temas importantes dentro del campo de la ingeniería de software conciernen al mejoramiento y a la mayor eficiencia de métodos y herramientas para el desarrollo y mantenimiento del mismo.

6. Soporte efectivo del software para sus usuarios.

Cuando se desarrolla un software, es importante dar soporte técnico a los usuarios que lo estén utilizando, pues si no están satisfechos con el sistema tratarán de evitarlo o expresarán nuevos requerimientos inmediatamente. Es importante estudiar al usuario minuciosamente, para así determinar los requerimientos apropiados y funcionales, además de otros aspectos cualitativos como la fiabilidad, sensibilidad y líneas amigables de usuarios, entre otros. Además el desarrollo del software implica algo más que la entrega del mismo, se deben entregar también manuales de usuario y material de aprendizaje, todo esto pensando en el ambiente en el que el

nuevo sistema va a ser instalado.

Con la lista anterior, se muestra que la ingeniería de software tiene muchas facetas y que la ingeniería de software no es lo mismo que programar, aunque la programación es parte de esta ingeniería.

Entre otras ciencias que dan apoyo a la ingeniería de software para el desarrollo de un proyecto podemos mencionar, las matemáticas ya que validan las prácticas de ingeniería para darle utilidad a los productos, la psicología juega un papel importante en la comunicación entre las personas y las máquinas, o entre otras personas. Además, el proceso de desarrollo debe ser administrado.

El término "ingeniería de software" sugiere un posible parecido entre la construcción de programas y la construcción de casas o puentes.⁶

FASES EN EL DESARROLLO DEL SOFTWARE

Al construir una casa, se empieza el proyecto con el análisis de los requerimientos y las posibilidades del cliente, además de informarse sobre algunos factores del medio ambiente que rodeará la casa. La arquitectura toma estos factores en consideración cuando se diseña una casa. Sólo después del diseño se inicia la construcción, siempre que ambas partes (diseñador y cliente) estén de acuerdo y además si el diseño va de acuerdo al medio ambiente.

Para el diseño de un software se requiere de lo mismo, se inicia analizando el problema y describiendo los requerimientos en forma muy precisa. Se realiza después un diseño basado en los requerimientos y finalmente se hace el programa de solución. Para la realización de un proyecto de esta índole, se distinguen las fases que se ven en la figura 3.1.1.

⁶ VAN VLIET, Hans. *"Software Engineering: Principles and Practice"*, Ed. John Wiley & Sons, 1993, England, pág.5.

El llamado **modelo de proceso** representado en esta figura, es bastante simple. En realidad, las cosas suelen ser más complejas, puesto que cada fase está visto de forma global, pero cada una tiene su modelo, distinguiéndose en algunas varias subfases. Los componentes básicos, sin embargo, se siguen como en la figura anterior. La fase de diseño del modelo del proceso se pasa a través de todo el proyecto. Dependiendo del tipo de proyecto y del ambiente del trabajo, puede ser necesitado un esquema más detallado.

En la fig. 3.1.1, las fases han sido representadas secuencialmente, para un proyecto no es necesario separar estas actividades tan estrictamente como se indica en la figura. Estas partes pueden superponerse. Por ejemplo, es totalmente posible que la implementación inicie de una parte del sistema mientras que otras partes no han sido diseñadas totalmente todavía. No hay una estricta progresión lineal desde el análisis de los requerimientos para el diseño, hasta el diseño de la implementación.

- **Análisis de los requerimientos.**

Aquí se da una descripción completa de los problemas a resolver y de los requerimientos planteados por el medio ambiente en el que el sistema va a funcionar. Estos últimos pueden incluirse en el hardware y el software de soporte, o el número prospecto de usuarios del sistema a desarrollar. Alternativamente, el análisis de los requerimientos puede conducir a requerimientos impuestos por el hardware ya adquiridos, o a la organización en el que el sistema está funcionando.

Parte del análisis de los requerimientos es un estudio factible, cuyo propósito es evaluar si existe una solución al problema económica y técnicamente posible.

El documento en el que el resultado de esta actividad es establecido, es llamado **especificación de requerimientos**.

- **Diseño.** Durante esta fase, se desarrolla un modelo de todo el sistema, en el que mediante algún lenguaje de programación se resuelve el problema del usuario. Para este fin, el problema es descompuesto en piezas manejables -módulos-; las funciones de éstos además de las interfaces entre ellos, son especificadas en forma muy precisa.

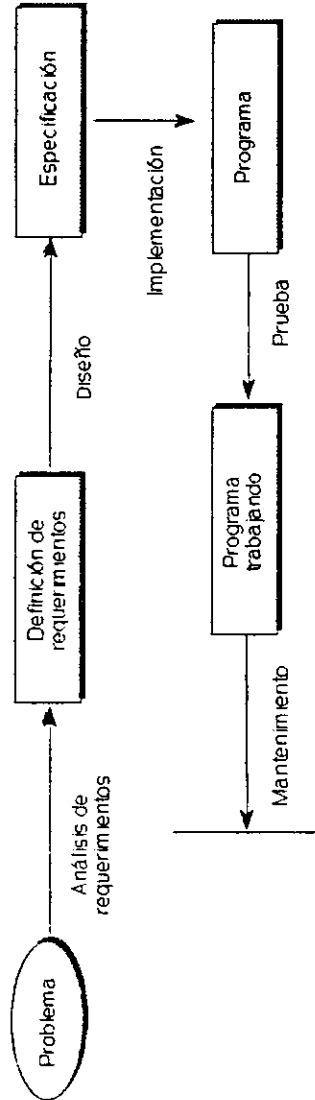
Durante la fase de diseño se trata de separar el qué del cómo. Se tiene que concentrar en el problema y no distraerse en la implementación. Métodos modernos de especificación se pegan a este principio. Ellos ofrecen posibilidades para formular las operaciones de un módulo matemáticamente. Estos métodos de especificación no producen algoritmos.

El resultado de la fase de diseño, es la especificación (técnica), que sirve como punto de inicio de la fase de la implementación. Si la especificación es de forma natural, también puede ser usada para pruebas de corrección.

- **Implementación.** En la fase de implementación, se concentra en los módulos individuales. El punto de inicio es la especificación del módulo. Es a menudo necesario que el paso de especificación de módulos para códigos ejecutables inicie muy grande en la introducción de una fase extra de "diseño", en cuyo caso, se puede tomar ventaja de algún lenguaje de alto nivel de programación, como pseudocódigo. (Un pseudocódigo es una forma de lenguaje de programación. Su sintaxis y semántica son en general menos estrictos que los de Pascal, así que esos algoritmos pueden ser formulados para un nivel más alto y abstracto).

El resultado de la fase de implementación es un programa ejecutable.

Figura 3.1.1. Modelo de Proceso.



- **Prueba.** Las pruebas no se aplican sólo después de la fase de implementación, sino a lo largo de todo el proyecto para poder detectar los errores tempranamente y que el costo por las correcciones resulte menor.

Existen dos límites en las pruebas: La prueba de la verificación, que es la que comprueba que de una transición desde la fase i a la fase $i+1$ siga correcta, y la de la validación, que es la que verifica que se continúa en el camino correcto considerando el desempeño de los requerimientos de los usuarios.

- **Mantenimiento.** Después de la entrega del software, aún se detectarán varios errores. Obviamente, estos errores se deben corregir. Además, el uso del sistema actual puede llevar a la solicitud de cambios y aumentos. A todos estos cambios se les denomina en conjunto, mantenimiento. Este concierne, por tanto, a todas las actividades necesitadas para mantener el sistema operacional después de ser entregado al usuario.

Existen algunos procesos no incluidos en la gráfica 3.1.1. que son igualmente importantes que los antes mencionados, sólo se hará mención a algunos de ellos como son:

- **La administración del proyecto** es una actividad que atraviesa todas las fases del proyecto. Como otros proyectos, el del software debe ser administrado correctamente para garantizar el producto solicitado en el tiempo de entrega y con el presupuesto acordado.
- **Documentación.** Los componentes clave de la documentación de un sistema incluyen el plan del proyecto, la calidad del plan, especificaciones de requerimientos, diseño, documentación y pruebas del proyecto.

Además del modelo simple existen otros modelos más complejos, entre ellos está el modelo del ciclo de vida del software.

Una de las dificultades de los usuarios es el expresar precisamente sus requerimientos. Este empieza a aclarar sus ideas a través de un prototipo que puede lograrse mediante el desarrollo de una interface adelantada para el usuario. El presunto usuario puede trabajar con un sistema que contenga los componentes interactuando, pero no mucho más extenso, al software que actualmente procesa la entrada del usuario. En este sentido, el usuario puede tener una buena impresión del futuro sistema que obtendrá, antes de que se haga una gran inversión en la actualización del sistema de este modo. El prototipo se convierte en una herramienta del análisis de requerimientos.

FASES EN LA PLANEACION DEL PROYECTO DEL DESARROLLO DEL SOFTWARE.

Antes de aceptar un proyecto de desarrollo de software, se debe cuidar su planeación. Hay que revisar, entre otras cosas las instalaciones que albergarán al sistema.

Como muchos otros aspectos del proyecto de desarrollo del software, la planeación no es una actividad corta. Más bien es de naturaleza altamente dinámica. El plan del proyecto puede ser establecido en un documento que servirá como una guía durante la ejecución del proyecto.

Cooper lista 13 puntos importantes y consecutivos para el plan del proyecto:⁷

1. **Introducción.** En la introducción del plan del proyecto, los antecedentes y la historia del proyecto son dados, junto con sus propuestas, los nombres de las personas responsables, y el sumario del proyecto, esto es parte de la

⁷ Op.cit., pág.24.

documentación del sistema.

2. **Participación del Usuario.** Las perspectivas del usuario serán de vez en cuando involucradas en el proyecto. El plan del proyecto tiene un estado cuya información, servicios, recursos y facilidades están suministrados por el usuario y otro cuando no están suministrados por éstos.
3. **Riesgos.** Los riesgos potenciales tienen que ser detectados lo antes posible. Siempre habrá riesgos: el hardware puede no ser entregado a tiempo, el personal calificado puede no estar disponible cuando se le requiere, información crítica falta cuando es requerida, etc. Es bastante ingenuo suponer que el proyecto de desarrollo del software corre tranquilamente. Regularmente, en cualquier campo existen problemas. Uno puede diagnosticar los riesgos de un proyecto de software tempranamente, y suministrar medidas para enfrentarlos. Conforme las dudas del proyecto empiecen a crecer, también lo harán los riesgos.
4. **Estándares, líneas de guía y procedimientos.** Los proyectos de software son proyectos grandes. Generalmente, mucha gente está involucrada. Por lo que se requiere una fuerte disciplina de trabajo, en la que cada persona involucrada sigue los estándares, líneas de guía y procedimientos sobreconvenidos. Además empiezan a manifestarse en papel, muchas de éstas pueden estar soportadas o reforzadas por herramientas.
5. **Organización del Proyecto.** Aquí, la relación con otros proyectos y la organización del proyecto mismo son enfrentados. Dentro del equipo del proyecto, varios papeles pueden identificarse: administrador del proyecto, comprobador, programador, analista, etc. Se tiene claramente delineado estos papeles y se identifica la responsabilidad de cada uno de ellos.
6. **Fases del Proyecto.** Existen muchas variaciones en el modelo del ciclo de

vida simple para el desarrollo del software. Para cada proyecto, se tiene que decidir sobre el modelo del proceso exacto a ser seguido, cuyas actividades serán emprendidas y cuyas partes serán identificadas, como acertemos si esas partes son alcanzadas, también lo serán sus caminos críticos. El proyecto debe dividirse en piezas manejables que puedan localizarse para los diferentes miembros del equipo. También tiene que estimarse, para cada fase y para todo el proyecto, el total del costo y del tiempo.

7. **Análisis de los requerimientos y diseño.** Debajo de este encabezado, los métodos y técnicas que será usadas durante el análisis de requerimientos, serán dados. Se debe indicar también los recursos y herramientas que se necesitarán para el soporte de estos métodos y técnicas.
8. **Implementación.** Se identifican los recursos y herramientas para el soporte de la implementación. Generalmente, también se describen aquí la forma de la versión y el control de la configuración para los componentes del software que están dados. Se debe tener cuidado con la documentación en esta fase, ya que aquí se produce una gran proporción de la documentación.
9. **Pruebas.** Describe las pruebas necesarias para el ambiente y para el equipo. Durante la prueba, normalmente se pone mayor presión en el equipo, por lo que esta actividad debe planearse cuidadosamente. Después de unir las pruebas, los diferentes componentes son integrados en el mismo orden.
10. **Recursos.** Durante la ejecución del proyecto, se necesitarán muchos recursos. Aquí son listados el hardware, ciclos de CPU y las herramientas que se necesitan para dar soporte al proyecto.
11. **Garantía de calidad.** Los diferentes aspectos de un Plan de Garantía de Calidad pueden también dividirse con documentos separados.

12. **Cambios.** Antes de ser instalado, los cambios son inevitables. Uno tiene que asegurarse que esos cambios estarán antes de dividirlos de una forma ordenada. Así se necesitan procedimientos claros cuyas propuestas estén a la mano. Cada propuesta de cambio tiene que registrarse y revisarse. Se tiene que ver el impacto en el costo estimado.
13. **Entrega.** Finalmente, los procedimientos del sistema son entregados al usuario final, junto con la documentación de los mismos.

Existen autores que no coinciden con el planteamiento anterior con respecto a la planeación de un proyecto de software, y plantean éste de la forma en que aparece en la tabla 3.1.2. que como se aprecia las diferencias son de forma y no de fondo.

Además de estos puntos, es importante el tener unas buenas anotaciones, ya que éstas pueden aclarar las relaciones e interacciones de interés.

Algunas consideraciones importantes para el desarrollo de un proyecto de software.

- Las anotaciones apropiadas proporcionan puntos de apoyo para la comunicación entre el personal del proyecto, además se tiene la posibilidad de emplear herramientas automatizadas para manipular las anotaciones y verificar su uso apropiado. Esto puede beneficiar a un proyecto particular, pero los mayores beneficios podrán resultar cuando se realicen esquemas de anotaciones bien definidos para ser adoptados por proyectos grandes.
- La calidad del software es una preocupación primaria de los ingenieros de software. Los atributos importantes de calidad para cualquier producto dependen de la naturaleza del mismo. En algunos casos, la utilización eficiente del espacio de memoria puede ser lo principal, mientras que para otros la

Tabla 3.1.2

| <u>Planeación de un proyecto de software</u> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Definición del problema. |
| 1. Desarrollar un informe definitivo para el problema a ser resuelto, redactado en la terminología del cliente. Incluir una descripción de la situación actual, limitantes del problema y un informe de las metas que se llevarán a cabo. |
| 2. Justificar una solución computarizada estratégicamente para el problema. |
| 3. Identificar las funciones que suministrarán por y con los límites planeados, los subsistemas de hardware, de software, y el de personas. |
| 4. Determinar las metas y los requerimientos del nivel de sistemas para el proceso de desarrollo y los productos de trabajo. |
| 5. Establecer los criterios de aceptación del alto nivel para el sistema. |
| Desarrollo de una estrategia de solución. |
| 6. Trazar distintas estrategias de solución, sin la consideración de restricciones. |
| 7. Llevar a cabo un estudio factible para cada estrategia. |
| 8. Recomendar una estrategia de solución, indicando porqué otras estrategias fueron rechazadas. |
| 9. Desarrollar una lista de prioridades para las características del producto. |
| Planeación de los procesos de desarrollo. |
| 10. Definir un modelo de ciclo de vida y una estructura computacional para el proyecto. |
| 11. Planear la configuración administrativa, garantía de calidad y actividades de validación. |
| 12. Determinar las fases dependientes de herramientas, técnicas y anotaciones a usarse. |
| 13. Establecer costos preliminares primarios para el desarrollo del sistema. |
| 14. Establecer un esquema de desarrollo preliminar. |
| 15. Establecer una lista preliminar del personal estimado. |
| 16. Desarrollo preliminar de los recursos computacionales requeridos para la operación y mantenimiento del sistema. |
| 17. Preparar un glosario de términos. |
| 18. Identificar los recursos de información, y referirlos a través del plan del proyecto. |

transportabilidad del producto de software entre las máquinas puede ser un atributo de suma importancia. Hay, sin embargo, algunos atributos fundamentales de calidad que cada producto de software debe poseer. Esto incluye *utilidad, claridad, fiabilidad, eficiencia, y costo-eficacia*.

El atributo más importante de la calidad de un producto de software es la *utilidad*; es decir, el producto de software debe satisfacer las necesidades del usuario.

- La **fiabilidad del software** está en "la habilidad de un programador para realizar una función requerida bajo condiciones estables en un periodo estable de tiempo."⁸ El grado de fiabilidad requerido para un producto particular puede expresarse en términos de la falla del producto.
- Los **productos de software deben estar claramente escritos y deben ser fáciles de entender**. Como se vio, las actividades de pruebas y mantenimiento consumen la mayoría del presupuesto del software. La clave para hacer una prueba al sistema y el mantenimiento es hacerlo entendible.
- Un **producto de software debe ser eficiente**, en los primeros días de la computación digital, el hardware era muy caro y muy lento a comparación de las máquinas estándares actuales, y un mayor énfasis fue puesto en exprimir el último bit y el último ciclo de memoria para cada programa; la eficiencia fue el primer atributo de calidad de los productos de computadoras. Como el software empezó a agrandarse y a ser más complejo, los atributos como utilidad, fiabilidad, y claridad asumieron mayor importancia para la mayoría de los productos de software.

⁸ FAIRLEY, Richard E. "Software Engineering concepts". Series in Software Engineering and Technology", Ed. McGraw-Hill, New York, México, c.1985, pág.6.

Para concluir este apartado, sólo se harán las siguientes consideraciones con respecto a la ingeniería de software.

El nivel de la tecnología utilizada en un proyecto de software explica factores como el lenguaje de programación, el ambiente de la máquina, las prácticas de programación, y las herramientas del software. Los lenguajes modernos de programación dan mayores facilidades para la definición y manejo de los datos, mejora en las estructuras para la especificación del control del flujo, facilidades en la modularización, definición al usuario del manejo de las excepciones, y facilidades para el programa utilizado.

El ambiente de las máquinas incluye la instalación del hardware y las facilidades del software disponible para el desarrollo, uso, y mantenimiento de un producto de software. La estabilización y disponibilidad del medio ambiente de la máquina ejerce una fuerte influencia en la productividad y calidad.

Se requiere de una planeación cuidadosa para el proceso de desarrollo y la obtención de productos de trabajo adecuados para evitar problemas como la pérdida del tiempo innecesariamente, costos fuera de presupuesto, baja calidad, y altos costos de mantenimiento para el software.

Cada vez hay mayor aceptación en los procedimientos y las técnicas en la ingeniería de software, pues al existir estas prácticas estándares se distingue una de las características de una disciplina profesional. Existen varias metas sistemáticas para el desarrollo y mantenimiento del software, pero debido a que esta ingeniería está en su infancia, muchos de los procedimientos y técnicas todavía no tienen madurez. Además es irrazonable esperar que un sólo método para el desarrollo y mantenimiento del software pueda ser aún adecuada para cubrir todas las situaciones, por lo que a menudo no es fácil distinguir el método que deba ser utilizado en cada situación en particular. Sin embargo, cualquiera de los métodos utilizados en la ingeniería de software son mejores que los métodos tradicionales.

3.2. Análisis de las diferentes estructuras computacionales que puedan integrar el sistema.

Entre las estructuras computacionales que pueden ser de utilidad para el SEIT se encuentran obviamente las redes, pero existen otras estructuras menos costosas y que pueden servir a los propósitos del Sistema, como la interconexión de computadoras por medio de algún programa de comunicación como el INTERLINK⁹, manejar una red mediante el sistema operativo UNIX, mandar la información vía módem teniendo varios centros de cómputo en cada centro de acopio, trabajar con PC's sin ninguna conexión, WINDOWS 95, etc.

Las estructuras computacionales que se tomarán en cuenta son las siguientes:

- Redes
- INTERNET
- Utilización de INTERLINK y
- Servicios en Línea (vía módem)

3.2.1. Redes.

El uso de redes hoy en día, permite compartir de forma eficiente un mismo hardware como son las impresoras o los fax, distribuyendo así los costos a más de un usuario, lo mismo sucede con el software, pues diferentes usuarios pueden compartir paqueterías, compiladores, etc., y lo que tal vez sea el mejor atributo en el manejo de redes, se puede compartir la información, como puede ser una base de datos o archivos tipo texto, gráficas, etc., pues la forma en que se comparte la información por este medio, facilita mucho las cosas al no tener que duplicar la

⁹ Interlink es una utilidad disponible en el MS-DOS v. 6.0 en adelante.

captura o ejecutar un programa muchas veces, o copiar en más de una ocasión el archivo texto o la base de datos donde se encuentra la información que se necesita. Tal vez se piense que esto no tiene importancia pues se podría resolver haciendo lo que ya se indicó, copiando la información o volviéndola a introducir por otros medios, pero hay que tomar en cuenta que dicha información puede ser alterada cada cierto tiempo y que la distancia en la que se encuentra una terminal de la otra puede ser bastante grande, o tal vez que la rapidez en la que se pueda acceder a esa información es vital, como en la venta de boletos de una línea aérea, por lo que en estos casos el manejo de una red es la mejor opción.

Una red puede ser tan simple, como cuando se conectan dos computadoras para compartir archivos e impresoras, o muy compleja al enlazar todas las computadoras de una compañía a nivel mundial.

Como es de imaginarse, esta estructura sería ideal para el SEIT en el manejo de los datos, pues su vía de acceso y de introducción de los datos es la forma más rápida y segura, aunque también tiene sus desventajas.

Se comenzará por definir los diferentes tipos de redes que existen actualmente.

- **"Red local (LAN)¹⁰**. - Se tomaron dos definiciones de redes locales, una red local *"Se entiende por un soporte de comunicación para la interconexión de sistemas en un área restringida, es decir, una cobertura geográfica de una red que está comprendida entre 0.1 km y 10 km"*¹¹. Ahora tomando la definición de Peter Norton, se tiene que: *"Básicamente una red de área local es una combinación de dos o más computadoras personales o estaciones de trabajo que están física y lógicamente conectadas entre sí. Una red puede incluir hasta un millar o más de*

¹⁰ Local Area Network.

¹¹ BELTRAO Moura, José Antão y PHILIPPE Sauvé, Jacques *"Redes Locales de Computadoras"*, México, 1a. Edición en Español, Ed. McGraw-Hill, 1992, pág.1.

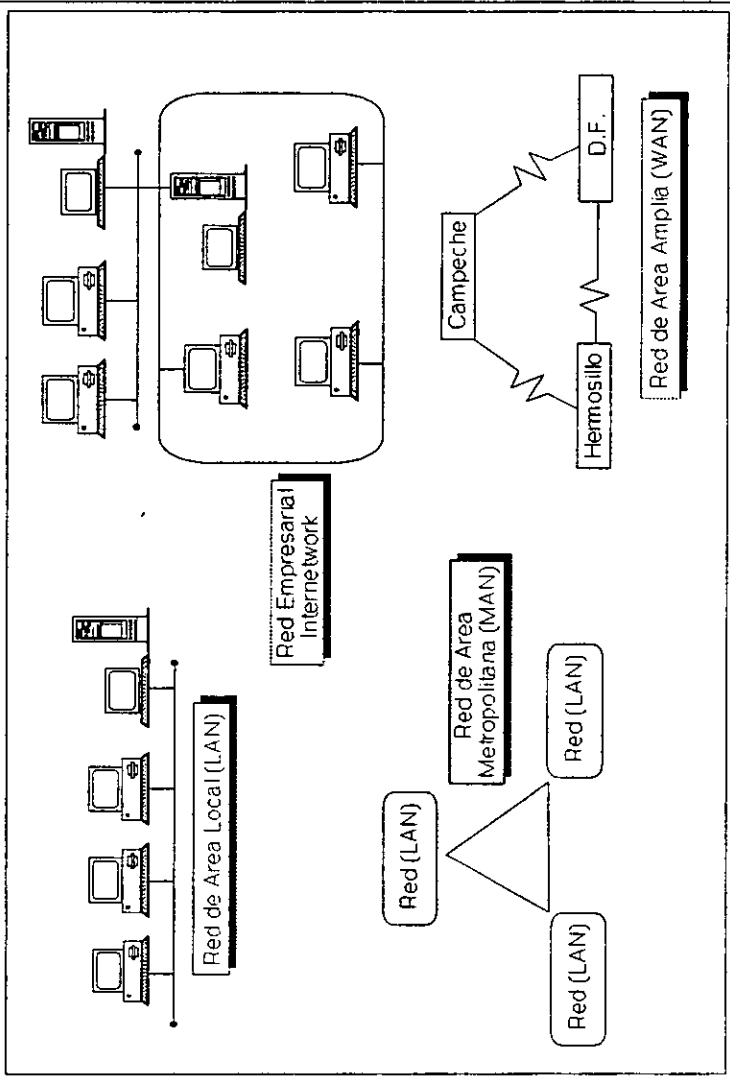
estaciones de trabajo o nodos".¹² Las LAN pueden encontrarse en el mismo edificio o en diferentes edificios siempre que estén cerca unos de otros, como se puede ver en la figura 3.2.1.

- **"Red de área amplia o larga distancia (WAN)"**.- Casi todas las WAN consisten de varias LAN físicamente separadas entre sí. Una WAN proporciona interconexión a grandes distancias, puede extenderse a varios edificios, ciudades, estados, y aun naciones. La interconexión entre LAN's se consigue por medio de líneas telefónicas, satélites, microondas y transmisión por fibra óptica. Las interconexiones de las WAN son mucho más lentas que las LAN pues generalmente la cantidad de tráfico es más baja que los segmentos normales de LAN. Una LAN típica transmite 10 Mb/seg, mientras que una WAN transmite 1 Mb/seg. Esto está cambiando con la introducción de la nueva fibra óptica de redes de comunicación nacional e internacional que provee más ancho de banda y velocidades de transferencia más rápidas. Esto se ve con mayor claridad en la figura 3.2.1.
- **"Red de área Metropolitana" (MAN)**.- Proporciona conexiones con redes y recursos a grandes distancias. Generalmente usan grandes velocidades conectando los segmentos de LAN con fibra óptica dentro de una área específica, como una universidad o una ciudad. Una MAN utiliza principalmente cables de alta velocidad que se conectan a los servidores directamente, como se puede ver en la figura 3.2.1.
- **"Red Empresarial"**.- Una red empresarial junta todos los sistemas computacionales dentro de una organización, no importa el sistema operativo que usen. Una red empresarial puede juntar minicomputadoras, mainframes, estaciones de trabajo de cualquier plataforma¹³ y otros recursos computacionales dentro de un

¹² NORTON, Peter. *"Periféricos y Accesorios para la IBM-PC, PS/2 y Compatibles"*. México, Ed. Prentice Hall, 1993, pág.208.

¹³ Una Plataforma es un Sistema Operativo.

Figura 3.2.1. Los rangos de las Redes.



sólo sistema interconectado. Ver figura 3.2.1.

- **"Mainframes"**.- Es una minicomputadora, o sea una computadora que consta de varias terminales. Estas terminales son llamadas "terminales tontas" porque no tienen procesador ni memoria propia, y tampoco cuentan con drives para poder guardar información en discos, así que el sistema operativo es el mismo para todas las terminales, además todos los servicios (paquetería, periféricos, etc.), procesos y los archivos con información o datos son almacenados en un mismo CPU. Los mainframes no son en sí redes, pero se pueden manejar como tales y/o pueden ser utilizados como sus elementos.

Generalmente, los recursos informáticos de los usuarios, esto es, equipo, software y la manipulación y producción de datos, tienen formatos incompatibles, por lo que son heterogéneos, debido a que normalmente son de diferentes proveedores y cuentan con características diferentes. Esto por un lado beneficia a los usuarios que pueden elegir entre diferentes sistemas para sus aplicaciones, dependiendo de sus necesidades y presupuesto, pero también se dificulta la interconexión de equipos de diferentes fabricantes.

La interconexión de redes aumenta el problema al existir diferentes servicios de transmisión en diferentes redes, por lo que se requirió de una serie de protocolos o normas para poder salvar esta dificultad, así se generó un modelo de arquitectura estratificada en categorías llamado modelo de referencia para la interconexión de sistemas abiertos ("Reference Model for Open Systems Interconnection" RM-OSI, Modelo de Referencia para los Sistemas Abiertos de Interconexión).

El término "Sistemas abiertos" se emplea a los productos diseñados para trabajar juntos, transfiriendo información entre terminales, personas, redes, etc., que estén abiertos a los demás, adoptando las normas OSI, modelo puesto en marcha por la Organización Internacional de Estandarización (ISO).

El diseño de una red es muy complejo pues se tiene que pensar en todos los dispositivos que se van a usar, el sistema operativo, el tipo de cableado, los recursos, etc. Para simplificarlo se pensó en el diseño global de una red con diferentes niveles, cada nivel superpuesto al inferior; al diseñar cada uno de ellos, se simplifica el trabajo de desarrollo y de mantenimiento.

Para efecto de poder compartir información entre redes, existen siete niveles en OSI, aunque el único enlace físico entre las computadoras existe en el nivel físico (nivel 1). La información realmente fluye del nivel 1 a otro nivel a través de las líneas de comunicación y luego atraviesa los niveles hasta el nivel de aplicación (nivel 7). Estos niveles son:

1. **Nivel físico:** Responsable de la transmisión de la información entre dos puntos en un circuito, tales como un anfitrión y una terminal. Define las características mecánicas, eléctricas, funcionales y de procedimientos para activar, mantener y desactivar conexiones físicas para la transmisión de bits entre las entidades del nivel de enlace de información (nivel 2), posiblemente a través de sistemas intermedios y controla la velocidad de la información a través de la línea.
2. **Nivel de enlace de información:** Proporciona los medios funcionales y de procedimiento para activar, mantener y desactivar una o más conexiones de enlace de datos entre unidades de niveles de red. Controla el cierre de circuitos en el nivel 1. Detecta y trata de corregir los errores del nivel 1, organizando los bits de información en bloques o marcos administrables. Envía estos bloques en secuencia a su destino y espera un acuse de recibo. En caso de que la información no se reciba correctamente en el otro extremo, retransmite los bloques corrompidos o malos. Este nivel puede ser manejado por software o hardware.

3. Nivel de red: Decide adónde debe ir la información y qué camino debe tomar a través de una red por medio de una trayectoria de conexión entre una pareja de entidades del nivel de transporte, pasando tal vez por intermediarios, además mantiene y termina esa conexión. También toma un mensaje y lo reformatea en paquetes, dirigiéndolo a su destino apropiado.

4. Nivel de transporte: Proporciona un servicio de transferencia entre entidades del nivel de servicios al dar la conexión lógica para la transferencia de la información entre los nodos de la red y controla el flujo de los paquetes de información, identificando a sus usuarios sólo por su dirección de transporte y no por su localización. Está encima del nivel de red para quitarles a los niveles superiores la tarea de transporte de datos entre ellos. Las conexiones que suministra, pueden establecerse en una única subred de comunicaciones o a través de un número de subredes en serie. El nivel (3) se encarga de todos los detalles de realización de paquetes, si la conexión ofrecida por el nivel de red es fiable y económica, las funciones en el nivel 4 quedarán proporcionalmente reducidas. La calidad del servicio ofrecido depende de la calidad del servicio de la red disponible, pero es responsabilidad del nivel 4. Todos los protocolos definidos en el nivel de transporte tienen significado fin-a-fin, es decir, de una entidad de transporte a otra correspondiente, sin sistema remoto.

5. Nivel de sesión: Los programas de software en este nivel administran la sesión completa de comunicación, estableciendo y terminando una sesión, y sincronizando el flujo de información entre entidades del nivel de presentación de comunicantes a través del uso de una conexión de transporte. También es responsable de los registros de entrada y de salida, y de las características de seguridad para el acceso, tal como procesamiento de contraseñas, y la transferencia de archivos entre nodos. Los servicios ofrecidos por este nivel son los primeros a realizar con las aplicaciones propiamente dichas, después del servicio de comunicación.

Los servicios del nivel de sesión se clasifican en dos categorías:

- a) "Servicio de administración de sesiones," une dos entidades para una relación y que más tarde las desune.
- b) "Servicio de diálogo de sesión," controla una transferencia de datos, delimita y sincroniza operaciones con los datos entre dos entidades de presentación.

6. Nivel de presentación: Maneja el formato y el despliegue gráfico de información con la generación de caracteres especiales para las impresoras, las trazadoras y las terminales CRT, como lo hace Microsoft Windows. También realiza los servicios que pueden ser relacionados por el nivel de aplicaciones para la interpretación de la sintaxis de los datos transmitidos. Esos servicios gestionan la entrada, transferencia, presentación y control de los datos estructurados. Este nivel resuelve problemas de diferencia de sintaxis entre sistemas abiertos en comunicación.

7. Nivel de aplicación: Sirve como una interfaz o ventana entre los usuarios de la red. Básicamente, todos los programas de aplicación (hojas de cálculo, bases de datos, etc.) forman parte de éste nivel produciéndose la transferencia de información para esos usuarios. El resto de los niveles existen para dar soporte a éste. Sus servicios son utilizados por los propios usuarios en el ambiente OSI. Cada usuario es representado para los demás por su entidad de aplicación correspondiente. Es importante señalar que no debe residir toda la aplicación final en este nivel, sólo la parte que necesita comunicarse con las entidades remotas, para ello utiliza los protocolos de aplicación. El resto de la aplicación no encuadra en el modelo de referencia para la interconexión de sistemas abiertos.

Complementando lo anterior, se puede mencionar que los protocolos están divididos en dos grandes ramas:

- Los **PROTOCOLOS DE BAJO NIVEL DE DESEMPEÑO**, que a su vez se dividen en protocolos de acceso al medio que controla las comunicaciones a través del medio, el protocolo de enlace que regula una comunicación entre interfaces y el protocolo de acceso a la red que especifica y supervisa las iteraciones entre una interfaz y su usuario¹⁴.

Estos protocolos son llamados también de alzada de una subred de comunicaciones.

- Los **PROTOCOLOS DE ALTO NIVEL DE DESEMPEÑO**, son los protocolos que definen y supervisan una comunicación entre usuarios. Estos protocolos además de dedicarse a los usuarios de la red, también lo hacen con sus operaciones y se aplican a la comunicación entre usuarios (entre los puntos "finales" de la comunicación).

Se puede decir que los protocolos de bajo nivel de desempeño son los que se encuentran en los tres primeros niveles de la red, y los de alto nivel de desempeño son los que se encuentran en el resto de los niveles.

Para la instalación de una red, es necesario contar con cierto tiempo para llevar a cabo las pruebas de su instalación, empezando con los nodos o terminales, su comportamiento con el servidor, la conexión de los cables, la forma en que se realizan los respaldos, etc., esto es para una red local, pero si se realiza una red a larga distancia, es un poco más complicado, pues se tiene que ver los medios en que se va a mandar la información, (teléfonos, módems, etc.)

Existe ahora una nueva tecnología llamada HPSN (High Performance Scalable Networking - Alto Desempeño en Redes Escalables), desarrollado por "3Com", esta

¹⁴ BELTRAO Moura, José Antão y PHILIPPE Sauvé, Jacques *"Redes Locales de Computadoras"*, México, 1a. Edición en Español, Ed. McGraw-Hill, 1992, p.p. 2-3.

tecnología permite que un producto entregado para operar sea flexible, administrable, controlado y económico, además de que cuente con capacidad para crecer paulatinamente. Para lograr el crecimiento progresivo de la red, la plataforma HPSN plantea tres escenarios, pudiendo poner en marcha cada uno de ellos, ya sea por separado o combinados, y en el orden que sus requerimientos exijan.

El primer escenario consta de tender una malla virtual a través de un backbone que por medio de enlaces privados de gran velocidad interconecta los despachos y sucursales regionales.

El segundo es aplicable a un modelo de empresa matriz a sucursales suministrando conectividad punto-multipunto, en la que los sitios remotos suelen tener redes locales que por lo general no poseen soporte propio del área de sistemas, por lo que se consideró poner equipos fáciles de instalar, usar y administrar en forma centralizada.

El tercer escenario es para usuarios móviles, pequeñas oficinas y usuarios en casa, los cuales requieren poder emplear todos los servicios de red de la casa matriz, utilizando módems y enlaces telefónicos. Para estos ambientes hay que contar con medios multiprotocolo y multiacceso para cada caso.

Al querer instalar una nueva red, se tienen que ver los costos de cada terminal, la tarjeta que se va a utilizar para su conexión con la red, los cables, el software, equipo de respaldo, etc.

Una red que consta de unas cinco terminales, tiene un costo promedio de más de 2,500 dólares, aparte del costo de las computadoras.

Como se puede ver, la instalación de una red consume demasiados recursos, además de requerir personal especializado para su instalación y operación. Para la realización de un sistema como el SEIT, se requiere de una red que esté funcio-

nando perfectamente, con bastante tiempo de anticipación, si se cuenta con una red, puede ser de gran utilidad, pero todavía existirían algunos ajustes que habría que realizar, como el hecho de contar con las terminales necesarias y el equipo de software que más se apegue a las necesidades del sistema, si se cuenta con todo esto, lo mejor es aplicar una red, pero si no es así, se tiene que hacer un estudio para calificar las opciones y ver qué estructura computacional es la más conveniente.

"INTERNET."

Actualmente, el INTERNET se encuentra de moda, ya casi todas las actividades cotidianas se pueden realizar por medio de una computadora conectada a INTERNET. Afortunadamente Internet cuenta con protocolos amigables de conexión y no con los complicados que son indispensables para poder acceder otra clase de redes cualquiera.

El hecho de que sea más fácil se debe a que es parte de una red a nivel mundial y que no se requiere de tanto equipo como lo puede ser una red que requiere de que funcione como servidor de procesos y una persona con los conocimientos suficientes para el diseño y manejo de ésta, convierte al Internet en una gran opción. De hecho, en las pasadas elecciones federales del 6 de julio de 1997, se pudieron observar los resultados obtenidos por el IFE conforme se realizaba el procesamiento de los datos, en un sistema que ellos llamaron PREP¹⁵.

Se sabe que una red se forma por medio de dos o más computadoras conectadas entre sí permitiendo el intercambio de información, si varias computadoras están enlazadas a manera de red, entonces todas pueden utilizar archivos y programas que tiene cada una por separado de manera simultánea. Ahora bien, como ya se

¹⁵ PREP: Programa de Resultados Electorales Preliminares.

explicó antes, existe una computadora que concentra los archivos principales, esta computadora es el "servidor" de la red, y las otras computadoras que se conectan a ella son los "clientes". Un servidor de una red se puede conectar a otro servidor de otra red, formándose así redes de redes o interredes.

De acuerdo a lo anterior, "INTERNET, es la suma de todas las interredes conectadas entre sí. Una computadora conectada a Internet, se puede conectar a las otras que se encuentran enlazadas a la red, sin importar el lugar geográfico en que se ubiquen."¹⁶

Haciendo un posible planteamiento para el SEIT, es tentadoramente factible resolver el problema de comunicación mediante INTERNET debido a que es completamente posible mandar archivos mediante E-MAIL'S o correos electrónicos, además como ya se mencionó, los resultados pueden ser publicados por este medio.

3.2.2. Utilización de INTERLINK.

Con esta opción se pueden conectar dos computadoras por medio de un puerto paralelo o serial, y permite a las computadoras compartir discos duros y el puerto para la impresora, además de controlar ambas desde el teclado de una.

Se pueden manejar varias PC'S con poca memoria RAM y con poca capacidad en disco duro, además de que se tiene la ventaja de que existe una gran diversidad de software que puede ser de gran utilidad. Otra gran ventaja es que no se necesita mucho tiempo para poner a funcionar varias computadoras bajo este esquema de trabajo, pues no se necesita estabilizar ninguna red, ni se necesita mucho tiempo para realizar pruebas, ni tanto equipo como para una red.

¹⁶ RENDÓN Ortiz, Gilberto. "INTERNET para principiantes", Cuarta reimpresión, De. SELECTOR, 1997, pág.23.

En caso de elegir esta opción, se necesitará además de las máquinas, varios multiplexores y cables, en el hardware.

El INTERLINK es una de las tantas utilerías del MS-DOS V. 6.0, y su manejo realmente es mucho más sencillo que el de una red, pues no se necesita estudiar tantos protocolos y paquetes nuevos, además de que se puede trabajar en un ambiente completamente conocido para la mayoría de los usuarios de PC's, lo cual es muy importante, debido a que no será necesario estudiar o familiarizarse con otro tipo de software, lo que permite planear y organizar el operativo en un tiempo mucho menor que en el caso de instalar una red, claro que también es necesario familiarizarse con este tipo de estructura, con sus comandos y variantes.

A continuación se describe la sintaxis del comando INTERLINK, en la versión del sistema operativo MS/DOS 6.0, además de aclarar algunos aspectos básicos para su operación.

Para poder ejecutar la utilería INTERLINK, es necesario que éste se localice por medio de un comando <DEVICE> o <DEVICEHIGH> en el archivo CONFIG.SYS de las computadoras en uso, la sintaxis de esta utilería es:

```
DEVICE = [drive:][path]INTERLNK.EXE[/DRIVES:n][/NOPRINTER][COM[:]  
[n|dirección]][/LPT[:][n|dirección]] [/AUTO][/NOSCAN][/LOW]  
[/BAUD:frecuencia][/V]
```

Los parámetros: [drive:][path], especifican la localización dentro del disco duro del archivo "INTERLNK.EXE", a continuación se realizará una breve revisión de los parámetros:

/DRIVES:n

Especifica el número de drives de redireccionamiento. Por default, el número es 3. Si especifica 0, Interlink redirecciona sólo impresoras.

/NOPRINTER

Especifica la impresora que no será redireccionada cuando instale INTERLNK.EXE.

/COM[:][n|dirección]

Especifica un puerto serial para usarlo en la transferencia de datos. El parámetro *n* especifica el número del puerto serial. El parámetro de *dirección* especifica la dirección del puerto serial. Si omite *n* o *dirección*, el Interlink del cliente buscará todos los puertos seriales y usará el primer puerto que encuentre conectado con el servidor.

/LPT[:][n|dirección]

Especifica un puerto paralelo para la transferencia de datos. Los parámetros *n* y *dirección* funcionan como en /COM, sólo que en vez de puerto serial es paralelo.

/AUTO

Instala el programa INTERLNK.EXE en la memoria sólo si el cliente puede establecer una conexión con el servidor cuando el cliente se da de alta.

/NOSCAN

Instala el programa INTERLNK.EXE en memoria, pero impide la estabilización de una conexión entre cliente y servidor durante el setup.

/LOW

Busca el programa de INTERLNK.EXE en la memoria convencional.

/BAUD: frecuencia

Pone una frecuencia máxima de baud para una comunicación serial.

Los valores válidos para las frecuencias son 9600, 19200, 38400, 57600 y 115200. Por default es de 115200.

N

Previene conflictos con el cronómetro de la computadora.

Es posible y necesario renombrar los drives y el disco duro de alguna de las máquinas conectadas, para facilitar el trabajo y para evitar conflictos a la hora de trabajar, éstos se hace de la siguiente manera:

INTERLNK [cliente [:] = [servidor][:]]

Los parámetros:

cliente

Especifica la letra del drive del cliente que es redireccionado para un drive del servidor del Iterlink.

servidor

Especifica la letra del drive en el servidor del Interlink que será redireccionado. El drive será enlistado en la columna del servidor de la pantalla del Interlink del servidor.

En caso de perder la corriente eléctrica, si es que no se cuenta con no-brakes¹⁷, existe la posibilidad de que no todas las bases se dañen, también es posible procesar la información y al mismo tiempo se continúa capturando en otras terminales. La desventaja principal de utilizar el INTERLINK, es que es lenta en comparación con una red o Internet, ya que la base de datos está fraccionada en diferentes máquinas y es necesario ejecutar algún procedimiento para unirlos en una base.

¹⁷ No-brakes.- Abastece a la computadora de corriente durante un lapso mínimo de tiempo en caso de que se corte la energía eléctrica. Ver capítulo 4.4.

3.2.3. Servicios en Línea (Vía módem).

Cuando se utilizan los servicios en línea, una computadora *A* se conecta con otra *B* por medio de un módem (MODulador DEModulador) y un programa de comunicaciones que conecta una PC a la otra usando una línea telefónica. La pantalla de la computadora *A* y su teclado actúan como una ventana en las operaciones de la computadora *B*. Conforme se introduzcan comandos en *A*, la otra máquina los llevará a cabo y le enviará el resultado para que sean mostrados en la pantalla de *A*. Así, a la computadora remota se le llama computadora anfitrión. En vista de que la máquina anfitrión realiza todo el procesamiento, no se requiere de mucho más que de un teclado y de una pantalla. Al dispositivo que ofrece estas funciones se le denomina terminal. Por lo que, todo lo que se necesita para tener acceso a una computadora anfitrión, es una terminal.

Aunque las máquinas que se utilizarán para este sistema son mucho más que simples terminales, ya que normalmente cuentan con memoria, discos y procesador propios. Sin embargo, cuando se conectan a una computadora anfitrión, éstas se comportan de la manera que el anfitrión espera. Para conseguirlo se utiliza un programa que hace que las PC "*terminales*" se comporten como tales. Se dice entonces que estas PC emulan terminales.

El enviar archivos de una computadora a otra es uno de los usos más comunes de las conexiones entre computadoras. Cuando la PC que se está usando envía un archivo a otra computadora, se dice que está "*cargando*" (uploading) el archivo, cuando la otra computadora envía un archivo a la computadora que se está utilizando, se dice que esta última lo está "*descargando*" (downloading).

Finalmente, cuando un programa de comunicaciones se está ejecutando y la computadora que se está utilizando está conectada con la otra, se dice que su PC

está en línea. Cuando no hay conexión, se dice que *está fuera de línea.*

Al igual que en el caso de las redes, para que no existan problemas con el envío de información entre computadoras, es necesario ciertos protocolos para que ambas computadoras envíen y reciban la información de la misma manera. Existen cuatro parámetros de comunicación básicos que describen el envío de la información. Cada computadora anfitrión espera que estos parámetros se establezcan de cierta manera, por lo que antes de contactar con un sistema remoto, se debe indicar al programa de comunicaciones qué parámetros utilizar. Estos son:

- Velocidad.
- Paridad.
- Número de bits de información.
- Número de bits final.

Casi todos los programas de comunicaciones mantienen un directorio de todas las computadoras con las que ha hecho conexión. En cada registro del directorio, almacenará el nombre del sistema y su número telefónico. Además, también incluye los parámetros por utilizarse al conectarse con ese anfitrión. Entonces, cuando se le pide al sistema que llame a un sistema en particular, el programa sabrá establecer los parámetros de comunicación correctos.

El primer parámetro, la **velocidad**, indica qué tan aprisa se envía la información entre los módems que interconectan a las dos computadoras. Este valor, depende del diseño tanto del módem que esté utilizando, como del módem anfitrión. La velocidad baja más común es de 2,400 bits por segundo (bps). La velocidad más común es de 9,600 bps, aunque algunos módems modernos ultrarrápidos funcionan a más de 28,800 bps. Ambos módems deben ajustarse para trabajar a la misma velocidad. El módem con velocidad más alta se autoajusta a la del de velocidad más baja.

El segundo parámetro, la **paridad**, describe un tipo primitivo de corrección de errores. Aunque existen varios posibles valores de paridad, los únicos con los que tal vez se tenga contacto son la paridad nula y la paridad par.

Cuando se interconectan dos PC's, casi siempre se utiliza paridad nula. Cuando se conecta con un anfitrión que no sea computadora personal o se conecte con una red que intercambie información en paquetes, es posible que a veces se tenga que usar paridad par.

El tercer parámetro es el **número de bits de información**, indica de cuántos bits consta cada carácter. En el caso de PC, cada carácter o byte consta de 8 bits, lo que da un total de 256 caracteres posibles. Sin embargo, en algunos sistemas, sólo se usan 7 bits. Estos sistemas utilizan sólo 128 caracteres.

Cuando se conecta con otra PC, normalmente se ajusta el número de bits de información a 8. Cuando se conecta con otros tipos de computadora, puede que se necesite cambiar el valor a 7. Usualmente 8 bits va junto con paridad nula, y 7 bits de información va junto con paridad par.

El último parámetro de comunicaciones, el **número de bits de final**, indica exactamente eso. Este parámetro es una señal utilizada para marcar el final de cada carácter. En comunicaciones con PC, siempre se utiliza 1 bit de detención o bit final.

Normalmente la velocidad de transmisión será de 2,400 o 9,600 bps, por lo que se refiere a los otros parámetros, se utiliza casi siempre alguna de las combinaciones siguientes:

- 1.- 8 bits de información, paridad nula, 1 bit de detención (referido como 8N1).
- 2.- 7 bits de información, paridad par, 1 bit de detención (referido como 7E1).

8N1 es el más común. Algunas computadoras anfitrión son lo suficientemente flexibles para aceptar cualquiera de estas combinaciones. Si necesita conectarse con una computadora remota y no se está seguro de qué juego de parámetros escoger, casi siempre se tendrá éxito al probar con 2,400 bps y 8N1. En caso de encontrar dificultades, trate con 7E1.

Como se puede observar, cada opción tiene sus ventajas y desventajas, todo depende del tiempo y del presupuesto disponibles, siempre evaluando que la opción seleccionada cumpla con los mayores requerimientos de seguridad, revisando punto por punto el funcionamiento de la estructura empleada, tal vez se piense que se está exagerando en precauciones, pero a la hora crítica del proceso, ninguna precaución sale sobrando, pues cualquier falla puede causar una gran pérdida. Lo más importante en la selección de la estructura computacional, tiene que ver con los conocimientos del diseñador o encargado del departamento de cómputo, ya que muchas veces se cuenta con el mejor equipo disponible en ese momento, pero si el diseñador o encargado de cómputo no tiene la experiencia suficiente para usar y manejar ese equipo, no será posible explotarlo apropiadamente, o será necesario un proceso de capacitación, lo cual es muy sano, salvo en el caso que no se disponga de tiempo para ello.

"TOPOLOGIAS."

Una topología es la disposición física de la red, puede utilizarse para cualquiera de las estructuras computacionales que se mencionaron anteriormente. La topología depende de la forma en que sus nodos (terminales y periféricos) estén conectados entre sí. Las principales topología son:

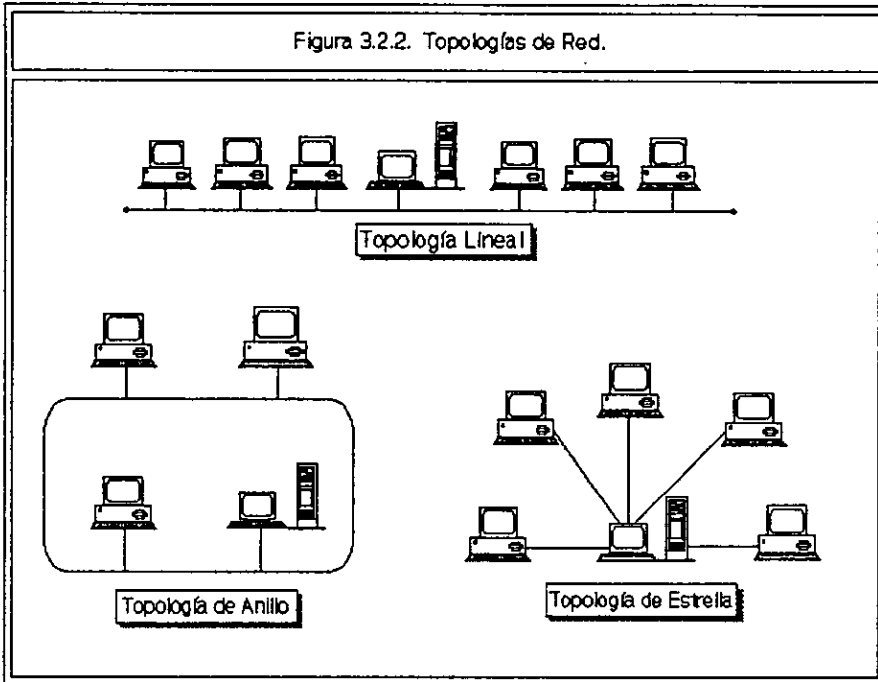
- **Topología de Bus o Lineal.**- En esta topología, un cable largo funciona como camino de la información. Todos los nodos se conectan al cable en distintas localidades. La información viaja hacia las estaciones de trabajo y desde ellas

inhabilita toda la red hasta que la ruptura sea corregida. Además, si varias estaciones de trabajo tratan de transmitir mensajes simultáneamente, la información puede chocar, teniendo que retransmitir la información.

- **Topología de anillo.-** Las estaciones de trabajo se conectan entre sí formando un círculo o anillo. En esta topología, la información se transmite de un nodo al siguiente. Cada estación de trabajo analizará la información, y si ésta no le corresponde, se transfiere a la siguiente. Como la información viaja sólo en una dirección, no existe colisión, pero cualquier ruptura en la conexión provocará la detención de toda la red.
- **Topología de estrella.-** Está construida de tal forma que todos los nodos se conectan a un núcleo central, por lo que no existe peligro de que se detenga toda la red si existe alguna ruptura, además de que se pueden añadir o quitar nodos, o cambiarlos de posición sin afectar la red y tampoco existe riesgo de colisión de información. Todo esto hace que esta topología resulte ser la más flexible.

En la figura 3.2.2, se esquematiza lo antes mencionado.

Figura 3.2.2. Topologías de Red.



3.3. Análisis del Software y de la Mejor Forma de Manejar los Datos y la Información.

En el caso del software, cada día se ven nuevos programas y paquetes computacionales, los cuales van incrementándose en sofisticación, por poner un ejemplo, hoy día existen procesadores de texto que tienen la capacidad de manejar bases de datos, graficadores, hojas de cálculo, etc., de forma clara y suficiente para cubrir todas las necesidades del usuario.

Para efecto de apoyar a la implementación del SEIT, se analizarán diferentes tipos de software, que puedan proporcionar una forma de trabajo eficiente para el proyecto.

Así como se vio que existen diferentes tipos de estructuras computacionales, también existen diferentes tipos de software, en materia de sistemas operativos, software para redes, bases de datos, graficadores, software para la comunicación, etc. Se tratará de hacer un recorrido por algunas opciones presentadas con estos elementos para tener un panorama general de las posibles aportaciones al sistema.

3.3.1. Sistemas Operativos.

Es necesario hacer un planteamiento del manejo de la información y conocer los recursos disponibles para poder seleccionar la estructura computacional y el sistema operativo, lo cual debe hacerse a la par.

- **MS-DOS.**

Es una de las opciones con más posibilidades, debido a que el personal que opera computadoras, normalmente está bien identificado con este sistema, además de que es muy amigable, y por otro lado, la disponibilidad de software es muy grande, por lo que es prácticamente fácil conseguir programas de comunicación, bases de datos, paquetes graficadores, etc., con lo que se pueden diseñar de forma práctica los procesos para tratar la información y presentar resultados; además de que los recursos a nivel computacional que se requieren son mínimos, este sistema se adapta a casi cualquier tipo de equipo.

Entre las desventajas que se pueden considerar, está el hecho de que es monotarea, es decir, no es posible ejecutar más de una aplicación a la vez, por otro lado, también existe el riesgo de contraer un virus que puede, en un

momento determinado, dañar las aplicaciones y llevar a mal término el proyecto.

- **WINDOWS 95.**

Este sistema operativo es muy eficiente en el manejo de gráficas, las aplicaciones para trabajar en él comienzan a hacer su aparición y se caracterizan por su calidad gráfica y facilidad de operación, además las aplicaciones que ya existían para versiones anteriores de Windows también pueden ser ejecutadas en este sistema. Por otro lado, con éste se emula un ambiente multiusuarios de modo eficiente, con lo que es permitido ejecutar varias aplicaciones al mismo tiempo, además de poder trabajar en red.

Aunque el riesgo de contraer un virus sigue latente y las especificaciones del equipo de cómputo no son tan básicas, se necesita de una computadora que tenga un procesador al menos 80486, 8 MB en memoria RAM, monitor mínimo VGA y más de 200MB de espacio en disco duro; normalmente las máquinas que operan bajo este sistema operativo miden la capacidad de disco duro en Giga Bytes, y el hecho de ser la primera versión de Windows con este concepto no está completamente libre de fallas, además de que el personal que domine completamente este sistema seguramente todavía es escaso.

- **UNIX.**

Este sistema es una de las mejores opciones de trabajo, es un sistema amigable, multiusuario, optimiza de forma completamente eficiente los recursos de la computadora, hasta el momento no se han dado casos de virus para dañar las aplicaciones en el sistema, aunque por otra parte no puede trabajar en una PC, se requieren equipos de cómputo muy grandes para poder operar, además de que el costo del mismo sistema es elevado, lo cual complica el planteamiento, pero afortunadamente existe un versión de UNIX, para PC's,

esta versión es LINUX, este sistema se autodenomina un "clon de UNIX", fue desarrollado por un estudiante europeo, y actualmente está liberado en Internet. Trabaja de forma muy parecida a UNIX, sólo que es capaz de ejecutarse satisfactoriamente en una PC, con un procesador 80386.

Las únicas desventajas que se le pueden encontrar a este sistema es que actualmente es muy escaso el personal capacitado en su operación, además de que las aplicaciones son escasas, algunas de las aplicaciones para MS-DOS, o Windows pueden ser ejecutadas en él, para lo cual deben ser consultadas las especificaciones del productor.

Cabe hacer la acotación de que si se decide trabajar en red, es posible que las terminales estando en operación y conectadas, se encuentren trabajando cada una en su propio sistema operativo.

3.3.2. Bases de Datos.

- **CLIPPER.**

Clipper es un lenguaje especializado para manejar bases de datos, está diseñado en lenguaje de programación "C", y tiene muchas similitudes en su codificación a la hora de programar, trabaja en ambiente MS-DOS, con mínimas exigencias de equipo de cómputo, con este lenguaje de programación es posible programar un sistema para administrar bases de datos de forma eficiente, además con un mínimo de esfuerzo se pueden hacer ejecutables. Es posible también relacionar varias bases de datos, y presentar la información de la forma necesaria, cuenta con una amplia librería de funciones matemáticas con la capacidad de realizar los cálculos necesarios, y se puede considerar que es una excelente opción para realizar algunos programas de apoyo al SEIT.

- **FOX PRO.**

Este software brinda una serie de facilidades para diseñar sistemas administradores de bases de datos, se cuenta con varias versiones, algunas para ambiente MS-DOS, y otras para Windows, los requerimientos dependen de la versión que se esté empleando, considerando la versión 2.6, se requiere de una PC, que tenga al menos 4 MB en memoria RAM, monitor VGA, 40 MB de espacio en disco duro y mínimo la versión 3.1 de Windows. Con este manejador de bases de datos es posible diseñar un sistema administrador de bases de datos, explotando las propiedades gráficas de Windows, como son la introducción de dibujos, el uso de ventanas, etc. De igual manera se pueden realizar programas ejecutables, si se cuenta con la versión profesional, y no presenta ninguna dificultad instalar la aplicación en una red.

- **DATA EASE.**

Este software representa una opción muy eficiente para realizar el proceso de captura, debido a que es demasiado fácil crear una pantalla de captura capaz de controlar una base con el número de registros que se requiera, la realización de reportes se puede diseñar mediante el DQL, que es una especie de diseñador automático de reportes, en el que el operador lo único que tiene que hacer es definir las posiciones de las variables, poner encabezados, direccionar la salida a una impresora, archivo, etc., trabaja de excelente forma en un ambiente compartido de red, también es posible relacionar bases de datos, automáticamente, tiene versiones para trabajar en Windows, en donde se diseñan las aplicaciones con las ventajas que esto brinda, realmente es una base de datos que cuanta con las características suficientes para diseñar una aplicación con los requerimientos necesarios, los requerimientos computacionales dependen de la versión que se esté manejando, en la versión para MS-DOS, éstos son los básicos.

- **DBASE.**

Este programa es de lo más utilizado para realizar revisiones a las bases de datos, crear estructuras, visualizar los datos sin necesidad de cargar algún paquete, lo cual brinda una opción fácil y barata en recursos para manipular la información. Además es factible diseñar un programa de captura, lo cual no se recomienda por lo laborioso y poco práctico que resulta.

Como se puede observar, las opciones en materia de Bases de datos son muchas y prácticamente todas cumplen con las características necesarias para apoyar al diseño del SEIT, realmente el factor que más podría influir es la experiencia que pueda tener el programador manejando dicha base, lo cual podría ser de vital utilidad a la hora de diseñar una aplicación con las características particulares y necesarias para poder ser empleada por el sistema.

3.3.3. Graficadores.

Para el caso de los graficadores, sucede más o menos lo mismo que en el caso de las bases de datos, la mejor opción es la que se conoce, aunque lo ideal es buscar el que más se acerque a las necesidades del sistema y presupuesto.

- **HARVARD GRAPHICS.**

En su tercera versión es una de las mejores alternativas si se trabaja en MS-DOS, ya que contiene un menú de fácil manejo, con opciones para la creación de gráficas que es muy versátil, además de tener un área de dibujo que permite arreglar y añadir gráficos para una mejor presentación. Facilita también el manejo de las presentaciones, da diferentes opciones para que ésta no se haga monótona, aunque quizá una de sus desventajas mayores sea que no genera programas

ejecutables para la presentación, por lo que es necesario ejecutar primero el paquete y después la presentación.

Una de las principales ventajas de este graficador es que permite la creación de procedimientos, y también es capaz de actualizar la información contenida en una gráfica de presentación desde un archivo en formato de DBASE o inclusive en código ASCII, manejando funciones de tiempo. Además los requerimientos del equipo de cómputo necesarios son realmente mínimos, con una computadora con procesador 80386, trabajando con MS-DOS V. 5.0 o mayor, 2MB de memoria RAM, con monitor VGA se puede correr la presentación de modo satisfactorio.

Actualmente existe una versión para WINDOWS, lo que puede ser útil si se cuenta con el equipo apropiado, ya que la información puede presentarse de forma más atractiva en este ambiente. Trabaja de modo muy parecido a la versión para DOS, con algunas ventajas en su presentación, como son más funciones de tiempo, mejor capacidad para el manejo de dibujos que se inserten en la gráfica para mejorar la presentación y se puede agregar sonidos, aunque no permite importar datos de DBASE.

Los requerimientos del equipo de cómputo como es de esperarse crecen, ya que para esta versión es necesario que el equipo de cómputo cuente con procesador 80486, trabaje en ambiente WINDOWS, tenga al menos 4MB en memoria RAM y monitor VGA para poder presentar de forma eficiente la presentación.

- **WORD PERFECT PRESENTATION.**

En su versión 2.0 también representa una buena opción, y en algunos aspectos llega a superar a HARVARD GRAPHICS, ya que inclusive permite limpiar imágenes digitalizadas, también contiene menú con diferentes tipos de gráficas, permite la importación de una mayor variedad de archivos, tiene la facilidad de

que los gráficos diseñados en este programa se integren al procesador de textos WORD PERFECT, además que es posible crear un archivo ejecutable con la presentación, aunque por otro lado, ocupa más memoria y más espacio en disco duro, y la presentación no se puede diseñar con la facilidad de HARVARD GRAPHICS, ya que si por un lado permite casi el mismo tipo de diseño, por otro lado, es necesario hacer una mayor cantidad de movimientos para su realización. Sus requerimientos son: una máquina 80386, trabajando con MS-DOS V. 5.0 o mayor, 2MB de memoria RAM como mínimo y monitor VGA.

- **COREL DRAW.**

Es un excelente graficador, de hecho su última versión CorelDraw 6 permite crear animaciones y efectos de perspectivas tridimensionales con una calidad similar a la realidad, casi fotografía. Además permite hacer presentaciones con elementos de multimedia, e incluye gráficos de barras, animaciones y sonido a las presentaciones. Este programa de diseño necesita de WINDOWS 95, por lo que aprovecha muchas de las características que éste ambiente le ofrece. Requiere también mínimo de una PC 80486 con 8 MB de RAM (16 MB recomendados), CD-ROM y video SVGA. La instalación completa ocupa 80 MB en disco duro, aunque si se desea correr las aplicaciones directamente desde el CD, se requieren de 20 MB, y lógicamente requiere de que la PC trabaje en WINDOWS 95.

- **EXCEL.**

Es un excelente graficador, permite la inserción de ilustraciones dentro de la gráfica y cambiar de paquete y de archivo de forma rápida, ya que se encuentra en ambiente WINDOWS, por lo que se puede intercambiar información de una base de datos a una hoja de cálculo o a un área de gráficos de forma rápida, los requerimientos computacionales para Excel varían mucho dependiendo de la versión que se utilice, van desde una máquina 80386 con WINDOWS V. 3.1. con 2MB en RAM y monitor VGA, hasta una computadora con procesador

Pentium, con WINDOWS 95, 8 o 16MB en memoria RAM, monitor ultra VGA, etc.

- **POWER POINT.**

En su versión 7.0 para WINDOWS 95 es una excelente opción, ya que con este software es realmente fácil diseñar presentaciones, gráficas, etc., con una muy alta calidad, debido a que es posible tomar diferentes diseños e insertarlos en las diapositivas, se pueden combinar colores al gusto del diseñador, etc. Una de sus desventajas con otros paquetes es que no cuenta con funciones de tiempo para la presentación, ni la importación de los datos. Los requerimientos computacionales varían dependiendo de la versión de igual forma que en el caso de EXCEL.

3.3.4. Programas de Comunicación.

Los programas de comunicación se encargan de proveer todo lo necesario para conectar una PC con otra computadora vía remota, cuando se realiza una sesión típica de conexión, se empieza ejecutando el programa de comunicación, normalmente este automáticamente inicializa el puerto serial y el módem, una vez que esto se ha realizado, el programa está listo para llamar a la computadora remota, esta clase de programas normalmente mantiene un directorio de números telefónicos, en donde aparte contiene información como los ajustes a los parámetros estándar de comunicación, que incluyen velocidad, paridad, bits de información, bits de final; entonces para llamar a un sistema remoto, lo que se tiene que hacer es indicar al programa el registro apropiado en el directorio, y el programa se encargará de todo. Una vez establecida la comunicación, el programa verifica si existe un guión o lista de instrucciones predeterminadas que deba ejecutarse, en esta lista puede existir información, como por ejemplo, el nombre de la computadora y alguna contraseña para accederla, etc., llegando a este punto el programa se desaparece de la pantalla y se ejecuta en asincronía pasándole el control a la

persona que opera la máquina que realizó la conexión. Las operaciones más comunes son: la transferencia de archivos, y la consulta de la máquina vía remota.

Normalmente todos los programas de comunicación tienen las mismas características, es probable que como función adicional algunos tengan ayuda en línea, bitácora de llamadas, cambio entre voz y datos, archivos de captura, etc. De entre los programas más conocidos para este tipo de trabajo, se pueden mencionar: Quick Link, Blast (U.S.Robotics), SmartCom (Hayes), Norton Utilities, PC Tools (ambos de Symantec).

Existe una forma altamente eficiente de trabajar esta parte del sistema, y es desarrollando una aplicación a la medida, que puede ser diseñada para trabajar en cualquier ambiente, según se requiera, para esta alternativa es necesario programar en cualquiera de las diversas opciones de lenguajes de programación disponibles. La ventaja de emplear esta opción, es el hecho de poder manejar bases de datos y graficadores al mismo tiempo y poder transportar la información que se obtiene al efectuar el procesamiento de los datos a la pantalla en donde se localiza la gráfica con los resultados, de forma inmediata.

3.4. Requisitos mínimos para armar la Estructura Computacional.

Como ya se comentó anteriormente, lo que más influye en la selección del software y hardware del sistema, es el presupuesto, el tiempo y los recursos con los que se cuenta. Para la implantación de una estructura computacional se necesita pensar en el tipo de software y el hardware en forma paralela, para poder hacer un planteamiento adecuado.

El software que se va a utilizar, la planeación de la estructura computacional tiene que diseñarse pensando en la cantidad de datos que van a ser introducidos en la base, la dispersión geográfica de los datos, el presupuesto con el que se cuenta, el lugar donde se van a procesar los datos, la calidad del trabajo, etc., todo esto influye en el diseño de la estructura.

En cualquiera de las opciones que se maneje, será necesario incluir equipo de cómputo suplementario como: impresoras, módems, etc., para disponer de impresiones de los resultados y gráficas, también poder transferir los resultados al lugar donde se requiere la información, etc.

A continuación se analizarán algunas de las opciones en la estructura computacional para poder implantarlas en el SEIT, se conjuntará la estructura desde el punto de vista hardware y software y se analizará cada una para evaluar qué tan apropiada puede ser para integrarse al sistema.

3.4.1. Redes y sus componentes.

Una red de área local (LAN), se forma básicamente de las siguientes partes:

- **Despachador o Servidor de archivos.** Corre el sistema operativo de la red. Actúa como punto central de procesamiento y almacenamiento para las aplicaciones de la red, dejándolas disponibles a cualquiera de sus usuarios. Los periféricos compartidos también están conectados a este Servidor. Es una computadora de gran potencia pues cuenta con un disco duro de gran capacidad de almacenamiento.

- **Tarjeta de interfaz de red.** Se encuentra una en la estación de trabajo y otra en el despachador, permite la comunicación entre el despachador y las estaciones de trabajo. Estas tarjetas son adaptadores, sin los cuales la computa-

dora no podría transmitir, recibir o interpretar ninguna información procesada en la red, y serían imposibles las comunicaciones entre los componentes de la red.

- Sistema de cableado que enlaza la tarjeta de interfaz de red del despachador con la de la estación de trabajo. Sin los cables de la computadora no puede existir la transferencia de la información entre los diferentes componentes de la red. Cada estación de trabajo tiene varios cables conectados a los accesos (salidas) en la parte trasera del CPU. Uno de ellos conecta el CPU al servidor de la red. El cable de la red conecta a la estación de trabajo con los recursos de la red.

- Estación de trabajo o nodo. *"Es una computadora personal (PC) autónoma que está conectada a la red y tiene acceso al sistema operativo de la misma."*¹⁸ *"Es en principio un punto de la red a través del cual se encadena de ésta algún dispositivo, ya sea una estación de trabajo, una impresora o un despachador de archivos"*.¹⁹ Las estaciones de trabajo pueden contar con sus propios periféricos o pueden no tenerlos, entre ellos los dispositivos de discos duros o flexibles; si no cuentan con discos sólo pueden procesar programas y datos almacenados dentro del sistema de la red. Estas estaciones pueden trabajar con cualquier sistema operativo, incluso diferente del que usa la red.

- Sistema operativo de la red, que reside en la memoria del despachador. *"Un sistema operativo controla la forma en que las aplicaciones utilizan la computadora y los periféricos conectados a la misma"*.²⁰ El sistema operativo de

¹⁸ WOODWARD, Jeff. *"ABC NOVELL NETWARE"*, México, 1a. Edición, Ed. SYBEX, 1991, pág.50

¹⁹ NORTON, Peter. *"Periféricos y Accesorios para la IBM-PC, PS/2 y Compatibles"*, México, Ed. Prentice Hall, 1993, p.p.208-209.

²⁰ WOODWARD, Jeff. *"ABC NOVELL NETWARE"*, México, 1a. Edición, Ed. SYBEX, 1991, pág. 229.

la red (NOS) se instala cuando todas las conexiones de la red están hechas, funciona como cualquier sistema operativo, pero además controla a las estaciones de trabajo que están a su cargo. Existen dos tipos básicos de NOS: **peer-to-peer**, que permite a los usuarios compartir entre ellos los recursos de sus computadoras, dejando a todos los sistemas en un mismo nivel en la red, esto es, que ninguno está bajo supervisión de otro; y **servidor dedicado**, donde una o más computadoras son destinadas sólo como servidores de archivos. Los usuarios accesan directorios y recursos compartidos de un servidor de archivos dedicado, no así a los directorios a los que no estén autorizados, brindando mayor seguridad que en las computadoras individuales.

- **Programas de entorno de la red**, son los responsables de redirigir los comandos y solicitudes relacionados con la red al sistema operativo de la red.

En una red, cada computadora accesa programas y archivos de un servidor central corriendo esos programas en su propia memoria y con su propio procesador. En un sistema de minicomputadora o mainframe el procesamiento es centralizado; por lo que las tareas de procesamiento de las terminales se ejecutan todas con un mismo procesador. En una red, el servidor de archivos no se sobrecarga con el procesamiento de tareas de computadoras individuales por lo que puede optimizar archivos y servicios de redes.

Con fundamentos anteriormente descritos se puede plantear sin temor a equivocación que si el SEIT descansa en la parte que corresponde a la estructura computacional en una red, el planteamiento puede ser el siguiente:

Desde un número "n" de terminales dedicadas a la captura de los datos cargando desde el servidor un programa diseñado en Data Ease en la versión deseada; no es necesaria una versión específica, debido a que como ya se mencionó, las versiones para WINDOWS y MS-DOS trabajan en forma similar y el único cambio real son las cualidades de presentación que da la versión para WINDOWS, se puede realizar el

proceso de captura a una sola base de datos de forma simultánea, desde otra terminal que sería la "n+1", se estarían realizando los cortes de información, bloqueando la base por períodos no máximos de 5 segundos, para ejecutar un procedimiento que duplique la base de datos en un formato de código ASCII, una vez liberada la base de datos, la captura puede continuar, mientras en esta máquina se realiza el proceso de la información con una rutina programada en Clipper que realizará los cálculos necesarios y dejará los resultados de estos en una base de datos en el servidor con el formato requerido, para que posteriormente desde la máquina "n+2", puede estar siendo ejecutada una rutina diseñada en Harvard Graphics para WINDOWS, que estará actualizando la información de la gráfica, con la base de datos, se hace la aclaración que la máquina que ejecuta la presentación puede estar conectada vía remota mediante Quick Link con algún otro equipo de cómputo para realizar la presentación de los resultados y sus actualizaciones.

INTERNET.

Este planteamiento puede ser el más práctico y sencillo de todos, el hardware que se necesita únicamente es: Terminales que cuenten con un CPU 386 o superior, con 4MB RAM o más, 20MB de espacio en disco duro, floppy de 1.44 MB, Monitor VGA, Windows 3.1 o más, Módem de 14.4K (28.8K recomendado). Líneas telefónicas con servicio telnet. Con respecto al software se requiere que todas las computadoras cuenten con los mismos: programas de captura, software de conexión y navegación en Internet.

Con el hardware y el software anterior es posible instalar centros de acopio de datos distribuidos a lo largo del territorio involucrado en el estudio, cada uno de estos "n" centros consistirá en una sola terminal que capturará la información de su área de influencia. Es posible plantear la posibilidad de realizar dos o tres cortes de información a lo largo del período de tiempo disponible, cada corte consistirá en la acción de mandar un correo electrónico, el cual tendrá como objetivo enviar la base

de datos capturada hasta el momento, ésta será mandada a un centro de acopio general, se estarán recibiendo estas bases en una terminal con las mismas características, conectada ésta ya sea mediante una red LAN o con un programa de comunicación, mandará dichas bases a otro equipo de cómputo, el cual con una rutina programada en Clipper realizará los cálculos matemáticos necesarios y dará formato a la base de datos con la información actualizada que se requiere para que en una tercera computadora en donde se está ejecutando una rutina en Harvard Graphics se estén actualizando los datos de la gráfica que estará exponiendo las tendencias estimadas.

3.4.2. Lo necesario para la utilización del programa INTERLINK.

Para poder diseñar una estructura computacional en base al INTERLINK, no se requiere de mucha capacidad en el equipo de cómputo, las terminales pueden por ejemplo, tener procesador 80386, con disco duro de 20Mb, con 2MB en RAM, monitor VGA, no se requiere mouse, drive de alta densidad, los cables paralelos para la interconexión, y varios multiplexores para realizar la transferencia de la información de una máquina a otra, un módem que se requerirá para la transferencia de los resultados al lugar donde sean requeridos.

La forma de conectar las computadoras es integrando el comando INTERLINK con sus diferentes parámetros al programa de configuración "Config.Sys" de las terminales de captura y a la máquina recolectora de bases de datos, como ya se vio con anterioridad. Este planteamiento consiste en conectar un número "n" de terminales a una que estará juntando la información, se requiere una conexión en estrella, y un multiplexor para poder operar. La conexión se logra, encendiendo primeramente la máquina concentradora, se coloca el multiplexor en la posición 1, direccionando la comunicación a la máquina 1, y se enciende ésta, al momento de iniciar, ésta manda un mensaje avisando si se logró la comunicación, si la comunicación fue exitosa, se repite el procedimiento con la segunda terminal, ubicando el multiplexor

en la posición 2, se repite este procedimiento hasta que todas las terminales estén conectadas con la computadora recolectora.

Cada una de las terminales cuenta con un programa de captura que puede estar diseñado en Data_Ease o Clipper, la única condición es que se encuentren operando en MS-DOS, cada una de éstas tendrá su propia base de datos, así a la hora de la captura, la persona que esté operando la máquina concentradora, avisará a la persona que esté capturando en la terminal 1 que se realizará un corte, ésta dejará de capturar y suspenderá la ejecución de la captura, el operador colocará el multiplexor en la posición 1, accesará el disco duro de ésta y copiará con comandos de DOS la base de datos al disco duro de la máquina concentradora, terminando este proceso (para el cual puede también sistematizarse con un programa con extensión .BAT). Este procedimiento se repetirá con las "n" terminales conectadas; cabe señalar que terminado el proceso por máquina ésta se puede integrar inmediatamente a la captura, el tiempo estimado de realización del proceso de copiado es de 30 segundos a un minuto por máquina.

Una vez que se ha recuperado la información de todas las terminales, con un proceso en Clipper se juntan las bases de datos y posteriormente se realizan los cálculos y se da formato a la información con los resultados del corte para su presentación en una tercera máquina, la cual estará ejecutando una macro diseñada en Harvard Graphics presentando una gráfica con las tendencias estimadas y el porcentaje de la muestra procesada.

Como se puede observar, las opciones son muchas, aquí sólo se presentaron algunas posibilidades y como resulta evidente, existen mejores opciones como pueden ser los sistemas con servicios en línea, que son los que ocupan los bancos para mantener el control en cuentas bancarias, por otro lado, el software de igual manera puede no ser el mejor, pero si es suficiente para soportar las necesidades del SEIT.

Inicialmente se hizo un repaso de lo que es la Ingeniería de Software para poder identificar las aportaciones de esta disciplina al SEIT, se revisó la filosofía de esta disciplina, sus características principales, las fases en el desarrollo del software, etc. Se procuró tener un panorama general de dicha ingeniería con la finalidad de identificar las aportaciones al sistema, se continuó con un análisis de las posibles estructuras computacionales que podrían emplearse para facilitar el manejo de la información, como se pudo observar, la mejor forma es mediante la instalación de una red LAN, sin embargo, como también se observó no es la única forma, ya que con la instalación de la red con programas de comunicación se puede lograr un eficiente manejo de la información a un costo mucho menor. También se presentaron diversas formas de procesar la información, con el software disponible, se estudiaron desde sistemas operativos, bases de datos, graficadores, programas de comunicación, etc., se dan solamente algunos ejemplos de programas que pueden ser de utilidad para el proceso de los datos, y finalmente, en este capítulo se conjunta lo antes visto, en cuáles podrían ser en este caso los requisitos mínimos para armar la posible estructura computacional, ya sea una red LAN, o mediante un programa como el INTERLINK.

CAPITULO IV

LA ORGANIZACION Y EL DESARROLLO DEL SISTEMA.

4.1. Metodología para la selección de la muestra aplicada al ejemplo general de esta investigación.

En este capítulo se verá poco de lo que en la práctica es todo lo antes mencionado, además de desarrollar un ejemplo completo de lo expuesto, debido a la seriedad del trabajo se analizará un ejemplo con datos reales, para ello se han seleccionado las pasadas elecciones presidenciales en la República Mexicana, llevadas a cabo el pasado 21 de Agosto de 1994. Debido a la naturaleza del trabajo algunos de los procesos sólo se describirán.

A lo largo de este capítulo y el siguiente se hará referencia al ejemplo antes mencionado, para describir los procesos.

Para la creación de la muestra es importante plantear con toda la precisión deseada algunos aspectos como son:

Marco muestral.

El marco muestral lo conforman las 63,610 secciones electorales que integran el sistema electoral del país.

En este caso vale la pena recordar que como ya se mencionó anteriormente, el IFE divide al país en 32 entidades federativas, de ahí se subdivide en distritos federales o locales, como el caso que ocupa a esta tesis es unas elecciones federales, casi resulta redundante aclarar que se tomarán en cuenta los distritos federales, de ahí nuevamente éstos se subdividen en secciones, y la siguiente división son los votantes, personas mayores de 18 años, con su credencial para votar, es decir toda

aquella persona que tenga la capacidad de sufragar.

Al tomarse como unidad del marco muestral a una sección, se está tomando a un grupo completo de votantes, debido a lo cual la muestra toma características de la técnica de muestreo del tipo de conglomerados.

El tamaño de la muestra.

Para calcular el tamaño de la muestra se recurrirá al tamaño de la muestra calculado con la fórmula planteada en el capítulo 2.2 con los siguientes parámetros:

- El parámetro $P = 0.5$ (parámetro de población).
- Error en la estimación del 5%.
- Con una población de 63,596 secciones.

Como se puede ver en el capítulo 2.2, se tiene un tamaño de muestra de 398 secciones, para efectos del diseño de la muestra se cerrará este número a 400, partiendo del hecho de que como la población es lo suficientemente grande, estadísticamente se puede decir que tiende a infinito. Es necesario hacer la aclaración que este tamaño de muestra es representativo a nivel nacional, si fuese necesaria una muestra representativa de cada región del país o cada estado, sería necesario volver a calcular el tamaño de la muestra, el cual seguramente se modificará dependiendo del criterio aplicado y de la representatividad deseada.

Una vez identificado el marco muestral y definido el tamaño de la muestra, se procederá a evaluar la técnica de muestreo apropiada, se hace la aclaración de que la población en estudio es una población heterogénea, por lo que se recomienda estratificar la muestra.

Primeramente se estratificará por estado, es decir, que cada estrato está delimitado por cada estado de la República, a cada uno de estos estratos se le asignará el número de secciones a ser observadas de acuerdo a la proporción del padrón que le corresponda.

En el cuadro 4.1.1 se puede analizar la situación de cada estado de la república, en este caso se analiza la relación que guardan éstos con respecto al padrón electoral nacional, como se puede observar en cada estado se tienen secciones que por sus características se clasifican en urbanas y rurales, debido a lo cual se hace necesaria otra estratificación dentro de cada estado, de donde se plantea una situación que pudiese de alguna forma crear confusión, es evidente que es necesario conservar equilibrio entre secciones urbanas y rurales, equilibrio que debe estar representado por la proporción que guarde la población votante dentro de cada estado.

El IFE dentro de su clasificación toma ciertos criterios para definir si una sección es rural o urbana, sólo que si en este caso se toman las proporciones que marcan las secciones como tales, se puede incurrir en un error, debido a que esta proporción no es representativa de la población votante, es decir, a nivel nacional el 59.47% de las secciones están clasificadas como urbanas y el 40.48% como rurales, sin embargo a nivel del padrón electoral el 70.35% de los votantes se ubica en secciones urbanas y el 29.62% en secciones que se clasifican como rurales.

Para esquematizar más esto, se tomará un ejemplo de lo que se plantea en el cuadro 4.1.1; en el estado de Durango el porcentaje de secciones urbanas es del 40.94%, y el porcentaje de personas que viven en secciones urbanas es del 60.67%, la diferencia es de aproximadamente 20 puntos porcentuales. Esto quizá se deba a algún fenómeno demográfico, de emigración y/o geográfico.

RESUMEN DE SECCIONES POR TIPOLOGIA, PADRON ELECTORAL Y DISTRIBUCION DE SECCIONES EN MUESTRA A NIVEL ESTATAL

| ESTADO | | Distritos | Secciones | Padron 1894 | % VS Padron | Secciones Urbanas | % VS Secc. Edo | Padron Secc. Urbanas | % VS Pad. Edo. | Secciones Rurales | % VS Secc. Edo. | Padron Secc. Rurales | % VS Pad. Edo. | # Secc. Mta. Est. |
|--------|-----------------------|-----------|-----------|-------------|-------------|-------------------|----------------|----------------------|----------------|-------------------|-----------------|----------------------|----------------|-------------------|
| 1 | AGUASCALIENTES | 2 | 488 | 426,878 | 0.90 | 353 | 72.83 | 315,545 | 73,919 | 133 | 27.37 | 111,333 | 26,081 | 4 |
| 2 | BAJA CALIFORNIA NORTE | 6 | 1,368 | 1,057,753 | 2.28 | 1,230 | 89.91 | 998,986 | 91,846 | 138 | 10.09 | 90,885 | 8,354 | 9 |
| 3 | BAJA CALIFORNIA SUR | 2 | 352 | 190,749 | 0.40 | 217 | 61.85 | 136,078 | 71,339 | 135 | 38.35 | 54,671 | 28,661 | 2 |
| 4 | CAMPECHE | 2 | 469 | 908,668 | 0.85 | 260 | 53.17 | 203,002 | 65,787 | 228 | 48.83 | 105,668 | 34,233 | 3 |
| 5 | COAHUILA | 7 | 1,520 | 1,142,170 | 2.41 | 1,085 | 71.38 | 853,758 | 83,504 | 435 | 28.82 | 188,412 | 16,496 | 10 |
| 6 | COLIMA | 2 | 338 | 263,701 | 0.56 | 198 | 58.93 | 171,139 | 64,907 | 138 | 41.07 | 92,541 | 35,093 | 2 |
| 7 | CHIHUAHUA | 8 | 1,928 | 1,714,198 | 3.61 | 616 | 31.83 | 872,139 | 39,210 | 1,313 | 68.07 | 1,042,057 | 60,790 | 14 |
| 8 | DISTRITO FEDERAL | 40 | 5,543 | 5,535,268 | 11.66 | 1,867 | 33.60 | 80,783 | 843 | 31.11 | 31.11 | 285,682 | 19,217 | 13 |
| 9 | DURANGO | 6 | 1,360 | 733,247 | 1.54 | 569 | 40.84 | 444,930 | 88,414 | 821 | 82 | 87,808 | 1,586 | 47 |
| 10 | GUANAJUATO | 13 | 3,005 | 2,188,005 | 4.63 | 1,685 | 58.07 | 1,400,124 | 63,758 | 1,320 | 43.83 | 289,317 | 38,321 | 6 |
| 11 | GUERRERO | 10 | 2,718 | 1,244,366 | 2.62 | 636 | 28.72 | 613,237 | 48,260 | 1,930 | 69.47 | 785,861 | 35,242 | 19 |
| 12 | HIDALGO | 6 | 1,707 | 1,037,024 | 2.18 | 432 | 25.31 | 406,892 | 39,217 | 1,275 | 74.69 | 631,188 | 50,720 | 10 |
| 13 | JALISCO | 20 | 3,309 | 3,026,460 | 6.37 | 2,157 | 65.19 | 2,271,935 | 75,069 | 1,152 | 34.81 | 754,525 | 24,931 | 25 |
| 14 | JALISCO | 34 | 5,916 | 6,050,969 | 12.74 | 4,364 | 73.77 | 4,891,542 | 60,639 | 1,552 | 26.23 | 1,159,427 | 18,161 | 51 |
| 15 | MICHHOACAN | 13 | 2,674 | 1,888,250 | 3.98 | 1,302 | 48.69 | 1,074,360 | 56,898 | 1,372 | 51.31 | 613,870 | 43,102 | 16 |
| 16 | MICHHOACAN | 4 | 906 | 769,597 | 1.62 | 603 | 66.56 | 544,143 | 70,705 | 303 | 33.44 | 225,454 | 28,295 | 6 |
| 17 | MORELOS | 3 | 878 | 475,775 | 1.00 | 464 | 52.85 | 301,331 | 63,335 | 414 | 47.15 | 174,444 | 36,665 | 4 |
| 18 | NAYARIT | 11 | 2,123 | 1,948,008 | 4.10 | 1,692 | 79.70 | 1,740,902 | 89,322 | 431 | 20.30 | 208,006 | 10,678 | 16 |
| 19 | NEVO LEON | 10 | 2,449 | 1,514,166 | 3.19 | 766 | 32.50 | 619,416 | 40,908 | 1,683 | 67.50 | 894,750 | 58,082 | 13 |
| 20 | OAXACA | 14 | 2,548 | 2,184,055 | 4.60 | 1,322 | 51.88 | 1,258,370 | 59,070 | 1,226 | 46.12 | 815,779 | 41,930 | 18 |
| 21 | PUEBLA | 3 | 687 | 620,240 | 1.31 | 348 | 50.86 | 342,161 | 55,169 | 339 | 49.34 | 278,059 | 44,831 | 5 |
| 22 | QUERETARO | 2 | 450 | 317,894 | 0.67 | 293 | 65.11 | 225,071 | 70,912 | 157 | 34.89 | 82,323 | 28,088 | 3 |
| 23 | SINALOA | 9 | 1,785 | 1,078,188 | 2.27 | 670 | 37.33 | 591,425 | 54,956 | 1,125 | 62.67 | 484,763 | 45,044 | 9 |
| 24 | SANTO LUIS POTOSI | 7 | 1,330 | 1,180,609 | 2.49 | 910 | 68.42 | 817,601 | 77,723 | 420 | 31.58 | 263,006 | 22,277 | 11 |
| 25 | SONORA | 5 | 1,133 | 866,039 | 1.82 | 478 | 42.18 | 1,065,128 | 47,170 | 655 | 57.81 | 457,529 | 52,830 | 7 |
| 26 | TAMAULIPAS | 2 | 605 | 450,089 | 0.95 | 363 | 60.00 | 332,042 | 73,771 | 242 | 40.00 | 308,801 | 22,478 | 12 |
| 27 | TLAXCALA | 23 | 4,711 | 3,822,708 | 7.63 | 2,363 | 50.16 | 2,049,172 | 56,565 | 2,348 | 49.84 | 1,573,538 | 43,455 | 41 |
| 28 | VERACRUZ | 4 | 1,059 | 745,587 | 1.57 | 708 | 66.95 | 558,552 | 74,915 | 350 | 33.05 | 187,029 | 23,085 | 6 |
| 29 | YUCATAN | 5 | 1,882 | 695,068 | 1.46 | 584 | 31.03 | 313,691 | 45,130 | 1,298 | 68.97 | 381,987 | 54,870 | 6 |
| 30 | ZACATECAS | 3 | 300 | 83,586 | 0.50 | 37,823 | 59.47 | 33,402,851 | 70,351 | 25,744 | 40.48 | 14,077,208 | 28,648 | 400 |
| TOTAL | | | | | | | | | | | | | | |

Cuadro 4.1.1.

Partiendo del hecho de que lo que se va a estimar es una cantidad de votos y que se deben tomar las proporciones correctas, se delimitarán los estratos por medio del padrón electoral, primeramente se asignará la cantidad de secciones a cada estado por su peso en padrón, como ya se vio en el cuadro 4.1.1, posteriormente se dividirá este número de secciones en urbanas y rurales según la proporción del padrón del estado, como se ve en el mismo cuadro.

Una vez definido el número de secciones por estado se seleccionarán aleatoriamente las secciones que formarán parte de la muestra. Para ello se apoyará en las rutinas diseñadas para este fin, se puede esperar que el equilibrio entre secciones rurales y urbanas se dará en base a las probabilidades de selección de cada una de éstas.

Después de analizar lo antes mencionado, cabe señalar que la técnica empleada para diseñar la muestra es mediante el Muestreo aleatorio estratificado. Aunque conserve algunas características de la técnica de muestreo por conglomerados, debido a que se está tomando un censo a nivel sección electoral, sin embargo, como se comentó en su momento, la sección electoral es la unidad de muestreo.

Solamente resta hacer hincapié en que la composición real de los estratos está dada por la cantidad de votantes en las secciones urbanas y rurales, y esta proporción la da sólo el padrón electoral, las secciones son únicamente una forma de agrupar a estos votantes, que en muchas ocasiones debido a situaciones geográficas y/o demográficas no es posible uniformar cuantitativamente dichas secciones, de ahí que las proporciones entre padrón y secciones varíen de esta forma.

4.2. El flujo de la información y algunas observaciones

Debido a que ya se cuenta con la metodología para el diseño de la muestra, es tiempo de analizar las posibles opciones, para agilizar al manejo de la información. Como se ya se ha hecho hincapié en esto, es vital que el manejo de la información sea lo más ágil posible, antes de proponer algunos procedimientos para lograr lo antes mencionado, al hacer un repaso de cómo a grosso modo, debe funcionar el SEIT, debido al tema que se tratará, será necesario el apoyo en algunas gráficas para esquematizar el flujo de la información.

Se describirá el proceso en su totalidad diferenciando las partes que no pueden ser controladas por el sistema pero que son parte del proceso, es decir delimitaremos las fronteras del sistema.

Cabe señalar que cuando se hace referencia a "la casilla", se hace referencia a las urnas de las secciones que se encuentran en la muestra.

La Jornada Electoral, primera parte del proceso.

(fuera del sistema):

- Se genera la información, se tomará como punto de partida la generación de la información, sin importar algún proceso preparatorio para llevar a cabo el escrutinio, debido a que éste no es el tema de la tesis, en esta parte, el sistema aún no entra en operación debido a que ubicando este punto dentro del proceso electoral, se podría decir que en este momento es cuando los votantes acuden a las urnas para votar, y que de alguna forma es el proceso

en estudio.

- Se realiza el conteo de votos, en cada casilla al término de la votación, se procede a realizar el conteo de las boletas y se asigna a cada opción (partido político), el número de votos obtenidos por casilla.
- La publicación de los datos, al terminar el conteo de votos y después de llevar a cabo cualquier tipo de revisión, negociación, etc., se llenan unos formatos que se exhiben en el lugar donde se llevo a cabo la votación, en los cuales se presentan los resultados obtenidos, y su asignación para cada uno de los partidos participantes.

La recolección de los datos, segunda parte del proceso.

(entra en función el sistema):

Aquí se comenzará por describir el proceso y posteriormente se darán algunas alternativas reales y operativas para llevar acabo el resto del proceso.

- Obtener la información, insumos del sistema. Después de publicados los resultados de la casilla, lo que procede es obtener estos resultados, la persona que está encargada de la observación de la casilla obtiene estos resultados y los envía a su punto de concentración.
- La recepción de la información, es necesario tener la capacidad de juntar los resultados de las casillas de alguna forma, se propone que sea en una base de datos para poder manipular los datos cuantas veces sea necesario inclusive procesar la información de la forma más rápida posible, con la ayuda de una computadora, para lo cual es necesario implantar un operativo de captura de datos.

- El procesamiento de la información, ya sea que se tenga la información en una base de datos o no, lo que continua es realizar el proceso de los datos, el cual consiste en el cálculo de las tendencias, y algunos estimadores que puedan dar a conocer algunas características de consistencia de la muestra.
- La presentación de los resultados, éste es un aspecto muy relativo, ya que esto dependerá del consumidor final del producto generado, es decir, si se requiere únicamente de un reporte, alguna gráfica, el nivel de detalle con el que se espera la información, si se requiere alguna presentación vía remota, o simplemente desea que se le notifique verbalmente los resultados.

Ahora se presentarán una serie de opciones para llevar a cabo la segunda parte de este proceso, es decir, algunas formas de armar este sistema, las opciones presentadas varían desde el punto de vista técnico hasta el económico.

Planteamiento Básico.

Esta es la posible conformación del SEIT, tomando en cuenta, el proceso que inicia en la obtención de los datos afuera de las casillas, hasta la presentación de los resultados, con esto no se pretende que se piense que ésta es la única forma de trabajar un proyecto de este tipo, ésta es únicamente una propuesta, de la cual se plantearán una serie de opciones reales, donde, de la misma manera, es necesario aclarar que no son ni las únicas ni las mejores opciones.

El SEIT entra en acción en el momento en que en las casillas de las secciones publican los conteos realizados a lo largo de la jornada electoral, en ese momen-

to un observador especialmente considerado para ello, copia dichos resultados, apoyándose en un formato previamente diseñado, después se comunicará a la central de acopio vía telefónica, y dictará los resultados a algún capturista para su incorporación a una base de datos, la situación aquí es que no se sabe qué tan grande puede llegar a ser, esto depende del tamaño de la muestra, para efecto de que la captura de los datos no se prolongue más de lo esperado, se plantea el uso de varias computadoras, como se puede pensar que "varias" es muy confuso, aclarando el punto, se necesitan "i" computadoras, donde "i" depende de la cantidad de información en el formato que se vaya a capturar, el número de formatos y el tiempo en que se espera que estén capturados, para lo cual se propone emplear el modelo matemático siguiente:

$$i = \text{int} \left(\frac{tn}{T} \right) + C^1$$

donde:

t = el tiempo en que se captura la información de una casilla, en segundos.

n = el tamaño de la muestra.

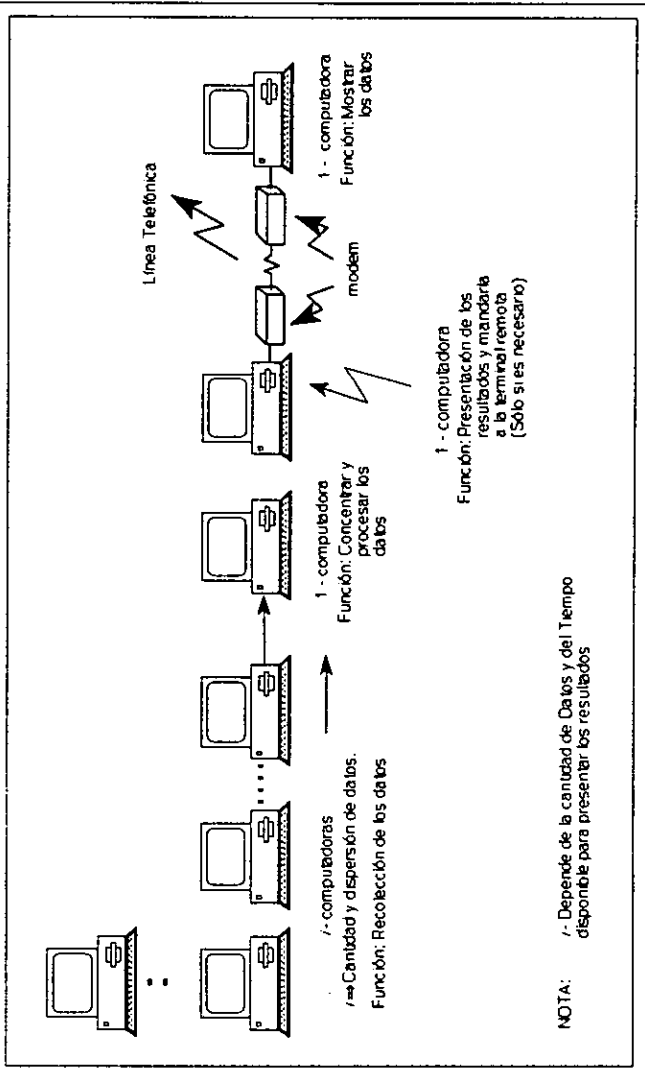
T = tiempo total del que se dispone para completar la captura en segundos.

C = número de terminales adicionales como margen de seguridad.

En la figura 4.2.1. se esquematiza lo antes dicho, después del proceso de captura es necesario juntar la información en una sola base de datos, esto sólo en el caso de que la estructura computacional no de la facilidad de realizar el llenado de modo compartido, después se procede a procesar los datos, actividad que se realizará en otra terminal, a continuación es necesario dar formato a la información obtenida, ya sea en la presentación de un reporte escrito o con la presentación de una gráfica que puede ser realizada en esta misma máquina.

¹ Modelo propuesto por los autores de esta tesis.

Figura 4.2.1 PLANTEAMIENTO BASICO PARA LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT



NOTA: i - Depende de la cantidad de Datos y del Tiempo disponible para presentar los resultados

El proceso antes mencionado se repetirá el número de veces que sea necesario, hasta tener cubierto el 100% de la muestra.

Si es necesario mandar la información vía remota a otra terminal, sobra mencionar que se necesitan un par de módems, una línea telefónica y la terminal receptora del informe. Se recomienda que la terminal que procesa y la que manda la señal no sean la misma, para no perder la posibilidad de seguir procesando.

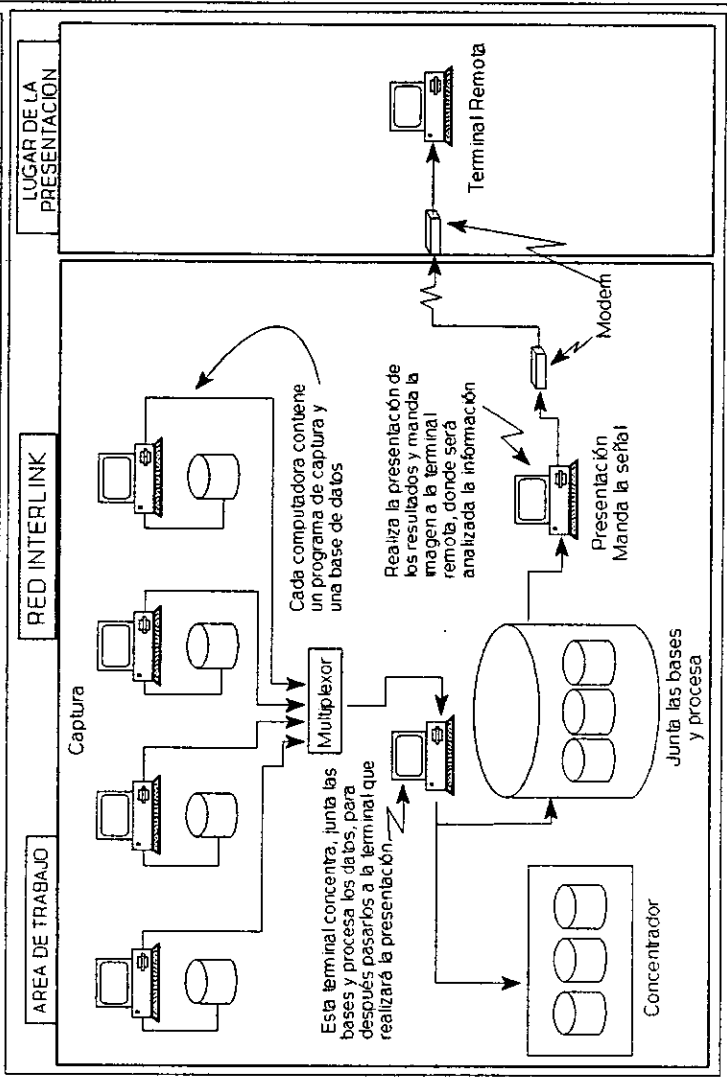
Segundo Planteamiento.

En este planteamiento se propone el uso de "i" computadoras independientes trabajando en modo caracter, en MS-DOS, cada una de estas máquinas con un programa de captura y su propia base de datos.

No se sabe qué tan grande puede ser "i", el número de computadoras necesarias para llevar a cabo la captura, pero previendo que este número crezca, no se puede pensar que se va a realizar el traslado de la información de una máquina a otra mediante la copia de la base en un disquete por lo tardado que pudiese llegar a ser esto, aunque si "i" no es tan grande no se puede despreciar esta opción, si es que esto trae consigo un beneficio en seguridad o rapidez.

Debido al ambiente, se abre la posibilidad del empleo de una de las utilerías de MS-DOS, que es el INTERLINK, utilería que ya se examinó con anterioridad, el manejo que se plantea es el copiado de bases de máquina a máquina, en este caso la topología recomendada es la de estrella, con este planteamiento, el modo de operación será el siguiente, cada lapso de tiempo determinado, una máquina concentradora se conectará con un multiplexor como enlace a cada una de las máquinas de captura, esto se esquematiza en la figura 4.2.2.

Figura 4.2.2. PLANTEAMIENTO DOS PARA LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT



La máquina que está concentrando las bases de las máquinas de captura, tiene la capacidad de realizar un corte y reportar en cualquier momento, esta máquina puede contar con varios procedimientos, se recomienda estén programados en la misma herramienta de software que las de captura, esto es con la finalidad de no perder tiempo en adaptar la información a otro paquete o lenguaje de programación, esta máquina lo primero que debe realizar es juntar las bases de datos que se hayan generado, depurar la base, procesar los datos y presentar un reporte.

Si se requiere de mandar la información vía remota a otra terminal, igual que en el planteamiento anterior, se recomienda el uso de una computadora especialmente destinada para realizar esta tarea.

Cabe señalar que la captura y el procesamiento de los datos se están manejando en ambiente MS-DOS, pero que la máquina que posiblemente esté mandando el reporte vía remota no necesariamente tiene que estar bajo este mismo ambiente, ya que ésta únicamente enviará la información, además Windows puede presentar algunas facilidades de comunicación e indudablemente la presentación será mejor.

Tercer Planteamiento.

La recolección de los datos no varía, el observador copiará los resultados de la cartulina que se presentará afuera de las casillas de las secciones en la muestra, en un formato previamente diseñado, posteriormente se comunicará vía telefónica a la central y dictará los datos a un capturista que en ese momento estará introduciendo éstos a una base de datos. Este planteamiento presenta una mejor forma de trabajar, debido a que considera el empleo de una red de

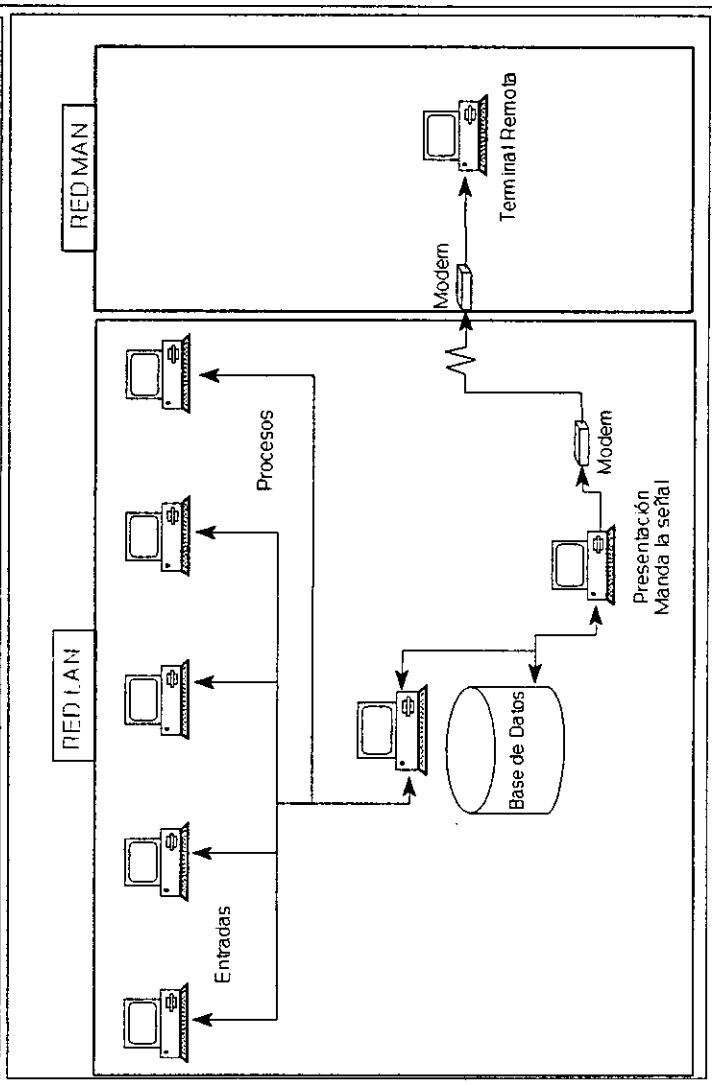
tipo LAN (Local Area Network), con este esquema se logra un gran ahorro de tiempo, debido a que de el número de terminales determinado se está complementando una sola base de datos, y la terminal destinada a realizar el proceso puede, en el momento deseado, acceder la base y realizar los cálculos necesarios, inclusive sin detener la captura, después se procede a realizar el reporte, que como ya se ha mencionado puede ser escrito, gráfico, etc.

A continuación, es posible mandar vía remota la información en caso de presentarse esa necesidad, debido a que trabajando en red se conecta alguna terminal mediante una línea telefónica y un par de módems, siendo la terminal remota parte de la red, lo cual simplifica la transmisión de la información. Este tercer planteamiento se esquematiza en la figura 4.2.3.

Con este planteamiento la red deja de ser una red de tipo LAN, y se convierte en una de tipo MAN (Metropolitan Area Network).

El software que se emplea, como en los planteamientos anteriores puede ser formado por varias combinaciones, primeramente para la conformación de la red, el software que proporciona Lantastic o Novell son ampliamente recomendables para este tipo de proyectos, los programas de captura pueden programarse en Data Ease, Fox Pro, Acces, Paradox, Oracle, Clipper, etc. Existen muchas bases de datos que dan la facilidad del trabajo en red, el programa para proceso de igual manera puede ser programado en alguno de estos paquetes, para la presentación de igual manera existe software suficiente para poder elegir el que más se acomode, éste va desde Harvard Graphics, Excel, Power Point, Freelans, etc.

Figura 4.2.3. PLANTEAMIENTO TRES DE LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT



Cuarto Planteamiento.

Este esquema no varía mucho del planteamiento anterior, sólo que en éste se secciona el trabajo en varios centros de acopio, quizá este planteamiento es un refuerzo del anterior, pensando que el trabajo y la captura crecieran mucho y un esquema como el anterior fuese insuficiente, se puede ampliar o modificar a este nivel.

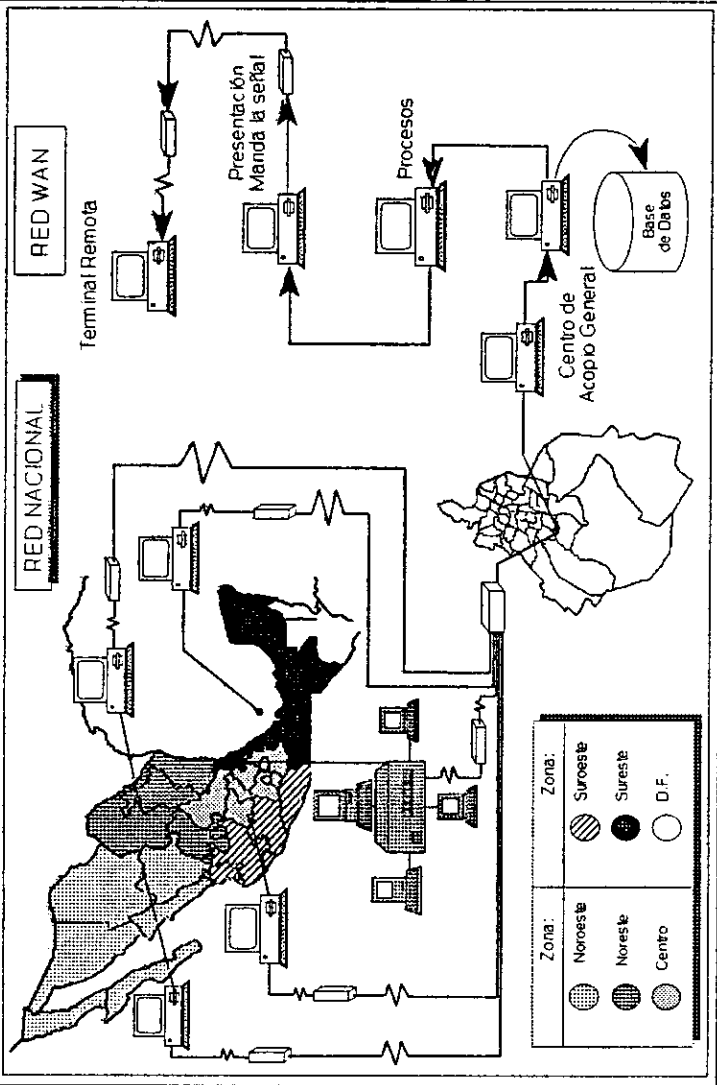
Aquí se propone la creación de varios centros de acopio, cada uno de los cuales recibirá información de una región determinada de modo que cada uno de estos centros concentre dicha información, la recolección de los datos se lleva a cabo de la misma forma que los planteamientos anteriores, únicamente que en este caso, la comunicación se realiza al centro de acopio correspondiente a la región o zona en donde geográficamente está ubicada la sección observada.

La información concentrada en estos centros regionales será transferida a un centro general de acopio, es decir, un centro de acopio que reciba la información de los centros regionales, que la estandarice, la junte, la procese y la reporte. En la figura 4.2.4. se esquematiza esta opción.

Cada uno de los centros de acopio exceptuando el general, no requiere más que una red de tipo LAN, que estará dedicada a la captura de la información y el centro general es un planteamiento idéntico a la tercera opción presentada. Este planteamiento lleva a la formación de una red del tipo WAN (Wide Area Network).

El Software que se podrá emplear en este caso, es el mismo que en caso anterior debido a que el planteamiento no tiene cambios substanciales.

Figura 4.2.4. PLANTEAMIENTO CUATRO DE LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT



Quinto Planteamiento.

En esta época de avances tecnológicos y en la que se comienza a disfrutar de los grandes beneficios que otorgan los servicios de comunicación, sería una falla muy grave no tomar en cuenta las facilidades que ofrece INTERNET.

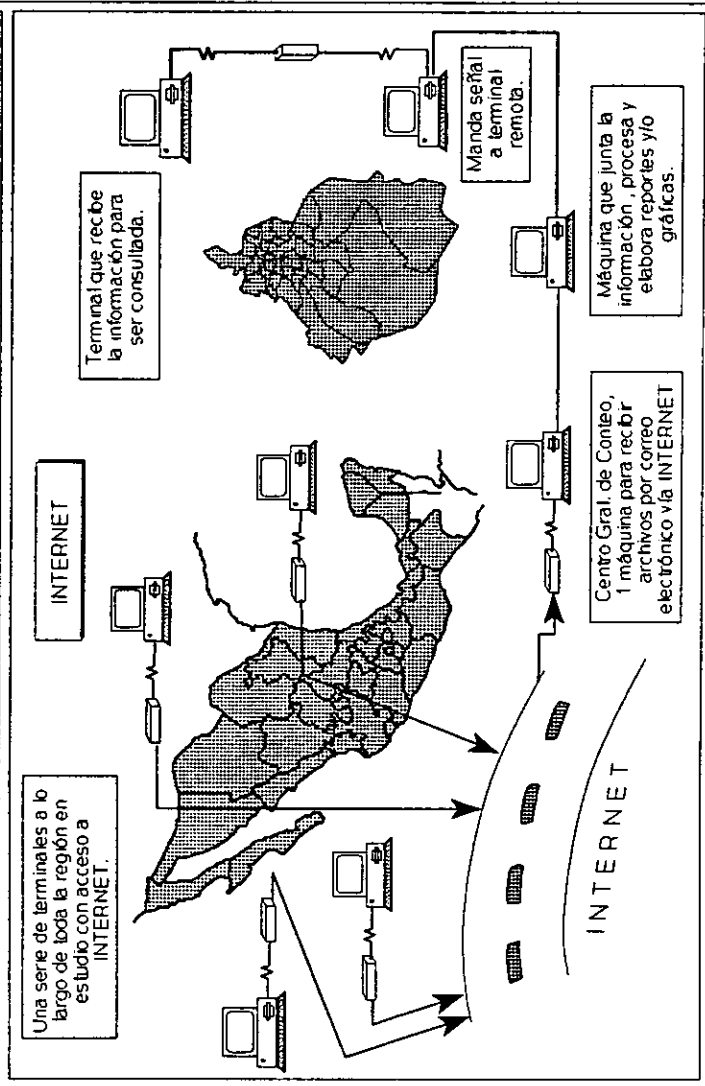
Aquí, el acopio de los datos no tiene ninguna diferencia al del cuarto planteamiento, debido a que también se plantea la opción de centros de acopio a lo largo de toda la región sobre la cual se está trabajando, cada uno de estos centros de operación cuenta con una sola computadora, la cual deberá estar conectada a Internet, por lo que es necesario que también cuente con un módem.

En dicha terminal se capturarán los resultados de un número de casillas asignado, una vez realizado esto, la base de datos que contiene los datos se mandará en un correo electrónico al centro de acopio general, todo esto a la brevedad posible.

En el centro de acopio general, se contará con una computadora que únicamente esté recibiendo la correspondencia y la mande a una terminal que se dedique a realizar los cortes de información, el cual consiste en unificar la base de datos, procesar y preparar un reporte ya sea gráfico o escrito, para que, posteriormente, este reporte sea mandado vía remota a alguna otra terminal para su consulta o pueda ser colocado en un WEB SERVER, para poder administrar las consultas vía INTERNET. La esquematización de esta opción se presenta en la figura 4.2.5.

El software que se emplearía en este planteamiento es básicamente el mismo de los anteriores, bases de datos, programas para capturar y para el procesamiento de la información, algún paquete para realizar gráficas, etc. Lo único que se

Figura 4.2.5. PLANTEAMIENTO CINCO PARA LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT



agregará es el software necesario para realizar la conexión a INTERNET.

Sexto Planteamiento.

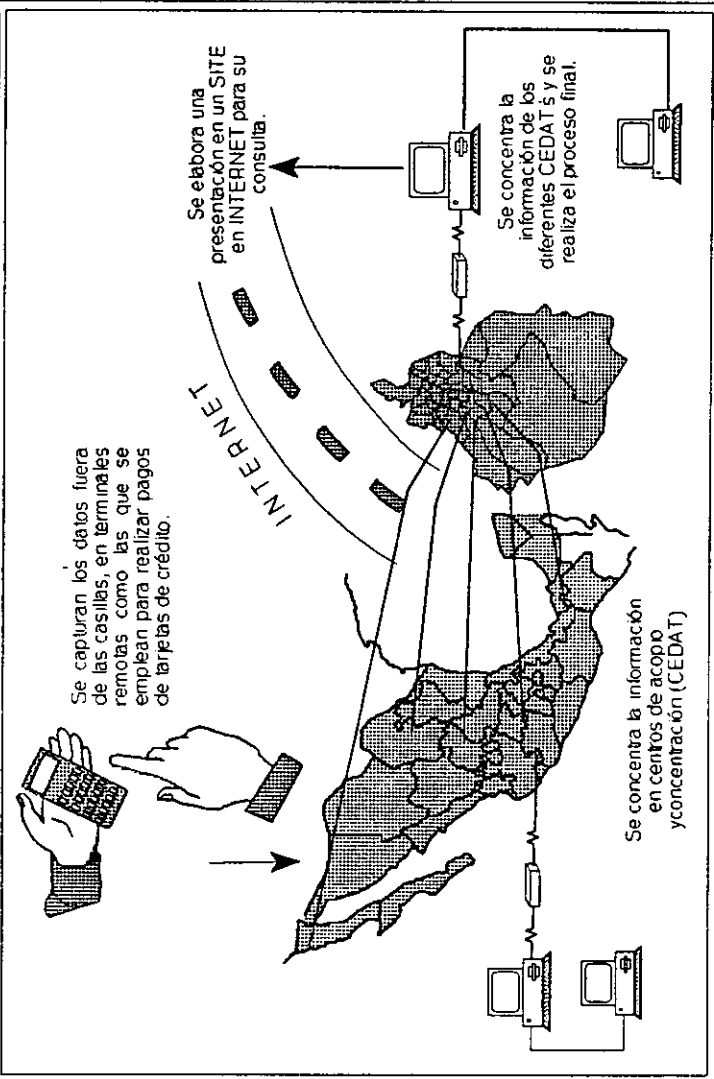
Este planteamiento es muy similar al que emplea el IFE para llevar a cabo el PREP (Programa de Resultado Electorales Preliminares), el cual consiste en consultar las publicaciones en las casillas después del escrutinio y cómputo, dichos datos son capturados en terminales similares a las que se emplean para realizar las compras con tarjetas de crédito, después de lo cual son mandados vía telefónica a centros regionales de concentración de la información, los cuales a su vez, mandan la información al centro nacional de concentración, en el que se lleva a cabo la unificación de las bases de datos y el proceso.

Después se actualiza la presentación disponible en INTERNET, en donde podrá ser consultada por el público en general.

El Software para realizar este planteamiento es básicamente el empleado en el anterior, lo que puede cambiar son las rutinas y el empleo de programas de comunicación que han sido mencionados con anterioridad.

En los planteamientos presentados se puede observar que son muy parecidos entre sí, y sin embargo, tienen diferencias en el manejo de la información, en el software e inclusive el hardware empleados, el emplear uno u otro depende de los recursos disponibles, tomando en cuenta que se puede lograr la eficiencia esperada en base a los recursos disponibles, es decir se puede disminuir el tiempo de manejo de la información, desde la unificación de la base y el proceso, si se trabaja en una red, que es lógicamente más costosa que una serie de máquinas trabajando independientemente, el tiempo de diferencia puede no ser

Figura 4.2.6. PLANTEAMIENTO SEIS PARA LA ESTRUCTURA COMPUTACIONAL REQUERIDA PARA EL SEIT CON INTERNET



mucho, esto depende del "i" número de terminales con el que se esté trabajando, pero la diferencia en costo y el riesgo puede ser muy grande.

En los planteamientos presentados se expuso la opción de mandar la información vía remota a alguna otra terminal, y en los últimos planteamientos se considera la opción poner los resultados en INTERNET para su consulta, sin que forzosamente sea necesaria algunas de las dos opciones, la situación se define ante el caso real y va de acuerdo con las necesidades del interesado.

4.3. Diseño de un modelo funcional para el SEIT.

Se analizarán las funciones de cada elemento del sistema, cabe mencionar que se denomina elemento del sistema a lo que en la práctica serían las diferentes áreas de trabajo.

1. Coordinación General.

Esta es el área de trabajo del sistema sobre la cual recae la mayor responsabilidad del proyecto, es aquí donde se coordina el esfuerzo de todas sus áreas involucradas, se evalúa el desempeño de cada una y se supervisa el cumplimiento satisfactorio de los objetivos, entre las funciones más importantes de esta área se encuentran:

1.1) Coordinar el trabajo de las diferentes áreas de trabajo involucradas.

- Desde la visión de la coordinación es posible conocer los productos esperados de cada una de las diferentes áreas de trabajo y es necesario estandarizar varios procesos para al final poder conjuntar los esfuerzos realizados.

1.2) Definir los objetivos generales del sistema, y de cada una de las áreas.

- Es definitivo que el objetivo general del sistema se define aquí, para lo cual se debe contar con el trabajo terminado de todas las áreas, es decir los objetivos de cada una de las áreas también debe definirse desde la coordinación.

1.3) Desarrollar la metodología del SEIT.

- Como ya se ha comentado, todo proyecto de esta especie para poder ser válido debe estar fundamentado en algún tipo de metodología, el determinar cual será la metodología del sistema y dar las pautas a seguir debe estar en manos de una persona que tenga la visión general del sistema.

1.4) Supervisar y calificar los productos generados por las diferentes áreas.

- Es una de las funciones más importantes de esta área, ya que es aquí en donde se evalúa el trabajo de las diferentes áreas y se determina si es lo que se espera para el buen cumplimiento de los objetivos.

1.5) Detectar posibles fallas en el sistema y corregirlas a tiempo.

- Si a la hora de calificar los objetivos de cada área se detecta alguna falla, en ese momento se toman las medidas necesarias para corregir el problema, esta labor debe llevarse periódicamente a cabo y con toda la anticipación posible.

1.6) Emplear el enfoque de sistemas y conjuntar el esfuerzo parcial de cada una de las áreas y enfocarlos al cumplimiento de los objetivos.

- Esta es la parte más fina del trabajo de la coordinación, ya que es aquí en donde se planean las funciones de cada una de las diferentes áreas, y se deja cimentada la fusión de los diferentes objetivos a fin de que el cumplimiento del objetivo final se de forma casi automática.

1.7) Generar los canales de comunicación entre las diferentes áreas del sistema.

- Es necesario dejar bien sentadas las bases de comunicación entre las diferentes áreas para así poder evitar las duplicidades de información o el trabajo innecesario, cabe señalar que es posible que se llegue a un punto dentro del proyecto en el cual la necesidad de información sea tal que se rompan estos canales y que esto tenga como consecuencia un desorden indeseado.

2. Area de Cómputo.

Se puede considerar que ésta es un área estratégica, ya que es aquí en donde se define la forma de manejar la información, y se realizan los diseños de la presentación, se programan las rutinas necesarias para procesar la información, etc. Todo lo que tiene que ver con el equipo de cómputo, desde su selección hasta la seguridad de su correcta operación. Entre sus funciones están:

2.1) Establecer un Monitoreo Electrónico.

- Esta parte consiste en mantenerse actualizado en materia de hardware y software, mediante publicaciones, revistas, artículos, etc. Se considera que esto es indispensable para poder seleccionar de forma segura cuál será la estructura computacional y el software adecuados para el proyecto.

2.2) Seleccionar la estructura computacional adecuada.

- Como ya se ha visto en capítulos anteriores, ésta es una de las tareas más importantes, es esta área la que decide cuál será la estructura computacional deseada. Aquí se debe evaluar la cantidad de información esperada y la forma de almacenar ésta para seleccionar la estructura apropiada.

2.3) Seleccionar el software apropiado para desarrollar las rutinas, programas y aplicaciones necesarias.

- Esta es otra fase muy importante de esta área, ya que aquí se definirá la forma de manejar la información, y cuáles serán las herramientas de software adecuadas.

2.4) Llevar a cabo la instalación del equipo de cómputo ó supervisar ésta.

- Siempre es necesario estar seguro que el equipo esté bien instalado, no está por demás estar seguro de ello, es posible que el mismo departamento de cómputo instale la estructura seleccionada, si se cuenta con el personal adecuado y el tiempo suficiente, por otro lado si no es así se tendrá que contratar los servicios de alguna empresa que preste este servicio, entonces se recomienda que alguna persona con conocimientos dentro del área supervise la instalación del equipo.

2.5) Realizar pruebas en el equipo de cómputo ya instalado con la anticipación necesaria para evitar una posible falla en el momento menos deseado.

- Es por demás decirlo, pero no debe pasarse por alto este gran detalle, aun después de la supervisión es necesario hacer pruebas con el equipo instalado a fin de poder detectar alguna posible falla.

2.6) Desarrollar las aplicaciones, programar y rutinas necesarias para apoyar al sistema.

- Siempre será necesario desarrollar sus propios programas ya que de esta forma se podrá tener la seguridad de que los procesos están siendo ejecutados de la manera correcta, y en algún momento determinado se podrán modificar o corregir si esto fuera necesario.

2.7) Realizar las pruebas necesarias con el software desarrollado para eliminar cualquier posible falla y verificar su perfecto desempeño.

- Todas las medidas de seguridad que puedan tomarse siempre serán pocas en esta clase de proyectos, el validar los programas es algo que siempre debe hacerse, debido a que es muy probable que cuando se programe se cometa algún error.

2.8) Establecer las medidas de seguridad necesarias para garantizar el desarrollo eficiente del proyecto.

- Aparte de las anteriores observaciones existen otras clases de riesgos que deben tomarse en cuenta, las posibilidades de un corte de energía eléctrica siempre están latentes, la posibilidad de que algún virus entre en el sistema es otro riesgo real y que debe prevenirse, la posible descompostura de alguna computadora puede resultar algo trágico, y como estas clases de problemas muchos más que deben tomarse en cuenta.

2.9) Apoyar a las diferentes áreas dentro del sistema.

- El área de cómputo es una área de apoyo, que por lo regular esta en constante comunicación con las diferentes áreas.

3. Departamento de Planeación y Desarrollo.

Este departamento es el responsable de que cualquier planteamiento propuesto cumpla con normas metodológicas válidas, aquí se analizan posibles alternativas para presentar los resultados, se valida la información,

se diseña la muestra, etc. Analizando sus funciones con más detenimiento se tiene que son:

3.1) Aportar elementos de apoyo par desarrollar la Metodología del SEIT.

- Aquí es donde se desarrollan los diferentes elementos que van a servir para fundamentar la metodología a seguir, y de aquí mandarla a la coordinación general para que se evalúe y se formalice.

3.2) Diseño de la Muestra.

- El diseño de la muestra es una de las partes más delicadas, ya que debe estudiarse la población y definir cual será la técnica de muestreo apropiada, cabe señalar que en la práctica, la muestra se diseña en el departamento de Planeación y Desarrollo, sin embargo es en el departamento de Operación Logística se le da el visto bueno, pues es ahí en donde se evalúa la posibilidad de que ésta se cubra físicamente bien, con esto se quiere decir que si en el diseño de la muestra entra por ejemplo una sección completamente aislada o de difícil acceso, probablemente no sea posible conservarla, por lo cual deberá ser reemplazada con el método correcto.

3.3) Diseño de la presentación.

- Debido a que este departamento es el que en un momento dado conoce más la información que se espera y la forma en la que debe presentarse, éste definirá el producto esperado del proceso de la información y lo diseñará con el apoyo del departamento de Cómputo.

3.4) Validación de la Información.

- Es posible que en algún proceso exista algún error a lo largo del proyecto, en este departamento debe darse el visto bueno a la información antes de que se de a conocer, en caso de que por alguna razón se considere que ésta no cubre los requisitos o se sospeche que no está correcta, se procederá a diagnosticar cuál podría ser el error.

3.5) Realización de documentos necesarios para cualquier tipo de gestión.

- En el caso de que alguna dependencia de tipo legal pida alguna especificación del sistema, que desee conocer alguno de los procesos, metodología etc., con el fin de permitir la publicación de los resultados, o autorizar alguna acción, etc. Será necesario crear algún documento que cumpla las especificaciones deseadas, de aquí que lo más común es que este departamento tenga los elementos necesarios para la creación de éste.

3.6) Establecer comunicación necesaria con las demás áreas del sistema.

- Como ya se ha visto, este departamento tiene mucha participación con las demás áreas, de aquí se desprende la necesidad de establecer comunicación de forma ágil y confiable.

3.7) Generar las directrices de trabajo para las demás áreas.

- Parece ser que este departamento es algo así como el cerebro de la operación, de aquí que se genera una serie de necesidades que deben satisfacerse, y bajo este esquema, los demás departamentos lo abastecen de

información, por esta razón, éste solicita insumos de las áreas y genera trabajo para las demás.

4. Departamento de Operación Logística.

En un proyecto como éste, es natural trabajar con un número bastante grande de personas pues es necesario personal para recoger información, para supervisar, para capturar, para la planeación, para hacer llamadas telefónicas, etc. De modo que siempre es necesario que exista orden, ésta es la función principal de este departamento; algunas de sus funciones específicas son:

4.1) Planear y organizar el Operativo en campo.

- Esta es una de las labores más complicadas ya que debe organizar a un número determinado de personas para el trabajo en campo, mismas que van a estar completamente dispersas en toda el área involucrada en el proceso.

4.2) Reclutar personal con el perfil necesario.

- Esto no parece tan complicado, reclutar a personas con un cierto perfil no debe ser tan difícil, pero hay que tomar en cuenta que para un proyecto como el que se tiene no es costeable reclutar gente en la central y mandarla a lugares alejados, por los costos que esto implica, y probablemente lo más adecuado es contactar gente en los puntos indicados y realizar el operativo a larga distancia.

4.3) Generar estrategias de campo para agilizar la obtención de los datos.

- Es posible diseñar alguna estrategia de campo para agilizar la recopilación de datos, de estas opciones hay que elegir la más eficiente y a menor costo.

4.4) Dirigir el Operativo en campo.

- Coordinar a todos los agentes recolectores de datos desde la central.

4.5) Supervisar el operativo en campo.

- Debe existir alguna forma de supervisar el operativo en campo, ver que la información que los agentes recolectores no sea falsa, y verificar que físicamente se cubran los puntos señalados por la muestra.

4.6) Diseñar formatos para estandarizar la información.

- Debido a que la información que se va a recoger es muy específica y que cada uno de los recolectores de información la puede interpretar a su forma, vale la pena estandarizar la recolección de la información, para ello es recomendable diseñar algunos formatos en donde sólo se ponga la información necesaria, y dé el criterio de rechazo a información redundante.

4.7) Dar solución a posibles problemas e imprevistos que surjan durante el operativo en campo.

- Algo que seguramente sucederá son los imprevistos en el trabajo en campo, por lo que es necesario tener el criterio estandarizado de qué se debe hacer en determinado caso, para esto es recomendable analizar cuáles

son los posibles problemas o imprevistos y dar solución anticipadamente.

4.8) Establecer comunicación con las diferentes áreas del sistema.

- Como ya se ha visto, siempre es necesario interactuar con otras áreas, en este caso esta área da y necesita apoyo, por lo cual es necesario como en los otros casos, apegarse a las bases establecidas para este propósito.

5. Departamento de Recursos Materiales y Administración.

Como en cualquier proyecto, siempre se involucran costos, y se asignan recursos, los cuales deben ser correctamente distribuidos. Por otro lado, existen requerimientos materiales por parte de las diferentes áreas de trabajo, que deben ser satisfactoriamente cubiertos, éstas y algunas otras necesidades son atendidas por este departamento, de aquí que entre sus principales funciones están:

5.1) Administrar los recursos disponibles.

- Como en todos los casos, un proyecto de éstos tiene un costo, para cubrirlo se destinan ciertos recursos, lo más recomendable es que sean administrados por un departamento completamente encargado de ello.

5.2) Abastecer de los recursos necesarios a las diferentes áreas.

- No sería práctico que, por ejemplo, el departamento de Cómputo esté comprando papel para impresora y otras cosas. Debido a esta clase de situaciones que pueden quitar tiempo y elevar los costos, es recomendable

que esta área que se encargue de ello.

5.3) Proporcionar instalaciones necesarias para la colocación del equipo que se requiera.

- Los espacios para operar, para instalar las computadoras, la línea eléctrica que se requiere para conectar todo el equipo, las líneas telefónicas para realizar la comunicación, etc., son elementos que pueden en un momento ser determinantes para el buen funcionamiento del sistema.

5.4) Realizar cualquier tipo de gestión necesaria.

- Seguramente existirá una serie de trámites que tendrán que realizarse, inclusive existen instancias a las cuales es necesario facilitar documentos con metodología, diseño de muestra, etc., para poder publicar los datos en algún medio masivo de comunicación.

5.5) Liquidar los adeudos generados durante el proyecto.

- Pagar a todo el personal involucrado en el proyecto, desde los recolectores de datos, los supervisores, proveedores, etc. Además de cualquier otro gasto que se derive durante el desarrollo del proyecto.

Esto es únicamente una propuesta de modelo funcional, como se puede observar, son muchos los aspectos que hay que tomar en cuenta, y por más extenso o completo que se diseñe este modelo siempre existirán detalles que faltarán por considerarse, sin embargo, normalmente no es complicado asignar esas tareas no consideradas aquí, al departamento adecuado.

4.4. Medidas de Seguridad para el Sistema.

Imagínese que ya tiene todo el sistema listo e instalado y que ya sólo falta introducir los datos para procesarlos y obtener los resultados tan esperados, pero que en ese momento llega a haber un corte de energía eléctrica, o logra introducirse un virus muy destructivo. Esto puede causar muchas horas de trabajo perdidas, si no se está preparado. En el caso de la luz, si se llega a ir por un instante no existe gran pérdida, sólo la información que se puede llegar a perder al no estar salvada. Pero si la luz se llega a ir por horas, todo el trabajo realizado puede llegar a ser una gran pérdida de tiempo, esfuerzo, inversión y prestigio.

En el caso de que se llegue a introducir un virus en el sistema, si el responsable del departamento de cómputo tuvo el cuidado suficiente para sacar respaldos de todo el sistema, lo peor que puede llegar a pasar es que los resultados se entreguen con cierto tiempo de retraso, cosa que crea ciertos conflictos, pues supuestamente se está hablando de un Sistema de Estimación Inmediata de Tendencias. Pero en el caso contrario no se logran obtener los resultados en el tiempo estimado (seguramente se obtendrían los resultados oficiales primero) y por lo tanto, el Sistema pierde validez al no conseguir sus objetivos. Como es de imaginarse, los gastos producidos para un evento como éste, son bastante altos.

Pueden llegar a presentarse otros problemas como el caso de que el sistema llegue a caerse, que algunas terminales dejen de funcionar, al igual que alguna línea telefónica, un módem, etc. Cualquier equipo de hardware puede llegar a dañarse.

En el caso de una red con topología bus si la primera terminal llegara a tener problemas, puede llegar a causar un caos, ya que las otras terminales no podrán

funcionar.

Existen otros problemas debido a la naturaleza del sistema, como el daño intencional, el robo de información confidencial o espionaje industrial, pues se maneja información clasificada sobre todo en aspectos políticos.

Hasta aquí se han planteado los problemas más comunes que pudieran surgir, a continuación se presentarán algunas posibles soluciones a los problemas antes mencionados.

Antes que nada, lo primero es una buena planeación del Sistema, ya que con ello se evitarán muchos problemas que pudieran surgir.

Energía Eléctrica.

Uno de estos problemas es el corte de suministro de energía eléctrica, siempre se está expuesto a que pueda suceder en cualquier momento, para evitar la pérdida de información y para asegurar el poder avanzar un poco más en la recolección y el procesamiento de datos, lo mejor es contar con equipos no-breakes. Estos equipos proporcionan un tiempo extra de luz en caso de que se corte el suministro, además de funcionar como regulador de voltaje. Existen diferentes tipos de no-breakes, con diferentes capacidades de tiempo de suministro de energía y para diferentes equipos (terminales, redes, etc.). Estos se deben de elegir dependiendo del tipo de sistema que se esté utilizando y del presupuesto que se tenga planteado.

Además de los no-breakes, se pueden contar con generadores de emergencia, pero como es de imaginarse, éstos son muy caros y si el proyecto solo va a realizarse por un tiempo, tal vez no sea conveniente, se tendría que analizar la conveniencia

de contar con uno al revisar la zona en donde se realizará el acopio y procesamiento de los datos (la periodicidad de estos eventos y su duración), además de analizar si los generadores pueden servir para trabajos posteriores.

Virus informáticos.

Otro problema que puede surgir es el de la introducción de algún virus informático. El primer virus conocido fue creado en 1983, y para 1987 ya eran una gran amenaza para todos los usuarios. Afortunadamente también existen antídotos y formas de protección contra los virus. Para proteger el sistema de ellos, primero se debe saber qué son y cómo funcionan, para ello se dará la siguiente definición.

"Virus Informático".

Alberto Rojas define en su artículo *¿Ya vacunó su PC?* a un virus informático como "Todo aquel código que al ser ejecutado altera la estructura del software del sistema y destruye programas o datos sin autorización ni conocimiento del operador".

Las definiciones más aceptadas en el libro **VIRUS en las computadoras** de G. Ferreyra² son la citada por Ralph Burger (1) y la del Club de Virólogos de Microcomputadoras de Guadalajara (2) que son las siguientes:

1. *"Un programa debe clasificarse como virus si combina los siguientes atributos:*

² FERREYRA Cortés, Gonzalo. *"VIRUS en las computadoras"*, México, D.F., Ed. Computec, 3a. Ed., p.p.50-51.

- *Modificación de códigos del software que no pertenece al propio programa virus, a través del enlace de las estructuras del programa virus con las estructuras de otros programas.*
 - *Facultad de ejecutar la modificación en varios programas.*
 - *Facultad para reconocer, marcándola, una modificación realizada en otro(s) programa(s).*
 - *Posibilidad de impedir que vuelva a ser modificado el mismo programa, al reconocer que ya está infectado o marcado.*
 - *El software modificado asimila los atributos anteriores para, a su vez, iniciar el proceso con otros programas en otros discos".*
2. *"Son programas que en forma prevista por sus autores, causan daños a otros programas, archivos, discos y otras partes de la computadora, y algunas veces se autorreplican completa o parcialmente".*

La definición de virus informático dada por Peter Norton es algo más entendible, *"Un virus biológico es un organismo simple, que se multiplica en células vivas y a menudo capaz de causar enfermedades. Un Virus de computadora se ajusta a esta descripción bastante bien. Suele ser un programa relativamente simple; puede que haya sido diseñado astutamente, pero por lo general es pequeño y difícil de detectar, a menudo de menos de 4KB. Los virus de computadora se multiplican en computadoras que están en operación y pueden causar daño. El nombre "virus", incidentalmente, fue utilizado originalmente por David Gerrold en una historia de ciencia ficción escrita a principios de los años setenta"³.*

Según Norton, un virus puede hacer cualquier cosa que se pueda realizar con los

³ Norton, Peter. *"Periféricos y Accesorios para la IBM-PC, PS/2 y Compatibles"*. México, Ed. Prentice Hall, pág. 301.

comandos de computadora, pues estos comandos también los puede realizar un programa, ya que un programa puede lograr copiarse a sí mismo en determinadas ocasiones. *"Por ejemplo, un virus en la sección de arranque de un disco en la unidad A "despertará a la vida" cuando una computadora se conecte. Un virus dentro de un programa ejecutable, puede iniciar su actividad cuando este archivo se corre. Una vez activo, el virus puede copiarse a sí mismo en cualquier disco cargado en la computadora, en cualquier otro sistema de computadoras conectado a través de un módem, o en otros programas ejecutables. Algunos virus inclusive buscan versiones anteriores de sí mismos y los actualizan; y otros son capaces de hacer mutaciones de sí mismos".*

"Hablando estrictamente, un virus es un fragmento de código que se inserta por sí mismo en un programa más grande, aunque actualmente el término se utiliza libremente para cubrir prácticamente cualquier programa que cause daño intencional".⁴

Los virus informáticos tienen muchas formas de operar. Los virus del sector de arranque se insertan a sí mismos en la sección de arranque de los discos flexibles o duros. Cuando se conecta la computadora, la primera parte que se transfiere de un disco a la memoria es el sector de arranque, así un virus pasa a la memoria antes que cualquier otra cosa. Los virus de sector de arranque se diseminan copiándose a sí mismos a los sectores de arranque de otros discos.

Otros virus se fijan a archivos ejecutables, como son los archivos .EXE y .COM, a los manejadores de dispositivos, y a los archivos de procesamiento por lotes. Cuando se ejecuta ese archivo, se ejecuta también el virus. Y los virus llamados *"furtivos"*, son algunos virus de los más recientes que obtienen acceso al BIOS del

⁴ Op. Cit., pág. 302.

sistema, y pasando a ser parte de las tablas vectorizadas de interrupción de DOS convirtiéndose en una de las peores especies.

Los virus pueden ser tan inofensivos como los que no causan más problema que el de quitar un poco de tiempo desplegando mensajes, hasta causar daños irreparables y sumamente costosos, como borrar archivos, impedir salvar la información que se encuentra en pantalla en ese momento, y peor aún, que se caiga el sistema, que destruyan archivos, reformateen el disco duro, o que lleguen a quemar el monitor o los drives.

Para protegerse de los virus, se pueden tomar algunas precauciones básicas:

- Antes de instalar un software nuevo, hay que colocar la protección en los discos flexibles, para evitar que los virus se propaguen de la PC a los discos.
- Algunos programas de instalación no funcionarán si se protege a los discos contra escritura. En este caso, hay que volver a arrancar la computadora con un disco libre de virus en la unidad de discos A. Inmediatamente después, instalar el software nuevo, sin cambiar a la unidad C.
- Instalar solamente software original. Devolver el software si la cubierta ha sido violada.
- Si se tiene que utilizar programas de dominio público o shareware, tener cuidado. Comprar los programas de empresas que usen algún sistema de prevención de virus para mantener discos limpios. Tener cuidado especialmente con programas a cuyo autor se desconozca.

- Instalar solamente el software que realmente necesite y utilice.
- Modificar los archivos ejecutables a sólo lectura.
- Comprar y utilizar software antivirus. Vacunar el software nuevo antes de utilizarlo.
- No permitir que los capturistas y gente ajena al equipo de cómputo instalen su propio software. Sólo consentir la instalación de software que se haya revisado primero. Se debe asegurar de que los visitantes no instalen software sin consentimiento.
- Utilizar una PC como una máquina aislada de "demostración" para prueba de software traído por empleados, vendedores y consultores.
- Si se cuenta con una red, se debe considerar la posibilidad de instalar estaciones de trabajo sin unidades de disco, a fin de evitar que los capturistas instalen su propio software.
- Debe asegurarse de que los capturistas transfieren archivos por vía de la red, y no transportando diskettes entre PC.
- Efectuar respaldos del sistema periódicamente.

Software antivirus.- Es una forma de prevenir la introducción de un virus al sistema, además en caso de llegar a infectarse, logra eliminar el virus sin causar mayor daño (en caso de detectarse a tiempo). Existen muchas versiones en el mercado, entre las más comunes o populares se encuentran Norton Antivirus, Mccafee, etc.

Los programas antivirus pueden proteger el sistema de diferentes maneras. Se puede cargar como un programa residente en la memoria de la computadora para vigilar continuamente que no entre algún virus y avisar si encuentra alguno; puede rastrear los discos, -tanto duros como flexibles- buscando virus y quitándolos o neutralizándolos antes de que puedan hacer daño alguno; y puede *"inmunizar"* los archivos, asegurándose de que no puedan ser infectados.

Hardware antivirus.- Existen opciones en hardware y combinaciones de hardware y de software que pueden ser utilizadas para proteger el sistema contra virus, por ejemplo PC-cillin que detecta infecciones y también salva los sectores de arranque y la información de la tabla de partición en memoria en un enchufe especial "inmunizador" conectado al puerto paralelo de la PC. Si el sector de arranque o la tabla de partición son dañados por un virus, el inmunizador puede volver a copiarlos al disco duro. Este producto también se combina con software que se carga en la memoria de la computadora para vigilar contra actividad viral y que rastrea los archivos en busca de virus.

Como se mencionó en el capítulo 3.1, si se utiliza el Sistema Operativo UNIX, no es necesario realizar todas estas precauciones, pues hasta ahora todavía no se han realizado virus que ataquen en este sistema operativo, ésta puede ser otra precaución que puede tomar.

Daño intencional, robo de información confidencial y espionaje industrial.

Estos pueden realizarse al dejar las computadoras sin supervisión, mediante inserción de discos, por comunicaciones vía módem, comunicaciones en redes y derivaciones de comunicaciones. Esta última forma de violación al sistema puede hacerse relativamente fácil al interceptar información transmitida por un módem conectado

a un teléfono celular. Si la transmisión es efectuada utilizando una línea telefónica normal, es más difícil aunque posible, el derivar directamente de la línea.

Este problema se puede evitar en gran medida controlando el acceso al equipo de cómputo, para ello existen diferentes formas de lograrlo:

- Evitando que personas no autorizadas se acerquen a las computadoras. Como es de imaginarse, esto no garantiza gran cosa en este sistema, ya que existirán muchas personas que manejen las computadoras, entre ellas se encontrarán todos los capturistas, que estarán hasta cierto punto autorizadas a manejar las computadoras.
- Existen algunos métodos simples que se pueden usar para impedir que intrusos lleguen hasta la información, aun con la computadora encendida y dentro del sistema. Por ejemplo, se puede utilizar el comando de DOS ATTRIB para que los archivos sean solamente de lectura o estén bien escondidos, a fin de que el intruso no pueda encontrarlos ni borrarlos. El entorno gráfico de DOS 5.0 tiene una contraseña que evita que usuarios no autorizados puedan arrancar aplicaciones de dicho entorno, aunque aún pueden hacerlo desde la línea de comando. Algunos menús de programas tienen la función de contraseña. Estos métodos son simples y no muy eficaces.
- Muchas aplicaciones permiten asignar contraseñas a archivos específicos, por lo que aunque se pueda arrancar dicha aplicación, el archivo en cuestión no puede ser abierto. Algunos programas de hecho cifran el archivo, por lo que, aun si se consigue abrirlo en un procesador de palabras ASCII o mediante un programa de utilerías de archivos como Norton Utilities, no se tendrá acceso a la información. Pero también el usuario tiene forma de obtener

la en caso de que realmente quiera obtenerla, a menos que sea un archivo encriptado, generalmente éstos no pueden ser violados, y algunos fabricantes no dirán cómo descifrarlos en caso de perder la contraseña, por lo que si se pierde la contraseña el archivo se perdió para siempre.

- Los archivos borrados pueden ser recuperados mediante diferentes técnicas, a menos que ya se haya escrito algo encima de éste. Para evitar que se recupere esta información se puede reformatear el disco, usar un programa de utilerías de archivo que escriba caracteres sin sentido en el área del disco duro donde estaba almacenado el archivo a borrar, o bien utilizar un programa de optimización o compresión de disco duro, que escriba basura en las porciones aún no utilizadas del disco.

Todos estos procedimientos son de uso limitado. Si realmente se desea proteger el sistema de manos de intrusos existe una variedad de métodos a utilizarse, ya sea de hardware o de software. Generalmente se basan en dos cosas:

1. Algo que sólo el encargado del área sabe, como una contraseña.
2. Algo que sólo el encargado del área posee, como una tarjeta o una llave.

Algunos de los procedimientos más avanzados se apoyan en otra característica. Algo que sólo el encargado del área es. Sistemas denominados biométricos analizan algo del encargado, huellas digitales, el tono de su voz, etc., para cerciorarse de que la persona encargada es quien dice que es. Estos dispositivos son generalmente costosos y poco frecuentes.

- Programas de control de acceso.- El software de seguridad puede limitar el acceso a la computadora o a archivos en especial, a directorios, a unidades de discos flexibles y a puertos externos. Algunos programas inclusive permi-

ten convertir las unidades de disco flexible en sólo de lectura, por lo que un intruso podrá estudiar los archivos pero no robárselos. Los productos de seguridad en software no son tan seguros como los de hardware, pero su gran ventaja es su bajo costo, lo que puede resultar atractivo en grandes instalaciones. Los programas varían en costos de 60 a 300 dólares, tal vez más. Muchos programas de utilerías de archivos, como Norton Utilities, también ofrecen cierto número de características de control de acceso, como ciframiento de datos y puesta bajo llave de archivos y directorios.

El software de control de acceso es especialmente útil para uso en redes. El administrador puede establecer diferentes clases de usuario. Algunos usuarios pueden tener derecho a un uso muy restringido. Otros usuarios pueden tener muchos más derechos, permitiéndoseles el uso de discos flexibles y puertos de comunicación, pero quizá aún negándoles acceso a ciertos directorios.

- Selección de contraseñas.- Las contraseñas son casi un chiste hoy en día. Son comunes las contraseñas fáciles de adivinar, con nombres de niños, nombres propios, iniciales, con términos relacionados con pasatiempos, y además, que se escriben, e inclusive se pegan a la pantalla de la computadora, a fin de que los colegas puedan arrancar la computadora, cuando sea necesario.

De hecho la forma más común de violar la seguridad de una red es adivinando contraseñas débiles. Los intrusos pueden simplemente adivinar contraseñas, utilizando conocimientos que poseen respecto a un usuario en particular, o pueden buscar en correo electrónico referencias a contraseñas. Algunos invasores utilizan inclusive diccionarios de contraseñas comunes, así como programas para descifrar contraseñas que puedan utilizar esos

términos.

Existe un producto que auxilia a los usuarios de redes a generar contraseñas seguras. Password Coach de Baseline Software es un producto para redes que analiza contraseñas, buscando alrededor de 45 contraseñas fáciles de violar: También puede generar un diccionario de contraseñas prohibidas. Si Password Coach encuentra una contraseña que sea débil, guiará al usuario a través de un procedimiento para crear una más segura.

- Hardware de control de acceso.- El hardware de control de acceso varía desde lo simple y barato hasta lo complicado y costoso. Una de las formas más sencillas de hardware de control de acceso es la cerradura. Muchas computadoras, quizás la mayoría en estos días tienen una o dos cerraduras incorporadas en su caja -una para desactivar el teclado y algunas veces otra para imposibilitar que se retire la cubierta de la computadora-. Las computadoras IBM utilizan cerraduras únicas, cada llave es diferente. Muchos fabricantes de clones, sin embargo, utilizan cerraduras baratas, y una llave puede abrir docenas de otras computadoras. Estas cerraduras impiden a personas no autorizadas utilizar la computadora o literalmente evitan que se lleven el disco duro. No son 100% seguras, naturalmente. Si puede abrir la caja de la computadora, usualmente se podrá desconectar la cerradura del teclado.

También se puede obtener cerraduras que impidan que alguien se lleve toda la computadora. Y otra forma de cerradura es una para la unidad de discos flexibles. Estas cuestan alrededor de 25-30 dólares. Impedirán que alguien que inserte un disco en la unidad, lo pueda retirar, y por lo tanto un intruso no puede cargar su propio software, copiando su información en un disco, o arrancar la computadora a partir de un disco flexible.

- **Candados.-** Son pequeñas cajas que se enchufan en el puerto paralelo de la computadora. El software comercial protegido contra copias es raro, pero sigue. Muchas compañías que producen software para utilización interna o bien software de alto costo para venta comercial limitada, desean proteger su inversión. Los candados ofrecen una forma de hacer difícil, aunque no imposible, el robo de software valioso.
- En caso de utilizar el planteamiento 2 propuesto en el capítulo 4.2., existe la facilidad de que cada computadora sólo contiene el sistema operativo, el programa ejecutable de captura y la base de datos, por lo que sería bastante difícil robar la información completa. De igual manera, si llega a dañarse la información, sólo pasará en esa computadora.

Respaldos.

Los respaldos se deben realizar para cualquier tipo de sistema que se haya elegido y para cualquier software. Los respaldos son el seguro con que el sistema cuenta para cualquier accidente (ya sea falla de energía eléctrica, introducción de virus, fallo en alguna máquina, robo de información, etc.), además de que sirve de respaldo para cualquier auditoría que pueda presentarse para la verificación de la fiabilidad de los resultados.

Al adquirir cualquier software se debe de realizar un respaldo (Backup) del mismo, además de respaldar la información que considere importante, en el caso del SEIT, se debe de tener respaldos de todo el software que se requiere (Sistema operativo, manejador de base de datos, graficadores, software de comunicación, vacunas, etc.), además de cada base de datos, los programas que obtendrán los resultados

finales, archivos con los resultados finales, archivos de machote de gráficas y presentación de resultados, etc.

Falla del equipo.

Además de todo lo anterior, se debe prever el caso de que en algún momento llegue a fallar una computadora, un teléfono, un módem, un cable, etc., o varios de ellos, por lo que es necesario contar con equipo extra.

Como se mencionó en el capítulo anterior, en el caso de usar alguna red, lo más conveniente en el caso del SEIT, puede ser el uso de la topología de estrella, ya que en caso de que alguna máquina llegara a fallar, el paso de la información no llega a afectarse, ya que cada una de las máquinas está conectada al servidor, y no como en una topología de anillo o lineal, en las cuales si alguna computadora llegara a fallar, las siguientes no podrían recibir ni mandar información.

Sintetizando lo que se vio en este capítulo, en la primera parte se describió detalladamente el procedimiento para el diseño de la muestra, se definieron de forma clara y específica aspectos importantes como son: marco muestral, tamaño de la muestra y el método mediante el cual se diseñó dicha muestra; todo esto aplicado a un caso práctico como lo fueron las elecciones federales de 1994, específicamente las elecciones para presidente. Todo lo anterior tomando como apoyo el padrón electoral vigente en 1994. A continuación se analizó la forma en que fluye la información durante el proceso señalando los límites del SEIT, para con ello plantear posteriormente algunas propuestas de cómo podría funcionar el SEIT. Las propuestas planteadas son seis, las cuales llegan a ser muy diferentes entre algunas de ellas pues se buscaron planteamientos económicos y eficientes, sin

embargo algunos de estos tienen un nivel de eficiencia mucho mayor que otros, aunque no tan económicos como seguramente se pudo observar. Posteriormente se propone un modelo funcional a nivel de área de trabajo, en donde se describen de forma breve las funciones de las diferentes áreas de trabajo, y finalmente, en este capítulo se abordó el tema de la seguridad, que como ya se vió es un aspecto muy importante y amplio, debido a la gran cantidad de posibles contingencias que pueden presentarse durante el proceso.

CAPITULO V

APLICACION, EVALUACION Y POSIBILIDADES DE EVOLUCION DEL SEIT.

5.1. Recopilación de los resultados de las casillas en la muestra.

En este capítulo se llevará a cabo una aplicación de lo que se ha visto anteriormente, se pretende diseñar un par de muestras representativas del número de casillas de votación del día 21 de agosto de 1994, de las elecciones federales para presidente en México, para poder evaluar las técnicas antes mencionadas, se compararán los resultados entre sí y entre los diversos sistemas de conteos rápidos realizados en esa fecha, además se presentarán algunas herramientas diseñadas para facilitar esta tarea mediante el uso de la computadora, y al final se comentará un poco respecto a las posibilidades de evolución del SEIT.

Para llevar a cabo esta tarea fue necesario realizar un extenso y laborioso trabajo, debido a que el IFE no puso a la disposición pública esta información hasta mediados de 1996, los resultados estaban disponibles en una publicación llamada: *"Estadística de las Elecciones Federales de 1994, Compendio de Resultados"*, con la única salvedad de que los resultados estaban a nivel distrito federal electoral y municipio, de modo que fue necesario acudir a los archivos del IFE, a consultar esta información, afortunadamente estaba disponible, con el inconveniente de que no era posible obtener copia fotostática de nada, y que tampoco era posible adquirir algún tipo de publicación que contuviera dicha información y estuviera a la venta, de modo que fue necesario copiar a mano los resultados de cada una de las casillas que previamente se habían seleccionado. Tarea que consumió varias semanas.

Para realizar el diseño de la muestra, fue necesario construir el marco muestral, lo cual como se mencionó en el capítulo 2.4., no fue difícil, ya que con el apoyo del cuadro 4.1.1 se puede observar que en cada estado se tiene un número determinado de secciones, las cuales están numeradas del uno a dicho número, por

lo tanto únicamente se diseñó una base de datos con los números de secciones electorales que tiene cada estado de la República, de esta manera se obtuvo una base de datos con 63,596 registros, cada registro contiene un campo con la clave del estado y otro con el número de la sección, después se procedió a seleccionar aleatoriamente el número de secciones por estado que deberían entrar en la muestra, es decir, se estratificó la muestra.

Se hace la aclaración de que por las características del estudio, cada estrato queda delimitado por cada estado de la República, el número n_i que denota el número de secciones por estrato fue calculado por la proporción de padrón electoral de cada estado con respecto al padrón nacional, tomando como base el tamaño de la muestra de 400 unidades, que son suficientes para lograr representatividad, lo cual ya se demostró con anterioridad, esta distribución se muestra en el cuadro 4.1.1, de modo que una vez determinado esto, se procedió a la selección de las secciones que entrarían en la muestra, para ello se diseñó una rutina en lenguaje de programación "Clipper", para que aleatoriamente seleccionara las secciones electorales por estado.

Un aspecto que se debe tomar muy en cuenta es el tipo de casilla del que se trata, se tiene la teoría de que los comportamientos electorales varían mucho de las casillas de tipo rural y urbano, motivo por el cual es necesario mantener un equilibrio entre ambos tipos de secciones. Como al diseñar el marco muestral, no se puede calificar el tipo de casilla que se trata, se realizó la selección dejando ese aspecto momentáneamente de lado.

Al momento de la recopilación afortunadamente para el sistema, las casillas tenían esta clasificación, la cual se muestra en los resúmenes de la muestra, en los anexos 1 y 2, y como se puede apreciar no existe problemas de representatividad.

Para efectos de validar lo antes mencionado, a la hora de recopilar los datos, se obtuvo el nombre del municipio dentro del cual se encontraba la sección, en los anexos antes mencionados se incluye también una lista de secciones en la muestra ubicadas por estado y municipio.

Para poder dimensionar la distribución de las casillas, en los siguientes cuadros se presenta un resumen de la muestra, desglosado por estado y tipo de casilla, ya sea rural o urbana, y su proporción en cada estado, para el caso de la primera muestra el resumen se localiza en el anexo 1, y para poder apreciar la cobertura geográfica dada por la dispersión de las secciones en muestra se debe observar la gráfica 5.1.1.

De la misma manera, para una segunda muestra se presenta el cuadro resumen, con las mismas características de la primera muestra, esto en el anexo 2.

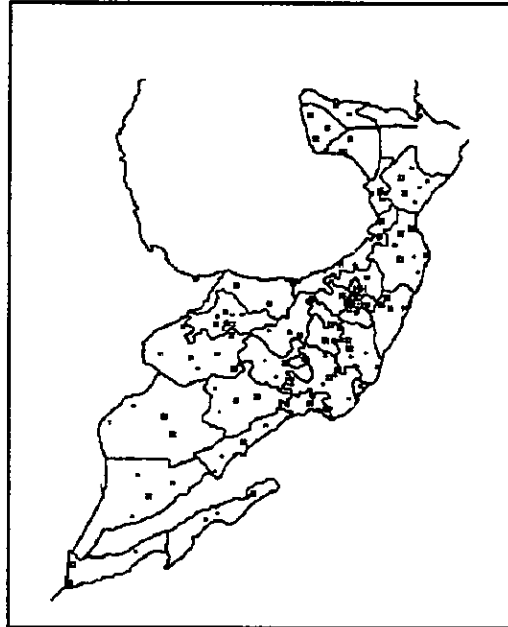
5.2. Organización y procesamiento de los datos obtenidos.

Una vez recolectados los datos se procederá a su análisis, para que puedan ser completamente manejables, es evidente que los medios magnéticos representan un sin fin de bondades, de esta forma la manera más adecuada de realizar los cálculos de esta información es mediante una base de datos, cuya estructura se puede observar en la figura 5.2.1.

Para efectos de agilizar la captura, pensando que durante el momento de efectuar dicha función, el tiempo es lo más importante y cualquier ahorro de

Gráfica 5.1.1.

COBERTURA GEOGRAFICA DE LA PRIMERA MUESTRA



éste es vital, se recurre al uso de claves para identificar ciertos elementos, como son por ejemplo, el tipo de casillas en donde en vez de capturar "Rural" o "Urbano", únicamente se captura una "R" o una "U" según sea el caso, el estado también se maneja de esta forma, las claves para identificarlos a cada uno, se asignaron realizando una lista, ordenando alfabéticamente los estados y numerándolos, de esta forma en lugar de capturar, por ejemplo, "Aguascalientes", únicamente se captura su clave que en este caso será "1". De la misma forma para identificar el municipio al cual pertenece cada una de las secciones, únicamente se introduce la clave, la cual fue tomada de las listas de donde se obtuvieron los datos de las secciones, por lo que en este caso asignamos las claves que tiene el IFE.

Lo anterior con la finalidad de poder validar la información, y tener identificada la procedencia de los datos, ya que ésta, no es necesaria para la presentación de los resultados, pero quizá posteriormente se requiera para algún tipo de estadística, validación o demostración, por lo tanto se recomienda conservarla.

Para la realización de la base y por motivos obvios, fue necesario realizar la captura de las 800 secciones en la muestra, recordando que fueron dos muestras de 400 secciones cada una, donde además se tuvo la oportunidad de poner a prueba la aplicación diseñada para satisfacer esta necesidad, las opciones de trabajo ubicadas dentro del SEIT, pueden ser las siguientes:

- **Una primera opción** es el diseño una sistema para la administración de la base de datos elaborado en Fox Pro, versión 2.6, para Windows. Este sistema puede trabajar en ambiente de RED, es decir, desde varias computadoras se puede acceder a la misma base de datos, sin generar ningún conflicto, además de gozar de la capacidad gráfica que ofrece Windows, esto con el fin de tener una pantalla de captura atractiva y de fácil manejo.

**ESTRUCTURA DE LA BASE DE DATOS
EMPLEADA PARA LA CAPTURA**

| Campo No. | Nombre del campo | Tipo de campo | Longitud del campo |
|-----------|------------------|---------------|--------------------|
|-----------|------------------|---------------|--------------------|

| | | | |
|----|----------|----------|---|
| 1 | ESTADO | Numerico | 2 |
| 2 | MPIO | Numerico | 3 |
| 3 | LOC_URB | Caracter | 2 |
| 4 | SECCION | Numerico | 4 |
| 5 | CASILLA | Numerico | 4 |
| 6 | HORAREP | Caracter | 5 |
| 7 | PAN | Numerico | 5 |
| 8 | PRI | Numerico | 5 |
| 9 | PPS | Numerico | 5 |
| 10 | PRD | Numerico | 5 |
| 11 | PFCRN | Numerico | 5 |
| 12 | PARM | Numerico | 5 |
| 13 | UNO | Numerico | 5 |
| 14 | PT | Numerico | 5 |
| 15 | PVEM | Numerico | 5 |
| 16 | NOREGIST | Numerico | 5 |
| 17 | VVALIDOS | Numerico | 5 |
| 18 | VNULOS | Numerico | 5 |
| 19 | VTOTAL | Numerico | 5 |

Cuadro 5.2.1.

El proceso de esta información se puede realizar, con una rutina que también puede estar elaborada en Fox Pro, debido a lo básico de los cálculos, es decir, no se necesitan funciones matemáticas sofisticadas.

- **Una segunda opción** está dada por un programa de captura de datos elaborado en CLIPPER v.5.1, en máquinas conectadas entre sí por un programa de comunicación, no necesariamente trabajando en RED, además sería bajo ambiente de MS-DOS, lo cual simplifica mucho la instalación, y elimina muchos riesgos de conexión, bajo este planteamiento, no se estaría trabajando en una sola base de datos, el procedimiento sería realizar cortes en las diferentes máquinas y mediante la conexión copiar la base de la máquina de captura a otra que estaría concentrando la información.

Una vez realizado el corte de varias computadoras con un proceso diseñado también en CLIPPER, juntar las bases, en una sola, en este planteamiento se necesitarían 6 computadoras conectadas entre sí, para capturar 400 registros y realizar la fusión de las bases.

Para el proceso de la información de la misma manera, puede llevarse a cabo con una rutina elaborada en CLIPPER, debido a que, como ya se mencionó los cálculos no son tan sofisticados.

Bajo cualquiera de las dos opciones presentadas, cada computadora tendría una carga de trabajo de 80 registros, si se estima que la hora en la que se empieza a transmitir la información es dos horas después de cerrar las casillas de votación, se podría decir que a las 20:00 hrs. se estarían recibiendo las primeras boletas, vía telefónica, y que en una computadora es posible capturar 80 registros, con esta cantidad de información, en aproximadamente 2 horas se pueden reportar las tendencias, por decir algo, a las 22:00 hrs. con el 100% de la muestra registrada.

Es evidente que esto en práctica no es posible, y no es el equipo de cómputo el que retarda la presentación de los resultados, en realidad ese problema puede tener sus orígenes en el trabajo de campo, debido a que pueden surgir un sin fin de contratiempos, como: que no existan teléfonos cerca de la sección, o que las personas que están llevando a cabo el conteo de los votos dentro de la casilla no se pongan de acuerdo y retarden la publicación de éstos, etc. Sin embargo, cabe señalar que desde el punto de vista estructura computacional, el problema está resuelto.

Con respecto al procesamiento de los datos, como se mencionó, no es un proceso matemáticamente sofisticado, debido a que la cantidad de los datos no es muy grande, además que la forma de presentar los datos está dada en porcentajes, debido a que es el lenguaje que el común de la gente conoce, además de ser el estándar de presentación en las encuestas de opinión pública, y de los estudios de este tipo, de esta forma, el proceso consiste en acumular los votos de los partidos en la contienda electoral y calcular sus proporciones, es decir, lo que hay que hacer son sumas, divisiones y multiplicaciones.

Las opciones para trabajar esta parte del sistema, como ya se presentó, son muchas y muy variadas, inclusive se puede pensar que con un nivel de eficiencia diferente. Estas opciones son solamente un ejemplo de cómo podría realizarse el manejo de la información, actualmente existe software tan potente y sofisticado que no representa ninguna dificultad manipular una base de datos de máximo 800 registros. Por otro lado, cabe señalar el papel que juega el ingenio de la persona que desarrolle los programas para procesar la información, pero de alguna forma lo que se pretende transmitir con esto, es la posibilidad de realizar este manejo de información de una forma fácil, eficiente y barata.

5.3. Presentación de los resultados.

Este subcapítulo es uno de los más interesantes, ya que es aquí en donde se podrán ver los resultados de todo el trabajo realizado. Donde, para empezar, se conocerán cuáles fueron los resultados que arroja el sistema, y se buscará la forma de presentarlos lo más claramente posible, además de que debe ser entendible para el común de las personas que puedan tener acceso a ella.

Se comenzará con el análisis de los resultados con respecto a la muestra número uno, como se puede ver en el cuadro 5.3.1 en donde se muestran para empezar el total de los votos que se obtuvieron para cada uno de los partidos, y en donde se estima el total de los mismos que es de 174'066 votos, en la información que se presenta es evidente que las tendencias ubican al PRI en primer lugar con un total de 80'876 votos y una diferencia de 26'107 votos con respecto a su inmediato rival, el PAN, que cuenta con un total de 54'769 votos y que a su vez mantiene una diferencia de 27'810 votos, con respecto al que políticamente se denomina tercera fuerza, que es el PRD, con un total de 26'959 votos, el siguiente contendiente es el PT con una diferencia de 21'447 votos, dándole un total de 5'512 votos.

Al decirlo así, ya son muy claros los resultados, sin embargo, no es suficientemente claro como para saber con qué margen de diferencia divide a la primera de la segunda fuerza electoral, y a ésta de la tercera, y así sucesivamente; matemáticamente se puede manejar el término proporciones, lo cual es más que claro para la persona que esté involucrada con el manejo y el concepto, es decir, resulta significativamente diferente mencionarlo de la siguiente forma. El PRI mantiene una proporción del 0.466 del total de los votos, y su inmediato seguidor el Pan obtuvo una proporción de 0.31, contra el 0.154 de la tercera fuerza que es

el PRD, quien a su vez es seguido por el PT con una proporción del 0.031 del total de la votación.

Sin embargo, la nomenclatura más común, que es incluso la que maneja el IFE, que es el órgano oficial en materia electoral, se puede mencionar que en los resultados del estudio el PRI obtuvo el 46.46% del total de los votos, contra el 31.46% que obtuvo el PAN, el PRD obtuvo el 15.48% del total de la votación y el 3.16% del PT. Lo cual parece ser un lenguaje más al alcance de la población en general.

De aquí que se procura manejar los resultados en términos de porcentajes, para efectos de realizar la presentación gráfica.

Por otro lado, los resultados de la segunda muestra, como se puede ver en el cuadro 5.3.2., realmente se mantienen muy parecidos a los de la primera muestra, las estimaciones son muy estables como se puede observar, analizando los resultados con detalle, para lo cual se emplearán las medias aritméticas, se puede observar que, por ejemplo el PRI, tiene una diferencia en las medias matemáticas de 6.59 votos a favor por casilla en la segunda muestra, y sin embargo, a nivel de tendencias éstas se mantienen con una diferencia de 0.41%, es decir en la primera muestra la tendencia estimada para el PRI es del 46.46% contra el 46.87% de la segunda muestra, para el caso del PAN la diferencia en sus medias aritméticas es prácticamente despreciable, pues sólo es de 0.247 votos más por casilla en la primera muestra, lo cual se refleja en las tendencias de la siguiente forma, en la muestra uno la tendencia fue de 31.46% contra la segunda que fue de 30.68% lo que significa una diferencia del 0.78%. Para el caso del PRD la diferencia en medias es de 4.128 votos más por casilla en la muestra 2, y a nivel de tendencias de la muestra 1 que es de 15.48% a 16.06% de la segunda muestra se tiene una diferencia de 0.58%. Y con respecto al PT, que sería la cuarta fuerza electoral se

MUESTRA 1

| Partido | Votos | Media Estimada | Tendencia | Media Estratificada | Tendencia M(st) | ECM M(st) |
|--------------|----------------|-----------------|-----------------|---------------------|-----------------|-----------|
| PAN | 54.769 | 136.9225 | 31.4645 | 137.3239 | 31.515 | 47.694 |
| PRI | 80.876 | 202.1900 | 46.4628 | 202.4880 | 46.470 | 4682.939 |
| PPS | 1.320 | 3.3000 | 0.7583 | 3.2905 | 0.755 | 0.436 |
| PRD | 26.959 | 67.3975 | 15.4878 | 67.2590 | 15.436 | 626.050 |
| PFCRN | 1.567 | 3.9175 | 0.9002 | 3.8888 | 0.892 | 0.658 |
| PARM | 670 | 1.6750 | 0.3849 | 1.6710 | 0.383 | 1.874 |
| UNO | 424 | 1.0600 | 0.2436 | 1.0720 | 0.246 | 0.219 |
| PT | 5.512 | 13.7800 | 3.1666 | 13.8227 | 3.172 | 2.182 |
| PVEM | 1.969 | 4.9225 | 1.1312 | 4.9249 | 1.130 | 0.055 |
| Total | 174.066 | 435.1650 | 100.0000 | 435.7408 | 100.0000 | |

Figura 5.3.1.

MUESTRA 2

| Partido | Votos | Media Estimada | Tendencia | Media Estratificada | Tendencia M(st) | ECM M(st) |
|--------------|----------------|----------------|------------|---------------------|-----------------|-----------|
| PAN | 54.807 | 136.6758 | 30.6875 | 136.8909 | 30.7086 | 53.862 |
| PRI | 83.721 | 208.7805 | 46.8770 | 208.9164 | 46.8660 | 3844.446 |
| PPS | 818 | 2.0399 | 0.4580 | 2.0477 | 0.4594 | 0.339 |
| PRD | 28.682 | 71.5262 | 16.0596 | 71.5532 | 16.0515 | 429.600 |
| PFCRN | 1.368 | 3.4115 | 0.7660 | 3.4212 | 0.7675 | 1.635 |
| PARM | 660 | 1.6459 | 0.3695 | 1.6382 | 0.3675 | 1.965 |
| UNO | 656 | 1.6359 | 0.3673 | 1.6378 | 0.3674 | 0.010 |
| PT | 6.009 | 14.9850 | 3.3646 | 14.9860 | 3.3618 | 0.099 |
| PVEM | 1.876 | 4.6783 | 1.0504 | 4.6823 | 1.0504 | 0.228 |
| Total | 178.597 | 445 | 100 | 445.7737 | 100.0000 | |

Figura 5.3.2.

MUESTRA GLOBAL

| Partido | Votos | Media Estimada | Tendencia | Media Estratificada | Tendencia M(st) | ECM M(st) |
|--------------|----------------|----------------|------------|---------------------|-----------------|-----------|
| PAN | 109.576 | 136.7990 | 31.0710 | 137.0622 | 31.1078 | 51.377 |
| PRI | 164.597 | 205.4894 | 46.6726 | 205.6189 | 46.6674 | 4264.234 |
| PPS | 2.138 | 2.6692 | 0.6062 | 2.6697 | 0.6059 | 0.002 |
| PRD | 55.641 | 69.4644 | 15.7774 | 69.3875 | 15.7482 | 524.067 |
| PFCRN | 2.935 | 3.6642 | 0.8322 | 3.6550 | 0.8295 | 1.092 |
| PARM | 1.330 | 1.6604 | 0.3771 | 1.6549 | 0.3756 | 1.919 |
| UNO | 1,080 | 1.3483 | 0.3062 | 1.3552 | 0.3076 | 0.034 |
| PT | 11.521 | 14.3833 | 3.2669 | 14.3984 | 3.2679 | 0.813 |
| PVEM | 3.845 | 4.8002 | 1.0903 | 4.8028 | 1.0900 | 0.128 |
| Total | 352.663 | 440 | 100 | 441 | 100.0000 | |

Figura 5.3.3.

tiene una diferencia en sus medias de 1.205 votos por casillas más en la muestra dos, y en tendencias la diferencia es de 3.167% de la primera muestra a 3.36% de la segunda muestra, es decir solamente un 0.19%.

El análisis para los demás partidos no se realizará, debido a que sus proporciones los colocan muy lejos de las primeras tres fuerzas, además como se puede ver en los cuadros 5.3.1 y 5.3.2, también se presentan con bastante estabilidad.

Únicamente para efectos de análisis se formará lo que podría llamarse una tercera muestra, la cual consistirá en juntar las dos muestras anteriores; debido a que las dos muestras cuentan con los criterios de dispersión requeridos, y ambas están diseñadas por el método del muestreo aleatorio estratificado. No existe el riesgo de que algún error de estimación se acrecente, por el contrario se puede pensar que dicho error disminuirá, debido a que el tamaño de la muestra se incrementa.

Como se puede ver, en la mayoría de los casos no existen cambios substanciales en cuanto a las tendencias de los nueve partidos, es posible observar que, en apariencia es como calcular promedios de las tendencias y presentarlos como una tercera muestra, el procedimiento que se realizó para calcular las tendencias fue exactamente idéntico al realizado en la evaluación de la primera y segunda muestra, ahora como se analiza, las diferencias entre la primera y la segunda muestra son mínimas.

Los resultados de la evaluación de la tercera muestra, se presentan en el cuadro 5.3.3, en donde como se puede observar, presentan una estabilidad bastante aceptable.

En los cuadros 5.3.1, 5.3.2, 5.3.3, se incluyen las estimaciones de la media estimada y de la media estimada estratificada, las tendencias estimadas también

presentadas en los cuadros, están calculadas en base a sus medias respectivamente, además se presenta el cálculo del error cuadrático medio (ECM), el cual fue realizado con los resultados finales del proceso, los cuales son presentados en el cuadro 5.3.5.

Analizando los resultados de las primeras dos muestras, se puede pensar que las estimaciones están muy cerca de la realidad, pues ¿cómo podría ser posible que con dos muestras completamente independientes se obtengan las mismas estimaciones de las tendencias?, la única explicación es: uno, que realmente se está muy cerca de la realidad y que la técnica de muestreo es apropiada y el diseño de la muestra es el adecuado, o dos, que en ambos casos se cometió el mismo error en el muestreo y los resultados son igualmente erróneos.

Por otro lado, tomando en cuenta lo que es la presentación gráfica de los resultados, en una computadora se tienen varias opciones, la aquí presentada es sólo una alternativa, ésta se da como una alternativa cómoda, fácil y cumple ampliamente con todos los requisitos necesarios para este objetivo, se trata de una rutina en Harvard Graphics, la versión que se empleará es la 2.00 para Windows.

El formato para presentar la información será mediante una gráfica de barras representando la proporción de votos para cada partido, para efecto de analizar el impacto que puede tener la presentación de esta forma se observa la gráfica 5.3.4, en donde se aprecian las tendencias estimadas en base a la primera muestra.

La rutina diseñada tiene entre sus características diferentes propiedades: para comenzar, es capaz de leer en un archivo del disco duro, la información contenida en código ASCII que es necesaria que para actualizar la gráfica desplegada, situación que puede ser explotada ya que del programa de captura es posible crear un reporte que quede grabado en el disco duro, en el formato que sea necesario.

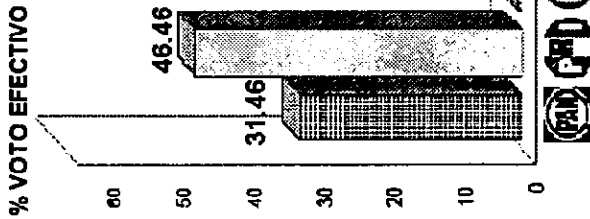
Otra de las ventajas de Harvard Graphics es la facilidad de emplear las funciones de tiempo disponibles en el paquete, con lo cual se puede controlar los lapsos de actualización de las gráficas, es decir, se pueden programar cortes cada determinado lapso de tiempo, para actualizar los datos desplegados en éstas, o si uno lo prefiere puede ser controlado desde el teclado, es decir, una vez que está lista la actualización, se puede pulsar una tecla y automáticamente se actualiza la información en la gráfica.

Esta opción es presentada debido a que es práctica y sencilla, por la facilidad de manipulación del paquete, y la preparación de la información para alimentar el proceso, pues las alternativas presentadas para realizar la captura, dan la facilidad de realizar con lujo de detalle un reporte, además de dar la facilidad de grabarlo en el disco duro.

Si es necesario mandar esta información a una terminal vía remota, es completamente posible, ya que lo único que debe hacerse es conectar las computadoras mediante un paquete de comunicación, de los que ya se han analizado con anterioridad, tomar el control de la máquina en donde se realizará la presentación, y desde la máquina en donde está la presentación, mandar la señal, con ello las personas que se encuentren observando la presentación no necesitarán operar la máquina y podrán observar las tendencias cada vez que se decida realizar un corte.

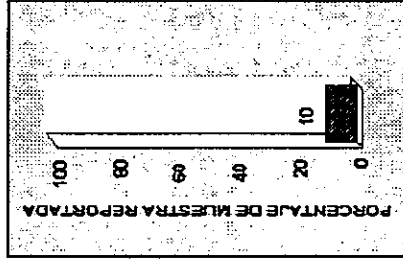
Gráfica 5.3.4

PRESIDENTE



HORA DE REPORTE:
9:47 pm

TOTAL DE MUESTRA



Resultados finales de las Elecciones Federales para Presidente de la Republica Mexicana en día 21 de Agosto de 1994

| Partidos | Votos | % Total de la Votación | Media |
|--------------|-------------------|------------------------|---------------|
| PAN | 9,146,841 | 26.72 | 144.23 |
| PRI | 17,181,651 | 50.19 | 270.92 |
| PPS | 166,594 | 0.49 | 2.63 |
| PRD | 5,852,134 | 17.09 | 92.28 |
| PFCRN | 297,901 | 0.87 | 4.70 |
| PARM | 192,795 | 0.56 | 3.04 |
| UNO | 97,935 | 0.29 | 1.54 |
| PT | 970,121 | 2.83 | 15.30 |
| PVEM | 327,313 | 0.96 | 5.16 |
| Total | 34,233,285 | 100.00 | 539.79 |

Figura 5.3.5.

El total de secciones es de 63,420.

Fuente: Estadísticas de las elecciones federales de 1994, Compendio de Resultados. IFE 1995

Otra alternativa es poner la gráfica en una página de internet, la cual puede ser consultada por las personas que lo deseen, opción que ofrece diferentes alternativas de consulta y que pueden llegar a ser de tipo internacional y sin límite de consultas simultáneas, es por demás decir que éste es un medio de comunicación electrónico masivo, por lo fácil que resulta realizar consultas en él.

En esta etapa el sistema ha llegado a su frontera¹, es decir, ésta es la frontera del SEIT, las gráficas que se presentan aquí son las salidas del sistema, con ello quedan cubiertos los objetivos, ésta es la presentación de las tendencias, el hecho de que sean inmediatas depende mucho de la agilidad con la que se recopilen los datos en campo, pues como ya se ha visto, una vez ingresados al sistema, el proceso y la presentación es prácticamente automática.

5.4. Comparación de los resultados obtenidos.

Una de las últimas partes de esta investigación es evaluar los resultados obtenidos, uno de los planteamientos básicos de la tesis fue la aplicación de un sistema capaz de medir tendencias de forma inmediata, y hasta ahora puede decirse que se han obtenido de esta forma, como ya se aclaró, lo que podría retrasar la presentación de los resultados es el trabajo en campo. Pero aun no se sabe si las tendencias son al menos cercanas a la realidad, afortunadamente los parámetros para calificar el SEIT en este caso particular existen, y se puede decir que para empezar, son los diferentes Conteos Rápidos realizados en la misma fecha, para las mismas elecciones.

Además la mejor forma de validar los resultados que arrojaron los diferentes sistemas ese día, es comparándolos con los resultados finales de dichas elecciones, los cuales se encuentran disponibles.

¹ Cuando se hace referencia al término "frontera", los autores lo hacen desde el enfoque de sistemas del que se habló en el capítulo 1.3.

Los datos obtenidos por algunos de los diferentes Conteos Rápidos ese día, se pueden observar en el cuadro 5.4.1, en donde también en la primera fila y sombreados se encuentran los resultados finales de dicho escrutinio, esta información se obtuvo directamente del IFE, en el caso de los resultados finales, de *"Estadística de las Elecciones de 1994, Compendio de Resultados"*, publicado en 1995, y en el caso de los resultados de los Conteos Rápidos del IFE, Alianza Cívica, CIRT-Televisa, TV-Azteca y COPARMEX, fueron tomados del periódico "La Jornada", del día Lunes 22 de Agosto de 1994, en primera plana, en donde además se menciona que el conteo rápido que patrocinó Alianza Cívica, se realizó en base a una muestra de 700 casillas, y en el periódico "Reforma" del día 25 de agosto de 1994, en la página 6-A, se publicaron los resultados de los conteos patrocinados por la Organización Nacional de Observación del Magisterio, Presencia Ciudadana, Partido Verde Ecologista de México, Partido Revolucionario Institucional, Partido Acción Nacional, Partido del Frente Cardenista de Reconstrucción Nacional, del mismo Periódico "Reforma" y del PREP, que dio a conocer el mismo IFE, entre otros.

Se cuenta con los resultados de 13 Conteos Rápidos, y los resultados finales del mismo proceso electoral para poder comparar la información proporcionado por el SEIT.

Como se puede ver en el cuadro 5.4.1, en la mayoría de los conteos, las estimaciones de las tendencias se presentan para las tres primeras fuerzas, debido a que son las de mayor interés, sólo en el caso de los conteos patrocinados por el PVEM, el PFCRN, la Organización Nacional de Observación Electoral del Magisterio, y el PREP presentan las tendencias estimadas para la cuarta fuerza, el PT.

A continuación se analizarán los resultados desde un punto de vista superficial, es decir, como cualquier ciudadano interesado los interpretaría, posteriormente se

CUADRO COMPARATIVO DE RESULTADOS, SEIT Y OTROS CONTEOS REALIZADOS EN EL MISMO PROCESO ELECTORAL.

ELECCIONES PARA PRESIDENTE DE LA REPUBLICA MEXICANA, EL 21 DE AGOSTO DE 1994.

| ORGANIZACION/MUESTRA | PAN | PRI | PRD | PT | PFCRN | PARM | UNO | PPS | PVEM |
|----------------------------------------------------------------|-------|-------|-------|------|-------|------|------|------|------|
| RESULTADOS DEFINITIVOS | 28.69 | 50.13 | 17.7 | 2.83 | 0.87 | 0.58 | 0.29 | 0.49 | 0.95 |
| SEIT (PRIMERA MUESTRA) | 31.46 | 46.46 | 15.49 | 3.17 | 0.9 | 0.38 | 0.24 | 0.76 | 1.13 |
| SEIT (SEGUNDA MUESTRA) | 30.69 | 46.88 | 16.06 | 3.36 | 0.77 | 0.37 | 0.37 | 0.46 | 1.05 |
| SEIT (MUESTRA TOTAL) | 31.10 | 46.70 | 15.80 | 3.30 | 0.80 | 0.38 | 0.30 | 0.58 | 1.10 |
| IFE-PREP | 26.81 | 50.08 | 17.03 | 2.85 | N/D | N/D | N/D | N/D | N/D |
| IFE-CONTEO RAPIDO | 31.35 | 47.14 | 15.49 | N/D | N/D | N/D | N/D | N/D | N/D |
| CURT-TELEVISIA | 27.00 | 50.00 | 16.00 | N/D | N/D | N/D | N/D | N/D | N/D |
| Organización Nacional de Observación Electoral del Magisterio. | 25.83 | 48.89 | 16.91 | 2.62 | N/D | N/D | N/D | N/D | N/D |
| ALIANZA CIVICA | 27.75 | 47.85 | 15.24 | N/D | N/D | N/D | N/D | N/D | N/D |
| TV AZTECA | 24.10 | 51.80 | 17.30 | N/D | N/D | N/D | N/D | N/D | N/D |
| COPARMEX | 27.90 | 49.10 | 13.70 | N/D | N/D | N/D | N/D | N/D | N/D |
| Presencia Ciudadana | 29.30 | 48.90 | 13.80 | N/D | N/D | N/D | N/D | N/D | N/D |
| Diario Reforma | 29.70 | 49.70 | 14.70 | N/D | N/D | N/D | N/D | N/D | N/D |
| PVEM | 31.59 | 46.88 | 15.11 | 3.25 | N/D | N/D | N/D | N/D | N/D |
| PRI | 25.00 | 50.00 | 15.00 | N/D | N/D | N/D | N/D | N/D | N/D |
| PAN | 34.00 | 43.00 | 12.00 | N/D | N/D | N/D | N/D | N/D | N/D |
| PFCRN | 22.70 | 43.70 | 22.10 | 2.70 | N/D | N/D | N/D | N/D | N/D |

Cuadro 5.4.1

analizarán con los elementos estadísticos adecuados para conocer qué tan buenos son los estimadores calculados.

Los resultados son de alguna forma sorprendentes. En primera instancia se observa que el PREP tiene la mejor aproximación, esto se puede explicar de la siguiente forma, el PREP como su nombre lo dice, es un *"Programa de Resultados Electorales Preliminares"*, no es exactamente una técnica de sondeo, consiste en la captura del total de las casillas del proceso, y se van dando resultados parciales, es decir *"hasta este momento, con los resultados de las casillas capturados, las tendencias son..."*. Además de capturar los resultados de las casillas estratégicamente dispersas, previamente seleccionadas, y de esta manera comenzar la captura total.

Las aproximaciones del PREP, son sorprendentes, para el caso de PAN, sólo existe una diferencia de 0.12% de error en la estimación, para el caso del PRI el error de estimación es aún menos, sólo el 0.05% de error, para el PRD, la diferencia es de 0.67% de error y para el PT el error sólo fue del 0.02%, se puede suponer que la base de datos tenía la información de más de 2500 casillas, pues con ese número de casos para la composición de una muestra, se puede alcanzar un error máximo del 2.00 con un 95% de confianza, según la fórmula que determina el tamaño de la muestra.

Observando el cuadro se ve que, en el caso del SEIT la estimación de las tendencias es aceptable para el caso de la primera muestra, la estimación para el PAN fue del 31.46%, mientras que el porcentaje final fue del 26.69%, es decir una diferencia de 4.77%, la tendencia estimada para el PRI fue del 46.46% siendo su resultado final un 50.13%, lo cual significa un error de estimación de 3.67%, el PRD obtuvo una estimación del 15.49% y su resultado final fue de 17.7%, lo cual significa una diferencia de 2.21%, y en el caso del PT al cual se le estimó un

3.17% del total de los votos, obtuvo en el conteo final el 2.83% es decir una diferencia de 0.34%.

Analizando la segunda muestra se observa que el PAN con una estimación de 30.69, tiene un error del 4.0%, el PRI con la tendencia obtenida del 46.88%, el error obtenido fue de 3.25%, con una estimación del 16.06%, el PRD sólo tiene un error de estimación del 1.64%, y el PT con una diferencia de estimación del 0.53%.

La tercera muestra se mantiene con la consistencia de las muestras anteriores, pues con la tendencia estimada para el PAN el error es de 4.41%, para el PRI es de 3.43, en el caso del PRD el error es de 1.9% y el PT tiene un error del 0.47%; en el caso de esta muestra, parece que son por demás obvios los resultados obtenidos, pues las dos primeras muestras mantienen estas tendencias, por lo tanto, no existe una variación significativa entre las dos primeras muestras y la tercera.

Para considerarse que las tendencias estimadas son buenas debido a que en ningún caso el error de estimación sobrepasa al 5%, que es aceptado por las ciencias sociales, y por otro lado cuando se empleó la fórmula para el cálculo del tamaño de la muestra, se aclaró que con el tamaño de la muestra se obtienen resultados con un nivel de confianza del 95% y con un margen de error del 5%.

De los demás resultados se ve que, en efecto, existen mejores aproximaciones que las del SEIT, sin contar al PREP, por lo antes mencionado, que son por ejemplo, los resultados obtenidos por el conteo rápido patrocinado por Organización Nacional de Observadores Electorales del Magisterio, debido a que el error máximo fue 1.24%, y fue en la tendencia para el PRI, después el patrocinado por CIRT-TELEVISIA, es uno de los que más se ajusta a los resultados reales, pues su error máximo es el

correspondiente al del PRD y es del 1.7%, de ahí le sigue el patrocinado por el PRI, su máximo error fue de 2.7%, para el caso del PRD, los demás, incluyendo al SEIT, con sus tres muestras, aparentemente están dentro del margen de error permitido.

Los conteos que más se alejaron de la realidad, fueron los patrocinados por el PFCRN, en donde el PAN tiene un error del 3.99%, lo cual no es tan grave como el error para el estimador del PRI, para el cual obtuvieron una tendencia del 43.70%, por lo tanto el error fue de 6.43%, y al PRD con una tendencia de 22.10% y un error en la estimación de 4.4%; y el patrocinado por el PAN, que se estimó un 34.0% del total de los votos con un error del 7.31%, al PRI le asignó un 43.0%, con un error del 7.13%, y al PRD un total de 12.0% con un error de 5.7%. En el caso de estos conteos se puede decir que las estimaciones obtenidas de alguna forma señalan las tendencias, es decir efectivamente colocan a las tres primeras fuerzas en el lugar que les corresponde, pero tienen errores en sus parámetros estimados mayores del 5%, lo cual inclusive puede ser válido, siempre y cuando se den a conocer, el nivel de confianza y el margen de error esperado.

Analizando estos resultados desde un punto de vista más estricto, empleando criterios matemáticos, se observa que existen ciertos detalles que saltan a la vista, si se calculan los estimadores puntualmente y literalmente se toma el margen de error de $\pm 5\%$ como ya se observó, las estimaciones son buenas y aparentemente están dentro de los márgenes permitidos. Sin embargo, si los parámetros fueran manejados como estimadores por intervalo, se puede observar que en la primera muestra, exceptuando al estimador del PAN, los demás están dentro de los intervalos de confianza estimados. Estos intervalos fueron estimados con los criterios proporcionados en el capítulo 2.4. que como se recordará están dados por:

$$\bar{y} \pm z_{\alpha} s_y$$

en donde \bar{y} es el estimador obtenido y z_{α} es el límite de la curva normal con un nivel de significación α , el cual como se mencionó con anterioridad, es de 0.05, es decir, con 5% de error, como se analizó anteriormente, y esto se deduce del teorema del límite central. El valor para z_{α} es 1.645, como se puede observar de la tabla 1 del apéndice de esta tesis, el valor $s_{\bar{y}}$, será tomado de la "tabla de estimadores puntuales insesgados", proporcionada por Mendenhall² en donde:

$$s_{\bar{y}} = \sqrt{\frac{pq}{n}}$$

p , es la proporción estimada y n es el tamaño de la muestra, $q = 1 - p$.

Con los criterios antes mencionados se construyó la tabla 5.4.2., en donde como ya se mencionó y se puede observar, el único intervalo que no incluye el resultado final es el caso del Partido Acción Nacional (PAN), siendo su diferencia de 0.0093, menos del 1% del Límite de Confianza Inferior (L.C.I.), y aunque en el caso del Partido Revolucionario Institucional (PRI), el resultado final está dentro del intervalo se puede observar su cercanía con el Límite de Confianza Superior (L.C.S.), fuera de estas dos observaciones, todos los demás intervalos incluyen al parámetro objetivo, (la proporción final de los votos para cada partido en la contienda).

Por otro lado, llama la atención que en el caso de la segunda muestra coincidentemente sucede lo mismo, en el caso del PAN, el parámetro objetivo se encuentra fuera de los intervalos de confianza estimados, la diferencia no es mucha, es sólo de 0.0017 del límite de confianza inferior, lo cual representaría sólo un 0.17%, para el caso del PRI, nuevamente se incluye el parámetro objetivo entre los límites de confianza, con la única observación que éste está muy cercano al

² MENDENHALL, William; SCHEAFFER, Richard L.; WACKERLY, Dennis D. "Estadística Matemática con Aplicaciones". Estados Unidos de América, Ed. Grupo Editorial Iberoamérica, 1986, pág.303.

límite de confianza superior, para los demás casos las estimaciones cumplen con las características esperadas, como se puede observar en la tabla 5.4.3.

Para el caso de la muestra total, se puede observar que para el caso del PAN, el parámetro objetivo nuevamente se encuentra fuera de los intervalos, y esta vez se aleja con 0.0031, esto es tan sólo 0.31%, este acortamiento en los intervalos se debe a que como se duplicó la muestra, el margen de error se disminuyó del 5% al 3.5%, con lo cual se espera más precisión en la estimación y por lo tanto los intervalos de confianza disminuyen como se observa en la tabla 5.4.4.

Como se nota en los cuadros anteriores, las estimaciones en su mayoría están dentro de sus correspondientes intervalos de confianza, exceptuando el caso del PAN, que en el caso de las tres muestras el parámetro objetivo estuvo por debajo del límite de confianza inferior por cantidades mínimas, y en la muestra total el caso del PRI, que estuvo por encima del límite de confianza superior por una cantidad prácticamente insignificante, sin embargo, llama la atención que con las primeras dos muestras que son completamente independientes la una de la otra, las tendencias se mantienen similares, y como consecuencia la muestra total también. Sin embargo, en los resultados finales, las estimaciones para las dos primeras fuerzas varían de esa forma, siendo éstos los únicos casos.

Cabe resaltar que en los casos del SEIT, el conteo patrocinado por el IFE y el PVEM, presentan similitudes en las tendencias obtenidas, una situación que cabe remarcar, es que el IFE tiene a su disposición la información deseada y con la prioridad que ellos marquen, de alguna forma ellos son dueños de la información, y por otro lado el conteo realizado por el PVEM, estuvo conformado por una muestra de casi 1600 casillas.

ANALISIS DE ESTIMACIONES CON RESPECTO A SUS LIMITES DE CONFIANZA MUESTRA 1

| Partido | Votos | Media Estimada | Tendencia Estimada | L.C.I. | L.C.S. | Resultados Finales |
|---------|---------|----------------|--------------------|---------|--------|--------------------|
| PAN | 54,769 | 136.9225 | 0.3146 | 0.2765 | 0.3528 | 0.2672 |
| PRI | 80,876 | 202.1900 | 0.4646 | 0.4236 | 0.5057 | 0.5019 |
| PPS | 1,320 | 3.3000 | 0.0076 | 0.0004 | 0.0147 | 0.0049 |
| PRD | 26,959 | 67.3975 | 0.1549 | 0.1251 | 0.1846 | 0.1709 |
| PFCRN | 1,567 | 3.9175 | 0.0090 | 0.0012 | 0.0168 | 0.0087 |
| PARM | 670 | 1.6750 | 0.0038 | -0.0012 | 0.0089 | 0.0056 |
| UNO | 424 | 1.0600 | 0.0024 | -0.0016 | 0.0065 | 0.0029 |
| PT | 5,512 | 13.7800 | 0.0317 | 0.0173 | 0.0461 | 0.0283 |
| PVEM | 1,969 | 4.9225 | 0.0113 | 0.0026 | 0.0200 | 0.0096 |
| Total | 174,066 | 435.1650 | 1.00 | | | 1.00 |

Figura 5.4.2.

ANALISIS DE ESTIMACIONES CON RESPECTO A SUS LIMITES DE CONFIANZA MUESTRA 2

| Partido | Votos | Media Estimada | Tendencia Estimada | L.C.I. | L.C.S. | Resultados Finales |
|---------|---------|----------------|--------------------|---------|--------|--------------------|
| PAN | 54,807 | 136.6758 | 0.3069 | 0.2689 | 0.3448 | 0.2672 |
| PRI | 83,721 | 208.7805 | 0.4688 | 0.4277 | 0.5098 | 0.5019 |
| PPS | 818 | 2.0399 | 0.0046 | -0.0010 | 0.0101 | 0.0049 |
| PRD | 28,682 | 71.5262 | 0.1606 | 0.1304 | 0.1908 | 0.1709 |
| PFCRN | 1,368 | 3.4115 | 0.0077 | 0.0005 | 0.0148 | 0.0087 |
| PARM | 680 | 1.6459 | 0.0037 | -0.0013 | 0.0087 | 0.0056 |
| UNO | 656 | 1.6359 | 0.0037 | -0.0013 | 0.0086 | 0.0029 |
| PT | 6,009 | 14.9850 | 0.0336 | 0.0188 | 0.0485 | 0.0283 |
| PVEM | 1,876 | 4.6783 | 0.0105 | 0.0021 | 0.0189 | 0.0096 |
| Total | 178,597 | 445 | 1.00 | | | 1.00 |

Figura 5.4.3.

ANALISIS DE ESTIMACIONES CON RESPECTO A SUS LIMITES DE CONFIANZA MUESTRA TOTAL

| Partido | Votos | Media Estimada | Tendencia | L.C.I. | L.C.S. | Resultados Finales |
|---------|---------|----------------|-----------|---------|--------|--------------------|
| PAN | 109,576 | 136.7990 | 0.3107 | 0.2809 | 0.3405 | 0.2672 |
| PRI | 164,597 | 205.4894 | 0.4667 | 0.4346 | 0.4988 | 0.5019 |
| PPS | 2,138 | 2.6692 | 0.0061 | 0.0011 | 0.0111 | 0.0049 |
| PRD | 55,641 | 69.4644 | 0.1578 | 0.1343 | 0.1812 | 0.1709 |
| PFCRN | 2,935 | 3.6642 | 0.0083 | 0.0025 | 0.0142 | 0.0087 |
| PARM | 1,330 | 1.6604 | 0.0038 | -0.0002 | 0.0077 | 0.0056 |
| UNO | 1,080 | 1.3483 | 0.0031 | -0.0005 | 0.0066 | 0.0029 |
| PT | 11,521 | 14.3833 | 0.0327 | 0.0212 | 0.0441 | 0.0283 |
| PVEM | 3,845 | 4.8002 | 0.0109 | 0.0042 | 0.0176 | 0.0096 |
| Total | 352,663 | 440 | 1.00 | | | 1.00 |

Figura 5.4.4.

Afortunadamente en este momento ya se conocen los resultados y se puede realizar este análisis, sin embargo, en el día de las elecciones, unas horas después, normalmente no se tiene tal información, motivo por el cual la primera publicación de tendencias se convierte en una verdadera angustia, pues ¿es posible imaginar el tipo de problema que puede surgir si es que se llegan a publicar tendencias que estén alejadas de la realidad?, y por otro lado, el papel no es tan decoroso si uno es de los últimos en presentar las tendencias obtenidas, pues aunque se ajusten al común de las publicadas, ya se perdió la oportunidad de ser los primeros.

5.5. Perspectivas y posibilidades de desarrollo del SEIT.

En la época actual, México está empezando a cambiar, día con día se moderniza, se empieza a tener una apertura, a diferentes grupo políticos con diferentes ideologías, la competencia electoral es cada día más compleja. En las pasadas elecciones en el país, se gastó poco más de 480 mil millones de pesos, éste es el precio económico de la democracia, precio que pagan todos los mexicanos, y aunque aparentemente sea muy alto, no es nada comparado con los beneficios que trae consigo, por ejemplo, la estabilidad social, la económica, etc. Movimientos como el de Chiapas, Guerrero y otros estados han sido provocados por la falta de credibilidad en sus gobernantes y en los procesos para elegirlos. Se piensa que son impuestos por las grandes cúpulas en el poder, sin embargo, se comienza a vivir la época de las comunicaciones, hoy día no puede ocurrir algo en cualquier lugar del mundo sin que todos se enteren.

Las encuestas de opinión pública, cuyos resultados se dan a conocer constantemente, son un medio para hacer participar a la población en muchos campos, uno de los más explorados son la mercadotecnia, la política, etc. Hoy día se pueden dar a conocer las preferencias sobre tal o cual producto, y dicha información puede llegar a todos los mexicanos, por los diferentes medios de comunicación, llámese radio, televisión, prensa, internet, etc.

Un sistema como el SEIT, tiene la función de hacer por ejemplo, que esos 480 mil millones de pesos, no sean mal invertidos, esta clase de proyectos dan certeza y credibilidad a los procesos electorales, y representan un avance en el uso de las técnicas estadísticas y su aplicación en la vida cotidiana del país.

El SEIT, como se presenta aquí, puede ser un proyecto, que bien organizado, no emplea muchos recursos, y no tiene que ser muy costoso, sin embargo, ya representa una herramienta muy poderosa en los procesos electorales, no sólo en México, si no también en E.U.A., Europa y Asia, y se emplea solamente para eso, para estimar tendencias, en el momento preciso en que se terminan las elecciones, la forma de trabajar es la misma, en todo el mundo.

Un aspecto que debe cuidarse mucho, en este tipo de técnicas, es emplear gente muy capacitada para construir una metodología teóricamente válida y aplicarla al sistema, últimamente se ha hablado mucho de los *"errores de muestreo"*, lo cual es realmente alarmante, pues no deja de ser cierto que muchas instituciones políticas, o de otro tipo, hacen mal uso de estas técnicas, desvirtuándolas. Son muy conocidos los casos en que se publican encuestas en donde los resultados presentados tienen la única intención hacer que la gente crea alguna situación que no es cierta, subestimando el criterio de la gente que recibe esa información, es evidente que si se incurre en este error, la credibilidad de estos proyectos y técnicas se irá perdiendo.

Por otro lado, una de las posibilidades de desarrollo del SEIT, se observa en el PREP, se puede considerar que éste es un SEIT, con la salvedad que no se tiene un número límite de casillas para la composición de la muestra, lo cual, como se vio anteriormente, resulta por demás benéfico, siempre y cuando cumpla con cierta metodología pues los resultados hablan por sí solos; como ya se vio, entre los conteos rápidos y el SEIT, el PREP fue el más cercano a la realidad, aunque cabe señalar que para conseguirlo, el IFE emplea una cantidad de recursos que difícilmente se podrán igualar.

El SEIT con las bondades que prestan las comunicaciones actualmente, brinda como resultado estimaciones aceptables de las tendencias, no es difícil pensar que evolucionando más las comunicaciones, dentro de pocos años, las tendencias queden atrás, y lo que se presente sean los resultados finales de las elecciones.

A lo largo de este capítulo se analizó una situación real, y se pusieron a prueba las partes del SEIT que así lo permitieron, en donde se pudieron evaluar los resultados obtenidos mediante el sistema. Primeramente se recopilaron los resultados de las casillas que pertenecían a la muestra, y se describió el proceso mediante el cual se obtuvo esta información, y se diseñaron unos cuadros resumen para poder observar de dónde era tomada la información, señalando el estado y el municipio e identificando las secciones por su tipo, ya sea rural o urbana. Posteriormente se describió el procesamiento de la información para el cual se dieron un par de opciones sencillas y eficientes, para después analizar los resultados, los cuales se presentaron en forma de porcentajes, se analizaron tres muestras, dos de ellas tomadas independientemente una de la otra, la tercera se conforma de la unión de las dos primeras. Los resultados se presentaron en términos de la media estimada y la media estratificada, las cuales no tuvieron una gran variación y se pudo constatar la consistencia de las tres muestras estudiadas. A continuación se

mostraron los resultados en una gráfica como se podrían presentar si se tratara de un caso real. También se compararon los resultados obtenidos con los resultados reales, en donde se hicieron dos análisis: el primero fue desde el punto de vista de cualquier lector, sin mayores conocimientos en esta clase de técnicas de sondeo. El segundo, empleando los límites de confianza previamente establecidos, en donde resulta interesante la comparación de dichos análisis. Finalmente se termina este capítulo con algunas reflexiones respecto a las perspectivas y posibilidades de desarrollo del SEIT.

CONCLUSIONES Y RECOMENDACIONES

Es posible ubicar al SEIT como una de las técnicas de inferencia estadística altamente difundidas actualmente, al nivel de las encuestas de opinión pública, las encuestas Exit-Pool, etc., siendo más específica su clasificación como Conteo Rápido, el cual en esta época se convierte en una herramienta frecuentemente empleada. Actualmente esta clase de métodos estadísticos se han convertido en una forma de hacer política en México, hoy día es imposible no saber que en las vísperas de algún proceso electoral se hayan publicado encuestas que muestren las preferencias de los electores, y ayuden a plantear estrategias de campaña. Ahora casi todos los contendientes a algún puesto de elección popular incluyen en su equipo de trabajo a sus propios encuestadores, para lo cual inclusive, se llega a destinar un presupuesto bien definido, pues la experiencia indica que el dejar de hacerlo puede traer consigo consecuencias desastrosas.

El tema que se abordó en esta tesis fue el sobre diseño y planeación de un Sistema de Estimación Inmediata de Tendencias, para lo cual fue necesario conocer de forma completa el problema al cual se quería dar solución y plantear alguna serie de posibles alternativas, por lo que se abordaron diferentes temas, a lo cual se puede concluir:

- 1) **El SEIT y la Teoría General de Sistemas:** como su nombre lo dice la forma de trabajar es como un sistema, lo cual resulta completamente apropiado debido a la complejidad de este tipo de proyectos, como ya se analizó, el esfuerzo es muy grande y únicamente fraccionándolo se puede estar seguro de su buen funcionamiento, la teoría general de sistemas da las pautas para poder dividir las tareas en diferentes subsistemas y así tener mayor control sobre los

resultados de cada uno de los subsistemas, para que en el momento de conjuntar los resultados parciales sea posible presentar las salidas del sistema de la forma requerida, y con el orden esperado.

Por otro lado, esta teoría brinda los elementos para poder realizar el análisis de los niveles de sistemas, es decir, cómo se ubica el sistema y cómo interactúa con el medio ambiente para poder definir cuáles son sus límites, y de esta forma poder llevar a cabo la identificación de los sistemas, subsistemas y suprasistemas que tienen alguna relación con el problema.

Es necesario hacer la observación de que el SEIT es un sistema que está compuesto por varios subsistemas, de los cuales algunos son conceptuales, otros computacionales, otros humanos, etc., no se trata de un sistema meramente mecánico. Es un sistema dinámico con muchas diferencias entre sus elementos.

El enfoque de sistemas, también puede contribuir con el SEIT para considerar la relación del problema con las condiciones del medio y para identificar los factores y variables que afectan el problema. Además de proporcionar un marco de trabajo útil para evaluar el desempeño de varios de sus subsistemas y del sistema global; también puede ser tomado como un factor de actualización del mismo sistema. Debido a su metodología se puede usar para rediseñar, comparar y poner a prueba los sistemas.

- 2) **Técnicas de Muestreo:** Las técnicas analizadas en este trabajo, fueron sólo algunas de las disponibles, la teoría de muestreo presenta una amplia diversidad de opciones donde es posible adoptar la más adecuada al problema que se está abordando, lo que es un aspecto positivo de esta teoría, esto

siempre y cuando se cuente con la experiencia suficiente para poder evaluar la técnica adecuada, de forma contraria, se convierte en un arma de doble filo, debido a que puede suceder que se incurra en graves errores de muestreo que lejos de ofrecer algún beneficio complique y/o desvíe la solución del problema.

Esta teoría presenta una serie de bondades para poder realizar este tipo de proyectos, debido a que como ya se mencionó, brinda la facilidad de poder conocer las características de una población con sólo una pequeña fracción de ésta, lo cual trae como consecuencia que dichos proyectos sean baratos, realizables, que pueda llevarse a cabo una supervisión minuciosa del trabajo en campo y oficina, motivo por el cual en algunas ocasiones puede asegurarse que estos trabajos pueden ser inclusive más certeros que un censo.

Así pues, no es posible señalar a una técnica de muestreo como la indicada para diseñar las muestras del SEIT, debido a que como es evidente, la técnica empleada será determinada por el estudio del problema al que se pretende dar solución, estudiando la población objetivo y determinando las características propias de la situación que se está abordando.

Sin embargo, es necesario hacer hincapié en la importancia de la muestra, debido a que por las características del sistema no es posible corregir algún error de diseño de la muestra sobre la marcha. Se mencionó que esta muestra es la parte ingeniosa del sistema, pero también es la parte más delicada, un error en su diseño y el sistema estaría destinado al fracaso.

- 3) **La determinación del tamaño de la Muestra:** se convierte en un punto importante, del cual dependen muchos aspectos de planeación y cargas de

trabajo, debido a que determina el número de personas que se requerirán para el trabajo en campo, el número de computadoras para realizar la captura, etc.; el tamaño de la muestra se calcula dependiendo de la técnica de muestreo que se pretenda emplear, cada una de estas técnicas tiene sus modelos propios para estimar este dato.

La estimación del tamaño de la muestra, de igual manera que la técnica de muestreo, debe ser responsabilidad de una persona con los conocimientos y el perfil adecuado, pues un mal cálculo de éste, puede acarrear complicaciones severas al sistema. De igual manera que en la selección de la técnica de muestreo, para el caso del SEIT, no es posible señalar un modelo único para calcular el tamaño de la muestra, éste se dará como resultado del análisis del problema y de la determinación de la técnica de muestreo adecuada.

Cabe mencionar que en el caso de medir diferentes características, el tamaño de la muestra debe calcularse para cada una de ellas, una vez realizado esto, se analizan los resultados y de preferencia se toma el mayor número, si es que los recursos disponibles son suficientes, de no ser así, es posible tomar algún número intermedio que vaya de acuerdo con el presupuesto disponible, siempre y cuando se haga la aclaración.

El tamaño de la muestra es un aspecto de suma importancia, debido a que si se desea calcular las tendencias de forma inmediata resulta ilógico trabajar de más, pues si por alguna razón el tamaño de la muestra es mayor al requerido, el costo del proyecto, los recursos y el tiempo para poder reportar los resultados finales serán mayores. Y por el lado contrario, en el caso de estimar el tamaño de la muestra menor al que se requiere, se corre el riesgo de que la muestra no sea representativa.

- 4) **Los Resultados del Monitoreo Electrónico:** fueron de gran ayuda para poder evaluar las diferentes opciones en la forma de trabajar el sistema, éste se llevó a cabo con diferentes publicaciones de carácter periódico, entre las más consultadas se pueden mencionar la revista *PC Semanal* de publicación semanal, la sección Interfase que se publica todos los lunes en el periódico *Reforma*, la revista *PC Magazine*, entre otras; de donde se obtuvieron muy buenos reportajes sobre equipo de cómputo y software, de los cuales se extrajeron una serie de ideas para poder manejar parte del SEIT.

Cabe hacer mención que el llevar a cabo este monitoreo es una de las mejores formas de mantenerse actualizado, pues si bien, era difícil poder encontrar algún artículo o producto de utilidad al sistema, existían un sin fin de elementos útiles para poder integrar otro tipo de funciones y una serie de novedades de tipo informático de gran ayuda para cualquier persona que esté relacionada con el medio.

Por poner un ejemplo de los resultados de este monitoreo, uno de los grandes descubrimientos fue encontrar la forma de trabajar del *Programa de Resultados Preliminares* (PREP), del Instituto Federal Electoral (IFE), nota que fue publicada el lunes 7 de julio de 1997 en la sección Interfase del periódico *Reforma*, el cual muestra de forma global la metodología empleada, lo que a su vez sirvió como base para el sexto planteamiento de trabajo del sistema propuesto en el capítulo cuatro.

- 5) **La estructura computacional requerida para el SEIT:** es uno de los aspectos de mayor importancia en el sistema, debido a que es aquí donde se puede agilizar

el manejo de la información, la estructura tiene que soportar el proceso de captura, además debe estar preparada para poder atender la demanda necesaria y así evitar que se convierta en el cuello de botella del sistema.

Es necesario que esta estructura tenga la capacidad para poder capturar la información procedente de los diferentes puntos de trabajo en campo, de la forma más ágil posible; no resulta práctico para el SEIT retrasar el proceso de captura por falta de terminales, por lo que se propuso un modelo práctico, para el cálculo del número de terminales de captura. También debe de contar con las terminales necesarias para poder llevar a acabo el procesamiento de la información, los cálculos necesarios, la presentación de los resultados, y la comunicación con alguna terminal vía remota en caso de ser necesario.

Se propusieron diferentes estructuras capaces de soportar el trabajo, cualquiera de ellas cumple con los requisitos en diferentes medidas. De las estructuras antes mencionadas se puede concluir que si es posible contar con una RED de tipo LAN, con la cual no exista ningún tipo de problema de instalación, con una topología de estrella y que esté completamente estabilizada, ésta sería la mejor forma de trabajar el SEIT, debido a que resulta una opción barata, confiable y rápida. Se debe destacar que los requisitos antes mencionados son muy importantes pues una falla en la RED, que pudiera causar la caída del sistema, podría tener consecuencias completamente negativas.

Por otro lado, se presentó una alternativa más accesible, la estructura conectada con un programa de comunicación, en este caso el INTERLINK, la cual no debe despreciarse pues representa una opción válida para llevar a cabo la captura de cualquier número de información y con la capacidad de procesarla y presentar los resultados obtenidos de forma eficiente; es aún más

barata, bastante estable y puede inclusive ser más segura que la red LAN, el único problema que puede presentar es la rapidez con la que se reúne y procesa la información, que si en un momento determinado su volumen no es tan extenso, esta estructura presenta una excelente alternativa para el manejo de la información y demuestra que no se requiere del mejor equipo para poder realizar un excelente trabajo.

Cabe mencionar que el diseño e instalación de la estructura computacional es una tarea que consume muchos recursos, y es una de las más complicadas, por lo que se recomienda poner especial atención en esta labor, todo tiempo de anticipación del que se disponga es necesario para poder realizar toda clase de pruebas con el equipo.

Cada una de las estructuras presentadas, tiene sus pro y sus contras, la utilización de una u otra depende de las necesidades y los recursos disponibles. Cabe aclarar que estas estructuras no son las únicas y que además pueden ser modificadas dependiendo del ingenio y los recursos con los que cuente la persona encargada de el diseño de la misma.

- 6) **La Ingeniería de Software:** aportó una serie de criterios muy valiosos para el desarrollo de los programas empleados en esta tesis, además que esta materia es la que de alguna forma asegura que el ciclo de vida del software se pueda extender a medida que sea necesario mediante el mantenimiento a la aplicación desarrollada.

Si bien es cierto que la Ingeniería de Software se aplica al desarrollo de programas grandes, también brinda la posibilidad de trabajar en equipo, desarrollando fracciones del software en módulos, además de proporcionar

una metodología bastante eficiente para la planeación de proyectos de software.

En el caso del SEIT, se desarrollaron una serie de programas en diferentes lenguajes y también una serie de aplicaciones en diferentes paquetes, éstos fueron creados de forma independiente para que posteriormente, se pusieran a trabajar de forma conjunta, lo cual brinda la posibilidad de actualizar algunas partes de los programas desarrollados sin necesidad de afectar a otros, de este modo por ejemplo, es posible esperar adelantos en materia de graficadores, y actualizar las presentaciones sin necesidad de trabajar toda la programación desarrollada.

La Ingeniería de Software podría prestar una ayuda inmensamente mayor al SEIT, si en un momento dado alguien decidiera desarrollar una aplicación integral para este tipo de proyectos y la comercializara.

- 7) **El Software empleado en el SEIT:** es muy variado, va desde el uso de sistemas operativos, bases de datos, programas para realizar los procesos y los cálculos necesarios, programas graficadores, programas de comunicación, etc. Como se mencionó con anterioridad, es necesario realizar la aplicación de forma modular, lo cual bien planteado facilita el problema, debido a que si se requiere una aplicación específica se puede emplear el software que brinde mayores facilidades para la tarea determinada.

A raíz del monitoreo realizado, se puede constatar que la diversidad de opciones de software es inmensa y que en la mayoría de los casos, las facilidades otorgadas por las diferentes opciones son similares, de aquí que se puede concluir que el mejor software para el diseño de las rutinas necesarias

para el SEIT, es aquel con el que está familiarizado el programador encargado de dar solución a los problemas planteados, realmente es innecesario esforzarse por programar ciertas rutinas en tal o cual paquete o lenguaje, si el manejo de éste no es dominado por el programador encargado, pues bajo este esquema es probable que se cometan errores de programación o que el desarrollo de la aplicación sea más tardado, además de no ser explotado de la forma debida.

Para efectos de esta tesis, se planteó el uso del sistema operativo MS-DOS, los programas de captura pueden desarrollarse en Data Ease, para juntar la información en caso de ser necesario, y el procesarla con una rutina elaborada en Clipper es suficiente. Para realizar la presentación se sugirió Harvard Graphics para Windows, y obviamente Windows, éstas son las herramientas necesarias a nivel de software, es evidente que existen mejores sistemas operativos, bases de datos, etc., pero para efectos de trabajar el SEIT con lo antes mencionado es suficiente.

En la actualidad, existen paquetes que sirven para realizar diferentes tareas, como la administración de una base de datos, la graficación, etc., sin embargo es difícil encontrar algún tipo de software que cuente con todas las facilidades y elementos contenidos en un paquete especializado, que brinde la posibilidad de manejar bases de datos y graficar, tal vez en el futuro se cuente con algún tipo de software que pueda realizar diferentes tareas con bastante buena calidad y por ende, el encargado de dirigir el desarrollo de la parte computacional puede decidirse por éste, aunque siempre se debe de tener en cuenta los recursos en tiempo y equipo que se requieran para poder operar dicho paquete.

- 8) **El flujo de la información:** es uno de los factores en los que hay que poner mayor atención en el SEIT, debido a que de éste dependen las estrategias que se emplearán en el trabajo en campo, y la cantidad de equipo de cómputo que se requerirá para poder operar.

Respecto a este asunto se propusieron 6 diferentes planteamientos para poder cubrir satisfactoriamente esta tarea, están diseñados para poder operar en proyectos de diferentes magnitudes, sus costos son muy variables, es decir, con cualquier tipo de presupuesto se puede diseñar un sistema como el SEIT, en donde es evidente que a medida de que el presupuesto sea menor, su forma de trabajar será menos eficiente, aunque siempre posible.

De lo anterior se puede concluir que el flujo de la información es un aspecto que estará completamente ligado al presupuesto, y que además no es necesario disponer de un gran capital para poder diseñar alguna metodología válida para sistemas como el SEIT.

Es importante mencionar que de la agilidad con que fluya la información, dependerá la velocidad en la que se presentarán los resultados, el flujo de la información comienza desde el momento en que se recaban los datos afuera de las casillas, hasta que sus resultados se presenten en una gráfica que estará indicando las tendencias estimadas, en base a esto y con lo antes analizado se puede concluir que, si se cuenta con una estructura computacional eficiente, el tamaño de la muestra está bien estimado, los procesos trabajan de forma correcta y llega a existir algún retraso en la presentación de la información, entonces el error puede estar en el flujo de la información en su primera etapa, que es el trabajo en campo, la recolección de los datos. Esta es la parte vulnerable del sistema, debido a la cantidad de

percances que pueden ocurrir en el operativo de campo, por poner un ejemplo: si en alguna sección contenida en la muestra no existen medios de comunicación como teléfonos, o se está en alguna comunidad rural que no cuenta con caminos o transporte, etc., seguramente el tiempo de comunicación para poder hacer llegar los datos al centro de acopio de información será mayor.

- 9) **Las medidas de seguridad:** se recomienda muy ampliamente poner especial atención en este aspecto, ya que la seguridad forma una de las partes de mayor importancia en el sistema, debido a la gran cantidad de contratiempos que pueden surgir, los cuales van desde la posibilidad de sufrir algún daño en el equipo de cómputo, falta de corriente eléctrica, la introducción de algún virus, etc., hasta la posibilidad de sufrir algún sabotaje, o ser víctimas de algún tipo de espionaje.

Quizá suene muy exagerado lo antes mencionado, pero en la actualidad las empresas que se dedican a realizar este tipo de estudios, dan mucha importancia a las medidas de seguridad, invirtiendo cantidades enormes en este aspecto. No es muy difícil pensar que algún equipo se dañe, o que exista un corte de energía eléctrica, en el momento crítico de operación del sistema, o que por la naturaleza del estudio existan grupos o sectores interesados en arruinar el buen término de éste.

- 10) **El modelo funcional:** es una herramienta que conviene plantear siempre que se trabaja en equipo, y con diferentes áreas de trabajo. Se recomienda diseñar un modelo funcional al inicio del proyecto para poder delimitar responsabilidades,

además de ayudar en la definición de subsistemas en el caso del SEIT, convirtiéndose así en un factor de orden dentro de éste.

Abordando la aplicación práctica del SEIT planteada al inicio de esta tesis, se puede comentar que fue un trabajo bastante laborioso, debido a que por la naturaleza de la información no era fácil el acceso a ella, a pesar de que se supone que los datos están a la disponibilidad de quien quiera consultarlos, los resultados arrojados en esta aplicación resultaron muy interesantes como se verá a continuación:

- a) **La metodología para el diseño de la muestra:** debido a que a la población del país se le puede considerar heterogénea en sus preferencias políticas, desde el inicio se planteó que el diseño de la muestra fuera con la ayuda del muestreo aleatorio estratificado. Se consideró a las 32 entidades federativas como los diferentes estratos, en donde el factor w_i , fue determinado por la relación que guardaba cada estado con el padrón electoral, y no por el número de secciones de cada estado, debido a la diferencia en proporciones observada. Procediendo después a seleccionar aleatoriamente el número de secciones asignado por estado.

El tamaño de la muestra fue determinado por el modelo que proporciona el muestreo aleatorio, debido a que el modelo propuesto para el muestreo aleatorio estratificado está en función de la desviación estándar de cada estrato, y no era posible contar con ese dato, sin embargo, partiendo de un razonamiento lógico, no puede funcionar el SEIT como se pretende si el tamaño de la muestra es muy grande, y la aproximación dada por la alternativa antes mencionada fue de 400 secciones en muestra, esto sin

tomar en cuenta la *cpf*, debido a que la población objetivo superaba las 63 mil unidades y el tamaño de la muestra está muy lejos de ser siquiera el 1%.

- b) **El proceso de la información:** primeramente se recolectó la información directamente de la biblioteca del IFE, posteriormente se capturó en un programa diseñado en Data Ease, del cual se ejecutó una rutina que dejó el archivo en código *ascii*, para posteriormente con una rutina programada en Clipper, se realizarán los cálculos necesarios, dejando los resultados en otro archivo igualmente en código *ascii*. El proceso anterior se realizó con un par de muestras diseñadas de la misma manera con la finalidad, de poder formar una tercera muestra con la unión de ambas, sin incurrir en algún error de tipo metodológico.

- c) **La presentación de los resultados:** se llevó a cabo de forma satisfactoria por medio de una rutina en Harvard Graphics, en la versión 2.0 para windows, ésta fue mediante una gráfica de barras, en la que se mostraban las tendencias obtenidas, la hora del último reporte y el porcentaje de la muestra cubierta, los datos para la alimentación de la gráfica de barras y el porcentaje de la muestra cubierta, fueron tomados del archivo que dejó en código *ascii* el programa en Clipper.

Se le dió el formato de porcentajes debido a que es un término que la mayoría de la gente domina, y el cual es muy fácil de interpretar.

Se presentaron los resultados del análisis de tres muestras, de las cuales las primeras dos fueron seleccionadas de forma muy independiente, y la tercera es el resultado de la unión de las dos primeras, cabe señalar que las tendencias estimadas fueron muy cercanas entre sí, y que la media

estratificada, no presenta diferencias substanciales con respecto a la media estimada.

En este caso se puede concluir que, si se toma como parámetro el *Error cuadrático Medio* para calificar las tendencias estimadas, entonces la mejor estimación la proporciona la segunda muestra. Aunque siguiendo este mismo criterio, se preferirá la tercera muestra debido a que la probabilidad de que el estimador esté en el intervalo señalado es mayor.

- d) **Llevando a cabo una comparación de los resultados**, con los resultados de otros sistemas y con resultados definitivos, se puede observar que la metodología empleada para el SEIT es bastante aceptable, y la prueba son los resultados que se obtuvieron, como se puede ver, las estimaciones para los nueve partidos políticos están dentro del margen permitido, que como se mencionó es del 5%.

Además, llama la atención la forma en que convergen con el SEIT, los resultados de los conteos patrocinados por el IFE, Alianza Cívica y PVEM, que como se comentó en su momento, las diferencias principales entre esos conteos rápidos y el SEIT, es el tamaño de muestra, ya que el IFE es prácticamente el dueño de la información, en el caso de Alianza Cívica, el tamaño de muestra fue mayor de 700 casillas y el PVEM tuvo una muestra de aproximadamente 1,600 casillas.

Comparando los resultados del SEIT, con los resultados finales del proceso, se puede observar que éstos cumplen con las características especificadas, ya que las tendencias estimadas son lo suficientemente aproximadas a los resultados finales.

En el análisis puntual que se realizó, se vio que ante los ojos del común de las personas, los resultados están completamente dentro de lo admisible, sin embargo, en el análisis de los estimadores por intervalos, se pudo observar que en el caso particular del PAN, el estimador objetivo no estuvo dentro de los intervalos, y el caso del PRI en la muestra total. Aunque en estos casos la cantidad con la que se desviaron fue mínima, éstos no estaban contenidos en dichos intervalos, ahora que, si bien es cierto que en el caso ideal se desea que el intervalo sea estrecho y contenga al parámetro objetivo, no se debe olvidar que también está en juego una probabilidad asociada, la longitud de dichos intervalos está dada en función de las mediciones de la muestra, pues la longitud del intervalo y la localización son cantidades aleatorias, motivo por el cual no es posible tener la completa seguridad de que el parámetro objetivo siempre esté contenido entre dichos intervalos, tratándose de una o dos muestras.

Por lo anterior, se puede concluir que el SEIT cuenta con una metodología a la altura de las necesidades. Además también se puede concluir que, el diseño de la muestra fue el adecuado y su tamaño fue suficiente; lo cual queda confirmado con el análisis de la tercera muestra, pues como se realizó duplicando el número de elementos en ésta, no existieron cambios importantes en el cálculo de las tendencias.

Es válido hacer la observación que en los resultados presentados por otros sistemas, específicamente en el caso de los conteos rápidos patrocinados por el PFCRN y el PAN, no se cumplieron los objetivos, porque aunque definen al ganador de los comicios y colocan en su posición a las cuatro primeras fuerzas políticas del país, el objetivo es estimar las tendencias, es importante

obtener porcentajes y éstos deben estar dentro del margen de error permitido o especificado por la metodología empleada.

Con lo antes mencionado, también es posible concluir, que para la realización de un proyecto de esta índole, es posible hacer planteamientos económicos y eficientes, para lo cual solo se necesita llevar a cabo una planeación adecuada.

Por otro lado se puede considerar válida la siguiente aclaración, los Conteos Rápidos, están ostentando un nombre que no se apega a sus funciones, estos sistemas se sustentan en metodologías como las del SEIT, y lo que realmente hacen es estimar tendencias, y no contar rápidamente el total de los votos del escrutinio, aunque en la actualidad sí existen estos sistemas, lo más parecido a ellos en México, es el instrumentado por el IFE y es el PREP. Este planteamiento parte del hecho de comenzar a contar los votos en el momento en que terminan las elecciones y a su vez va mostrando las tendencias hasta el momento. Este sistema debe operar con mucha cautela, ya que si sólo se toman los resultados de una región, las tendencias pueden ser engañosas.

Un aspecto que vale la pena mencionar son las virtudes multidisciplinarias del SEIT, como ya se comentó, son varias las disciplinas que contribuyen a su conformación y una de las disciplinas más importantes en este tipo de sistemas, son las ciencias sociales, pues son éstas las que dan un uso práctico al esfuerzo realizado dentro del sistema, son las que contextualizan y capitalizan los resultados obtenidos.

Otro punto en el que se debe poner especial atención es el diseño de la muestra, pues como se ha venido hablando últimamente, los errores de muestreo son el principal factor de desprestigio de esta clase de técnicas.

También es de vital importancia para todos los profesionales relacionados con estos proyectos, el evitar el mal uso de estas técnicas, pues se han dado casos en los que se manejan los resultados con la intención de manipular a la opinión pública en lugar de servir como un foro de expresión ciudadana, menospreciando la capacidad crítica de las personas, lo cual traerá como única consecuencia la falta de credibilidad en esta clase de procesos.

Con respecto a lo que se espera del SEIT en el futuro, con los avances técnicos que surgen día a día, seguramente el flujo de la información irá creciendo en cantidad, calidad y rapidez, además de brindarle una mayor cobertura. Se contará con mejores recursos, ya que el software crece diariamente y seguramente surgirán paquetes que ayuden a la realización de un proyecto de esta índole para ser más eficientes a un menor costo. En medida que sistemas como el SEIT vayan ganando popularidad, se pensará en crear software especializado para esta clase de proyectos.

APENDICE

Valores esperados y varianzas para algunos estimadores puntuales comunes

| Parámetro- objetivo q | Tamaño de la(s) muestra(s) n | Estimador puntual $\hat{\theta}$ | $E(\hat{\theta})$ | σ_{θ}^2 |
|-------------------------------|-----------------------------------------|----------------------------------------|-------------------|---------------------------------------------------|
| m | n | \bar{Y} | m | $\frac{\sigma^2}{n}$ |
| p | n | $\hat{p} = \frac{Y}{n}$ | p | $\frac{pq}{n}$ |
| $m_1 - m_2$ | n_1 y n_2 | $\bar{Y}_1 - \bar{Y}_2$ | $m_1 - m_2$ | $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ |
| $p_1 - p_2$ | n_1 y n_2 | $\hat{p}_1 - \hat{p}_2$ | $p_1 - p_2$ | $\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$ |

* σ_1^2 y σ_2^2 son las varianzas de las poblaciones 1 y 2, respectivamente.

ANEXO 1 RESUMEN DE LA PRIMERA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Secc. | Proporcion Mta. Edo |
|-------------------------|-------------------------------|--------|-----------------|------------------------|
| 1 AGUASCALIENTES | 1 AGUASCALIENTES | Urbana | 3 | 0.75 |
| | 3 CALVILLO | Rural | 1 | 0.25 |
| | TOTAL | | 4 | 1.00 |
| 2 BAJA CALIFORNIA NORTE | 2 MEXICALI | Urbana | 5 | 0.63 |
| | 4 TIJUANA | Urbana | 3 | 0.38 |
| | TOTAL | | 8 | 1.00 |
| 3 BAJA CALIFORNIA SUR | 3 PAZ, LA | Urbana | 2 | 1.00 |
| | TOTAL | | 2 | 1.00 |
| 4 CAMPECHE | 1 CAMPECHE | Urbana | 2 | 0.67 |
| | 2 ASIENOS | Rural | 1 | 0.33 |
| | TOTAL | | 3 | 1.00 |
| 5 COAHUILA | 9 FCO. I MADERO | Rural | 1 | 0.10 |
| | 33 SAN PEDRO LAS COLONIAS | Rural | 1 | 0.10 |
| | 35 TORREON | Urbana | 8 | 0.80 |
| | TOTAL | | 10 | 1.00 |
| 6 COLIMA | 1 COLIMA | Urbana | 2 | 0.67 |
| | 4 CUAUHEMOC | Rural | 1 | 0.33 |
| | TOTAL | | 3 | 1.00 |
| 7 CHIAPAS | 27 CHIAPA DE CORZO | Rural | 2 | 0.14 |
| | 77 SAN CRISTOBAL DE LAS CASAS | Rural | 3 | 0.21 |
| | 87 SUCHIAPA | Rural | 2 | 0.14 |
| | 102 TUXTLA GUTIERREZ | Urbana | 6 | 0.43 |
| | 111 ZINACANTAN | Rural | 1 | 0.07 |
| TOTAL | | 14 | 1.00 | |
| 8 CHIHUAHUA | 4 AQUILES | Rural | 1 | 0.08 |
| | 17 CUAUHEMOC | Rural | 1 | 0.08 |
| | 19 CHIHUAHUA | Rural | 7 | 0.54 |
| | 37 JUAREZ | Urbana | 4 | 0.31 |
| | TOTAL | | 13 | 1.00 |
| 9 DISTRITO FEDERAL | 2 AZCAPOTZALCO | Urbana | 3 | 0.06 |
| | 3 COYOACAN | Urbana | 1 | 0.02 |
| | 5 GUSTAVO A. MADERO | Urbana | 6 | 0.13 |
| | 6 IZTACALCO | Urbana | 2 | 0.04 |
| | 7 IZTAPALAPA | Urbana | 5 | 0.11 |
| | 10 ALVARO OBREGON | Urbana | 5 | 0.11 |
| | 12 TLALPAN | Urbana | 2 | 0.04 |
| | 13 XOCHIMILCO | Urbana | 2 | 0.04 |
| | 14 BENITO JUAREZ | Urbana | 6 | 0.13 |
| | 15 CUAUHEMOC | Urbana | 6 | 0.13 |
| | 16 MIGUEL HIDALGO | Urbana | 4 | 0.09 |
| | 17 VENUSTIANO CARRANZA | Urbana | 5 | 0.11 |
| | TOTAL | | 47 | 1.00 |

ANEXO 1 RESUMEN DE LA PRIMERA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Secc. | Proporcion Mta. Edo |
|---------------|-------------------------------|--------|-----------------|------------------------|
| 10 DURANGO | 5 DURANGO | Urbana | 4 | 0.67 |
| | 8 GUADALUPE VICTORIA | Rural | 2 | 0.33 |
| | TOTAL | | 6 | 1.00 |
| 11 GUANAJUATO | 7 CELAYA | Urbana | 3 | 0.17 |
| | 15 GUANAJUATO | Urbana | 1 | 0.06 |
| | 17 IRAPUATO | Rural | 2 | 0.11 |
| | 20 LEON | Urbana | 4 | 0.22 |
| | 28 SALVATIERRA | Rural | 2 | 0.11 |
| | 31 SAN FRANCISCO DEL RINCON | Rural | 2 | 0.11 |
| | 37 SILAO | Urbana | 4 | 0.22 |
| TOTAL | | 18 | 1.00 | |
| 12 GUERRERO | 1 ACAPULCO DE JUAREZ | Urbana | 4 | 0.40 |
| | 28 CHILAPA DE ALVAREZ | Rural | 2 | 0.20 |
| | 29 CHILPANCINGO DE LOS BRAVO | Urbana | 2 | 0.20 |
| | 62 TIXTLA DE GUERRERO | Rural | 2 | 0.20 |
| TOTAL | | 10 | 1.00 | |
| 13 HIDALGO | 39 REAL DEL MONTE | Rural | 2 | 0.22 |
| | 47 PACHUCA DE SOTO | Urbana | 3 | 0.33 |
| | 75 TULA DE ALLENDE | Rural | 2 | 0.22 |
| | 76 TULANCINGO | Rural | 2 | 0.22 |
| TOTAL | | 9 | 1.00 | |
| 14 JALISCO | 41 GUADALAJARA | Urbana | 9 | 0.36 |
| | 47 IXTLAHUACAN | Rural | 1 | 0.04 |
| | 55 LAGOS DE MORENO | Rural | 2 | 0.08 |
| | 69 PUERTO VALLARTA | Urbana | 4 | 0.16 |
| | 72 EL SALTO | Rural | 2 | 0.08 |
| | 96 TEUCHITLAN | Rural | 1 | 0.04 |
| | 120 ZAPOPAN | Urbana | 5 | 0.20 |
| | 124 ZAPOTLANEJO | Rural | 1 | 0.04 |
| TOTAL | | 25 | 1.00 | |
| 15 MEXICO | 13 ATIZAPAN DE ZARAGOZA | Urbana | 2 | 0.04 |
| | 20 COACALCO | Urbana | 2 | 0.04 |
| | 25 CUAUTITLAN DE ROMERO RUBIO | Urbana | 3 | 0.06 |
| | 26 CHALCO | Rural | 1 | 0.02 |
| | 28 CHAPULTEPEC | Rural | 1 | 0.02 |
| | 32 CHIMALHUACAN | Urbana | 1 | 0.02 |
| | 34 ECATEPEC | Urbana | 10 | 0.20 |
| | 40 IXTAPALUCA | Rural | 1 | 0.02 |
| | 55 METEPEC | Urbana | 3 | 0.06 |
| | 58 NAUCALPAN | Urbana | 7 | 0.14 |
| | 60 NEZAHUALCOYOTL | Urbana | 8 | 0.16 |
| | 61 NICOLAS ROMERO VILLA | Rural | 1 | 0.02 |
| | 73 SANTA MA. RAYON | Rural | 1 | 0.02 |
| | 91 TENANGO | Rural | 1 | 0.02 |
| | 105 TLALNEPANTLA | Urbana | 6 | 0.12 |
| 107 TOLUCA | Urbana | 1 | 0.02 | |

ANEXO 1 RESUMEN DE LA PRIMERA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Secc. | Proporcion Mta. Edo |
|-----------------|--------------------------------|--------|-----------------|------------------------|
| 15 MEXICO | 110 TULTITLAN | Rural | 1 | 0.02 |
| | 119 ZINACANTEPEC | Rural | 1 | 0.02 |
| | TOTAL | | 51 | 1.00 |
| 16 MICHOACAN | 33 CUPATIZIO | Rural | 1 | 0.06 |
| | 54 MORELIA | Urbana | 6 | 0.38 |
| | 67 PATZCUARO | Rural | 2 | 0.13 |
| | 103 URUAPAN | Urbana | 4 | 0.25 |
| | 109 ZAMORA | Rural | 1 | 0.06 |
| | 113 ZITACUARO | Rural | 2 | 0.13 |
| | TOTAL | | 16 | 1.00 |
| 17 MORELOS | 7 CUERNAVACA | Urbana | 4 | 0.67 |
| | 11 JUTEPEC | Rural | 1 | 0.17 |
| | 17 PUENTE DE IXTLA | Rural | 1 | 0.17 |
| | TOTAL | | 6 | 1.00 |
| 18 NAYARIT | 17 TEPIC | Urbana | 4 | 1.00 |
| | TOTAL | | 4 | 1.00 |
| 19 NUEVO LEON | 6 APODACA | Urbana | 2 | 0.13 |
| | 19 SAN PEDRO GARZA GARCIA | Urbana | 1 | 0.06 |
| | 26 GUADALUPE | Urbana | 4 | 0.25 |
| | 34 LINARES | Rural | 2 | 0.13 |
| | 40 MONTERREY | Urbana | 5 | 0.31 |
| | 47 SAN NICOLAS DE LOS GARZA | Urbana | 1 | 0.06 |
| | 48 SANTA CATARINA | Urbana | 1 | 0.06 |
| | TOTAL | | 16 | 1.00 |
| 20 OAXACA | 5 ASUNCION CUYOTEPENJI | Rural | 1 | 0.08 |
| | 6 ASUNCION IXTALTEPEC | Rural | 1 | 0.08 |
| | 43 JUCHITAN DE ZARAGOZA | Urbana | 3 | 0.23 |
| | 66 OAXACA DE JUAREZ | Urbana | 2 | 0.15 |
| | 76 SALINA CRUZ | Urbana | 1 | 0.08 |
| | 182 SAN JUAN BAUTISTA TUXTEPEC | Rural | 3 | 0.23 |
| | 515 SANTO DOMINGO TEHUANTEPEC | Rural | 2 | 0.15 |
| | TOTAL | | 13 | 1.00 |
| 21 PUEBLA | 15 AMOZOC | Rural | 2 | 0.11 |
| | 19 ATLIXCO | Urbana | 2 | 0.11 |
| | 115 PUEBLA | Urbana | 5 | 0.28 |
| | 133 SAN MARTIN TEXMELUCAN | Rural | 3 | 0.17 |
| | 141 CHOLULA | Rural | 3 | 0.17 |
| | 155 TEHUACAN | Urbana | 3 | 0.17 |
| | TOTAL | | 18 | 1.00 |
| 22 QUERETARO | 14 QUERETARO | Urbana | 5 | 1.00 |
| | TOTAL | | 5 | 1.00 |
| 23 QUINTANA ROO | 1 CANCUN (BENITO JUAREZ IFE) | Urbana | 3 | 1.00 |
| | TOTAL | | 3 | 1.00 |

ANEXO 1 RESUMEN DE LA PRIMERA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Secc. | Proporcion Mta. Edo |
|--------------------|-------------------------------|--------|-----------------|------------------------|
| 24 SAN LUIS POTOSI | 20 MATEHUALA | Rural | 1 | 0.11 |
| | 28 SAN LUIS POTOSI | Urbana | 5 | 0.56 |
| | 35 SOLEDAD DE GRACIANO SANCHE | Rural | 1 | 0.11 |
| | 37 TAMAZUNCHALE | Rural | 1 | 0.11 |
| | 56 ZARAGOZA | Rural | 1 | 0.11 |
| | TOTAL | | 9 | 1.00 |
| 25 SINALOA | 1 AHOME | Rural | 1 | 0.09 |
| | 6 CULIACAN | Urbana | 5 | 0.45 |
| | 12 MAZATLAN | Urbana | 3 | 0.27 |
| | 18 NAVOLATO | Rural | 2 | 0.18 |
| | TOTAL | | 11 | 1.00 |
| 26 SONORA | 49 HERMOSILLO | Urbana | 3 | 0.30 |
| | 59 CAJEME | Urbana | 7 | 0.70 |
| | TOTAL | | 10 | 1.00 |
| 27 TABASCO | 4 VILLA HERMOSA, (CENTRO IFE) | Urbana | 4 | 0.57 |
| | 5 COMACALCO | Rural | 1 | 0.14 |
| | 8 HUIMANGUILLO | Rural | 2 | 0.29 |
| | TOTAL | | 7 | 1.00 |
| 28 TAMAULIPAS | 3 ALTA MIRA | Rural | 1 | 0.08 |
| | 9 MADERO | Urbana | 3 | 0.25 |
| | 38 TAMPICO | Urbana | 3 | 0.25 |
| | 41 VICTORIA | Urbana | 5 | 0.42 |
| | TOTAL | | 12 | 1.00 |
| 29 TLAXCALA | 10 CHIAUTEMPAN | Rural | 1 | 0.25 |
| | 33 TLAXCALA | Urbana | 3 | 0.75 |
| | TOTAL | | 4 | 1.00 |
| 30 VERACRUZ | 16 ANGEL R. CABADA | Rural | 2 | 0.06 |
| | 29 BOCA DEL RIO | Urbana | 2 | 0.06 |
| | 34 CATEMACO | Rural | 2 | 0.06 |
| | 41 COATZACOALCOS | Urbana | 1 | 0.03 |
| | 42 COATZINTLA | Urbana | 1 | 0.03 |
| | 46 CORDOBA | Urbana | 2 | 0.06 |
| | 89 XALAPA | Urbana | 4 | 0.13 |
| | 109 MINATITLAN | Rural | 2 | 0.06 |
| | 119 ORIZABA | Urbana | 2 | 0.06 |
| | 125 PAPANTLA | Rural | 3 | 0.10 |
| | 132 POZA RICA DE HIDALGO | Urbana | 1 | 0.03 |
| | 142 SAN ANDRES TUXTLA | Rural | 2 | 0.06 |
| | 188 TUXPAN | Urbana | 2 | 0.06 |
| | 192 VERACRUZ | Urbana | 5 | 0.16 |
| | TOTAL | | 31 | 1.00 |
| 31 YUCATAN | 50 MERIDA | Urbana | 6 | 1.00 |
| | TOTAL | | 6 | 1.00 |

ANEXO 1 RESUMEN DE LA PRIMERA MUESTRA

| <i>ESTADO</i> | <i>MUNICIPIO</i> | <i>TIPO</i> | <i>No. de Secc.</i> | <i>Proporcion Mta. Edo</i> |
|---------------|------------------|-------------|-------------------------|--------------------------------|
| 32 ZACATECAS | 10 FRESNILLO | Urbana | 1 | 0.17 |
| | 17 GUADALUPE | Rural | 2 | 0.33 |
| | 50 VETA GRANDE | Rural | 1 | 0.17 |
| | 56 ZACATECAS | Urbana | 2 | 0.33 |
| | TOTAL | | 6 | 1.00 |

ANEXO 2 RESUMEN DE LA SEGUNDA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Proporción | |
|-------------------------|-------------------------------|--------|-------------------|----------|
| | | | Secc. | Mta. Edo |
| 1 AGUASCALIENTES | 1 AGUASCALIENTES | Urbana | 3 | 0.75 |
| | 3 CALVILLO | Rural | 1 | 0.25 |
| | TOTAL | | 4 | 1.00 |
| 2 BAJA CALIFORNIA NORTE | 2 MEXICALI | Urbana | 4 | 0.44 |
| | 4 TIJUANA | Urbana | 5 | 0.56 |
| | TOTAL | | 9 | 1.00 |
| 3 BAJA CALIFORNIA SUR | 3 PAZ, LA | Urbana | 2 | 1.00 |
| | TOTAL | | 2 | 1.00 |
| 4 CAMPECHE | 1 CAMPECHE | Urbana | 2 | 0.67 |
| | 2 ASIENTOS | Rural | 1 | 0.33 |
| | TOTAL | | 3 | 1.00 |
| 5 COAHUILA | 33 SAN PEDRO LAS COLONIAS | Rural | 3 | 0.30 |
| | 35 TORREON | Rural | 7 | 0.70 |
| | TOTAL | | 10 | 1.00 |
| 6 COLIMA | 1 COLIMA | Urbana | 2 | 1.00 |
| | TOTAL | | 2 | 1.00 |
| 7 CHIAPAS | 12 BERRIOZABAL | Rural | 1 | 0.07 |
| | 27 CHIAPA DE CORZO | Rural | 1 | 0.07 |
| | 77 SAN CRISTOBAL DE LAS CASAS | Rural | 5 | 0.36 |
| | 87 SUCHIAPA | Rural | 2 | 0.14 |
| | 102 TUXTLA GUTIERREZ | Rural | 4 | 0.29 |
| | 111 ZINACANTAN | Rural | 1 | 0.07 |
| TOTAL | | 14 | 1.00 | |
| 8 CHIHUAHUA | 17 CUAUHTEMOC | Rural | 2 | 0.15 |
| | 19 CHIHUAHUA | Rural | 8 | 0.62 |
| | 37 JUAREZ | Urbana | 3 | 0.23 |
| | TOTAL | | 13 | 1.00 |
| 9 DISTRITO FEDERAL | 2 AZCAPOTZALCO | Urbana | 3 | 0.06 |
| | 3 COYOACAN | Urbana | 3 | 0.06 |
| | 4 CUAJIMALPA DE MORELOS | Urbana | 1 | 0.02 |
| | 5 GUSTAVO A. MADERO | Urbana | 8 | 0.17 |
| | 6 IZTACALCO | Urbana | 3 | 0.06 |
| | 7 IZTAPALAPA | Urbana | 4 | 0.09 |
| | 8 MAGDALENA CONTRERAS, LA | Urbana | 2 | 0.04 |
| | 10 ALVARO OBREGON | Urbana | 1 | 0.02 |
| | 11 TLAHUAC | Urbana | 3 | 0.06 |
| | 12 TLALPAN | Urbana | 2 | 0.04 |
| | 13 XOCHIMILCO | Urbana | 1 | 0.02 |
| | 14 BENITO JUAREZ | Urbana | 3 | 0.06 |
| | 15 CUAUHTEMOC | Urbana | 6 | 0.13 |
| | 16 MIGUEL HIDALGO | Urbana | 3 | 0.06 |
| | 17 VENUSTIANO CARRANZA | Urbana | 4 | 0.09 |
| | TOTAL | | 47 | 1.00 |

ANEXO 2 RESUMEN DE LA SEGUNDA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Proporción | |
|--------------------|------------------------------|--------|-------------------|----------|
| | | | Secc. | Mta. Edo |
| 10 DURANGO | 5 DURANGO | Urbana | 4 | 0.67 |
| | 8 GUADALUPE VICTORIA | Rural | 2 | 0.33 |
| | TOTAL | | 6 | 1.00 |
| 11 GUANAJUATO | 7 CELAYA | Urbana | 2 | 0.11 |
| | 15 GUANAJUATO | Urbana | 4 | 0.21 |
| | 17 IRAPUATO | Rural | 2 | 0.11 |
| | 20 LEON | Rural | 7 | 0.37 |
| | 28 SALVATIERRA | Rural | 2 | 0.11 |
| | 31 SAN FRANCISCO DEL RINCON | Rural | 1 | 0.05 |
| | 37 SILAO | Rural | 1 | 0.05 |
| TOTAL | | 19 | 1.00 | |
| 12 GUERRERO | 1 ACAPULCO DE JUAREZ | Urbana | 4 | 0.40 |
| | 28 CHILAPA DE ALVAREZ | Rural | 1 | 0.10 |
| | 29 CHILPANCINGO DE LOS BRAVO | Rural | 3 | 0.30 |
| | 62 TIXTLA DE GUERRERO | Rural | 2 | 0.20 |
| | TOTAL | | 10 | 1.00 |
| 13 HIDALGO | 39 REAL DEL MONTE | Rural | 2 | 0.22 |
| | 47 PACHUCA DE SOTO | Urbana | 2 | 0.22 |
| | 75 TULA DE ALLENDE | Rural | 3 | 0.33 |
| | 76 TULANCINGO | Rural | 2 | 0.22 |
| | TOTAL | | 9 | 1.00 |
| 14 JALISCO | 120 ZAPOPAN | Urbana | 1 | 0.04 |
| | 41 GUADALAJARA | Urbana | 10 | 0.40 |
| | 55 LAGOS DE MORENO | Rural | 3 | 0.12 |
| | 69 PUERTO VALLARTA | Urbana | 3 | 0.12 |
| | 72 EL SALTO | Urbana | 1 | 0.04 |
| | 96 TEUCHITLAN | Rural | 1 | 0.04 |
| | 99 TLAQUEPAQUE | Rural | 1 | 0.04 |
| | 120 ZAPOPAN | Urbana | 4 | 0.16 |
| | 124 ZAPOTLANEJO | Rural | 1 | 0.04 |
| | TOTAL | | 25 | 1.00 |
| 15 MEXICO | 13 ATIZAPAN DE ZARAGOZA | Urbana | 2 | 0.04 |
| | 20 COACALCO | Urbana | 2 | 0.04 |
| | 25 CUAUTILAN DE ROMERO RUBIO | Urbana | 2 | 0.04 |
| | 26 CHALCO | Rural | 1 | 0.02 |
| | 28 CHAPULTEPEC | Rural | 1 | 0.02 |
| | 32 CHIMALHUACAN | Urbana | 2 | 0.04 |
| | 34 ECATEPEC | Urbana | 6 | 0.12 |
| | 38 HUIXQUILUCAN | Rural | 1 | 0.02 |
| | 40 IXTAPALUCA | Rural | 1 | 0.02 |
| | 43 IXTLAHUACAN | Urbana | 1 | 0.02 |
| | 55 METEPEC | Rural | 3 | 0.06 |
| | 58 NAUCALPAN | Urbana | 6 | 0.12 |
| | 60 NEZAHUALCOYOTL | Urbana | 11 | 0.22 |
| | 61 NICOLAS ROMERO VILLA | Rural | 1 | 0.02 |
| 73 SANTA MA. RAYON | Rural | 1 | 0.02 | |

ANEXO 2 RESUMEN DE LA SEGUNDA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Proporción | |
|---------------|--------------------------------|--------|-------------------|----------|
| | | | Secc. | Mta. Edo |
| 15 MEXICO | 91 TENANGO | Rural | 1 | 0.02 |
| | 105 TLALNEPANTLA | Urbana | 4 | 0.08 |
| | 107 TOLUCA | Urbana | 3 | 0.06 |
| | 110 TULTITLAN | Rural | 1 | 0.02 |
| | 119 ZINACANTEPEC | Rural | 1 | 0.02 |
| | TOTAL | | 51 | 1.00 |
| 16 MICHOACAN | 33 CUPATIZIO | Rural | 1 | 0.06 |
| | 54 MORELIA | Urbana | 8 | 0.50 |
| | 67 PATZCUARO | Rural | 2 | 0.13 |
| | 103 URUAPAN | Urbana | 2 | 0.13 |
| | 109 ZAMORA | Urbana | 2 | 0.13 |
| | 113 ZITACUARO | Rural | 1 | 0.06 |
| TOTAL | | 16 | 1.00 | |
| 17 MORELOS | 7 CUERNAVACA | Urbana | 2 | 0.33 |
| | 11 JIUTEPEC | Rural | 2 | 0.33 |
| | 17 PUENTE DE IXTLA | Rural | 2 | 0.33 |
| | TOTAL | | 6 | 1.00 |
| 18 NAYARIT | 17 TEPIC | Urbana | 4 | 1.00 |
| | TOTAL | | 4 | 1.00 |
| 19 NUEVO LEON | 6 APODACA | Urbana | 2 | 0.13 |
| | 19 SAN PEDRO GARZA GARCIA | Urbana | 1 | 0.06 |
| | 26 GUADALUPE | Urbana | 3 | 0.19 |
| | 34 LINARES | Rural | 2 | 0.13 |
| | 40 MONTERREY | Urbana | 4 | 0.25 |
| | 47 SAN NICOLAS DE LOS GARZA | Urbana | 2 | 0.13 |
| | 48 SANTA CATARINA | Urbana | 2 | 0.13 |
| | TOTAL | | 16 | 1.00 |
| 20 OAXACA | 6 ASUNCION IXTALTEPEC | Rural | 2 | 0.15 |
| | 43 JUCHITAN DE ZARAGOZA | Urbana | 3 | 0.23 |
| | 66 OAXACA DE JUAREZ | Urbana | 2 | 0.15 |
| | 76 SALINA CRUZ | Urbana | 2 | 0.15 |
| | 182 SAN JUAN BAUTISTA TUXTEPEC | Rural | 2 | 0.15 |
| | 515 SANTO DOMINGO TEHUANTEPEC | Rural | 2 | 0.15 |
| TOTAL | | 13 | 1.00 | |
| 21 PUEBLA | 15 AMOZOC | Rural | 2 | 0.11 |
| | 19 ATLIXCO | Urbana | 3 | 0.17 |
| | 115 PUEBLA | Urbana | 5 | 0.28 |
| | 133 SAN MARTIN TEXMELUCAN | Rural | 3 | 0.17 |
| | 141 CHOLULA | Rural | 2 | 0.11 |
| | 155 TEHUACAN | Urbana | 3 | 0.17 |
| TOTAL | | 18 | 1.00 | |
| 22 QUERETARO | 14 QUERETARO | Urbana | 5 | 1.00 |
| | TOTAL | | 5 | 1.00 |

ANEXO 2 RESUMEN DE LA SEGUNDA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Secc. | Proporcion Mta. Edo |
|--------------------|-------------------------------|--------|-----------------|------------------------|
| 23 QUINTANA ROO | 1 CANCUN (BENITO JUAREZ IFE) | Urbana | 3 | 1.00 |
| | TOTAL | | 3 | 1.00 |
| 24 SAN LUIS POTOSI | 20 MATEHUALA | Rural | 1 | 0.11 |
| | 28 SAN LUIS POTOSI | Urbana | 5 | 0.56 |
| | 35 SOLEDAD DE GRACIANO SANCHE | Rural | 1 | 0.11 |
| | 37 TAMAZUNCHALE | Rural | 1 | 0.11 |
| | 56 ZARAGOZA | Rural | 1 | 0.11 |
| | TOTAL | | 9 | 1.00 |
| 25 SINALOA | 1 AHOME | Rural | 2 | 0.18 |
| | 6 CULIACAN | Urbana | 5 | 0.45 |
| | 12 MAZATLAN | Urbana | 3 | 0.27 |
| | 18 NAVOLATO | Rural | 1 | 0.09 |
| | TOTAL | | 11 | 1.00 |
| 26 SONORA | 49 HERMOSILLO | Urbana | 3 | 0.30 |
| | 59 CAJEME | Urbana | 7 | 0.70 |
| | TOTAL | | 10 | 1.00 |
| 27 TABASCO | 4 VILLA HERMOSA, (CENTRO IFE) | Urbana | 4 | 0.57 |
| | 5 COMACALCO | Rural | 1 | 0.14 |
| | 8 HUIMANGUILLO | Rural | 2 | 0.29 |
| | TOTAL | | 7 | 1.00 |
| 28 TAMAULIPAS | 3 ALTA MIRA | Rural | 2 | 0.17 |
| | 9 MADERO | Urbana | 4 | 0.33 |
| | 38 TAMPICO | Urbana | 3 | 0.25 |
| | 41 VICTORIA | Urbana | 3 | 0.25 |
| | TOTAL | | 12 | 1.00 |
| 29 TLAXCALA | 10 CHIAUTEMPAN | Rural | 2 | 0.50 |
| | 33 TLAXCALA | Urbana | 2 | 0.50 |
| | TOTAL | | 4 | 1.00 |
| 30 VERACRUZ | 16 ANGEL R. CABADA | Rural | 2 | 0.06 |
| | 29 BOCA DEL RIO | Urbana | 1 | 0.03 |
| | 34 CATEMACO | Rural | 1 | 0.03 |
| | 35 CAZONES DE HERRERA | Rural | 1 | 0.03 |
| | 41 COATZACOALCOS | Urbana | 2 | 0.06 |
| | 42 COATZINTLA | Urbana | 1 | 0.03 |
| | 46 CORDOBA | Urbana | 3 | 0.10 |
| | 89 XALAPA | Urbana | 4 | 0.13 |
| | 109 MINATITLAN | Rural | 3 | 0.10 |
| | 119 ORIZABA | Urbana | 2 | 0.06 |
| | 125 PAPANTLA | Rural | 2 | 0.06 |
| | 132 POZA RICA DE HIDALGO | Urbana | 1 | 0.03 |
| | 142 SAN ANDRES TUXTLA | Rural | 3 | 0.10 |
| 188 TUXPAN | Urbana | 2 | 0.06 | |
| 192 VERACRUZ | Urbana | 3 | 0.10 | |
| | TOTAL | | 31 | 1.00 |

ANEXO 2 RESUMEN DE LA SEGUNDA MUESTRA

| ESTADO | MUNICIPIO | TIPO | No. de Proporción | |
|--------------|----------------|--------|-------------------|----------|
| | | | Secc. | Mta. Edo |
| 31 YUCATAN | 50 MERIDA | Urbana | 6 | 1.00 |
| | TOTAL | | 6 | 1.00 |
| 32 ZACATECAS | 10 FRESNILLO | Urbana | 2 | 0.33 |
| | 17 GUADALUPE | Rural | 2 | 0.33 |
| | 50 VETA GRANDE | Rural | 1 | 0.17 |
| | 56 ZACATECAS | Urbana | 1 | 0.17 |
| | TOTAL | | 6 | 1.00 |

"BIBLIOGRAFIA."

BEIZER, Boris; BOEHM, Barry W. *"Software Engineering Economics"*, New Jersey, Ed. Prentice Hall.

BELTRAO Moura, José Antáo y PHILLIPPE Sauvé, Jacques *"Redes Locales de Computadoras"*, México, 1a. Edición en Español, Ed. McGraw-Hill, 1992, p.p. 1-4,7-21.

BLALOCK, Hubert M. *"Social Statistics"*, New York, 2a. Edición, Ed. McGraw-Hill, 1979.

BUSTAMANTE Terreros, Lenin. *Investigación. "Las demandas ciudadanas y tendencias electorales, como marco contextual en el análisis y formulación de políticas públicas del Gobierno del D.F."*, Acatlán, Edo. de México, diciembre de 1994.

CARPIDO, Jorge. *"ALGUNAS CONSIDERACIONES SOBRE EL PROCESO FEDERAL ELECTORAL DE 1994"*, Versión Estenográfica del Mensaje Pronunciado el 12 de septiembre de 1994, México, Secretaría de Gobernación, 1994.

COCHRAN, William G. *"Técnicas de Muestreo"*, México, 1a. Edición, Ed. CECSA, 1980.

CHURCHMAN, C.W.; AUERBACH, I. y SADAN, Simcha. *"Thinking for Decissions Deductive Quantitative Methods"*. Palo Alto California: Science Research Associates, 1975.

CORTEZ, Fernando. *"Métodos Estadísticos Aplicados a la Investigación en Ciencias Sociales: Análisis de Asociación"*, México, Ed. El Colegio de México, 1987.

DES, Raj. *"Teoría de Muestreo"*, México, 1a. Edición, Ed. Fondo de Cultura Económica, 1980.

Enciclopedia *"Hombre Ciencia y Tecnología"*. Tomo 8, México, Ed. Encyclopedia Británnica de México, S.A de C.V., 1982.

FAIRLEY, Richard E. *"Software Engineering concepts"*, Series in Software Engineering and Technology", Ed. McGraw-Hill, New York, México, c.1985.

FERREYRA Cortés, Gonzalo. *"VIRUS en las computadoras"*, México, 3a. Edición, Ed. Computec - Alfaomega, 1994, p.p. 47-58, 211-216.

FOWLER, Floyd J. Jr. *"Survey Research Methods"*, 2a. Ed., SAGE Publications. Newbury Park, 1993. (Applied Social Research Methods Series).

HAMACHER, V. Carl, Zvonko G. y ZAKI, Safwat G. *"Organización de Computadoras"*, México, 2a. Edición (1a. Ed. en español), Ed. McGraw-Hill, 1988.

JACOBSON, Ivar, CHRISTERSON, Magnus, JONSON, Patric y OVERGAARD Gunnar. *"Object-Oriented software engineering. A use case Driven Approach"*, U.S.A., Ed. Addison-Wesley, 1992, 1ª reimpresión e impresión. p.p. 429-457.

KEYNES, Milton. *"The Open University, Systems Performance. Human factors and systems failures"*, Inglaterra; The Open University Press, 1976, p.p. 8-11.

KISH, Leslie. *"Muestreo de Encuestas"*, Ed. Trillas, México, 1975.

KREYSZIG, Erwin. *"Introducción a la Estadística Matemática"*, México, Ed. Limusa, 1987.

LARSEN, Glenn H. y LARSEN, Kristopher A. *"Domine HARVARD GRAPHICS 3"*, E.U.A., Ed. Addison Wesley Iberoamericana, S.A., 1993.

MENDENHALL, William; SCHEAFFER, Richard L.; WACKERLY, Dennis D. *"Estadística Matemática con Aplicaciones"*, Estados Unidos de América, Ed. Grupo Editorial Iberoamérica, 1986.

MENDEZ R. Ignacio, NAMIHIRA G., Delia, MORENO A., Laura y SOSA de M., Cristina. *"El protocolo de INVESTIGACION: Lineamientos para su elaboración y análisis"*, México, Ed. Trillas, 3a. Reimpresión 1994.

NORTON, Peter. *"Periféricos y Accesorios para la IBM-PC, PS/2 y Compatibles"*, México, Ed. Prentice Hall, 1993, p.p. 85-89, 121-123, 129-131, 208-221, 236-237, 300-322.

RENDÓN Ortiz, Gilberto. *"INTERNET para principiantes"*, Cuarta reimpresión, Ed. SELECTOR, 1997.

TAKETT, Jack Jr.; GUNTER, David y BROWN, Lance. *"Linux Edición Especial"*, México, Ed. Prentice Hall Hispanoamericana, 1996.

TANENBAUM, Andrew S. *"Sistemas Operativos Modemos"*, México, Ed. Prentice Hall, 1993.

THOMPSON, Steven K. *"Sampling"*, New York, Ed. J.Wiley, 1993.

VAN Gigch, John P. *"Teoría General de Sistemas"*, México, 2a. Edición, Ed. Trillas, 1987.

VAN VLIET, Hans. *"Software Engineering: Principles and Practice"*, England, Ed. John Wiley & Sons, 1993, p.p.1-15, 21-39, 45-47.

WEINBERG, G.M. *"An Introduction to General System Thinking"*, Nueva York, Ed. Wiley, 1975, p.p. 62-63.

WOODWARD, Jeff. *"ABC NOVELL NETWARE"*, México, 1a. Edición, Ed. SYBEX, 1991.

HEMEROGRAFIA

BROWING, Dave. *"Bases de Datos por Relación: Nueva sangre, nueva potencia"*, en Revista **PC MAGAZINE** en Español, Volumen 4, Número 9, México, julio de 1993, p.p.25-47.

CISNEROS Stoianowski, Gerardo. *"Urge Formar Recursos Humanos en Ciencia Computacional"*, en Periódico **EXCELSIOR**, 8 de mayo de 1995.

CORNFIELD, Jerome. *"Modern Methods in Sampling of Human Populations. The Determination of Sample Size"*, E.U.A., American Journal Publication Health, 41. p.p. 654-661.

FUTRELL, David. *"Ten reasons why surveys fail"*. **Quality Progress**, Vol.27, Número 4, abril de 1994, p.66.

GAL, Avigdor y ETZION, Opher. *"Maintaining Data-Driven Rules in Databases"*, en Revista **Computer**, Vol.28, Número 1, E.U.A., enero de 1995, p.p.28-39.

GREENFIELD, David. *"Servidores de Archivos: PARA SERVIR Y PROTEGER"*, en Revista **PC MAGAZINE en Español**, Volumen 4, Número 8, México, agosto de 1993, p.p.45-64.

HOLTZMAN, Jeff. *"Computer Conections: Ubersoft uber alles"*, en Revista **Electronics Now**, Vol.66, Número 3, E.U.A., marzo de 1995, p.p.92-96.

KHORRAMI, F., JAIN, S. y TZES, A. *"Experimental Results on Adaptive Nonlinear Control and Input Preshaping for Multi-link Flexible Manipulators"*, en Revista **Automatica**, Vol.31, Número 1, enero de 1995, p.p.83-98.

LARSSON, P. y EDEFORS. *"Optimising a high-speed serial/parallel sum-of-products hardware structure with respect to bus utilisation"*, en Revista **IEE Proceedings Computers and Digital Techniques**, Vol.142, Número 1, E.U.A., enero de 1995, p.p.77-80.

López, Ernesto. *"Ofrecen opciones `baratas`"*, en Periódico **Reforma**, Sección A, Interfase, Lunes, 27 de febrero de 1995, pág.1A.

López, Ernesto. *"Realizan Ciberelecciones"*, en Periódico **Reforma**, Sección A, Interfase, Lunes, 7 de julio de 1997, pág.1A.

PROSISE, Jeff. *"DOS 6, ¿La colección ideal de servicios?"*, en Revista **PC MAGAZINE en Español**, Volumen 4, Número 7, México, julio de 1993, p.p.58-74.

PROSISE, Jeff. *"MS-DOS Q & A"*, en Revista **Microsoft systems journal**, Vol.10, Número 1, 1995, p.p.95-60.

RICCIARDI, Sal. *"Paradox 4.0"*, en Revista **PC MAGAZINE en Español**, Volumen 4, Número 2, México, febrero de 1993, p.10.

SALEMI, Joe. "Sistema Operativo OS/2 2.1", en Revista **PC MAGAZINE** en Español, Volumen 4, Número 9, México, julio de 1993, p.p.83-87.

STEVENSON, Daniel, HILLERY, Nathan y BYRD, Greg. "Secure Communications in ATM Networks", **ACM PRESS**, Vol.38, Número 2.

VARHOL, Peter D. "Servers: Upsizing with SMP: Don't Ignore the OS", en Revista **Datamation**, Vol.41, Número 1, E.U.A., 15 de enero de 1995, p.p.59-63.

VELA del Bosque, Humberto. "Informática no hará magia en elecciones", en Periódico **Reforma**, Sección A, Interfase, Lunes, 30 de junio de 1997, pág.2A

VELA del Bosque, Humberto. "Logra la informática buen papel en elecciones", en Periódico **Reforma**, Sección A, Interfase, Lunes, 14 de julio de 1997, pág.2A.

VETTER, Ronald J. "ATM, Concepts, Architectures, and Protocols", **ACM PRESS**, Vol.38, Número 2, E.U.A., febrero de 1995, p.p.30-39.