



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ANÁLISIS DE LOS NIVELES DE SERVICIO DEL FONDO DE EMERGENCIAS 133  
DE CARABINEROS DE CHILE

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL INDUSTRIAL

LORETO JAVIERA ELGUETA ESPINOSA

PROFESOR GUÍA:  
MARCELO OLIVARES ACUÑA

MIEMBROS DE LA COMISIÓN:  
PATRICIO CONCA KEHL  
RICHARD WEBER HAAS

SANTIAGO DE CHILE  
MARZO 2013

## ANÁLISIS DE LOS NIVELES DE SERVICIO DEL FONOS DE EMERGENCIAS 133 DE CARABINEROS DE CHILE

El Fono de Emergencias 133 de Carabineros es un servicio telefónico que atiende las emergencias reportadas por la ciudadanía para luego derivarlas a un procedimiento policial que las resolverá en terreno.

Está demostrado que una gestión inteligente del servicio de emergencias en una ciudad puede aumentar la seguridad objetiva y mejorar la percepción de inseguridad, y así incidir positivamente en la calidad de vida de la población [1]. Este trabajo de título investiga el proceso de servicio que brinda la unidad Nivel 133 de la Central de Comunicaciones (CENCO), en la Región Metropolitana. El objetivo principal es medir el nivel de servicio del sistema actual.

El desafío consiste en cuantificar los parámetros operativos del sistema y medir el desempeño a través del indicador más importante en este tipo de servicios, los tiempos de espera. Hoy en día, los tiempos de espera de las personas que llaman al 133 no se registran y no hay manera de observarlos. Tampoco se puede conocer el número de llamadas que el sistema rechaza cuando está lleno. Existe un consenso en cuanto a que las llamadas inoficiosas<sup>1</sup> son elevadas y pueden ser la causa de la alta congestión de las líneas<sup>2</sup>. Una publicación oficial señala que en el primer semestre del 2011 más de un 70 % de las llamadas contestadas fueron de esta naturaleza [6]. Lamentablemente, las campañas publicitarias en contra de las llamadas inoficiosas son contraproducentes, se ha visto que aumentan mientras la campaña está en curso. Tal es el nivel de llamadas inoficiosas, que han sido clasificadas en seis categorías. A su vez, existen 103 tipos de llamadas de emergencia, cada uno corresponde a un delito o falta, que si se reporta, Carabineros debe concurrir.

Para lograr el objetivo, se sigue una metodología que comienza procesando datos desagregados de llamadas atendidas, a partir de distintas fuentes, para determinar el comportamiento del sistema. Luego, se puede cuantificar la demanda de los distintos tipos de llamadas de emergencia e inoficiosas, en cuanto a volumen y también tiempos de servicio. A través de teoría de colas, se desarrolla un modelo incremental de varias etapas, que integra parámetros que reflejan cada vez mejor la realidad. De esta manera, se estima la tasa de abandono, pero aún así existen limitaciones en dichos modelos. El análisis final del nivel de servicio y evaluación de las medidas de rendimiento se hace a través de simulación, lo que permite flexibilizar los supuestos Markovianos de la teoría de colas, ya que éstos nuevos modelos aceptan comportamientos aleatorios de los parámetros del sistema, como las distribuciones del tiempo de servicio, cantidad de operadores y tasas de llegada.

Ésta investigación muestra que los niveles locales de servicio están por debajo de los estándares internacionales. Un fin de semana cualquiera, menos del 3 % de las emergencias se atienden en menos de 10 segundos. En el caso del 911, en Estados Unidos, la meta de servicio es atender al 90 % antes de 10, y en general se cumple o se supera. No es trivial decidir acciones a tomar para mejorar el servicio, ya que la dotación, el problema de volumen de llamadas inoficiosas, y la estructura de los datos, deben ser resueltos en paralelo. Algunas propuestas se evalúan y se obtienen mejoras considerables en los niveles de servicio. En días hábiles, si se logra disminuir la llegada de llamadas inoficiosas, se pasaría de un 47 % a un 63 % de emergencias atendidas en menos de 30 segundos.

Según las metas que se formulen, se debe evaluar un proyecto de mejora que considere varias propuestas, y mida su impacto antes de ser implementado. Para medir las consecuencias de cambios en parámetros del sistema, tales como: operadores, tasas de llegadas, tasas de servicio o capacidad tecnológica, se podría usar la metodología y herramientas que se usaron en este trabajo.

---

<sup>1</sup>Llamadas inoficiosas son las que generan mal uso del servicio

<sup>2</sup>La descripción del Fono 133, en el sitio web de Carabineros comienza diciendo "Cuando se disca el Fono 133 y existe un alto tráfico o congestión de llamadas...", luego indican que un 78,21 % de las llamadas fueron inoficiosas en 2011.

# Agradecimientos

A mi familia por el amor incondicional, el apoyo y la confianza que me han brindado siempre. A Ivan por acompañarme y ayudarme a tener convicción, aún en momentos difíciles. A mis amigos, que me han dado ánimo y razones para perseverar.

A los profesores que me guiaron en esta tesis y a quienes me han inspirado a través de la vocación que demuestran en el aula. Al Centro de Investigación Aplicada para el Análisis y Modelamiento de la Seguridad, CEAMOS, que me ofreció trabajar en este tema tan interesante y me apoyó durante todo el proceso.

Agradezco también a Carabineros de Chile y especialmente a la Academia de Ciencias Policiales, ACIPOL, por depositar su confianza en mí y respaldar este estudio.

# Tabla de contenido

<b>I</b>	<b>Contenido</b>	<b>1</b>
<b>1.</b>	<b>Antecedentes Generales</b>	<b>2</b>
1.1.	Introducción . . . . .	2
1.2.	Descripción del Problema . . . . .	3
1.3.	Justificación del Proyecto . . . . .	3
1.4.	Objetivos . . . . .	4
1.4.1.	Objetivo General . . . . .	4
1.4.2.	Objetivos Específicos . . . . .	4
<b>2.</b>	<b>Marco Teórico</b>	<b>5</b>
2.1.	Fundamentación . . . . .	5
2.1.1.	Importancia de los Call center de Emergencia . . . . .	5
2.1.2.	Estándares de Servicio . . . . .	6
2.2.	Descripción del Proceso Metodológico . . . . .	7
2.3.	Herramientas Utilizadas . . . . .	8
2.3.1.	Modelamiento de Procesos . . . . .	8
2.3.2.	Teoría de Colas . . . . .	9
2.3.3.	Simulación . . . . .	22
2.3.4.	Muestreo . . . . .	23
2.4.	Alcances . . . . .	23
<b>3.</b>	<b>Descripción del Sistema</b>	<b>25</b>
3.1.	El Nivel 133 . . . . .	26
3.1.1.	CAD . . . . .	26
3.1.2.	Variables de estado . . . . .	26
3.1.3.	Operadores . . . . .	27
3.2.	Tipos de Llamadas Entrantes . . . . .	28
3.2.1.	Llamadas Inoficiosas . . . . .	28
3.2.2.	Llamadas de Emergencia . . . . .	29
3.3.	Modelos de Procesos . . . . .	29
3.3.1.	Proceso de Tratamiento de Llamadas . . . . .	30
3.3.2.	Proceso de Atención de Llamadas . . . . .	30
3.4.	Patrones de Comportamiento . . . . .	31
3.4.1.	Demanda Total . . . . .	32
3.4.2.	Llamadas Entrantes . . . . .	32
3.4.3.	Tipos de Llamadas Atendidas . . . . .	36
3.4.4.	Llamadas Rechazadas . . . . .	39
3.4.5.	Tasas de Llegada . . . . .	40
3.4.6.	Tiempos de Servicio . . . . .	41
3.4.7.	Operadores Conectados . . . . .	43
3.4.8.	Resumen . . . . .	44

<b>4. Análisis y Evaluación</b>	<b>47</b>
4.1. Validez de la Simulación	47
4.1.1. Aplicación del Modelo $M/M/c$	48
4.1.2. Aplicación del Modelo $M/M/c/K$	49
4.2. Estimación de la Paciencia	50
4.2.1. Validez de la Paciencia Estimada	51
4.3. Comparación de Métodos	52
4.4. Modelamiento del Sistema Real	53
4.5. Nivel de Servicio	56
4.5.1. Días hábiles	56
4.5.2. Fin de Semana	57
4.6. Discusión	57
<b>5. Recomendaciones y Conclusión</b>	<b>59</b>
5.1. Dotación Óptima	59
5.2. Disminuir la Demanda	60
5.2.1. Sancionar Llamados Inoficiosos	60
5.2.2. Mejorar Nivel 139	61
5.2.3. Rediseño para Atención de Alarmas Activadas	61
5.3. Estudio de Tráfico	62
5.3.1. Modelo de Datos	62
5.4. Conclusiones	63
<b>Bibliografía</b>	<b>64</b>
<b>II Apéndices</b>	<b>65</b>
<b>A. Procesamiento de Datos</b>	<b>66</b>
A.1. Proceso KDD	66
A.2. Proceso CRISP-DM	67
A.3. Datos Disponibles	67
A.3.1. Sistemas de Información	67
A.3.2. Cantidad de llamadas Abandonadas	68
<b>B. Parámetros en Detalle</b>	<b>69</b>
B.1. Ajuste de Tasa de Servicio	69
B.1.1. Tiempos de Servicio de Emergencias Ponderados	70
B.2. Resultados del Muestreo	70
B.2.1. Tiempo de Espera en Cola	71
<b>C. Salidas Calculadora Erlang-X</b>	<b>73</b>
C.1. Horario 1	73
C.2. Horario 2	74
<b>D. Salidas Simulación</b>	<b>75</b>
D.1. $M/M/c$ Horario 1	75
D.2. $M/M/c/K$ Horario 2	75
D.3. $M/M/c/K$ 24 Horas	76
D.4. Simulación con Abandonos	77
<b>E. Resultados Nivel de Servicio</b>	<b>78</b>
E.1. Semana	78
E.2. Fin de Semana	79

<b>F. Información Istitucional</b>	<b>80</b>
F.1. Fono 133 . . . . .	80
F.2. Sistemas de Atención de llamadas . . . . .	81
F.3. Tipificación de Emergencias . . . . .	82

# Índice de figuras

2.2.1. Metodología utilizada (basada en CRISP-DM).	7
2.3.1. Transiciones entre estados de un Proceso de Nacimiento y Muerte Generalizado	10
2.3.2. Esquema de una cola M/M/1	13
2.3.3. Transiciones entre estados de una cola M/M/1 [13]	14
2.3.4. Esquema de una cola M/M/c	15
2.3.5. Transiciones entre estados de una cola M/M/c. [13]	15
2.3.6. Esquema de una cola M/M/c/K	17
2.3.7. Transiciones entre estados de un sistema de colas M/M/c/K	17
2.3.8. Esquema de una cola con abandonos	19
2.3.9. Transiciones entre estados de una cola con abandonos, M/M/c+M	19
2.3.10. Comparación entre Erlang-A y Erlang-C en cuanto a probabilidad de espera.	19
2.3.11. Comparación entre Erlang-A y Erlang-C en cuanto a tiempo promedio de espera.	20
2.3.12. Campos de entrada y salida de Erlang-X Calculator	22
3.0.1. Proceso de atención de emergencias en CENCO	25
3.1.1. <i>Layout</i> del Nivel 133	27
3.3.1. Modelo del proceso Tratamiento de Llamadas	30
3.3.2. Modelo del proceso Atención de Llamadas	31
3.4.1. Diagrama descriptivo de composición de la demanda	32
3.4.2. Volumen de llamadas atendidas y abandonadas por día. Primeras semanas de Agosto.	33
3.4.3. Volumen de llamadas atendidas y abandonadas por día. Primeras semanas de Noviembre.	33
3.4.4. Volumen de llamadas entrantes por minuto en días hábiles	34
3.4.5. Volumen de llamadas entrantes por minuto en fines de semana	34
3.4.6. Porcentaje de llamadas de emergencia e inoficiosas por día.	36
3.4.7. Porcentajes promedio de llamadas inoficiosas y emergencias por día.	36
3.4.8. Porcentajes de cada clase de llamada inoficiosa por día.	37
3.4.9. Tipos de llamadas atendidas durante el día en días hábiles	37
3.4.10. Tipos de llamadas atendidas durante el día en fin de semana.	38
3.4.11. Tipos de emergencias atendidas por día.	39
3.4.12. Tiempo entre atención de llamadas días de semana.	40
3.4.13. Llamadas atendidas por minuto	41
3.4.14. Frecuencias vs. Duración de llamadas.	41
3.4.15. Promedio de operadores por minuto en días hábiles.	43
3.4.16. Promedio de operadores por minuto en fin de semana.	43
3.4.17. Distribución de llamadas entrantes, atendidas y cantidad de operadores un día de semana.	44
3.4.18. Porcentaje de llamadas abandonadas versus nivel de ocupación.	46
4.1.1. Modelo en Arena para simulación del sistema M/M/c y M/M/c/K.	48
4.4.1. Modelo en Arena para simulación con dos tipos de llamadas.	53
A.1.1. Etapas de un proceso KDD	66
A.2.1. Fases de el modelo CRISP-DM para Data Mining	67
B.1.1. Salida Stat Fit. Ajuste de tiempos de servicio.	69
B.2.1. Distribución de tiempos de espera en Cola en días de semana.	71

B.2.2	Distribución de tiempos de espera en Cola en días de semana. . . . .	71
C.1.1	Captura de pantalla. Resultados Erlang-X en Horario 1. . . . .	73
C.2.1	Captura de pantalla. Resultados Erlang-X en Horario 2. . . . .	74
D.1.1	Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c en horario 1. . . . .	75
D.1.2	Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c en horario 1. . . . .	75
D.1.3	Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c en horario 1. . . . .	75
D.1.4	Tiempo en el sistema ( $W$ ). Salida simulación M/M/c en horario 1. . . . .	75
D.2.1	Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c/K en horario 2. . . . .	75
D.2.2	Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c/K en horario 2. . . . .	76
D.2.3	Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c/K en horario 2. . . . .	76
D.2.4	Tiempo en el sistema ( $W$ ). Salida simulación M/M/c/K en horario 2. . . . .	76
D.3.1	Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c/K . . . . .	76
D.3.2	Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c/K. . . . .	76
D.3.3	Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c/K. . . . .	76
D.3.4	Tiempo en el sistema ( $W$ ). Salida simulación M/M/c/K. . . . .	76
D.3.5	Contador de llamadas M/M/c/K . . . . .	77
D.4.1	Largo promedio de cola ( $L_q$ ). Salida simulación con Abandonos . . . . .	77
D.4.2	Tiempo de espera en cola ( $W_q$ ). Salida simulación con Abandonos. . . . .	77
D.4.3	Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación con Abandonos. . . . .	77
D.4.4	Tiempo en el sistema ( $W$ ). Salida simulación con Abandonos. . . . .	77
D.4.5	Contador de llamadas. Simulación con Abandonos y Capacidad Finita. . . . .	77
F.1.1	Captura de Pantalla. Descripción del Fono 133 . . . . .	80
F.2.1	Panel de recepción de llamada en Power 911 . . . . .	81
F.2.2	Ingreso de datos de la llamada en Power 911 . . . . .	81

# Índice de tablas

2.3.1. Notación BPMN . . . . .	9
3.1.1. Horarios de Turnos de los Operadores . . . . .	28
3.2.1. Clasificación de llamadas inoficiosas . . . . .	29
3.2.2. Ejemplo Tipificación de Emergencias . . . . .	29
3.4.1. Cantidad de llamadas atendidas por día. . . . .	35
3.4.2. Resumen Muestreo. . . . .	39
3.4.3. Tiempos de servicio promedio según tipo de llamada. . . . .	42
3.4.4. Ejemplo: Tiempo promedio de servicio de llamadas de emergencia. . . . .	42
3.4.5. Tramos horarios diferenciados en días de semana. . . . .	44
3.4.6. Tasas Generales de Llegada y de Servicio. . . . .	45
3.4.7. Tasas de Llegada y Servicio de llamadas de Emergencia. . . . .	45
3.4.8. Tasas de Llegada y Servicio de llamadas Inoficiosas. . . . .	45
3.4.9. Porcentaje de llamadas abandonadas en días hábiles por horario. . . . .	46
4.1.1. Resultados de Aplicación del Modelo M/M/c en Horario 1. . . . .	48
4.1.2. Resultados Modelo M/M/c/K en Horario 2. . . . .	49
4.1.3. Resultados Simulación con Capacidad Finita en Horario 2. . . . .	50
4.2.1. Tiempos de Paciencia Estimados usando Teoría de Colas . . . . .	50
4.2.2. Porcentaje de abandonos por horario. . . . .	51
4.2.3. Porcentaje de abandonos por horario. Resultado de la simulación con menos operadores en horarios 1 y 3. . . . .	52
4.3.1. Resultados Simulación vs. Teoría de colas para el modelo M/M/c en horario 1. . . . .	52
4.3.2. Resultados Simulación vs. Teoría de colas para el modelo M/M/c/K en horario 2. . . . .	52
4.3.3. Resultados Simulación con Capacidad Finita. . . . .	53
4.3.4. Medidas de rendimiento obtenidas por simulación, en cada horario por separado. . . . .	53
4.3.5. Medidas de rendimiento obtenidas por simulación, considerando los estados transientes. . . . .	53
4.4.1. Porcentaje de abandonos por horario simulando tiempos de servicio lognormales. . . . .	54
4.4.2. Medidas de rendimiento simulando tiempos de servicio lognormales. . . . .	54
4.4.3. Parámetros de servicio para cada tipo de llamada simulando tiempos de servicio lognormales. . . . .	54
4.4.4. Porcentaje de abandonos por horario asumiendo variabilidad en el número de operadores. . . . .	55
4.4.5. Medidas de rendimiento con tiempos de servicio exponencial asumiendo variabilidad en el número de operadores. . . . .	55
4.4.6. Parámetros de servicio para cada tipo de llamada asumiendo variabilidad en el número de operadores. . . . .	55
4.4.7. Porcentaje de abandonos por horario asumiendo variabilidad en distribución de llegadas. . . . .	55
4.4.8. Medidas de rendimiento con tiempos de servicio exponencial asumiendo variabilidad en distribución de llegadas. . . . .	55
4.4.9. Parámetros de servicio para cada tipo de llamada asumiendo variabilidad en distribución de llegadas. . . . .	55
4.5.1. Nivel de Servicio Actual para Llamadas Atendidas. Días Hábiles. . . . .	56
4.5.2. Nivel de Servicio en llamadas de Emergencia. Días Hábiles . . . . .	57
4.5.3. Nivel de Servicio Actual. Fin de Semana. . . . .	57
4.5.4. Nivel de Servicio en llamadas de Emergencia. Fin de Semana. . . . .	57

A.3.1 Llamadas entrantes atendidas y abandonadas por día . . . . .	68
B.1.1 Ejemplo. Cálculo de tiempo de servicio ponderado. . . . .	70
B.2.1 Promedio Tiempo de Espera en Cola. . . . .	71
B.2.2 Tiempo de espera en cola por horario. . . . .	72
E.1.1 Nivel de servicio para todas las llamadas. Días hábiles. . . . .	78
E.2.1 Nivel de servicio para todas las llamadas. Fin de Semana . . . . .	79
F.3.1 Tipos de Emergencias. . . . .	82

Parte I  
Contenido

# Capítulo 1

## Antecedentes Generales

### 1.1. Introducción

El Fono de Emergencias 133 es un servicio que atiende llamadas de todo el país. Una llamada al número 133 hecha desde teléfono fijo, rural, público o teléfono móvil será atendida por operadores preparados en la zona geográfica más cercana, derivándose un proceso acorde a las necesidades de la persona que llama.

Este servicio telefónico de emergencias está a cargo de Carabineros de Chile, y se administra en la Central de Comunicaciones (CENCO), la cual es considerada, por la misma institución, como el motor de sus operaciones [6]. Su principales actividades son administrar y coordinar las telecomunicaciones que permiten realizar procedimientos de emergencia, seguridad y operativos policiales de todo tipo, día a día y en todo Chile. Este estudio se acota a la CENCO Metropolitana, pero cabe destacar que existen 29 sub-centrales de CENCO en el país, estando entre ellas interconectadas, listas para atender las necesidades de la ciudadanía.

En la Región Metropolitana, el Nivel 133 se encuentra dentro de CENCO. Su misión es atender llamadas de emergencia a través de un proceso complejo, compuesto por decisiones e interacciones que son expuestas en este trabajo. Después de la atención de una llamada de emergencia en el Nivel 133, en la unidad Alpha I se administra el operativo policial requerido (ver figura 3) y luego, la situación de la persona afligida es resuelta en terreno.

En la actualidad Carabineros necesita de manera urgente medir los niveles de servicio que se entregan a la ciudadanía a través del Fono 133, debido a que hay un consenso en la institución acerca de que las llamadas inoficiosas<sup>1</sup> son elevadas y que existe un alto nivel de congestión en las líneas<sup>2</sup>. Esta misma percepción es reflejada con frecuencia en la prensa. Una publicación oficial de Carabineros señala que en el primer semestre del 2011, más de un 70 % de las llamadas contestadas fueron inoficiosas [6]. Ya que la cantidad de este tipo de llamadas es alta, es indispensable incluirlas en los análisis.

Entonces el problema que se debe resolver con este trabajo es la realización de un diagnóstico del proceso de servicio, estimando la demanda real del sistema y otras variables que influyen en el nivel de servicio. Típicamente los *call centers* se estudian como sistemas de servicio con colas de espera, para los cuales se mide el tiempo de espera en cola y tiempo total en el sistema de un cliente, con esto se fijan metas de servicio para cumplir con las expectativas de los clientes.

En los servicios de emergencia estas metas deben estar muy claras. Por ello aquí se examina, lo más minuciosamente posible, el proceso completo de atención de llamadas; se presenta el estado actual del sistema; luego se miden y evalúan los niveles de servicio; y finalmente se describen propuestas de metas de servicio y acciones para alcanzarlas.

---

<sup>1</sup>Llamadas inoficiosas son las que generan mal uso del servicio

<sup>2</sup>La descripción del Fono 133, en el sitio web de Carabineros comienza diciendo “Cuando se disca el Fono 133 y existe un alto tráfico o congestión de llamadas...”, luego indican que un 78,21 % de las llamadas fueron inoficiosas en 2011.

## 1.2. Descripción del Problema

En todos los servicios de emergencias, se reciben “falsas alarmas”. Cuando éstas constituyen una gran proporción del total de requerimientos, como es el caso del 133, es necesario y urgente tomar algún tipo de medida. Actualmente Carabineros no conoce la magnitud exacta de este tipo de llamadas, ni tampoco la congestión del sistema. La percepción en la institución y en la ciudadanía es que el problema de congestión de las líneas es grave.

El fundamento de este estudio es la preocupación de Carabineros de Chile por el nivel de servicio que brinda el 133. Se reporta una gran cantidad de denuncias en la prensa y de reclamos por parte de personas que dicen que llaman y no les contestan. Además, se sabe que más del 70% de las llamadas atendidas son inoficiosas [6], basta con observar un día cualquiera el trabajo de los operadores. Es preocupante ver que la descripción del Fono 133, en el sitio web institucional de carabineros, dice: “Cuando se disca el Fono 133 y existe un alto tráfico o congestión de llamadas, se escuchará una guía de voz que indica que el exceso de llamados no permite atenderle en ese momento...”, luego indican que un 78,21% de las llamadas fueron inoficiosas en 2011. Esto sigue ocurriendo, a pesar de que muchas campañas comunicacionales han sido lanzadas para disminuir este tipo de llamadas. Lamentablemente las campañas en contra de las llamadas inoficiosas son contraproducentes, éste tipo de llamadas aumentan mientras hay campañas publicitarias en curso.

Para medir los niveles de servicio actuales es necesario comprender cómo afectan las llamadas inoficiosas, cuantificar la demanda total del sistema y determinar cuantitativamente la demanda no atendida.

¿Por qué hasta ahora no se han medido los niveles de servicio? En los sistemas que manejan las llamadas, hay mucha información censurada o inexistente, otra que no es fácil de acceder, y además hay datos que simplemente no se registran. Es el caso de las llamadas que el sistema rechaza, son las que reciben tono ocupado cuando el sistema está congestionado. No hay ningún tipo de registro que cuantifique éstas llamadas, ni menos, detalles sobre la hora en que esto ocurre, quién realiza las llamadas o si éstas personas vuelven a llamar.

Actualmente no hay datos que permitan medir directamente el tiempo de espera de los usuarios. Sólo existen registros de las llamadas que son atendidas, se conoce la fecha y hora de atención, pero no cuánto tiempo esperaron antes de ser atendidas. Cuando todos los operadores están ocupados, las personas reciben un mensaje que les pide esperar en línea, formándose una cola virtual, pero puede que las personas agoten su paciencia y abandonen dicha cola. Se conoce el volumen bruto de las llamadas que abandonaron. No se tiene registro detallado de aquellas llamadas ni de las que fueron rechazadas, lo que ocurre cuando se supera la capacidad de la cola de espera.

Está claro, entonces, que las llamadas atendidas, de las cuáles sí hay registros detallados, no corresponden a la demanda total del sistema. Luego, no es posible plantear metas de servicio, ya que ni siquiera se ha podido medir la calidad de servicio actual; ¿Cuántas personas reciben tono ocupado y no logran reportar su emergencia? ¿Cuántas personas abandonan la cola de espera? ¿Cuál es el tiempo de espera promedio antes de ser atendido? ¿Si hubiese más operadores contestando, se atenderían más emergencias?

## 1.3. Justificación del Proyecto

Para medir la calidad del servicio, es indispensable la comprensión del funcionamiento actual del sistema a través de una descripción detallada del proceso. Además de una investigación de campo por medio de observaciones y entrevistas, requiere una investigación práctica mediante la aplicación de metodologías adecuadas que permitan identificar aspectos fundamentales para cumplir el objetivo planteado.

Es necesario utilizar herramientas, por ejemplo, para presentar el proceso de flujo de llamadas, estimar la demanda total del sistema y los tiempos de espera, e identificar los patrones de comportamiento de todas variables del sistema. Algunos métodos estadísticos descriptivos, modelos<sup>3</sup> analíticos y modelos

<sup>3</sup>Un modelo es una abstracción simplificada de un sistema, que permite identificar las características importantes de

de simulación deben ser aplicados para resolver este tipo de problemas, generalmente en el ámbito de la ingeniería de operaciones.

El procesamiento preliminar de los datos, es en sí complejo, ya que se cuenta con diferentes bases de datos, en distintos formatos, existen algunos inconvenientes para su extracción e incongruencias entre algunos reportes e incluso entre los propios atributos de las llamadas. Luego, se pueden realizar las descripciones cuantitativas para caracterizar el sistema, hacer un análisis cuantitativo y luego una interpretación y evaluación de los resultados.

Los resultados de este estudio pueden ser cruciales para abrir líneas de investigación en pos de proponer mejoras, basadas en argumentos cuantitativos concretos, para la seguridad de la ciudadanía y de la percepción de ella, que podría resultar en mejoras en calidad de vida. Para el Estado, la seguridad ciudadana es una prioridad. En la Estrategia Nacional de Seguridad Pública (ENSP), se plantea la necesidad de definir metas de gestión de recursos policiales, entre otros, y esto exige que Carabineros recopile información para realizar análisis fundamentales como éste, que faciliten la toma de decisiones y justifiquen el uso de recursos.

Carabineros de Chile debe planificar la destinación de recursos para mejorar la gestión del Nivel 133. Por ello se desea utilizar los resultados de este estudio como argumentos para proponer leyes en contra de las llamadas inoficiosas, entre otras medidas de gestión que se puedan tomar para mejorar el servicio.

## 1.4. Objetivos

### 1.4.1. Objetivo General

Realizar un diagnóstico del servicio Fono Emergencias 133 de Carabineros a través de métricas que permitan evaluar los niveles de servicio con el fin orientar la toma de decisiones operacionales, tácticas y estratégicas en en la Central de Comunicaciones de Carabineros de Chile.

### 1.4.2. Objetivos Específicos

- Presentar el proceso actual desde la llegada de una llamada hasta el despacho del requerimiento policial en CENCO, enfocándose en el servicio del *call center*, Nivel 133.
- Modelar el proceso de atención de llamadas considerando el registro de los atributos como parte del proceso.
- Caracterizar los tipos de llamadas, los patrones de llegada y servicio de cada una de ellas.
- Estimar el volumen de llamadas que no logran entrar al sistema.
- Medir el rendimiento del proceso actual, comparar con medidas de estándar internacional y evaluar el desempeño actual del servicio.
- Identificar cuellos de botella y otros problemas en el proceso.
- Proponer iniciativas justificadas cuantitativamente para mejorar el servicio.

---

un proceso y hacer predicciones más concretas que algunos experimentos o mediciones.

# Capítulo 2

## Marco Teórico

### 2.1. Fundamentación

En Chile, por primera vez se está haciendo un análisis del nivel de servicio entregado por el Fono de Emergencias Policiales. Mientras en otras regiones del mundo, esto se ha estudiado al menos desde 1970 . El servicio 133 atiende en promedio cerca de 18.000 llamadas diariamente (detalle en sección 3.4.1), siendo el primer canal de comunicación entre la ciudadanía y Carabineros de Chile.

Antunes y Scott [1] afirman que las llamadas a los servicios policiales influyen directamente en las demandas de la ciudadanía al gobierno. Por lo tanto, es una prioridad del gobierno conocer las operaciones de este servicio.

En general, se estudia la eficiencia de los operadores y los niveles de servicio. Los operadores tienen la responsabilidad de ser los únicos que reciben, procesan y transmiten la información de la emergencia reportada por un ciudadano. En gran parte, de ellos depende el tiempo de atención de una llamada. Por otro lado, los niveles de servicio se miden. Una manera de entregar un buen servicio de asistencia y percepción de seguridad a la ciudadanía, es que se sepa que la policía entrega una respuesta oportuna a las emergencias que se reportan en cada momento.

Por estos motivos, el presente estudio parte con un análisis de los niveles de servicio, proponiendo estudios posteriores que permitan mejorar la respuesta en el Fono 133 y mejorar la percepción que tiene la ciudadanía de éste. Es importante conocer cómo se maneja en el mundo este tipo de servicios, a continuación se entrega una breve reseña de esto.

#### 2.1.1. Importancia de los Call center de Emergencia

Antunes y Scott [1] llaman a los operadores “burócratas en terreno”. Son ellos quienes atienden, categorizan y determinan la respuesta policial que recibe el problema reportado en cada llamada. Los mismos autores señalan que la mayoría de las llamadas corresponden a solicitudes de información o asistencia, reporte de molestias, problemas de tráfico vehicular o solicitud de intervención en discusiones interpersonales. Solo el 20 % de las llamadas tienen que ver con crimen en curso, de cualquier tipo. Además, se sostiene que la respuesta típica de un operador es la promesa de despachar una patrulla o recursos policiales al lugar.

En la Unión Europea, el número 112 es discado gratuitamente por los ciudadanos para recurrir a los servicios de emergencia [11] . Yarpuzlu et al. [11] explican que la Comisión Europea monitorea el funcionamiento del 122 en los estados miembros. Existe un Comité de Comunicaciones que recopila información de de las autoridades representantes de cada estado, responsables de las comunicaciones electrónicas. La función del Comité es fomentar y asegurar un funcionamiento satisfactorio del servicio 112 en la Unión Europea. Regularmente se intercambian visiones entre las autoridades de los estados miembros para definir mejores prácticas a implementar en este servicio de emergencias.

En 2005 se creó el Grupo de Expertos en Acceso de Emergencias para trabajar en todos los problemas prácticos que enfrenta cada región miembro, con la misión de proveer un servicio eficiente y efectivo a la ciudadanía. Además, utilizan la tecnología *Global Systems for Mobile* (GSM) para redirigir las llamadas discadas al 112, al servicio de emergencia más cercano dentro de 61 países identificando la localización aproximada de la llamada. Es decir, si se marca el 112 en Norte América, la llamada se redirige al 911, si es en Australia, se redirige al 000.

En el caso de Estados Unidos, existe The Industry Council for Emergency Response Technologies (Consejo Industrial de Tecnologías de Respuesta para Emergencias), iCERT. Esta asociación representa la voz del sector comercial en el campo de las comunicaciones de emergencia. Se estableció el año 2005 originalmente como 9-1-1 *Industry Alliance*. iCERT, juega un importante rol asistiendo a los políticos de gobierno que proponen regulaciones en seguridad y emergencia<sup>1</sup>.

### 2.1.2. Estándares de Servicio

El año 2010, el estándar de servicio del 911, en Estados Unidos era responder el 90 % de las llamadas antes de 10 segundos. El periódico *Los Angeles Times*, afirma que en California, un 94 % de las llamadas se contestan antes de 10 segundos, en el año 2012<sup>2</sup>. Esto significa, que el 911 mide el tiempo de espera de los llamantes, y tiene noción de la cantidad de personas que esperan en cola y cuánto es el tiempo que esperan antes de ser atendidos. En los servicios de emergencia como éste, los segundos marcan una gran diferencia.

iCERT, en 2011 publica un estudio que indica que el 15 % de los llamados, en New Orleans, no son contestados debido a la gran sobrecarga del 911. 1 de cada 5 llamadas son puestas en espera, y no contestadas inmediatamente. Además, 1 de cada 4 llamadas abandonan la cola de espera en California, debido a los largos tiempos de espera<sup>3</sup>. Para hacer frente a estos problemas, ellos recomiendan, entre otras medidas, tener los callcenters con una capacidad de operadores óptima.

Un punto a destacar, es que las centrales del 911 permiten registrar todas las llamadas que discan el 911, aunque éstas abandonen la cola mientras esperan atención. La información del llamante se presenta como una llamada abandonada y se despliega como tan en una lista de llamadas abandonadas. Por ley, se devuelve la llamada para determinar si el llamante necesita asistencia.

80 % de las llamadas provienen de celulares, esto es un problema cuando muchas personas llaman desde una misma localización para reportar el mismo incidente. Por ejemplo, un accidente en determinada carretera puede colapsar las líneas de emergencia, muchas llamadas provienen de la misma localización<sup>4</sup>.

Las llamadas inoficiosas, llamadas que no son de emergencia, llegan a 3.600 por mes, lo que corresponde aproximadamente a un 1,5 % de las llamadas recibidas. Para el 911 esto representa una gran sobre carga y consumo de recursos extra, por lo tanto varias organizaciones se dedican a proponer y planificar soluciones (iCERT, NENA, 911 Institute, entre otros) <sup>5</sup>.

---

<sup>1</sup>Fuente: <http://www.theindustrycouncil.org/about/index.cfm> Consultado en Noviembre 2011.

<sup>2</sup>Fuente: <http://articles.latimes.com/2010/nov/28/local/la-me-911-calls-20101129> Consultado en Noviembre 2012.

<sup>3</sup>Fuente: <http://www.theindustrycouncil.org/publications/9-1-1overloadfactsheet.pdf> Consultado en Noviembre 2012.

<sup>4</sup>Fuente: <http://www.radioresourcemag.com/onlyonline.cfm?OnlyOnlineID=265> Consultado en Noviembre 2012.

<sup>5</sup>Fuente: <http://911calloverload.com/> Consultado en Noviembre 2012.

## 2.2. Descripción del Proceso Metodológico

Se utiliza una metodología adaptada de un proceso CRISP-DM, desarrollado por IBM, basado en *Knowledge Discovery in Databases* (KDD)<sup>6</sup>. La naturaleza de este trabajo requiere pasar por algunas etapas que no están presentes en el proceso KDD, ni en CRISP-DM, por eso se describe una nueva metodología, que es cíclica y se describe a continuación.

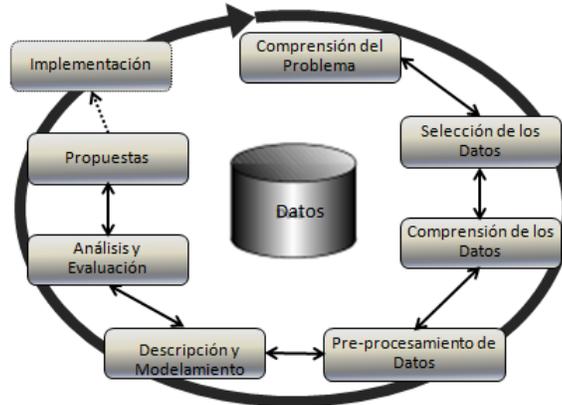


Figura 2.2.1: Metodología utilizada (basada en CRISP-DM).

La metodología CRISP-DM se puede considerar similar a la propuesta, en cuanto a que es aplicable transversalmente en cualquier tipo de industria [17]. Pero el modelo propuesto para este trabajo va más allá del solo análisis y procesamiento de datos, requiere también una comprensión del sistema completo; requiere observaciones, mediciones y entrevistas, como parte no solo de la etapa de comprensión del problema de negocio, sino también de las siguientes etapas de comprensión y procesamiento de datos. Por otro lado, el proceso a utilizar no termina con un desarrollo, pues esta etapa está fuera del alcance del proyecto, pero sí propone seguir el proceso de mejora continua para evaluar periódicamente los cambios. Las etapas que se llevarán a cabo se basan en las descripciones de Herirte y Hipo [17] además de las consideraciones especiales que se requieren en este trabajo.

**Comprensión del problema** En este caso, la comprensión del problema pasa primero por una subetapa de estudio de la bibliografía acerca de servicios de emergencia y *call centers*. Luego se participa en reuniones con el alto mando de distintas unidades de la institución y con el personal. Esto tiene que ver con la investigación documental y de campo que se requiere para comprender el problema de negocio a abordar.

**Selección de los Datos** Al tener una comprensión básica del problema de negocio, se puede definir qué tipo de datos se necesitan para trabajar y sus características, eligiéndose las variables más influyentes en el problema. Se debe seleccionar la unidad básica de análisis, sus atributos y el periodo de tiempo a estudiar según los objetivos. En este caso, existen datos censurados (o no disponibles), pero fundamentales para el análisis, entonces se hace un estudio de campo y un estudio estadístico para luego procesar esto y poder llegar a la etapa de descripción.

**Comprensión de los Datos** Es una fase que consiste en verificar la validez y el significado de los datos que inicialmente se consiguen. Se procede a realizar acciones que permitan familiarizarse con los datos, identificar problemas en la calidad de los datos, tener un acercamiento a la información que se podría adquirir, o detectar subconjuntos de las hipótesis que se plantearon en la primera fase.

**Pre-procesamiento de Datos** Consiste en todas las actividades de preparación de los datos para construir el set de datos final. Mediante el pre-procesamiento, se limpian los datos eliminando valores incorrectos, no válidos o desconocidos. Entre las actividades, se seleccionan los atributos, se limpian los

<sup>6</sup> *Knowledge Discovery in Databases* (KDD) es un proceso lineal que consiste en cinco etapas consecutivas, éstas son: selección, pre-procesamiento, transformación, aplicación de técnicas de minería de datos e interpretación.

datos, se construyen nuevos atributos, se seleccionan muestras, se transforman los formatos de las bases de datos y calculan nuevos atributos según los requerimientos del análisis. Estas actividades pueden ser realizadas múltiples veces y no en un orden estricto.

**Descripción y Modelamiento** En esta etapa se lleva a cabo un proceso de descubrimiento de relaciones, patrones y tendencias, usando varias técnicas y herramientas descriptivas y de modelamiento. Para hacer esta búsqueda se utilizan métodos de reconocimiento de patrones, a través de técnicas matemáticas y estadísticas. Se utilizan en mayor medida, modelos analíticos y modelos de simulación.

En esta etapa es frecuente volver a revisar la anterior para generar nuevos set de datos para poder observar en detalle algunos elementos o nutrirse de elementos que se van descubriendo.

**Análisis y Evaluación** Antes de interpretar y evaluar los resultados que arrojan los modelos, es necesario validarlo para comprobar que las conclusiones que arroja son robustas y cumplen con los objetivos. Al obtener resultados de varios modelos mediante el uso de distintas técnicas, se deben comparar los modelos e identificar el que mejor se ajuste al problema. Si ninguno de los modelos falla o no considera algún aspecto importante del problema, se debe regresar a alguna de las etapas anteriores para generar nuevos resultados.

Los modelos por sí solos, entregan sólo información. Un analista debe interpretar los resultados para transformar esa información en conocimiento y proceder a evaluar el rendimiento del proceso actual de negocio.

**Propuestas** A partir del análisis y evaluación se proponen cambios y se ponen a prueba. Este proceso puede repetirse varias veces y volver a cualquiera de las etapas anteriores para obtener resultados distintos garantizar mejoras. Luego de finalizar esta etapa, el proceso queda listo para pasar a la etapa de desarrollo de una solución e implementación de esta.

## 2.3. Herramientas Utilizadas

Se usan varios métodos de descriptivos, de análisis y evaluación como herramientas para cumplir cada una de las etapas de la metodología presentada. Las etapas Pre-procesamiento de Datos y Descripción son bastante complejas y requieren utilizar muchas herramientas en conjunto. Además no son especialmente atractivas para el lector, por eso no se describen en el cuerpo de este informe. A continuación se explican las herramientas que son menos intuitivas, entre varias que se utilizaron.

### 2.3.1. Modelamiento de Procesos

El modelamiento del proceso consiste en hacer un diagrama esquemático, que muestra las interacciones entre las etapas del procesos y los flujos de información, después de haber realizado el levantamiento de procesos en todos los niveles de la organización. Como una manera de comprender a fondo el problema de negocio y cumplir el primer objetivo del proyecto, se utilizará una herramienta basada en *Business Process Notation* (BPN). Se utilizará la metodología que sugiere Weske [15], la cual se compone de etapas y patrones precisos aplicables a todo tipo de procesos.

**Levantamiento de Procesos** El levantamiento de procesos es la primera actividad, cuya finalidad es presentar un Levantamiento de la Situación Actual, identificando entidades o actores que participan y describiendo sus actividades, tareas e interacciones en las distintas etapas que les permiten llevar a cabo el objetivo del proceso en que participan. Se deben identificar los *Inputs* y *Outputs* del proceso, las decisiones presentes y subprocesos.

Esta etapa implica investigación en los niveles estratégico, táctico y operacional del negocio. Como investigación de campo, se recopilan testimonios objetivos y subjetivos que permitan conocer los sistemas informáticos utilizados, la dotación de recursos humanos y las interacciones entre estos y otros actores o entidades, siempre teniendo en cuenta que se debe cuantificar la capacidad de las distintas etapas del proceso. Se realiza investigación documental a partir del estudio de manuales, reportes y publicaciones. Se realiza mediante la observación del funcionamiento normal del Nivel 133 en CENCO, se complementa

con información recopilada directamente en la empresa de telefonía que administra el sistema y con los datos de las llamadas procesados en durante este trabajo.

**Business Process Modeling Notation (BPMN)** Se utiliza la notación BPMN para representar las distintas actividades de un proceso. BPMN se basa en una técnica de diagrama de flujo, creando la representación de un proceso mediante un conjunto de elementos gráficos denominados "eventos". BPMN permite la estandarización de los procesos dentro de la organización y también que amplía el campo de acción para que estos puedan ser compartidos y entendidos entre los diferentes actores del de negocio, desde los agentes tomadores de decisiones hasta los analistas y los desarrolladores [16].

Al ser capaces de tener una amplia mirada de los procesos, se puede reconocer la complejidad del sistema en estudio y definir la unidad básica de análisis.

El software a utilizar es Bizagi Process Modeler 2.3.0.5. Los eventos que describen los proceso, en este estudio, se describen a continuación.

Evento	Descripción	Evento	Descripción
	Es el evento de inicio del flujo. Indica qué es lo que inicia el flujo del proceso.		Es el evento que indica que el inicio del flujo lo provoca un mensaje externo que llega en cualquier momento.
	Es un evento de decisión. Representa una pregunta en cierto punto del flujo que permite al actor decidir por dónde continuar.		Es un evento temporizador. Representa un momento en que el actor debe esperar a que algo suceda. Determina lo que luego suceda en el proceso.
	Es un evento que representa a una tarea que realiza un actor en el proceso y se requiere para continuar el flujo.		Es un evento que indica que un subproceso ocurre dentro del proceso principal. El <i>output</i> del subproceso continúa el proceso principal.
	Es un evento de fin. Marca el final del proceso.		Es un evento de fin que incluye un mensaje. Representa el final de un proceso que finaliza con un mensaje o envío de una señal para ser utilizada en otro proceso.
	Es un evento de fin que termina el proceso definitivamente.		Es un <i>pool</i> , lo que hay dentro, son todas las actividades que realiza un participante o entidad en el proceso.

Tabla 2.3.1: Notación BPMN

### 2.3.2. Teoría de Colas

En investigación de operaciones, la teoría de colas permite modelar analíticamente los sistemas con colas de espera y obtener medidas de su rendimiento. Los resultados pueden ayudar a minimizar inconvenientes y maximizar la eficiencia de los recursos limitados en negocios de servicio, procesos industriales, de transporte y telecomunicaciones, entre otros.

Un sistema de colas consiste en clientes<sup>7</sup> que llegan a un servicio, esperan si el servicio no es inmediato, y abandonan el sistema. El comportamiento del sistema depende de varias características de los clientes, los servidores y del *layout* del proceso. Por ejemplo, las colas de espera se forman esperando la luz verde del semáforo, esperando ser atendido en el banco o esperando en línea para ser atendido por teléfono.

A principios del siglo IX, Agner Karup Erlang desarrolló teorías estadísticas con aplicaciones en centrales telefónicas, mientras trabajaba en la Central Telefónica de Copenhague. En su investigación,

<sup>7</sup> El término "cliente" se usa para referirse a cualquier tipo de usuario del sistema de colas, no necesariamente humano. Por ejemplo, puede ser una piedra esperando una máquina pulidora, un avión esperando el aviso de despegue, entre otros.

mostró que se puede esperar que la llegada de las llamadas siguen una distribución de Poisson en un intervalo de tiempo. A partir de este supuesto, demostró expresiones para calcular el promedio de tiempo de espera en cola y otras medidas básicas de rendimiento de los sistemas de cola [7].

Entre las contribuciones más importantes de Erlang, se destaca el reconocimiento de los procesos estocásticos como origen del fenómeno de congestión en los sistemas telefónicos. Si las llamadas llegan a intervalos regulares y tomaran el mismo tiempo en ser atendidas, entonces, sería trivial determinar la demanda y ajustar la capacidad del sistema para anular la espera y la pérdida de llamadas.

### 2.3.2.1. Procesos de Markov

El comportamiento estocástico de un sistema, muchas veces, se puede representar matemáticamente describiendo los escenarios o estados en que se puede encontrar y cómo cambia de un estado a otro.

Muchas veces el comportamiento de dichos sistemas se puede representar a través de un Proceso de Markov con el fin de poder conocer en qué estado se encuentra el sistema en un dado momento  $t$ .

Los Procesos de Markov son procesos estocásticos compuesto por una familia de variables aleatorias  $\{X(t), t \in S\}$ . El espacio de estados  $S$ , es el conjunto de estados posibles del sistema, puede ser continuo o discreto y el proceso puede ser continuo o discreto, respectivamente.

Cuando  $S$  es un espacio discreto, el proceso se puede representar como una Cadena de Markov. En esta ocasión nos enfocaremos solamente en las Cadenas de Markov, ya sea de tiempo continuo o discreto.

**Proceso de nacimiento y muerte generalizado** Un sistema en donde llegan clientes a un servicio, es un ejemplo de un proceso discreto, en el cual el espacio de estados es el número de clientes. Cuando hay  $n$  clientes, el proceso está en el  $n$ -ésimo estado. En ese momento, los clientes llegan a tasa  $\lambda_n$ , y son atendidos a tasa  $\mu_n$ . Este es un tipo de proceso de Markov llamado de nacimiento y muerte, porque se comporta como el proceso de crecimiento de una población.

En un proceso de nacimiento y muerte, se dice que ocurren transiciones sólo entre los vecinos más cercanos. Es imposible que haya una muerte cuando la población es  $n = 0$  (i.e.  $\mu_0 = 0$ ), pero sí puede haber un nacimiento cuando la población es cero (i.e.  $\lambda_0 = 0$ ).

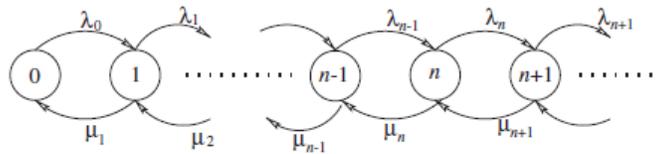


Figura 2.3.1: Transiciones entre estados de un Proceso de Nacimiento y Muerte Generalizado

Las transiciones entre estados ocurren instantáneamente y sólo depende del estado actual, no de la historia de transiciones. Esta es la propiedad de pérdida de memoria que cumplen los procesos de Markov.

Sea  $p_n(t)$  la probabilidad de que el sistema se encuentre en el estado  $n$ , en el tiempo  $t$ . Se define  $p_n$  como la probabilidad que el sistema contenga  $n$  clientes después de haber alcanzado el estado estacionario, cuando no importa en qué estado haya comenzado el proceso. Si la probabilidad conjunta de las transiciones es invariante en el tiempo, entonces el proceso es estacionario.

Cuando la tasa de nacimiento es  $\lambda_n$  y la tasa de muerte es  $\mu_n$ , la probabilidad de que el sistema llegue al estado  $n$  en tiempo  $t$  es

$$\lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t)$$

mientras la tasa de salida del estado  $n$  es

$$(\lambda_n + \mu_n)p_n(t)$$

La probabilidad efectiva del flujo al estado  $n$  es la resta de ambas ecuaciones, i.e.,

$$\frac{dp_n(t)}{dt} = \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t) - (\lambda_n + \mu_n)p_n(t), \quad n \geq 1$$

y

$$\frac{dp_0(t)}{dt} = \mu_1 p_1(t) - \lambda_0 p_0(t)$$

**Estado estacionario** Asumiendo que existe el estado estacionario, entonces

$$\frac{dp_n(t)}{dt} = 0$$

la probabilidad efectiva resulta

$$0 = \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1} - (\lambda_n + \mu_n) p_n, \quad n \geq 1$$

$$0 = -\lambda_0 p_0 + \mu_1 p_1 \implies p_1 = \frac{\lambda_0}{\mu_1} p_0$$

donde  $p_n$  se define como la probabilidad de que el sistema se encuentre en estado  $n$ , después de que la influencia del estado inicial ha sido olvidada. Esto se cumple cuando el siguiente límite existe

$$p_n = \lim_{t \rightarrow \infty} p_n(t)$$

Para encontrar los valores del estado estacionario, se resuelven las ecuaciones de balance global. Se demuestra, en el texto de Stewart [13], que la probabilidad de estar en el estado inicial es

$$p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \prod_{i=1}^n \rho_i}$$

Notar que el estado estacionario existe solamente si  $p_0 > 0$ , o si existe un entero  $k_0$  y una constante  $A < 1$  tal que

$$\frac{\lambda_0}{\mu_{k+1}} \leq A \leq 1 \quad \forall k \geq k_0$$

Estas son las condiciones que se deben cumplir para que el sistema se vaya vaciando y la población no crezca sin límites.

### 2.3.2.2. Ley de Little

Sea  $R$ , el tiempo de respuesta, la variable aleatoria que describe el tiempo que pasa cada cliente en el sistema.

La ley de Little establece que el promedio de clientes,  $N$ , en un sistema de colas en estado estacionario es igual al producto de la tasa de llegadas,  $\lambda$ , y el tiempo promedio en el sistema,  $W$  [5].

$$E[N] = \lambda E[R]$$

$$L = \lambda W \tag{2.3.1}$$

La ecuación 2.3.1 se conoce como la Ley de Little.

### 2.3.2.3. Medidas de Rendimiento

El propósito de analizar las colas es obtener medidas de las propiedades de un sistema. Por ejemplo, nos gustaría conocer el número de clientes en el sistema, el tiempo de espera de un cliente, la duración de un periodo en que todos los agentes están ocupados (*busy period*) o en que alguno está desocupado (*idle period*) [13], todas estas son variables aleatorias. La teoría de colas nos permite calcular los primeros momentos (media, varianza, etc.), estimando el comportamiento del sistema en estado estacionario.

Sin embargo, para este trabajo no es suficiente conocer los tiempos promedios de espera. Las siguientes medidas, que entrega la teoría de colas permitirán validar un modelo de simulación, que es finalmente lo que mejor puede representar el comportamiento real del sistema. Más que nada, interesa observar la distribución de los tiempos de espera.

**Cantidad de Clientes en el Sistema y Cantidad de Clientes en Cola** Sea  $N$  el número de clientes en estado estacionario y  $p_n$  la probabilidad de que en estado estacionario haya  $n$  clientes.

Entonces,  $p_n = \text{Prob}\{N = n\}$  y el número promedio de clientes en el sistema, en estado estacionario, es  $L = E[N] = \sum_{i=0}^{\infty} np_n$

Sea  $N_q$  el número de clientes esperando en cola, denotaremos  $L_q = E[N_q]$ .

**Tiempo en el Sistema y Tiempo en Cola** Cada cliente, llega al sistema y recibe el servicio en un tiempo de respuesta  $R$ . En promedio, el tiempo total que un cliente está en el sistema, desde su llegada a la cola hasta la salida del servicio, se denota  $W = E[R]$ . Parte de este tiempo es el tiempo de espera en cola, denotado  $W_q$ .

**Nivel de Utilización** La fracción de tiempo en que un operador esta ocupado, se llama nivel de utilización, se denota  $\rho$ . En una cola con un solo servidor ( $c = 1$ ), tasa de llegada  $\lambda$  y tasa de servicio  $\mu$ , la utilización se define como  $\rho = \lambda/\mu$ . Más adelante se verá que la utilización es distinta cuando hay más de un servidor y cuando algunos clientes son rechazados debido a que la capacidad del sistema es insuficiente (clientes perdidos).

### 2.3.2.4. Tipos de Colas

Según los académicos Gross y Harris [5], son seis los parámetros que permiten describir adecuadamente un sistema de colas. La notación de Kendall permite identificar los tipos de sistema de colas, usando un estándar bastante sencillo, en el que cada etapa de servicio del sistema se caracteriza por símbolos. Son cinco seis parámetros los que permiten reconocer un sistema de colas, A/B/C/X/Y/Z.

**Patrón de llegada (A)** La llegada de los clientes al sistema se puede medir como número de llegadas promedio en una unidad de tiempo o tiempo entre llegadas. Usualmente, el patrón de llegada es un proceso aleatorio que puede ser caracterizado por un proceso estocástico o distribución de probabilidad con el fin de modelar las llegadas. Es importante observar si este patrón es estacionario o si cambia en el tiempo (proceso no-estacionario).

Por otro lado cuando hay clientes impacientes<sup>8</sup>, ellos podrían decidir no entrar a la cola si no están satisfechos con el tiempo de espera que estiman, podrían abandonar después de haber esperado o podrían cambiarse a otra línea de espera, si hay más de una.

**Patrón de Servicio (B)** Los patrones de servicio se pueden describir como una tasa (número de clientes atendidos en una unidad de tiempo) o como un tiempo (tiempo requerido para atender a cada cliente). Puede haber uno o más servidores en sistema que otorguen un servicio individual, o en lotes (varios clientes atendidos simultáneamente por el mismo servidor).

A diferencia del patrón de llegadas, el patrón de servicio está asociado a una distribución condicional, dependiendo de si los servidores están ocupados o no. La tasa de servicio puede depender también del

---

<sup>8</sup>Los clientes impacientes son aquellos que no están dispuestos a esperar más de un cierto tiempo en cola. Después de cumplido su límite de paciencia, abandonan la cola.

número de clientes que estén esperando en cola. Si los servidores ven que la cola crece, podrían comenzar a atender más rápido o, por el contrario, podrían ponerse nerviosos y atender más lento, es decir, el estado del sistema puede afectar el tiempo de servicio.

**Número de Servidores (C)** Un sistema con más de un servidor, es un sistema multicanal. Cada servidor puede tener una cola o puede haber menos colas que servidores. Los servidores pueden ser idénticos o distintos, es decir, el patrón de servicio puede ser variar entre los servidores.

**Capacidad del sistema (X)** Si el sistema tiene un espacio limitado para mantener clientes en espera, esto se llama capacidad finita. Si la capacidad está completa, los nuevos clientes que arriban son rechazados por el sistema, sólo puede entrar un nuevo cliente después de completada una atención.

**Población total de clientes (Y)** Se refiere al tamaño total de la población de clientes.

**Disciplina de la cola (Z)** La manera en que los clientes reciben el servicio se llama disciplina de la cola. La manera más común de servir a los clientes es FIFO (*first in, first out*). En algunos casos, sobre todo cuando se trata de inventario no perecible, se sirve primero al último en llegar, LIFO (*last in, first out*). RSS (*random service selection*) es una disciplina de cola que selecciona aleatoriamente a un cliente. Existen varios tipos de sistemas que seleccionan a los clientes para atender primero a los que tienen mayor prioridad, entre otros.

Los tres primeros parámetros siempre deben conocerse. Algunas distribuciones posibles para las características A y B son M (Markov), E (Erlang), G (general) y D (determinística). La capacidad del sistema y población total se consideran de tamaño infinito por defecto. La disciplina de la cola siempre se considera FIFO, a menos que se indique lo contrario.

Por otro lado, cuando el cliente debe pasar por varias etapas consecutivas para completar el servicio, se trata de un sistema multietápico. También puede haber etapas en que se decida enviar al cliente a una etapa anterior, en teoría de colas esto se llama reflujo o reciclaje. Para analizar completamente este tipo de procesos, se necesita conocer las seis características de cada subproceso.

### 2.3.2.5. Colas M/M/1

Las fórmulas para analizar este tipo de colas provienen del estudio de cadenas de Markov nacimiento y muerte. Son las colas más sencillas, en las que los patrones de llegada y de servicio son Markovianos.

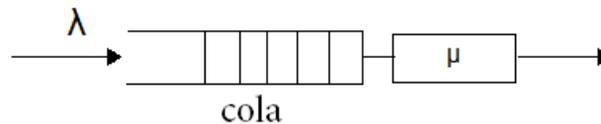


Figura 2.3.2: Esquema de una cola M/M/1

Usualmente se dice que las llegadas siguen un proceso de Poisson de tasa  $\lambda_n = \lambda$  y el tiempo de servicio sigue una distribución exponencial de tasa  $\mu_0 = \mu \forall n$ , hay un sólo servidor atendiendo, la disciplina de la cola es FIFO, la población y la capacidad son infinitas. Las probabilidades de transiciones entre estados se diagraman como una cadena de Markov, como muestra la figura

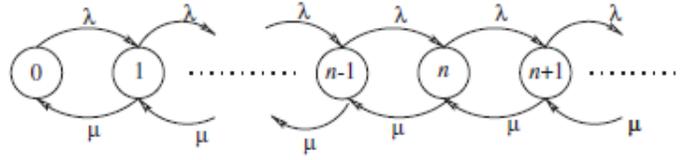


Figura 2.3.3: Transiciones entre estados de una cola M/M/1 [13]

Para encontrar las probabilidades de encontrar el sistema en estado  $n$  se resuelven las mismas ecuaciones del proceso generalizado, pero con la tasa de nacimiento  $\lambda$  y de muerte  $\mu$ . Con  $\rho_n = \rho = \lambda/\mu$ , se obtiene

$$p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n}$$

Recordar que la existencia de la solución estacionaria demanda que  $\rho = \lambda/\mu < 1$ , entonces la sumatoria  $\sum_{n=0}^{\infty} \rho^n$  es una serie geométrica que cumple  $|\rho| < 1$ . Luego,  $\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho}$ .

$$p_0 = 1 - \rho$$

Entonces, la solución en estado estacionario está dada por

$$p_n = \rho^n (1 - \rho) \text{ para } \rho = \lambda/\mu < 1$$

### MEDIDAS DE RENDIMIENTO

**CANTIDAD DE CLIENTES EN EL SISTEMA** Sea  $N$  la variable aleatoria que describe el número de clientes en el sistema en estado estacionario, y  $L = E[N]$ .

$$L = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n \rho^n (1 - \rho) = \rho (1 - \rho) \sum_{n=0}^{\infty} n \rho^{n-1}$$

$$L = \rho (1 - \rho) \frac{1}{(1 - \rho)^2} = \frac{\rho}{(1 - \rho)} = \frac{\lambda}{\mu - \lambda}$$

**CANTIDAD DE CLIENTES EN COLA** Sea  $N_q$  la variable aleatoria que describe el número de clientes en la cola en estado estacionario, y  $L_q = E[N_q]$ . Entonces,

$$L_q = 0 \times p_0 + \sum_{n=1}^{\infty} (n-1) p_n = \sum_{n=1}^{\infty} n p_n - \sum_{n=1}^{\infty} p_n = L - (1 - p_0)$$

$$L_q = \frac{\rho}{(1 - \rho)} - \rho = \frac{\rho^2}{(1 - \rho)} = \rho L = L - \rho = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

**TIEMPO EN EL SISTEMA** Si  $R$  es el tiempo de respuesta, entonces  $E[R] = W$  es el tiempo en el sistema. Según la ley de Little, como se vio en la sección 2.3.2.2,  $E[N] = \lambda E[R]$ .

$$E[N] = L = \lambda E[R]$$

Entonces,

$$W = \frac{L}{\lambda} = \frac{1}{\lambda} \frac{\rho}{(1 - \rho)} = \frac{1}{\mu(1 - \rho)} = \frac{1}{\mu - \lambda}$$

**TIEMPO DE ESPERA EN COLA** A partir de la ley de Little

$$L_q = \lambda W_q$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{\rho}{(\mu - \lambda)}$$

**2.3.2.6. Colas M/M/c**

Una cola con c servidores idénticos, llegadas que siguen un proceso de Poisson de tasa  $\lambda$  y el tiempo de servicio que sigue una distribución exponencial de tasa  $\mu$ .

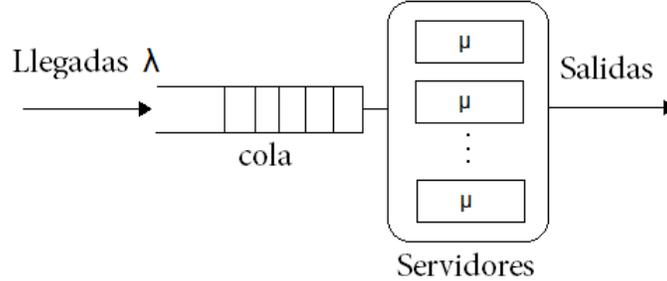


Figura 2.3.4: Esquema de una cola M/M/c

Este tipo de cola se puede modelar también como un proceso de Markov de nacimiento y muerte, como muestra la figura 2.3.5. También es llamado modelo Erlang-C.

El proceso se puede analizar considerando los valores:

$$\lambda_n = \lambda$$

$$\mu_n = \min(n, c) \mu$$

Cuando el número de clientes  $n$ , es mayor que  $c$ , todos los servidores están ocupados y la tasa de salida de clientes será  $c\mu$ .

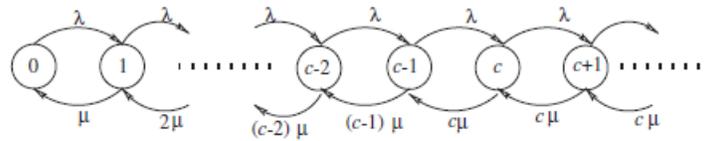


Figura 2.3.5: Transiciones entre estados de una cola M/M/c. [13]

Para despejar  $p_0$  se usa la ecuación encontrada para el proceso generalizado

$$p_n = p_0 \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i},$$

para las colas M/M/c, se convierte en

$$p_n = p_0 \prod_{i=1}^n \frac{\lambda}{i \mu} = p_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!} \text{ si } 1 \leq n \leq c,$$

y

$$p_n = p_0 \prod_{i=1}^c \frac{\lambda}{i \mu} \prod_{i=c+1}^n \frac{\lambda}{c \mu} = p_0 \left(\frac{\lambda}{\mu}\right)^n \frac{1}{c!} \left(\frac{1}{c}\right)^{n-c} \text{ si } n \geq c,$$

En este caso se define  $\rho = \lambda / (c\mu)$ , y para que el sistema sea estable, se requiere  $\rho < 1$ . Usando la igualdad  $c\rho = \lambda/\mu$ , las ecuaciones anteriores son

$$p_n = p_0 \frac{(c\rho)^n}{n!} \quad \text{para } n \leq c$$

$$p_n = p_0 \frac{(c\rho)^n}{c^{n-c}c!} = p_0 \frac{(c\rho)^n}{c!} \quad \text{para } n \geq c$$

Ahora,

$$1 = p_0 + \sum_{n=1}^{\infty} p_n = p_0 \left[ 1 + \sum_{n=1}^{c-1} \frac{(c\rho)^n}{n!} + \sum_{n=c}^{\infty} \frac{(c\rho)^n}{c!} \right],$$

$$p_0 = \left[ 1 + \sum_{n=1}^{c-1} \frac{(c\rho)^n}{n!} + \sum_{n=c}^{\infty} \frac{(c\rho)^n}{c!} \right]^{-1}$$

Las probabilidades, en este sistema en estado estacionario están dadas por

$$p_n \begin{cases} \frac{(c\rho)^n}{n!} p_0 & \text{si } 0 \leq n \leq c-1 \\ \frac{(c\rho)^n}{c!c^{n-c}} p_0 & \text{si } n \geq c \end{cases}$$

$$p_0 = \left[ \sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \frac{1}{1-\rho} \right]^{-1}$$

### MEDIDAS DE RENDIMIENTO

**CANTIDAD DE CLIENTES EN COLA** En este tipo de colas, se distingue que cuando  $n \geq c$ , todos los servidores están ocupados y, los clientes deben esperar en cola. El largo promedio de la cola,  $L_q$  es

$$L_q = \sum_{n=c}^{\infty} (n-c)p_n$$

reemplazando el valor de  $p_n$  cuando  $n \geq c$ , se obtiene,

$$L_q = \frac{(c\rho)^{c+1}/c}{c!(1-\rho)^2} p_0 = \frac{(\lambda/\mu)^c \lambda \mu}{(c-1)!(c\mu - \lambda)^2} p_0$$

**TIEMPO DE ESPERA EN COLA** Se usa  $L_q = \lambda W_q$  para encontrar  $W_q$

$$W_q = \left[ \frac{(\lambda/\mu)^c \mu}{(c-1)!(c\mu - \lambda)^2} \right] p_0$$

**TIEMPO EN EL SISTEMA** El tiempo en el sistema se calcula como el tiempo promedio de espera en cola más el tiempo de servicio de un cliente.

$$W = W_q + \frac{1}{\mu}$$

$$W = \left[ \frac{(\lambda/\mu)^c \mu}{(c-1)!(c\mu - \lambda)^2} \right] p_0 + \frac{1}{\mu}$$

**CANTIDAD DE CLIENTES EN EL SISTEMA** Usando la ley de Little se despeja  $L$ .  $L = \lambda W$ , entonces

$$L = \left[ \frac{(\lambda/\mu)^c \lambda \mu}{(c-1)!(c\mu - \lambda)^2} \right] p_0 + \frac{\lambda}{\mu}$$

**FÓRMULA ERLANG C:** En estado estacionario, la probabilidad de que un cliente tenga que esperar en cola es igual a la probabilidad de que todos los servidores estén ocupados, y está dada por  $C\left(c, \frac{\lambda}{\mu}\right)$ , una función de  $c$ ,  $\lambda$  y  $\mu$ , conocida como la fórmula Erlang C.

$$\lim_{t \rightarrow \infty} P(X(t) \geq c) = \sum_{n=c}^{\infty} p_n = C\left(c, \frac{\lambda}{\mu}\right)$$

$$C\left(c, \frac{\lambda}{\mu}\right) = p_0 \frac{(c\rho)^c}{c!(1-\rho)}$$

Esta fórmula permite también calcular las medidas de rendimiento anteriormente expuestas. Por esto, el modelo es llamado Erlang-C. Por supuesto, cuando  $c = 1$ , el resultado es aplicable para colas M/M/1.

**2.3.2.7. Colas M/M/c/K**

$K$  representa la capacidad de la cola, que en este caso es finita. Estos sistemas también se conocen como de canales paralelos con truncamiento, o sistema con pérdidas. Se modela como una cadena de Markov (ver figura 2.3.7).

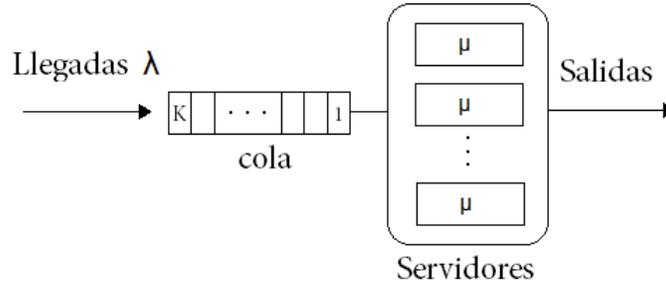


Figura 2.3.6: Esquema de una cola M/M/c/K

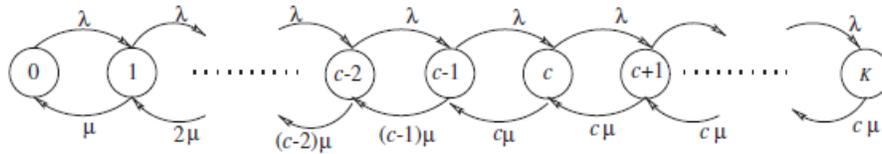


Figura 2.3.7: Transiciones entre estados de un sistema de colas M/M/c/K

Asumiendo la tasa efectiva de llegadas como

$$\lambda_n \begin{cases} \lambda & 0 \leq n < K \\ 0 & n \geq K \end{cases}$$

y de servicio

$$\mu_n \begin{cases} n\mu & 1 \leq n < c \\ c\mu & c \leq n \leq K \end{cases}$$

y usando la ecuación fundamental

$$p_n = p_0 \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i},$$

Se obtiene

$$p_0 = \left[ \sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \sum_{n=c}^K \frac{(c\rho)^n}{c^{n-c}c!} \right]^{-1} \quad 9$$

Un caso especial de este modelo es el modelo Erlang-B, también llamado *Erlang's Loss Model*, y es cuando  $K = c$ , es decir, cuando todos los servidores están ocupados, no hay capacidad para que los clientes esperen en cola.

### MEDIDAS DE RENDIMIENTO

**CANTIDAD DE CLIENTES EN COLA** El largo promedio de la cola,  $L_q$  cambia a

$$L_q = \sum_{n=c}^K (n-c)p_n$$

$$L_q = p_0 \frac{(c\rho)^c \rho}{c!(1-\rho)^2} [1 - \rho^{K-c+1} - (1-\rho)(K-c+1)\rho^{K-c}]$$

**CANTIDAD DE CLIENTES EN EL SISTEMA** Para calcular  $L$ , sólo basta sumar el número de servidores que están ocupados. Entonces,

$$L = L_q + \left( c - \sum_{n=0}^{c-1} (c-n)p_n \right) = L_q + c - \sum_{n=0}^{c-1} (c-n) \frac{(\lambda/\mu)^n}{n!} p_0$$

**TIEMPOS DE ESPERA PROMEDIO** A partir de la ley de Little se despejan los tiempos promedios de espera promedio en cola  $W_q$ . El tiempo  $W$  se calcula, al igual que en el modelo anterior, el tiempo de espera en cola más el tiempo de servicio. Los resultados son válidos sólo cuando  $\rho = \lambda/c\mu < 1$ , como explica MacGregor [12].

#### 2.3.2.8. Colas con Abandonos

**Modelo  $M/M/c+M$**  El modelo  $M/M/c+M$ , también llamado Erlang-A, es sensible a los abandonos. Asume una tiempo de paciencia (tiempo antes del abandono) distribuido exponencialmente. Se basa en el modelo  $M/M/c$ , o Erlang-C, pero se agrega el parámetro Markoviano de abandonos. El tiempo de paciencia promedio de un cliente es  $1/\theta$ , mientras la tasa de abandonos es  $\theta$ . Brown et al. [3], aplican este modelo a un call center de alto tráfico, estimando la paciencia como el tiempo que está dispuesto a esperar un cliente,  $R$ . Además usa el porcentaje de abandonos y el tiempo de espera  $W_q$  calculado a partir del modelo Erlang-C.

$$\% \text{ Abandonos} = \frac{W_q}{R}$$

Donde  $R$  será el estimador de  $1/\theta$ , el tiempo promedio de paciencia de un cliente.

En la realidad, los sistemas de colas, sobre todo los call centers, presentan abandonos. El modelo Erlang-A es el que, hasta ahora, mejor representa este tipo de sistemas, a pesar de que tiene limitaciones.

<sup>9</sup>Notar que haciendo tener  $K$  a infinito, se obtiene el mismo resultado que para las colas  $M/M/c$

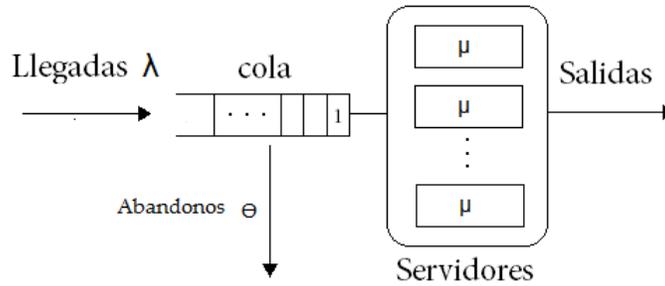


Figura 2.3.8: Esquema de una cola con abandonos

Este modelo sigue asumiendo llegadas de  $Poisson(\lambda)$  y servicio  $exp(\mu)$ , pero además considera que los clientes tienen tiempos de paciencia  $\tau$ , que distribuyen  $exp(\theta)$  i.i.d., siendo estos tres procesos independientes entre sí[10].  $\tau$  se define como el tiempo que cliente está dispuesto a esperar en cola por el servicio. De esta manera, el tiempo que esperará un cliente es

$$W_q = \min\{W, \tau\}$$

El diagrama, como proceso de Markov, es entonces

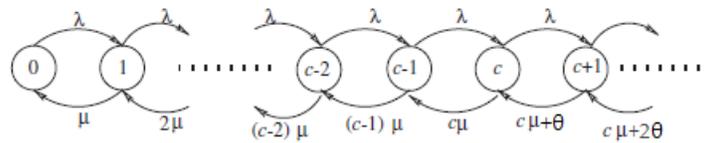


Figura 2.3.9: Transiciones entre estados de una cola con abandonos, M/M/c+M

**IMPORTANCIA** La importancia de incorporar un comportamiento de abandonos y paciencia a los modelos analíticos, permite estudiar sistemas que son estables, a pesar de que la tasa de llegada sea mayor a la de servicio. Además, es mucho más preciso que los modelos anteriores en cuanto los resultados que se obtienen para el problema de cantidad de servidores. Lo demuestran Mandelbaum y Zeltyn en 2005, comparando los resultados obtenidos a partir de la aplicación de los modelos Erlang-A y Erlang-C, con y sin abandonos, respectivamente. Esto está dado por la disminución en la carga del sistema que constituyen los clientes que abandonan la cola.

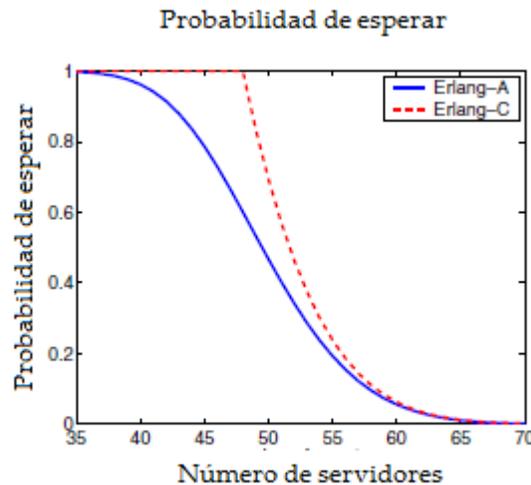


Figura 2.3.10: Comparación entre Erlang-A y Erlang-C en cuanto a probabilidad de espera.

Se aplican los modelos a un sistema que recibe 48 llamadas por minuto, el tiempo de servicio promedio es un minuto y la paciencia promedio es 2 minutos. Se observa en la figura 2.3.10 que el modelo sin abandonos estima que la probabilidad de esperar es menor que 1 sólo cuando hay más de 50 servidores. En cambio, esta probabilidad es menor que 1 cuando hay 40 servidores. Por ejemplo, si una meta de la empresa fuera tener menos de 80 % de probabilidad de espera, el modelo Erlang-C recomendaría contratar más de 50 servidores, y el modelo Erlang-A recomendaría contratar a 45.

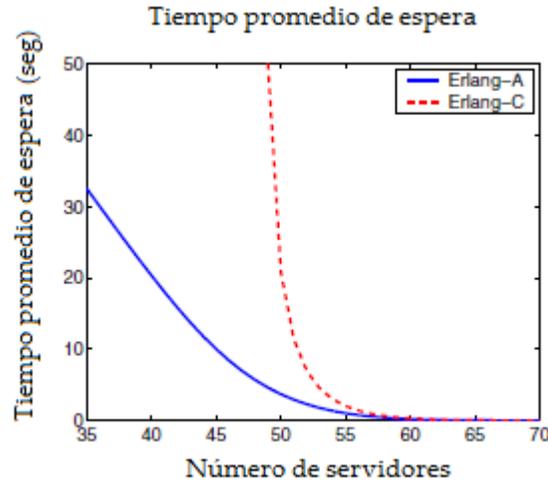


Figura 2.3.11: Comparación entre Erlang-A y Erlang-C en cuanto a tiempo promedio de espera.

A su vez, los resultados de tiempo promedio de espera son bastante distintos al aplicar los dos modelos. Por ejemplo, al haber 35 servidores, el tiempo de espera es cercano a los 30 segundos, para un cliente que llega al sistema, pero con el mismo número de operadores, si se calculara con el modelo Erlang-C, el resultado indicaría que el tiempo es infinito. Por lo tanto, Mandelbaum y Zeltyn confirman la importancia de considerar los abandonos en los sistemas de colas, pues esto reduce la carga del sistema y permite estimar mucho mejor el desempeño del sistema.

Para aplicar este modelo, se requiere conocer los parámetros  $\lambda$ ,  $\mu$ ,  $\theta$  y  $c$ . Típicamente se usan estimadores basados en datos históricos.

## MEDIDAS DE RENDIMIENTO

**CANTIDAD DE CLIENTES EN COLA Y CANTIDAD DE CLIENTES EN EL SISTEMA** Cuando se cumple  $\rho > 1$ , la cantidad de clientes se calcula con el modelo Erlang-C. Para calcular  $L_q$  y  $L$ , cuando hay abandonos, existen métodos numéricos bastante complejos que permiten cuantificar estas medidas de servicio cuando la condición de estacionariedad no se cumple.

**TIEMPOS DE ESPERA PROMEDIO** Con los tiempos de espera ocurre lo mismo. Ya que la ley de Little no se cumple cuando  $\rho > 1$ , además ésta no considera abandonos y capacidad limitada de la cola, lo que se hace en las investigaciones científicas que requieren estas estimaciones, es acudir a las calculadoras programadas con fórmulas parecidas a las de Erlang-A/Palm, sin nunca explicarlas en detalle, ya que los métodos numéricos subyacentes son bastante complejos.

En aplicaciones del modelo Erlang-A, como lo muestran Avramidis and L'Ecuyer [2], Brown et al. [3], Mandelbaum et al. [9], Mandelbaum [10], no aplican directamente las fórmulas, si no que utilizan alguna herramienta desarrollada para resolverlas, como CCC (*Call Centre Calculator*) del Software 4CallCenters.

En este caso los exactos de los cálculos son más similares a los resultados de la calculadora Erlang-X que a los que arroja CCC. Se utiliza entonces, Erlang-X Calculator, descrita más adelante en la sección 2.3.2.9.

### LIMITACIONES DE LA TEORÍA DE COLAS

A pesar de que Erlang comenzó a desarrollar la Teoría de Colas basándose en lo que observó en una central telefónica, rara vez se puede aplicar directamente a los callcenters. Algunas características de los callcenters como: capacidad finita de la cola, abandono de la cola de parte de los clientes, llamadas reiteradas, diferencia en tiempos de servicio entre los servidores y otras, dificultan el análisis directo con teoría de colas.

Todos los modelos revisados tienen el problema de asumir un dado tipo de distribución de llegadas y de tiempo de servicio, además asumen que todos los servidores tienen habilidades idénticas y que llega un solo tipo de cliente con las mismas características. Una de las principales limitaciones del modelo Erlang-C es que no considera capacidad finita de la cola ni la impaciencia de los clientes, quienes tienen cierto comportamiento de abandono asociado [3], pues con esto se sobrestiman los tiempos de espera y probabilidad de espera en cola de los clientes. La principal tarea de los investigadores de esta área es desarrollar modelos que superen dichas deficiencias [7].

En el caso de este estudio, no es posible la aplicación directa de teoría de colas porque se cuenta con dos tipos de llegadas de clientes con patrones bastante distintos, cada tipo de llamada además requiere un tiempo de servicio diferente, mientras los servidores atienden a todas las llamadas sin distinción.

Si intentáramos analizar las llamadas de tipo 1 como un sistema  $M/M/c$ , estaríamos asumiendo que todos los servidores están atendiendo llamadas de este tipo y obtendríamos resultados erróneos. Lo mismo con el otro tipo de llamadas.

Para validar la simulación y verificar que el modelo refleja la realidad, se va a ajustar los datos para poder modelar una cola sencilla  $M/M/c$ , la simulación debería arrojar resultados parecidos a los del modelo analítico.

#### 2.3.2.9. Calculadora Erlang-X

La calculadora Erlang-X es una herramienta desarrollada por el académico Ger Koole, de Vrije University Amsterdam, Amsterdam, Holanda. En su libro “Call Center Optimization” [8], describe detalladamente las especificaciones y las limitaciones de las calculadoras que desarrolló, Erlang-C y Erlang.X, y las pone a disposición en la web<sup>10</sup>. Se puede ver un ejemplo de su uso en el apéndice C.

Las calculadoras se basan en CCmath, un software y paquetes de algoritmos desarrollados para resolver problemas de call centers, pero también usados en otras industrias. Principalmente permiten planificar la fuerza laboral, realizar pronósticos, cálculos de staff óptimo, ruteo de llamadas y turnos de empleados<sup>11</sup>.

Como se dijo antes, todas las calculadoras de medidas de rendimiento en call centers, están basadas en este tipo de paquetes o librerías. Se debe aclarar que la librería gratuita CCmath, programada en C, para Linux es algo totalmente distinto, esta es una librería de simples funciones matemáticas.

<sup>10</sup>Más información en <http://www.gerkoole.com/CCO/>

<sup>11</sup>Más información en <http://www.ccmath.com/en/about-ccmath.html>

The screenshot shows the 'Erlang-X Calculator' interface. It has a title bar 'Erlang-X Calculator' and a 'Data' section. The 'Data' section contains several rows of input fields, each with radio buttons for 'In', 'Out', and 'Ignore'. The parameters and their units are: Arrivals (per minute), Service time (seconds), Number of agents (integer required), Number of lines (integer required), Blocking (%), Average time until abandonment (seconds), Abandonments (%), Average waiting time (seconds), and Service level (% waits less than 20.00 seconds). A 'compute the missing values' button is located at the bottom.

Figura 2.3.12: Campos de entrada y salida de Erlang-X Calculator

Fuente: Captura de pantalla de la interfaz en línea de <http://www.math.vu.nl/~koole/ccmath/ErlangX/index.php>

### 2.3.3. Simulación

El proceso de atención de llamadas en el Fono 133, es un sistema que funciona bajo mucha incertidumbre. Los patrones de llegada de clientes y los de servicio varían en el tiempo. Esta realidad supera los límites de la capacidad de de la teoría de colas. Para analizar procesos con estas características, la simulación es una herramienta suficientemente flexible, que soporta expresiones complejas que describen aquellos patrones.

Construir un modelo de simulación, permite simular la lógica del proceso y animarlo. La observación del estado transiente, cuando el sistema no ha alcanzado el estado estacionario, permite identificar cuellos de botella y otros inconvenientes que puedan existir en el proceso, además de poder medir niveles de rendimiento o parámetros relevantes, en el momento deseado.

Todos los callcenters operan bajo condiciones de enorme incertidumbre, dada porque que varían los patrones de llegada de llamadas, el tipo de servicio que requiere cada cliente, el número de servidores y otros parámetros que tornan complejo en análisis del sistema. La simulación permite analizar el impacto de ciertos cambios en dichos parámetros o en el flujo de los procesos y sin tener que aplicarlas en la realidad.

Se utiliza el software Arena Simulation. Arena es un software de simulación desarrollado por la compañía Estadounidense Rockwell Automation Inc. La industria de servicio al cliente es una de las que más usa Arena para modelar procesos de los más simples hasta procesos altamente complejos<sup>12</sup>.

Cada producto de Arena viene cargado con *templates* que son grupos de módulos, denominados *flowcharting modules*, que permiten construir los flujos de procesos. Cada *template* tienen funcionalidades para modelar los procesos y pueden generar outputs estadísticos detallados, según los requerimientos del investigador. El modelamiento de procesos en Arena es similar a BPM, pero se agregan módulos intermedios que para asignar atributos a los clientes y así diferenciar su comportamiento. La Edición Profesional (*Professional Edition*) de Arena soporta, entre otras cosas, simulaciones con más de 50 mil entidades participando en el proceso y permite crear expresiones complejas para simular la lógica del modelo, escribir archivos con datos importantes para el investigador.

<sup>12</sup>Fuente: [http://www.arenasimulation.com/Solutions\\_Contact\\_Center.aspx](http://www.arenasimulation.com/Solutions_Contact_Center.aspx)

### 2.3.3.1. Validación

Se corre el riesgo de cometer errores, ya que hay muchos detalles en el proceso que se deben administrar, en cada paso. Por eso, es necesario validar la simulación y verificar que los resultados hagan sentido. Para eso se usarán los resultados obtenidos mediante Teoría de Colas. La teoría de colas entrega resultados bastante precisos cuando se trata de sistemas que alcanzan el estado estacionario, por lo tanto se simulan escenarios de este tipo y se comparan las medidas de rendimiento obtenidas con las de los modelos de Erlang, para validar la simulación.

### 2.3.4. Muestreo

Cuando los registros disponibles no contienen todos los datos necesarios para contestar alguna pregunta, es necesario recurrir a otra fuente de acceso a la data. En este caso, las variables a observar son cuantitativas aleatorias, no se pueden predecir con exactitud, pero se puede estimar su distribución.

Algunas de las variables relevantes en este estudio, son los tiempos de espera de los clientes. El tiempo de espera en cola es una de las medidas de rendimiento que se estima usando teoría de colas, pero es muy importante tener una idea de la magnitud de este parámetro para verificar si los resultados de la aplicación del modelo son coherentes con la realidad.

Después de verificar que no existe registro de los tiempos de espera de los clientes del 133, se toma una muestra aleatoria, con la autorización de Carabineros de Chile. El mismo muestreo permite observar la cantidad de llamadas que reciben tono ocupado, las que son rechazadas por el sistema cuando este colapsa. Los resultados del muestreo detallan en el apéndice B.2.

## 2.4. Alcances

El proyecto se enmarca en la unidad Nivel 133 de la CENCO central, en la Región Metropolitana. Esta CENCO recibe llamadas desde las 32 comunas de Santiago y de las provincias Cordillera, Maipo, Melipilla y Talgante. Además, se filtran algunas llamadas de celulares provenientes de la V Región, de una cierta compañía de telefonía móvil.

Se tiene acceso a los datos de llamadas entrantes de 17 meses, desde que se instaló el sistema de administración de llamadas que se usa actualmente. Esto corresponde a 8 millones de entradas, sin embargo, el total de llamadas que discan el 133, que sería la demanda total, no está disponible, por eso se hacen estimaciones. Tampoco está disponible un estudio de detallado de la eficiencia de cada operador, por eso se asume que sus tiempos de atención son homogéneos. En el futuro, será fácil replicar esta metodología incluyendo más información sobre estos dos tipos de datos que hoy son censurados.

Existen 29 sub-centrales de CENCO en regiones, sin embargo no todas cuentan con el sistema Viper CAD<sup>13</sup>, u otro similar, que permita hacer un seguimiento de cada llamada, estudiar sus atributos y patrones de comportamiento. Por lo tanto, un estudio como este es aplicable a las centrales que tengan ese tipo de tecnología, o al menos puedan solicitar estudios de tráfico a su compañía proveedora de telefonía.

El proceso a observar es el de Tratamiento y el de Atención de llamadas, dejando fuera lo que ocurre después de que el operador cuelga la llamada. Cuando una llamada es de emergencia y requiere un procedimiento policial, esto se administra en la unidad Alpha I, en donde se despliega aun proceso complejo, absolutamente distinto, que se compone de actores, interacciones y capacidades que se deben estudiar en otro proyecto como este.

El Nivel Alpha I tiene una cantidad fija de operadores, quienes se contactan por radio con las patrullas y otros recursos policiales en terreno para informar los procedimientos policiales requeridos por la ciudadanía. Una vez estudiado y resueltos los problemas del Nivel 133, se debe medir el rendimiento del Nivel Alpha I y evaluar su capacidad para coordinar acciones y decisiones que permitan mejorar el

<sup>13</sup>Sistema de administración de llamadas. Se describe en detalle en la sección 3.1.1

proceso completo.

Es importante aclarar que existen otros números de Carabineros de Chile que prestan servicios a la ciudadanía. En CENCO, se encuentra el Fono de Informaciones Policiales 139. En otra unidad, llamada CENCO Familia, funcionan el Fono Drogas 135, Fono Niños 147 y Fono Familia 149. Ninguno utiliza un sistema de información para manejar las llamadas, por lo tanto no existen datos registrados adecuados, hasta el momento, para incluirlos en el estudio.

## Capítulo 3

# Descripción del Sistema

El Fono de Emergencias 133 es un servicio telefónico de emergencia que atiende llamadas de todo el país. Una llamada al número 133 hecha desde teléfono fijo, rural, público o teléfono móvil, será atendida por operadores preparados en la zona geográfica más cercana, derivándose un proceso acorde a las necesidades de la persona que llama.

Este proceso se lleva a cabo en la Central de Comunicaciones (CENCO), la cual es considerada el motor de las operaciones de Carabineros [6], su misión es administrar y coordinar las telecomunicaciones que permiten realizar procedimientos de emergencia, seguridad y todo tipo de los operativos policiales, día a día en todo Chile.

En el Nivel 133, las llamadas son atendidas mediante un complejo proceso. Luego, en otra unidad, llamada Alpha I, se administran los procedimientos policiales que son requeridos, hasta que la emergencia es resuelta (figura 3). El trabajo se remite al estudio de las operaciones del Nivel 133 en la CENCO de la Región Metropolitana.



Figura 3.0.1: Proceso de atención de emergencias en CENCO

Fuente: Adaptación de ilustración, Revista Carabineros de Chile, Edición N°676

Este capítulo describe en detalle en Nivel 133, desde su simple descripción hasta la cuantificación de cada tipo de llamadas que se atienden. Primero se describe el sistema cualitativamente, se explica quiénes son las entidades participantes (operadores y sistemas de información) y se enumeran las variables de estado que permiten entender la capacidad del sistema. En segunda instancia, se describen los dos grandes grupos de llamadas que se reciben, cada uno de ellos se subdivide en categorías muy importantes que se consideran en el desarrollo del trabajo.

En la tercera parte de este capítulo se presenta el flujo de las llamadas en CENCO, a través del modelamiento de los procesos pertinentes. Se comparan dichos modelos con el diagrama de flujo de procesos que actualmente se maneja en en Nivel 133. Finalmente, se describe la demanda total y el detalle la cantidad de llamadas atendidas, identificando cada tipo de llamadas, estudiando los patrones de llegada de llamadas, distribución del tiempo de servicio y la distribución de la cantidad de operadores conectados, para luego resumir el comportamiento del sistema en cinco horarios diferenciados, con marcados patrones de comportamiento.

### 3.1. El Nivel 133

En el Nivel 133 es un servicio destinado a la atención de llamados de emergencia de la ciudadanía. Es un *call center* que recibe llamadas durante las 24 horas del día. Su funcionamiento depende de un cierto número de líneas telefónicas, un sistema de *Computer Assisted Dispatch* (CAD), de las llamadas entrantes y de los operadores, quienes se encargan directamente de entregar el servicio.

Para entender los procesos que se describen en este informe, se debe tener en cuenta que existen algunas variables de estado, que son medibles en cada momento del tiempo, y permiten caracterizar el sistema, como veremos a continuación. Más adelante se describe también el funcionamiento del sistema CAD y el trabajo de los operadores, a grandes rasgos.

#### 3.1.1. CAD

Carabineros utiliza un sistema CAD (*Computer Assisted Dispatch*) que permite derivar los procedimientos policiales a los despachadores en Alpha I, cuando esto corresponda. Actualmente se usa Viper CAD y Power 911 de manera integrada.

A través del sistema Power 911, el operador puede contestar una llamada, registrar la información del usuario, localizar la dirección en un mapa, clasificar las llamadas inoficiosas y tipificar las emergencias. Al finalizar este proceso, el operador envía los requerimientos de procedimientos a Alpha I usando la interfaz del Viper CAD (detalles de la interfaz de los sistemas Power 911 y Viper CAD en apéndice F.2).

Al contestar una llamada se activa el panel para registrar los datos del llamante (ver figura F.2.2). Se tipifica la emergencia eligiendo o digitando uno de los 103 tipos de emergencias que ofrece el sistema. Si la llamada no genera procedimiento, el operador procede a clasificar la llamada en una de las seis categorías de llamadas inoficiosas. Luego se libera la llamada sin crear requerimiento.

Las aplicaciones Contact Center o Symposium y Power MIS, se utilizan para acceder a reportes de información agregada. Algunos informes entregan, por ejemplo: la cantidad de llamadas atendidas en cierto periodo, número de operadores conectados y procedimientos administrados en Alpha I, entre otros.

Este tipo de herramientas permiten obtener datos agregados, y algunos detalles de las llamadas atendidas, lo que permitió procesar los datos para obtener información relevante.

#### 3.1.2. Variables de estado

**Número de Líneas ( $K$ )** Hay una capacidad de 60 líneas contratada por CENCO, resultantes de dos centrales PABX<sup>1</sup> de 30 líneas cada una.

**Número de Usuarios en el Sistema ( $L$ )** Es el número total de usuarios que puede estar en el sistema.

$$L \leq 60$$

$$L = \text{usuarios siendo atendidos} + \text{usuarios esperando en cola}$$

**Número de Servidores ( $c$ )** El número de operadores conectados, atendiendo en una consola o módulo, es variable durante el día. Actualmente hay 34 módulos de atención telefónica, por lo tanto  $c < 34$ . La dotación actual y el manejo de los turnos se explican en la sección 3.4.7.

---

<sup>1</sup>PABX (*Private Automatic Branch Exchange*) Central telefónica que conecta terminales internos de una organización, usualmente conectada a la red telefónica pública a través de una línea troncal.

### 3.1.3. Operadores

Los servidores de este sistema, quienes entregan el servicio al cliente, son llamados operadores. Son los agentes que atienden las llamadas telefónicas, cuya misión es entregar información y respuesta oportuna. Los operadores tienen el primer contacto con la persona afligida, tienen un entrenamiento técnico y psicológico para dar una respuesta tranquilizadora, recopilar antecedentes rápidamente y distribuir las llamadas a los despachadores en Alpha I. Muchas veces atienden a personas alteradas o casos de pánico y riesgo de la ciudadanía.

El Nivel 133 cuenta con 34 módulos de atención telefónica. Como se mencionó anteriormente, hay 60 líneas telefónicas disponibles para la atención de llamados, pero sólo 34 podrían ser utilizadas, pues no hay más capacidad física o espacio disponible. Además, uno de los módulos presenta desperfectos técnicos, entonces en teoría solamente se podrían conectar 33 operadores.

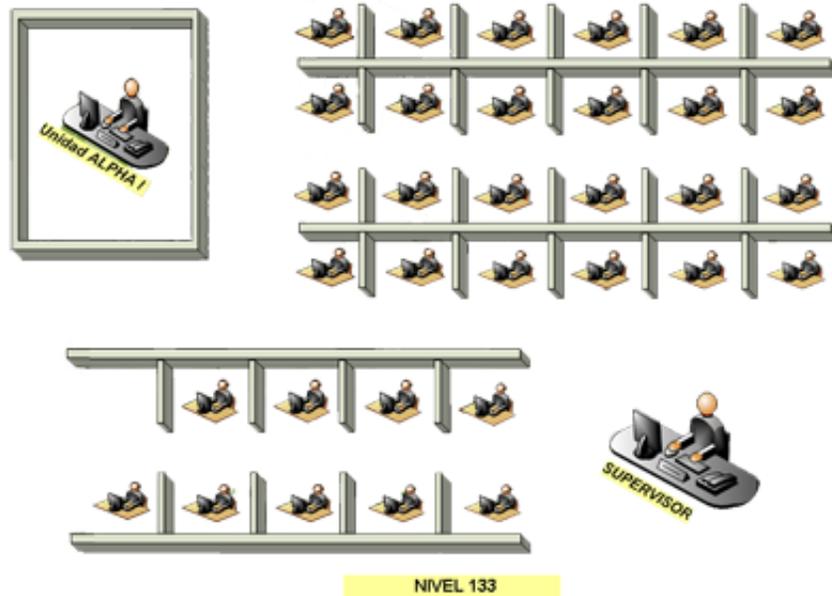


Figura 3.1.1: *Layout* del Nivel 133

Fuente: Adaptación de un documento Visio de Carabineros

Cada módulo está equipado con *hardware*<sup>2</sup> que permite atender las llamadas, y están conectados al sistema Viper CAD. Además cuentan con documentación para consultar teléfonos y direcciones de comisarías, u otros servicios, que solicitan los usuarios.

#### 3.1.3.1. Dotación y Turnos

CENCO contempla una fuerza laboral efectiva de 171 personas, quienes trabajan por turnos. La dotación total<sup>3</sup> es 271, pero de éstos, sólo 100 están disponibles para atender en el Nivel 133, esto ocurre porque otros cargos en CENCO se consideran de mayor prioridad.

En el Nivel 133 hay un total de 20 C.P.R.<sup>4</sup>, quienes según el Código del Trabajo, no pueden trabajar más de 45 horas semanales ni por más de 10 horas diarias. La mayoría de los operadores son mujeres, 12 tienen permiso para ausentarse por permiso de alimentación a sus hijos menores de 2 años y tienen fuero maternal, lo que no permite que trabajen fines de semana ni festivos y 2 personas están con licencia

<sup>2</sup>Dos pantallas, teclados, teléfonos, entre otros.

<sup>3</sup>Dotación CENCO. Fuerza Efectiva correspondiente al día 13 de Noviembre de 2012.

<sup>4</sup>Contratado por Resolución. De acuerdo a la ley 20490, Carabineros de Chile contrata a civiles para realizar tareas en la institución.

médica de pre-natal<sup>5</sup>.

Actualmente los operadores cumplen turnos como se describe en la siguiente tabla, según días de semana o fines de semana. Para los días festivos se asignan turnos de acuerdo a los horarios del fin de semana.

Lunes a Viernes	Sábado y Domingo
07:00 a 13:59 hrs.	08:00 a 19:59 hrs.
14:00 a 20:59 hrs.	20:00 a 07:59 hrs.
21:00 a 06:59 hrs.	-

Tabla 3.1.1: Horarios de Turnos de los Operadores

Generalmente no es posible acatar los turnos tal como son programados, dado que los carabineros son asignados a otros servicios que tienen mayor prioridad que el Nivel 133. Día a día ocurren imprevistos que provocan bajas en el número de operadores. Algunas causas se explican a continuación.

Cuando ocurren eventos masivos que afectan a una porción importante de la ciudadanía, como partidos de fútbol clásicos, marchas o protestas, entonces la unidad Comando de Control requiere más personal y algunos operadores del 133 deben abandonar sus módulos. Casi todos los días se requieren 2 personas para realizar la guardia del edificio Norambuena, donde está ubicado CENCO, entonces deben salir operadores del Nivel 133. Otro motivo es que algunos operadores se conectan a los módulos para contestar llamadas, e inmediatamente son asignados a Alpha I, por diversas razones, ya que tener a los operadores en Alpha I es prioritario.

## 3.2. Tipos de Llamadas Entrantes

Dado que la misión del Nivel 133 es atender llamadas de emergencia, se podría pensar que éste tipo de llamadas es el principal foco de este trabajo. Sin embargo, más del 70 % de las llamadas que se reciben a diario, no son emergencias [6].

Las llamadas atendidas se dividen en dos grandes grupos: llamadas de emergencia y llamadas inoficiosas. Estos dos grupos se subdividen en diferentes categorías, como se explica a continuación.

La distinción entre estos distintos grupos de llamadas tiene relevancia al medir las tasas de llegada, tiempos de servicio y otros indicadores importantes que se deben medir en un call center de este tipo, sobre todo porque cada tipo de llamada utilizará de manera distinta los recursos del call center.

### 3.2.1. Llamadas Inoficiosas

Las llamadas inoficiosas son aquellas que no generan procedimientos policiales, pues no son emergencias, y de hecho, no deberían llegar al 133. Son llamadas que utilizan inadecuadamente los recursos y se quiere evitar que sigan llegando.

Este tipo de llamadas ha alcanzado un volumen tan alto, que han sido categorizadas en seis clases e incorporadas en el sistema CAD que se instaló en 2011 para que los operadores puedan registrar qué tipo de llamada inoficiosa contestan, cuando corresponde.

<sup>5</sup>Permiso de Alimentación. Fuente:ORD. N°4135/088 - Normativa laboral. Dirección del Trabajo. Gobierno de Chile

N°	Clasificación	Descripción
1	Broma	Llamada efectuada por individuo solo para burlarse o bromear.
2	Cortada	Llamada que no alcanza a generar audio y se termina, el usuario cuelga.
3	Equivocado	Llamada que no concierne al nivel de emergencias 133.
4	Inaudible	Llamada que no es audible por el operador.
5	Información	Llamada de consulta de antecedentes u orientación policial.
6	Insulto	Llamada en que el usuario se expresa ofensivamente en contra de Carabineros o del operador que atiende.

Tabla 3.2.1: Clasificación de llamadas inoficiosas

La llegada de llamadas inoficiosas se atribuye a muchas causas. Algunas son: 1) Campañas incorrectas: se percibe que al lanzar campañas informativas las llamadas inoficiosas aumentan, 2) Hay falta de conciencia de la importancia del servicio, 3) Ignorancia: las personas llaman para resolver temas que no son emergencias policiales, 4) Malas intenciones: algunas personas llaman para congestionar el sistema.

### 3.2.2. Llamadas de Emergencia

Las llamadas que efectivamente son emergencias, requieren que se derive un procedimiento policial, lo cual se administra en Alpha I, pero es el operador del Nivel 133 quién está a cargo de reconocer qué tipo de procedimiento se requiere.

Las emergencias están categorizadas en 103 tipos de incidentes. Hoy en día existe un campo en el sistema en el sistema CAD que permite al operador seleccionar el tipo de emergencia que un usuario reporta. Durante la atención de una llamada, el operador debe elegir una opción de dicha lista, lo que se llama "tipificar" o asignar una "tipificación de la llamada" para luego poder presionar un botón que envía el requerimiento al despachador de la zona correspondiente. De esta manera el despachador, en Alpha I, sabe de qué tipo de incidente o delito se trata y qué tan grave o urgente es, de acuerdo a la prioridad que viene asignada a la tipificación.

En el sistema CAD, cada tipificación tiene una prioridad preasignada y el operador tiene la capacidad de modificarla de acuerdo a la información que entrega el llamante. La prioridad se indica con un dígito del 1 al 7, siendo 1 la prioridad más alta y 7 la más baja.

El listado completo de emergencias categorizadas y su respectiva prioridad se encuentra en el apéndice F.3.1. A continuación se muestra un ejemplo.

Tipificación	Prioridad
Accidente de Tránsito Sólo Daños	5
Accidente de Tránsito con Lesionados y Muertos	2
Accidente de Aviación	1
Accidente Ferroviario	1

Tabla 3.2.2: Ejemplo Tipificación de Emergencias

Se observa en la tabla, que un accidente de tránsito con daños, tiene una prioridad 5, siendo esta una prioridad baja a media. Mientras un accidente donde hay lesionados o muertos tiene prioridad 2, siendo mucho más alta. Un accidente masivo como puede ser el de aviación o ferroviario, puede implicar que hay muchas vidas que salvar o identificar las causas puede ser muy importante, por lo tanto estas últimas emergencias tienen prioridad 1.

## 3.3. Modelos de Procesos

El proceso completo de servicio que se lleva a cabo en en Nivel 133, se describe mediante el modelamiento de dos procesos. El proceso principal es llamado, en CENCO, "Tratamiento de llamadas". Si

se cumplen ciertas condiciones, las llamadas entran al subproceso llamado "Atención de llamadas".

A continuación se describen ambos procesos de servicio y se modelan empleando BPMN, mediante el uso del software Bizagi.

### 3.3.1. Proceso de Tratamiento de Llamadas

#### DESCRIPCIÓN DEL PROCESO

El proceso comienza cuando un usuario llama al 133. Las llamadas se dirigen automáticamente a las centrales PABX de la CENCO más cercana a la dirección que registra la compañía operadora del usuario llamante. El sistema asigna la llamada al operador que lleva menos tiempo ocupado en proporción al tiempo de conexión. Dependiendo del estado del sistema, la llamada del usuario puede seguir tres caminos:

1.  $N_s = 60$ , No hay capacidad en la cola, el usuario recibe tono ocupado y cuelga.
2.  $N_s \leq c$ . El número de usuarios en el sistema es menor que el número de operadores que están conectados, el usuario pasa directamente al proceso de atención en el Nivel 133.
3.  $c \leq N_s \leq 60$ . Hay más usuarios en el sistema que operadores, y menos de 60 usuarios (usuarios en cola y capacidad en la cola), entonces el usuario entra a la cola de espera y espera hasta que un operador le conteste. Si su tiempo de paciencia se sobrepasa, el usuario abandona la cola antes de ser atendido.

Finalmente, si el usuario es atendido y su emergencia es real, los recursos policiales que requiere se despachan a través del nivel Alpha I. Si la llamada es inofensiva, el proceso termina dentro del subproceso de atención, no se envía el requerimiento a Alpha I.

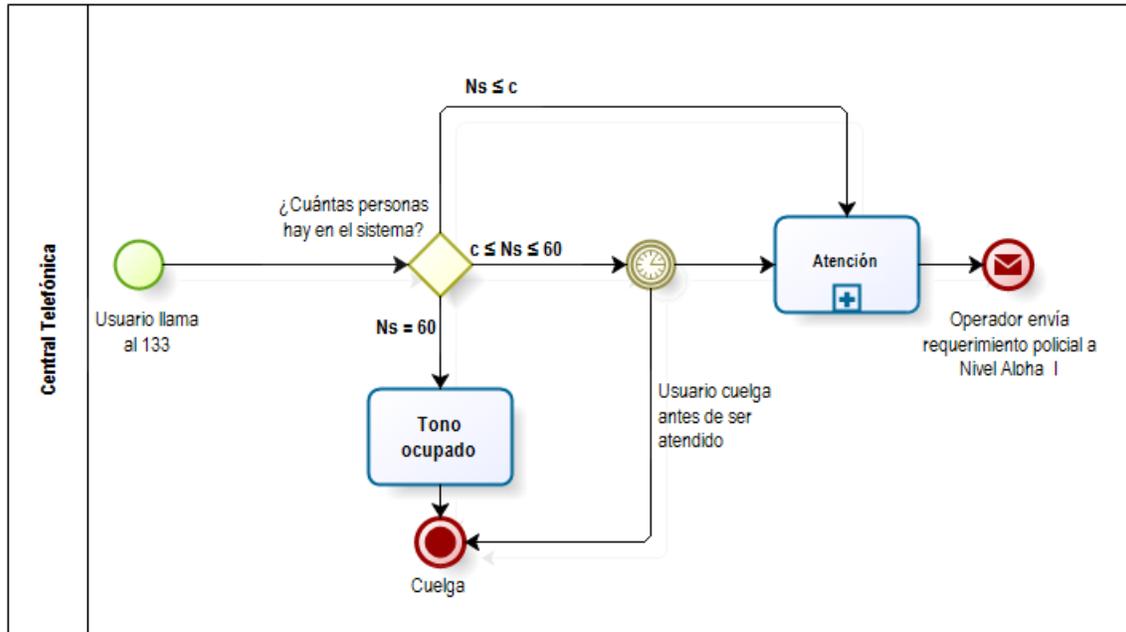


Figura 3.3.1: Modelo del proceso Tratamiento de Llamadas

### 3.3.2. Proceso de Atención de Llamadas

A continuación se expone paso a paso las tareas que realiza un operador durante el servicio de atención de una llamada. El modelo del proceso se puede ver ampliado en el anexo.

### DESCRIPCIÓN DEL PROCESO

Tras la llegada una llamado, el operador contesta y el usuario escucha un saludo del operador que ha sido grabado en el momento de conexión. Si el operador nota que no es una llamada de emergencia, debe atender la llamada de todas maneras, el operador debe clasificar la llamada y luego termina la atención. Existen seis clasificaciones de llamadas inoficiosas, se describen en la siguiente sección.

Si la llamada es realmente un caso que requiere un procedimiento policial, lo primero que hace un operador es averiguar dónde se necesita la presencia de Carabineros. Si por alguna razón el operativo se necesita en otra región, entonces manualmente el operador selecciona la CENCO correspondiente y el despachador regional recibe directamente en su pantalla el requerimiento.

Si el operativo es requerido en la Región Metropolitana, el operador conversa con el usuario y puede realizar varias tareas paralelamente, según cómo el usuario relate la emergencia. Entre estas tareas, el operador ingresa la dirección que entrega el llamante. El sistema permite validar la dirección encontrando la ubicación en el mapa. Luego se clasifica la emergencia en uno de los 103 tipos de incidentes que están listados en el sistema, cada uno con una prioridad preasignada y modificable.

Una vez que la emergencia ha sido tipificada, el operador puede hacer click en su pantalla para que el requerimiento aparezca en la pantalla del despachador de Alpha I. A la vez, o posterior al envío del requerimiento a Alpha I, el operador toma notas de la emergencia y de los datos del llamante. La llamada termina cuando el operador corta para quedar disponible y recibir otra llamada.

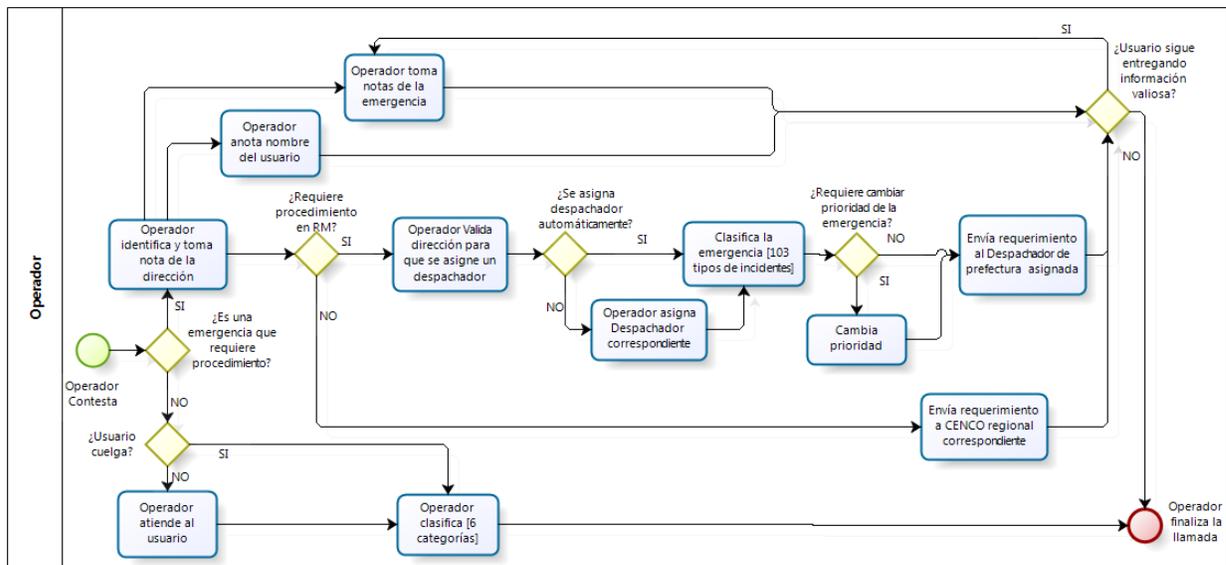


Figura 3.3.2: Modelo del proceso Atención de Llamadas

## 3.4. Patrones de Comportamiento

Ya se presentó la capacidad del sistema en cuanto a líneas telefónicas. Para aplicar métodos matemáticos como la Teoría de Colas (ver 2.3.2), se necesita estudiar y determinar los patrones de comportamiento del sistema y la capacidad en cada etapa del proceso.

Se debe comprender la composición de la demanda, se cuenta con información de las llamadas entrantes, pero es importante reconocer que hay una porción de la demanda total que corresponde a las llamadas que reciben tono ocupado, y no conocemos el volumen que representa.

Las llamadas atendidas pueden ser emergencias o inoficiosas. La tasa de llegada de cada tipo de llamada cambia, dependiendo del día y hora. El número de operadores conectados, es diferente también dependiendo del día y la hora en que se observe el sistema. Por otro lado, los operadores atienden cada tipo de llamadas en una cierta cantidad de tiempo, y esto también se puede describir mediante una tasa de servicio. Estos patrones de comportamiento del sistema, son esenciales para medir los niveles de servicio en el capítulo siguiente.

### 3.4.1. Demanda Total

La demanda total del sistema consiste en todas las llamadas que la gente hace al 133. Entre éstas, algunas son atendidas inmediatamente, algunas deben esperar en cola y luego son atendidas, otras cuelgan mientras esperan en cola y otras son rechazadas por el sistema, recibiendo tono ocupado cuando todas las líneas están ocupadas. Se puede describir la demanda total como

$$\text{Demanda Total} = \text{Llamadas Entrantes} + \text{Llamadas Rechazadas}$$

Solo se cuenta con información de las llamadas *Entrantes*, son aquellas que son atendidas inmediatamente o que esperan en cola. Esto ocurre cuando todos los operadores están ocupados, algunos clientes esperan, pero otros abandonan la cola. Entonces las llamadas Entrantes se dividen en dos tipos.

$$\text{Llamadas Entrantes} = \text{Llamadas Atendidas} + \text{Llamadas Abandonadas}$$

Las llamadas *Atendidas* son aquellas contestadas por un operador, que pasan por el proceso de servicio descrito en 3.3.2. Las llamadas *Abandonadas* son las que abandonan la cola de espera antes de ser contestadas. Los datos de las llamadas abandonadas no se registran, pero existe un reporte que permite conocer al menos el volumen de llamadas abandonaron, desagregado hasta 15 minutos (detalle en apéndice A.3.2).

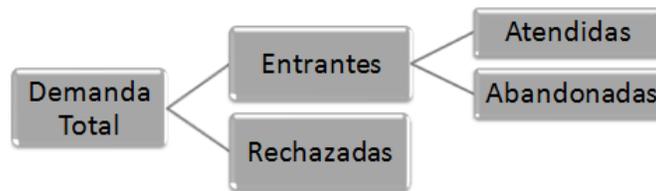


Figura 3.4.1: Diagrama descriptivo de composición de la demanda

*Rechazadas*, son las llamadas que no entran al sistema, esto pasa cuando todas las líneas están ocupadas. El cliente sólo recibe tono ocupado y debe colgar si quiere volver a marcar. Lamentablemente, de éstas últimas no se tiene ningún tipo de registro, pues las centrales no tienen las ventajas necesarias para recopilar datos de éstas llamadas. No existe ningún reporte, ni siquiera del volumen total que estas llamadas representan, ni menos de los detalles. Por lo tanto, no se conoce el volumen de la demanda total del sistema.

Como una manera de estimar el volumen de llamadas rechazadas, se realiza un muestreo durante una semana.

### 3.4.2. Llamadas Entrantes

Un reporte del sistema Contact Center permite obtener los datos del total de llamadas que entran, el número de llamadas atendidas y el número de llamadas abandonadas. Esto permite observar, de manera agregada por día, el volumen de llamadas atendidas y abandonadas.

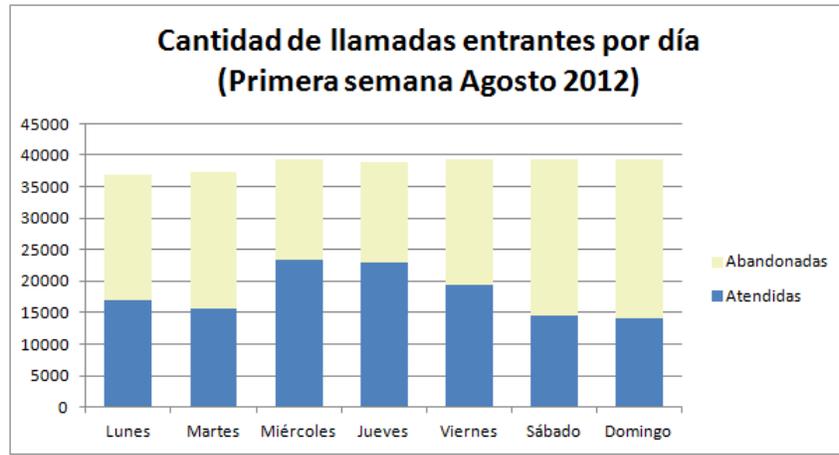


Figura 3.4.2: Volumen de llamadas atendidas y abandonadas por día. Primeras semanas de Agosto.

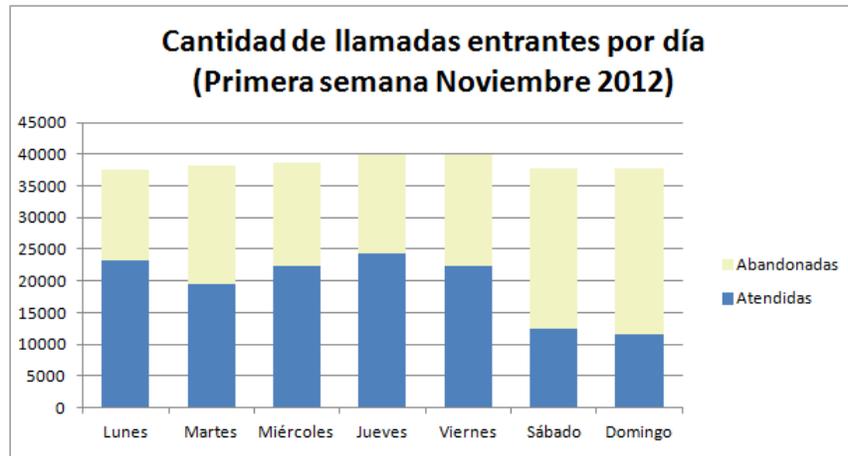


Figura 3.4.3: Volumen de llamadas atendidas y abandonadas por día. Primeras semanas de Noviembre.

Ya que los días de semana son similares entre sí, y lo mismo ocurre los fines de semana. El volumen de llamadas en promedio, es diferente separando días de semana y fines de semana. En la búsqueda de un patrón que defina las llegadas de llamadas, se encuentra que el volumen de llegadas varía de manera similar comparando los días de semana. En promedio, los días hábiles, la tasa de llamadas entrantes cambia durante el día según muestra el siguiente gráfico.

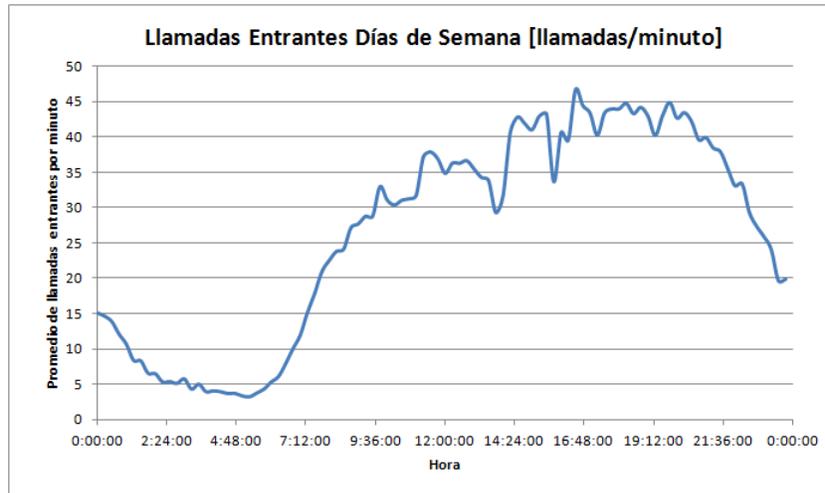


Figura 3.4.4: Volumen de llamadas entrantes por minuto en días hábiles

Durante los días Sábado y Domingo, se observa que la tasa de llegada varia de manera diferente a los días de semana, pero similar entre sí. La cantidad promedio de llamadas entrantes se muestra en el gráfico siguiente.

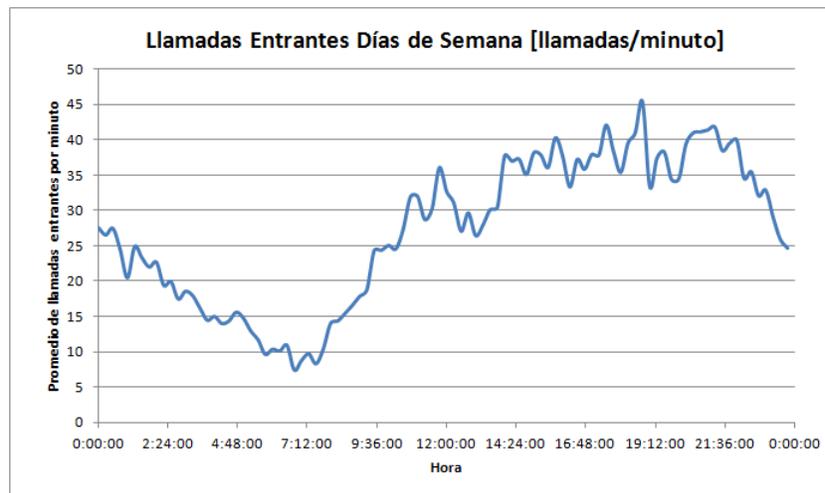


Figura 3.4.5: Volumen de llamadas entrantes por minuto en fines de semana

En promedio, el máximo de llamadas por minuto no supera las 50 llamadas. Esto no significa que no se estén ocupando las 60 líneas, pues cada llamada tiene un tiempo de atención asociado. El tiempo que consume cada llamada se estimará más adelante, luego de describir el tipo de llamadas que llegan al 133.

En la madrugada de los fines de semana, el volumen de llegadas es mayor a la misma hora los días de semana. Otra observación que aparece de los gráficos de llamadas entrantes es que en ambos casos, días de semana y fin de semana, las llamadas bajan al rededor de las 13 horas, en donde se observa un *peak* mínimo.

Los datos se pueden consultar desagregados hasta por 15 minutos, a modo de resumen de estos datos se muestran agregados por día, para una muestra de dos meses, Agosto y Noviembre en el apéndice A.3.2.

### 3.4.2.1. Llamadas Atendidas

Los atributos de las llamadas atendidas se registran desde Marzo de 2011, como se describe en la sección 3.2, algunos se registran durante el proceso de atención y otros los guarda automáticamente el sistema.

Entre los datos que se pueden acceder, se encuentra la tipificación de la emergencia en un reporte, y la clasificación de las llamadas cuando son ineficaces, en otro reporte. Los reportes que indican los detalles de una llamada de emergencia son bastante difíciles de obtener, pues son archivos muy pesados y requiere que se seleccione cada llamada a consultar. Actualmente no existe una manera de filtrar estos reportes, de manera de obtener solamente las llamadas requeridas, que son las de emergencia recibidas a través del 133. Éste último reporte se consiguió solo para la semana del 01 al 07 de Agosto.

Un reporte de todas las llamadas atendidas, muestra la fecha, hora de atención, tiempo de procesamiento y operador que atiende. Se pudo obtener este reporte para un periodo de 17 meses, comenzando del 13 de Marzo de 2011. Sin embargo, este reporte no permite reconocer el tipo de llamada. A partir de este reporte se calcula un promedio de llamadas mensual y semanal, lo que permite comprobar que la semana del 01 al 07 de Agosto es una semana representativa en cuanto a número de llamadas.

Intervalo de Fechas	Promedio [ $\frac{\text{llamadas atendidas}}{\text{día}}$ ]	Desviación Estándar [ $\frac{\text{llamadas atendidas}}{\text{día}}$ ]	Total [días]
13-Marzo-2011 a 30-Septiembre-2012	18158.5	2732.1	333
01-Agosto-2012 a 07-Agosto-2012	18190.7	3769.5	7

Tabla 3.4.1: Cantidad de llamadas atendidas por día.

En el cálculo del promedio de atendidas en los 17 meses, se excluyen los días en que el sistema se haya caído o haya sido desactivado por más de 15 minutos, en esos casos el sistema no registra el total de llamadas entrantes. Durante la semana del 01 al 07 de Agosto de 2012 todas las llamadas contestadas fueron registradas por el sistema, es decir que no se presentaron fallas ni días feriados, en los cuales cambia el volumen de llamadas.

Se debe tener en cuenta que para las llamadas atendidas, se cuenta con los siguientes datos: Fecha y Hora de atención, Tiempo de servicio, Tiempo total de procesamiento (incluye tiempo de espera desde que la llamada se asigna a un operador y éste la contesta), Identificación del usuario, Identificación del operador, Clasificación (llamadas ineficaces) y Tipificación (llamadas de emergencia).

### 3.4.2.2. Llamadas Abandonadas

Una característica importante de los *call centers* es que sus clientes abandonan la cola cuando el tiempo de espera sobrepasa el límite de su paciencia, y en este caso no se conoce cuál es ese tiempo máximo que están dispuestos a esperar.

Por otro lado, la ausencia de datos de llamadas no entrantes o rechazadas por el sistema, exigen profundizar el análisis de los datos disponibles para poder inferir de alguna manera el volumen de demanda que esto representa y caracterizar el comportamiento de abandono de los clientes.

Un análisis preliminar del porcentaje de abandonos en comparación con el nivel de ocupación de los operarios, permite observar, como es natural pensar, que las llamadas abandonadas aumentan a medida que los operadores están más ocupados (ver figura 3.4.18). Esto se debe a que aumentan los tiempos de espera en cola por el sólo hecho de que haya más llamadas.

No es trivial calcular el tiempo máximo que están dispuestos a esperar los clientes, pues los clientes y sus emergencias son distintas, entonces su paciencia no es necesariamente homogénea.

### 3.4.3. Tipos de Llamadas Atendidas

Entre las llamadas atendidas, algunas son inoficiosas y otras son emergencias. Un análisis descriptivo del tipo de llamadas que se atienden, muestra que el porcentaje de llamadas inoficiosas es más alto de lo que se percibe en la institución. En los siguientes gráficos se observa que hay claras diferencias entre el porcentaje de llamadas inoficiosas que llegan los días hábiles y los fines de semana.

A continuación se muestra la proporción de llamadas inoficiosas en comparación con las llamadas que generaron procedimientos.

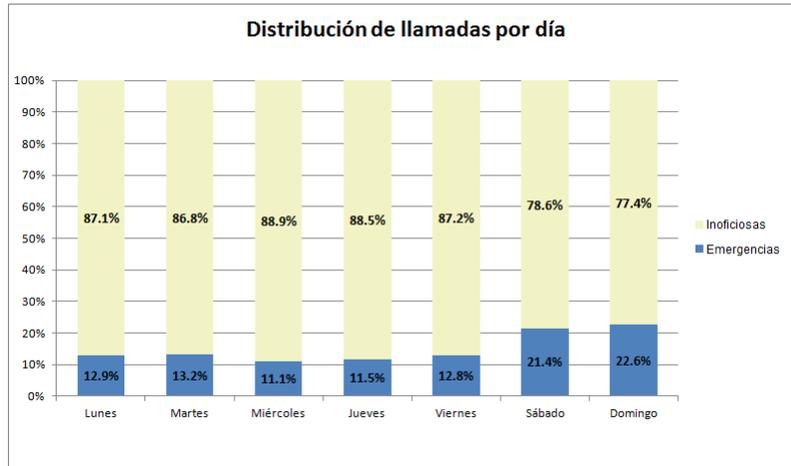


Figura 3.4.6: Porcentaje de llamadas de emergencia e inoficiosas por día.

Considerando los días de semana y fines de semana, como se muestra en la figura 3.4.7, existe una diferencia de casi un 10 % entre el número de inoficiosas que llegan en la semana y los fines de semana. El porcentaje de emergencias que generan procedimientos policiales, es del 12.9 % en la semana y 22 % en fines de semana.

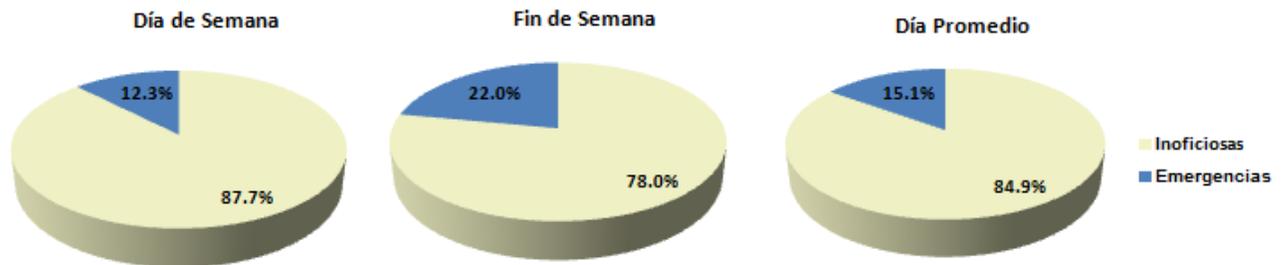


Figura 3.4.7: Porcentajes promedio de llamadas inoficiosas y emergencias por día.

En promedio, el número de llamadas inoficiosas alcanza el 84,9 % un día cualquiera. Esto significa que se atienden aproximadamente 1750 emergencias al día a través del Nivel 133, y 15.450 llamadas atendidas son inoficiosas.

#### 3.4.3.1. Llamadas Inoficiosas en Detalle

Sabemos que el 15,9 % de las llamadas, aproximadamente, son inoficiosas. Pero además cada clase de llamada inoficiosa genera una carga distinta, ya que una broma o una llamada cortada ocupan de manera distinta los recursos del sistema. Además nos interesa saber cómo se distribuye la llegada de llamadas durante el día.

Se calculó el porcentaje promedio de llamadas atendidas cada día según clase de llamada inoficiosa (ver figura 3.4.8). Cabe notar que las llamadas cortadas suman el 50 % de las llamadas inoficiosas en la semana, en cambio el fin de semana sólo completan el 30 %. En el caso de las bromas, pasa lo contrario,

el fin de semana aumentan a un 24%, en promedio y en la semana sólo un 18%, pero se observa una tendencia decreciente en el porcentaje de bromas en los días de semana.

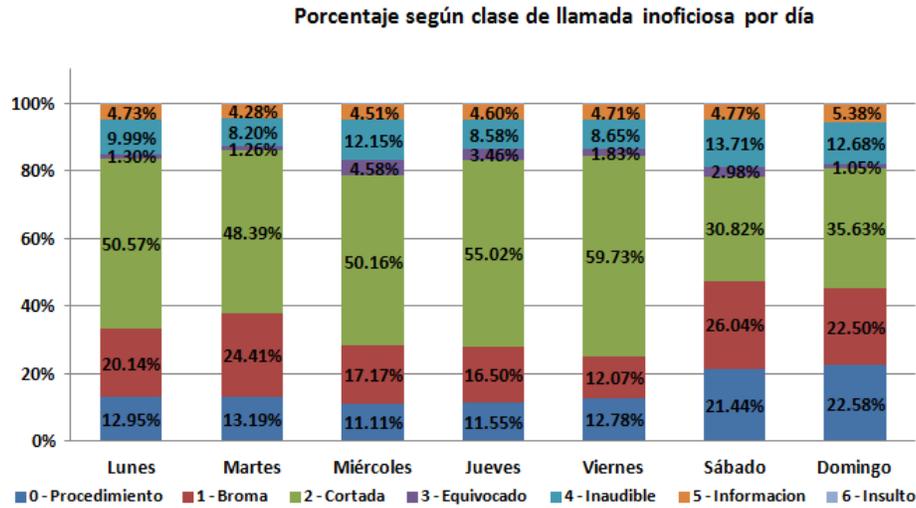


Figura 3.4.8: Porcentajes de cada clase de llamada inoficiosa por día.

Dado que existe una clara diferencia entre la carga de llamadas inoficiosas en días de semana, se procede a analizar estos días por separado, como en la sección anterior.

Podemos observar cómo se distribuyen las llegadas de llamadas inoficiosas durante un día hábil y un fin de semana. La línea gris representa el total de llamados inoficiosos y la roja, el total de procedimientos generados por llamadas de emergencia (ver figuras 3.4.9 y 3.4.10).

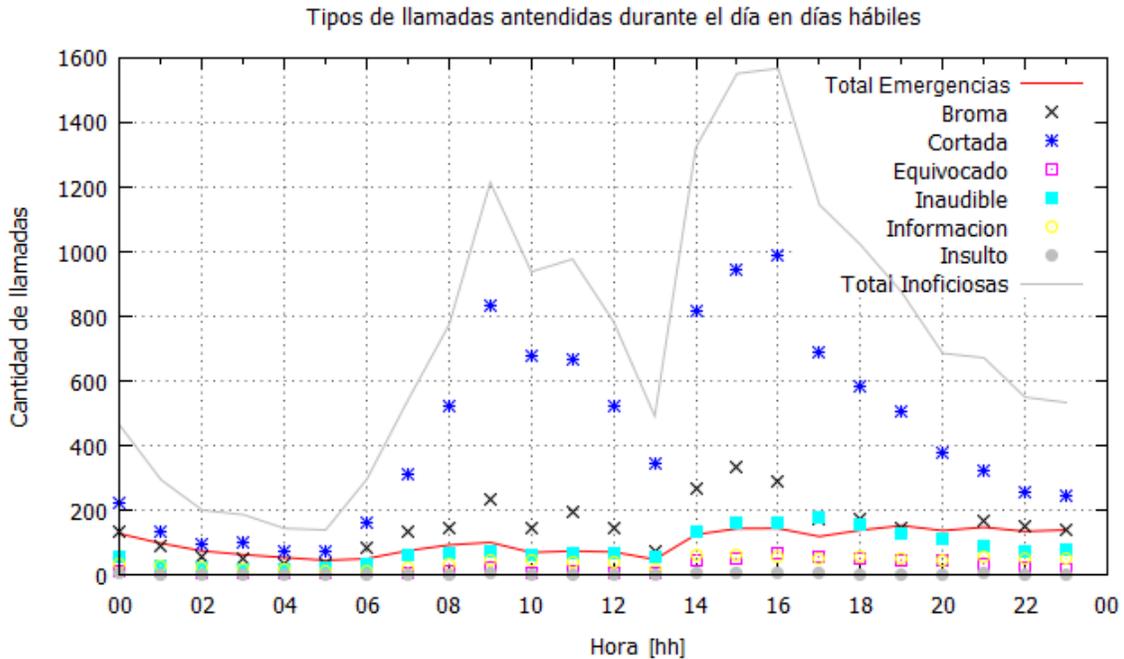


Figura 3.4.9: Tipos de llamadas atendidas durante el día en días hábiles

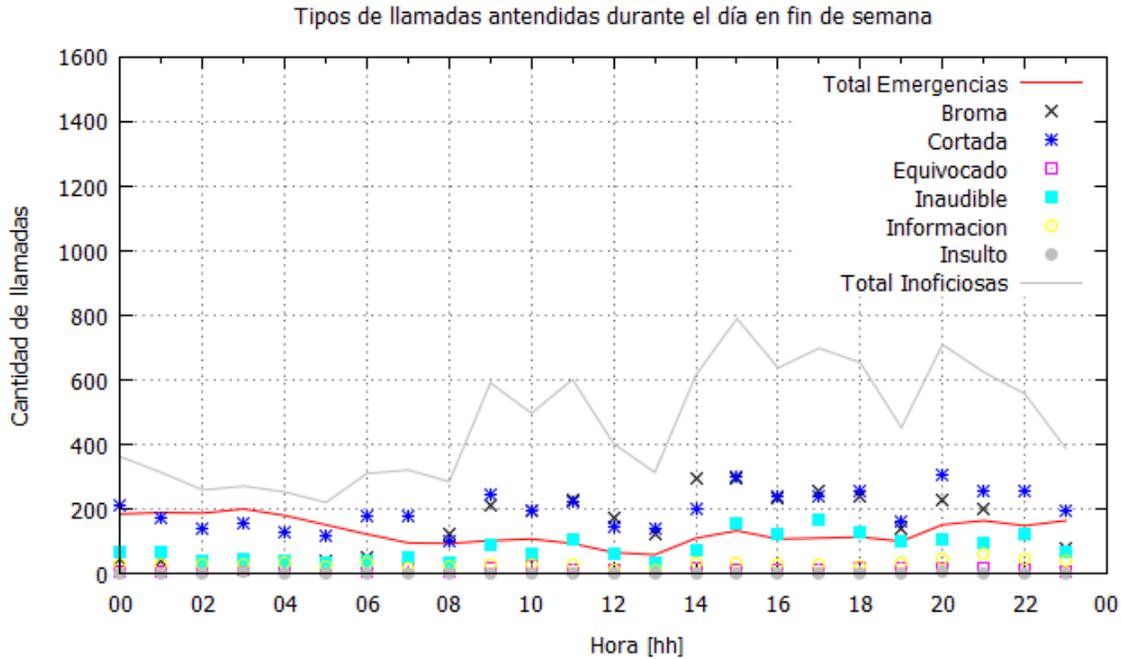


Figura 3.4.10: Tipos de llamadas atendidas durante el día en fin de semana.

### 3.4.3.2. Llamadas de Emergencia en Detalle

Las llamadas que reportan emergencias, y requieren presencia policial inmediata, son las únicas que realmente debieran llegar al 133. Estas llamadas ocupan más tiempo del operador y es muy importante caracterizarlas.

Este análisis es substancial, ya que en CENCO no se ha podido determinar el número de procedimientos que efectivamente se generan en el 133. Recordar que en Alpha I se administran procedimientos provenientes de variadas fuentes. Los procedimientos provenientes del sistema CAD, pueden ser generados en otros servicios, como el 147, 149 y 135, como también generados manualmente en por un operador o supervisor, sin necesariamente recibir una llamada. Fue necesario identificar, usando distintas fuentes de datos, los procedimientos que fueron efectivamente generados por llamadas al 133.

Se realiza un análisis por días de semana y fin de semana separadamente, ya que se observa un comportamiento heterogéneo al igual que en los análisis anteriores. Se observa que el 60 % del total está compuesto por los 10 tipos de emergencias que llegan con más frecuencia. El 80 % está compuesto por 20 tipos diferentes de emergencias. Recordar que existen 103 tipos de emergencias.

Notar en la figura 3.4.11 que más del 16 % de las emergencias atendidas corresponden a la tipificación A12, Alarma Activada. La mayoría de estas llamadas son hechas por telefonistas de empresas privadas, prestadoras de servicios de alarmas particulares.

En segundo lugar, los ruidos molestos es la emergencia más relevante, llegando a un 7.6%. Pero se observa que los fines de semana el porcentaje de estas llamadas se eleva a 15.3 %, siendo ésta emergencia la más frecuente. Es lógico, pues los fines de semana las personas están en sus casas y acusarán ruidos desagradables. En cambio en la semana las personas no están en sus casa, la mayoría está trabajando, y por eso también es esperable que las alarmas se activen con más frecuencia los días hábiles.

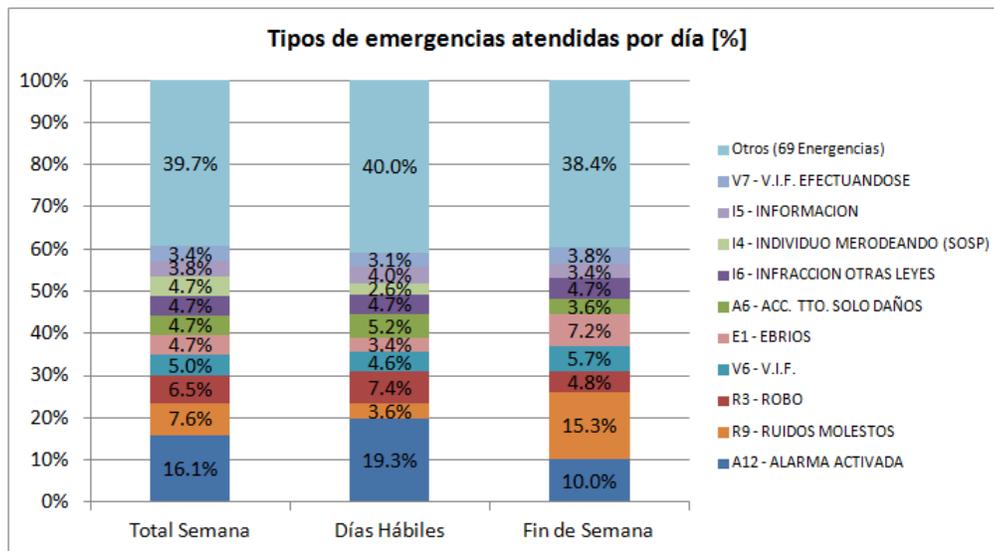


Figura 3.4.11: Tipos de emergencias atendidas por día.

Las llamadas que reportan individuos merodeando, también son mayores los días de semana, mientras al fin de semana aumenta el porcentaje de llamadas que acusan ebrios y violencia intrafamiliar.

#### 3.4.4. Llamadas Rechazadas

Como se explicó al comienzo de este capítulo, no se puede cuantificar precisamente el volumen de llamadas rechazadas. Las llamadas rechazadas son todas las que reciben tono ocupado al marcar el 133, lo que ocurre solamente cuando todas las líneas están ocupadas. La percepción en la institución y en la ciudadanía es que el volumen de estas llamadas es alto.

Con la autorización de Carabineros de Chile se realiza un muestreo, descrito en detalle en el apéndice B.2. El estudio permite obtener una aproximación a los tiempos de espera en cola y el porcentaje de llamados que fueron rechazados. Eso se refiere al tiempo que debe esperar un cliente hasta que le contestan y los llamados que recibieron tono ocupado, respectivamente.

Día	Total Muestra	Tono Ocupado	Promedio Tiempo de Espera en Cola [mm : ss]	Desviación Estándar Tiempo de Espera en Cola [mm : ss]
Semana	159	15.00 %	01:03	01:00
Fin de Semana	80	1.89 %	02:17	01:27

Tabla 3.4.2: Resumen Muestreo.

Durante el fin de semana, un 1.89 % de las llamadas son rechazadas. En la semana, un 15 % de las llamadas fueron rechazadas, pero éstas se concentraron en tres horarios. De 0 a 3 hrs, de 13 a 14 hrs y de 17 a 22 hrs.

### 3.4.5. Tasas de Llegada

Ya que no se conoce la demanda total, porque faltan los datos de las llamadas no entrantes, no se puede encontrar la tasa de llegada de llamadas con exactitud. Sin embargo, se conoce el volumen de llamadas entrantes y la hora de atención de cada llamada, se tiene una estimación de las llamadas abandonadas y se sabe que el volumen de las llamadas no entrantes es despreciable (ver sección 3.4.4). Por lo tanto, la tasa de llegadas se asume igual a la tasa de atención de llamadas.

El tiempo que transcurre entre una atención y otra es una de las medidas importantes de la demanda. Si aquel tiempo es corto, muchas llamadas entran en un cierto periodo, entonces la demanda es alta. Si el tiempo es largo entre la llegada de una llamada y otra, entonces se puede suponer que la demanda es baja, pues la frecuencia es baja. Notar que los datos no contienen hora de llegada de las llamadas, si no solamente la hora de atención.

Se grafican los tiempos entre atención de llamadas para comparar visualmente si los días se parecen entre sí, y si hay diferencias con los fines de semana y festivos. Efectivamente, se encuentran similitudes entre las llegadas de llamadas los días de semana. Los días festivos se comportan como los días de fin de semana. A partir de este tipo de gráficos, más adelante, se definen horarios del día con distintos niveles de demanda.

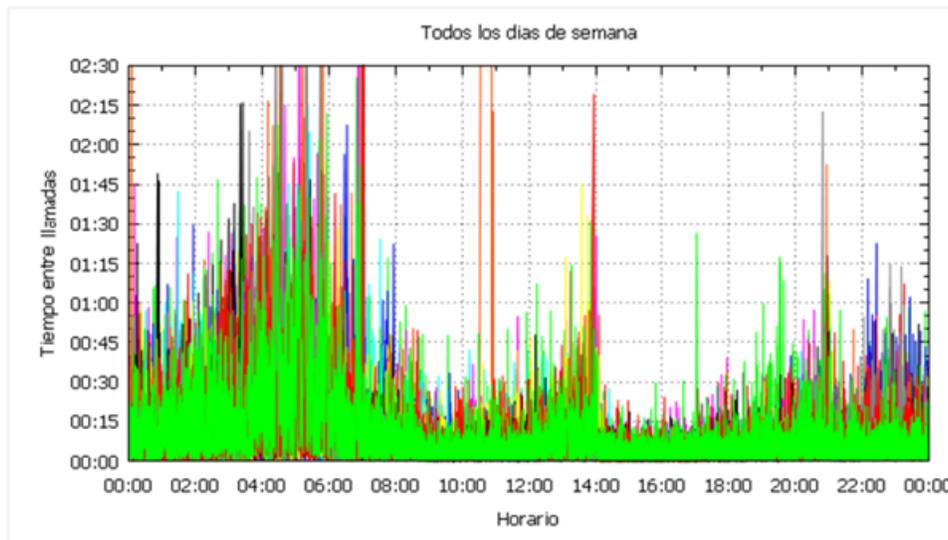


Figura 3.4.12: Tiempo entre atención de llamadas días de semana.

La figura 3.4.12 muestra el gráfico de tiempo entre atención de llamadas de todos los días de semana del mes (excluyendo un día festivo). Un determinado color representa los datos de un cierto día, entonces se puede ver cómo los distintos días presentan distribuciones parecidas.

Se debe advertir que el tiempo entre llamadas que se observa no representa exactamente el tiempo entre llegada de las llamadas, pues los clientes pueden haber estado esperando en cola. Luego, cuando la llamada se asigna a un operador, queda registrada esa hora (hora de llegada al operador). No hay registros de la hora de llegada de las llamadas al sistema. Las llamadas de clientes rechazados por el sistema y de los que abandonan la cola, no dejan registro de llegada.

Teóricamente, como propone Erlang, se podría asumir una distribución de Poisson para los tiempos entre llegadas de usuarios al sistema, pero no se puede asumir una distribución de Poisson del tiempo entre llamadas que estamos observando porque esta variable no es independiente, sólo cuando un operador está desocupado puede entrar otra llamada.

Los mismos datos nos permiten observar la cantidad de llamadas que son atendidas por minuto, como muestra el siguiente gráfico. Notar que es la distribución inversa a la del gráfico anterior.

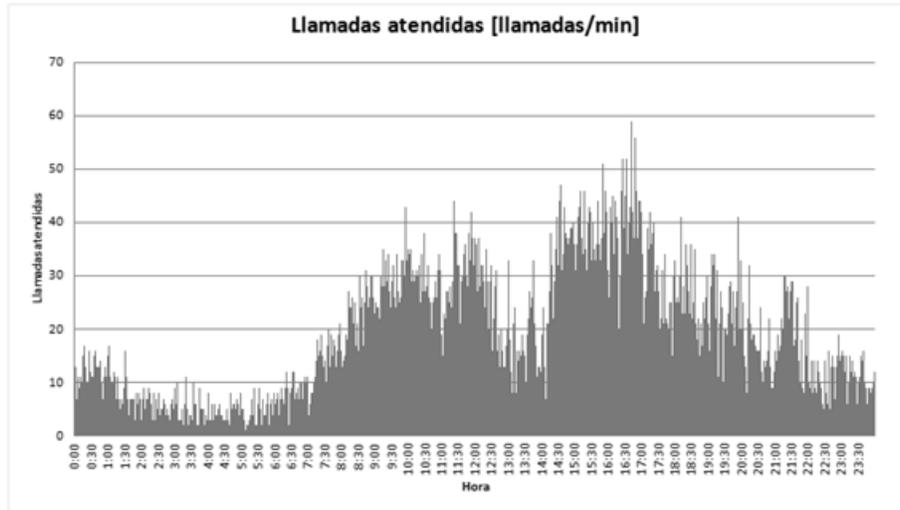


Figura 3.4.13: Llamadas atendidas por minuto

### 3.4.6. Tiempos de Servicio

El tiempo de servicio es un dato presente en los registros de llamadas atendidas. Se aprecia que existe un claro patrón de comportamiento, dejando de lado las diferencias entre tipos de llamadas o el operador que la atiende.

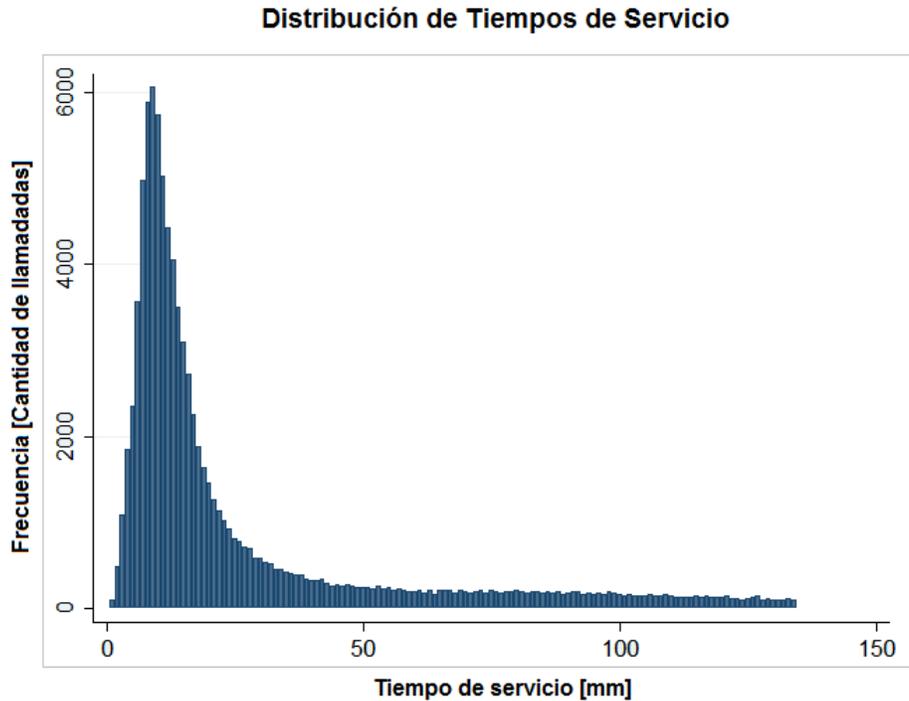


Figura 3.4.14: Frecuencias vs. Duración de llamadas.

Como explican Brown et al. [3], la mayoría de las aplicaciones de teoría de colas en *call centers* asumen tiempos de servicio exponencialmente distribuidos. Viendo la figura 3.4.14, es claro que la distribución de tiempos de servicio no es exponencial. En efecto, luego de separar las llamadas con tiempos más cortos (80% de las llamadas), se puede probar que el tiempo de servicio se ajusta a una distribución Lognormal (ver apéndice B.1.1).

Conocer la distribución de tiempos de servicio de las llamadas, sin importar el tipo de llamada, nos permite modelar usando teoría de colas y simulación de forma simplificada, como si tuviéramos un solo tipo de llamada y comparar los resultados, como se ve más adelante en el capítulo 4.

### 3.4.6.1. Tiempos de Servicio por Tipo de Llamada

Dado el gran volumen de llamadas cortadas, inaudibles y otras inoficiosas que toman muy pocos segundos del operador, es esperable que exista una gran diferencia entre los tiempos de servicio de éstas y los de las llamadas de emergencia.

Tipo de Llamada	Tiempo Promedio de Servicio [ <i>segundos</i> ]	Desviación Estándar [ <i>segundos</i> ]
Inoficiosa	29.273	59.176
Emergencia	203.431	125.084

Tabla 3.4.3: Tiempos de servicio promedio según tipo de llamada.

En la tabla 3.4.3 se puede observar que el tiempo de servicio promedio de las llamadas de emergencia es siete veces mayor que el de las llamadas inoficiosas, además la desviación estándar es de mayor magnitud. Esto se debe a que hay algunas llamadas de emergencia muy frecuentes, que tienen bajos tiempos de servicio y otras menos comunes con muy altos tiempos de servicio, como se muestra en la tabla 3.4.4.

Tipo de Emergencia	Frecuencia [ ]	Tiempo de Servicio Promedio [ <i>mm : ss.sss</i> ]
Alarma Activada	19.3	02:30.041
Robo	7.4	03:36.001
Imputado en Tienda Comercial	2.9	02:19.891
Robo Armado Efectuándose	0.1	07:19.770
Estafa	0.5	05:42.389

Tabla 3.4.4: Ejemplo: Tiempo promedio de servicio de llamadas de emergencia.

Las Alarmas Activadas son la emergencia reportada más frecuente, con un tiempo de servicio promedio de 2 minutos y medio. Luego viene el Robo, aunque con un tiempo de servicio mucho mayor. Este ejemplo indica que no se puede simplemente resumir el tiempo de servicio de las llamadas de emergencia como un promedio.

Para modelar el comportamiento del sistema asumiendo que existen varios tipos de llamadas, se utiliza entonces una medida ponderada de los tiempos de servicio.

### 3.4.7. Operadores Conectados

Los operadores son los agentes o servidores del sistema, como se describió en la sección 3.1.3. El número de operadores conectados en el sistema es un factor determinante de la demanda atendida. No así, de la demanda total, pues la llegada de llamadas es independiente. De todas maneras, es un parámetro de servicio importante para describir el sistema y hacer posteriores cálculos según la teoría de colas.

Como se explicó en 3.4.7, el número de operadores conectados en cada instante, no tiene relación directa con el número de turnos programados. A continuación se muestra la distribución del número de operarios promedio que se encuentra en un día hábil.



Figura 3.4.15: Promedio de operadores por minuto en días hábiles.

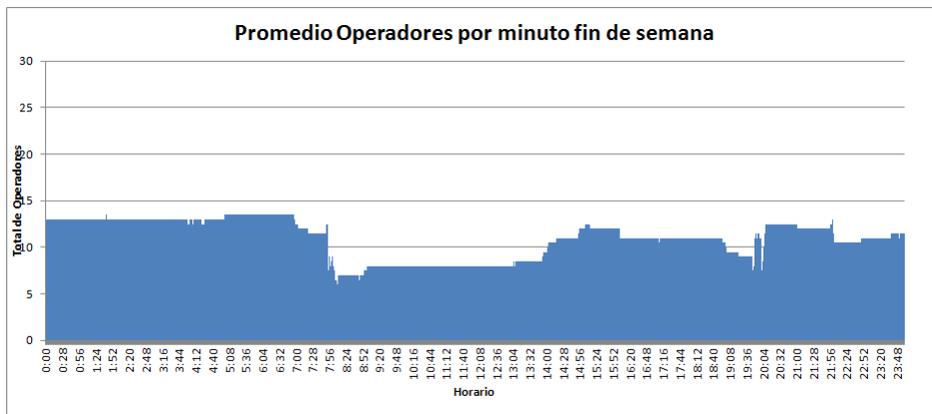


Figura 3.4.16: Promedio de operadores por minuto en fin de semana.

El menor número de operadores conectados los fines de semana, se explica porque la dotación es más baja. Esto es por las leyes que rigen los contratos de los C.P.R. y el gran número de mujeres con fuero maternal y pre-natal (ver en detalle la sección 3.1.3).

### 3.4.8. Resumen

A partir del análisis de los patrones observados se puede determinar, a simple vista, que existen tramos horarios diferenciados. El siguiente gráfico superpone la cantidad de llamadas entrantes, llamadas atendidas y operadores conectados (tal como se mostró por separado en las figuras 3.4.13, 3.4.4 y 3.4.15).

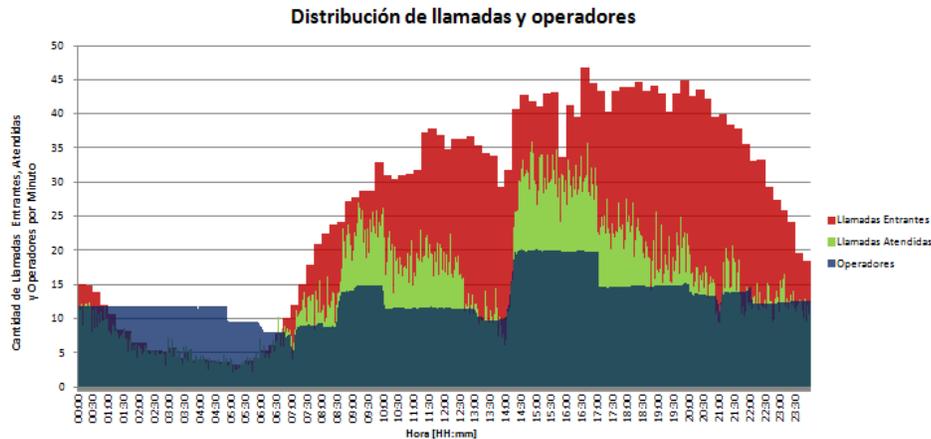


Figura 3.4.17: Distribución de llamadas entrantes, atendidas y cantidad de operadores un día de semana.

Las llamadas entrantes menos las atendidas corresponde al número de llamadas abandonadas. Esta cantidad aumenta en los horarios de 10 a 14 hrs. y de 16 a 0 hrs. Al disminuir el número de operadores aumentan las llamadas abandonadas. Esto podría sugerir que existe una relación importante entre los tiempos los abandonos y el número de operadores. Se puede observar que en el horarios en que aumenta el número de operadores (a las 9 y a las 14.30), aumentan las llamadas atendidas y disminuyen los abandonos.

El poder diferenciar horarios permite simplificar el análisis matemático a través de la teoría de colas y simulación, ya que estos métodos requieren conocer los patrones de comportamiento del sistema, que varían durante el día. Para identificarlos, se divide el día en cinco horarios.

Horario	Tramo Horario
1	00:00 a 06:00
2	06:00 a 13:00
3	13:00 a 14:00
4	14:00 a 17:00
5	17:00 a 00:00

Tabla 3.4.5: Tramos horarios diferenciados en días de semana.

Para poder modelar analíticamente el sistema, usando la teoría de colas, se deben conocer los patrones de llegada de llamadas, tiempo de servicio y número de operadores. Para esto se divide el día en cinco horarios, y se buscan los patrones que se ajustan a cada uno. Se puede estudiar el comportamiento del sistema en cada horario por separado y también su funcionamiento las 24 hrs. Para aplicar la teoría de colas se deben asumir tasas Markovianas y existe un solo tipo de llamadas.

Los parámetros que se encuentran detallados la tabla 3.4.6, son los promedios ponderados que describen la distribución de llegadas Poisson y tasa de servicio Exponencial. Recordar que el tiempo entre llamadas entrantes, para efectos prácticos será considerado como el tiempo entre llegadas.

Horario	N° Operadores	Tiempo entre Llegadas [segundos]	Tasa de Llegada [ $\frac{\text{llamadas}}{\text{segundo}}$ ]	Tiempo de Servicio [segundos]	Tasa de Servicio [ $\frac{\text{llamadas}}{\text{segundo}}$ ]
1	11	8.930	0.112	75.018	0.013
2	11	2.325	0.430	44.733	0.022
3	10	1.806	0.554	45.108	0.022
4	18	1.452	0.689	45.110	0.022
5	14	1.639	0.610	61.204	0.016

Tabla 3.4.6: Tasas Generales de Llegada y de Servicio.

Para una modelar un poco mejor el comportamiento real del sistema, usando simulación, se deben considerar al menos dos tipos de llegadas: las llamadas inoficiosas y las emergencias, ya que hay una gran diferencia entre la carga que representan cada una para el sistema. El número de operadores es un parámetro independiente. Los tiempos entre llegadas y de servicio en cada horario son los siguientes.

Horario	Tiempo entre Llegadas Emergencias [segundos]	Tasa de Llegada Emergencias [ $\frac{\text{llamadas}}{\text{segundo}}$ ]	Tiempo de Servicio Emergencias [segundos]	Tasa de Servicio Emergencias [ $\frac{\text{llamadas}}{\text{segundo}}$ ]
1	36.765	0.027	202.917	0.005
2	26.587	0.038	203.417	0.005
3	20.579	0.049	214.045	0.005
4	16.807	0.059	212.875	0.005
5	10.044	0.100	212.052	0.005

Tabla 3.4.7: Tasas de Llegada y Servicio de llamadas de Emergencia.

Horario	Tiempo entre Llegadas Inoficiosas [segundos]	Tasa de Llegada Inoficiosas [ $\frac{\text{llamadas}}{\text{segundo}}$ ]	Tiempo de Servicio Inoficiosas [segundos]	Tasa de Servicio Inoficiosas [ $\frac{\text{llamadas}}{\text{segundo}}$ ]
1	11.795	0.085	33.985	0.029
2	2.547	0.393	29.530	0.034
3	1.980	0.505	28.855	0.035
4	1.589	0.629	29.245	0.034
5	1.959	0.511	31.786	0.031

Tabla 3.4.8: Tasas de Llegada y Servicio de llamadas Inoficiosas.

Ya que no se sabe cuánta paciencia tiene una persona que llama al 133 para reportar una emergencia, es muy importante buscar parámetros que permitan estimar ese tiempo. Más adelante se muestra cómo la paciencia tiene relación con las llamadas abandonadas.

La curva que muestra la figura 3.4.18 entrega una idea de la paciencia que los clientes tienen, según datos reales. Si la curva es más pronunciada, la paciencia es menor. A medida que la pendiente disminuye, la paciencia de los clientes de un sistema, aumenta.

Dado que no se conoce el tiempo de paciencia de los clientes, el gráfico 3.4.18 puede servir para comparar la paciencia estimada a través del modelamiento con la real. Esto permitiría verificar, más adelante, si la paciencia de los clientes aumenta o disminuye al cambiar el número de operadores o incluso el proceso mismo de atención de llamadas.

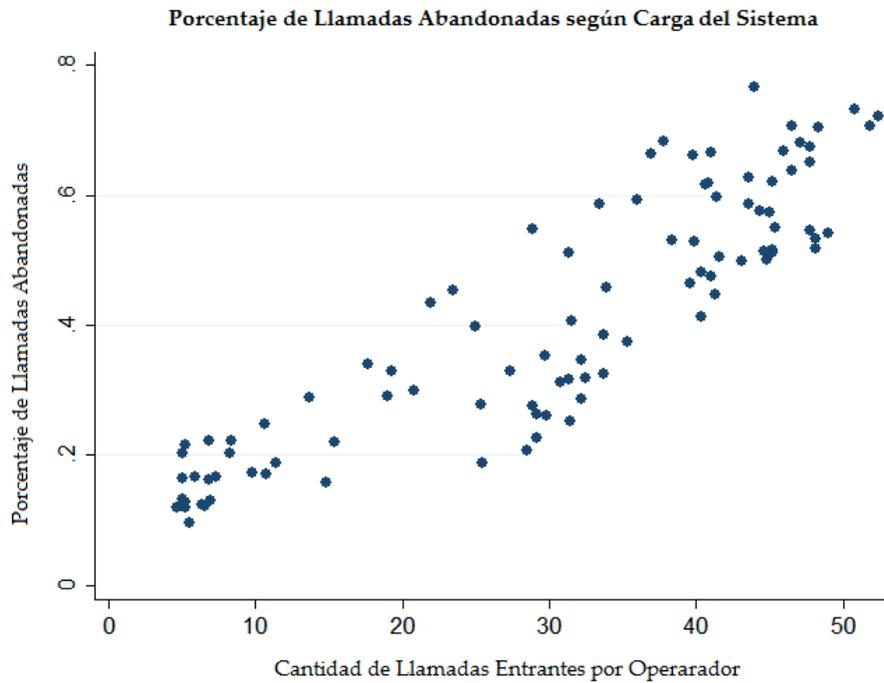


Figura 3.4.18: Porcentaje de llamadas abandonadas versus nivel de ocupación.

El comportamiento de abandono de los clientes también se debe establecer para poder estudiar el comportamiento real del sistema. El único dato que permite acercarse a la tasa de abandono es el volumen de abandonos que entrega un reporte del Contact Center. Estudiando estos datos en cada horario, se puede aplicar el modelo Erlang-A, de teoría de colas, para estimar la tasa de abandonos  $\theta$ , asumiendo que el tiempo de paciencia de los clientes sigue una distribución exponencial de tasa  $\theta$ .

Horario	Llamadas Abandonadas [%]
1	21.99
2	43.97
3	72.07
4	39.68
5	60.58

Tabla 3.4.9: Porcentaje de llamadas abandonadas en días hábiles por horario.

Los abandonos, entonces, se observan en cada horario por separado. El promedio de llamadas abandonadas, en un día de semana, es 48.65 % . En el próximo capítulo se desarrolla el estudio que permite estimar los tiempos de espera usando todos estos parámetros, y con eso, determinar los niveles de servicio actuales.

## Capítulo 4

# Análisis y Evaluación

En este capítulo se utilizan los parámetros hallados para modelar el comportamiento real del sistema y así poder medir el nivel de servicio actual. Se debe recordar que muchos datos relevantes son desconocidos.

Después de estudiar el comportamiento del sistema en detalle, en la sección 3.4, se reconoce que la teoría de colas no es suficiente para analizar un proceso tan complejo. Las tasas de llegada y servicio son variables durante el día, lo mismo ocurre con la cantidad de operadores, además hay varios tipos de llamadas. Un operador no se dedica a atender solo emergencias o solo inoficiosas, atiende la llamada que el sistema le asigna.

Por otro lado, la teoría de colas es ideal para estimar tiempos de espera y cantidad de clientes cuando  $\rho = \lambda/\mu c \leq 1$ , si se asumen las condiciones apropiadas. Cuando la condición se cumple, las medidas se comparan con las que arroja la simulación, ya que al construir los modelos se puede haber cometido algún error. De esta manera se construye un modelo de simulación válido.

Si bien la teoría de colas asume un solo tipo de llegadas, es un buen punto de partida y una excelente herramienta para estimar el comportamiento de abandonos, dados los datos disponibles. Las tasas de abandono son determinantes en los tiempos de espera y los niveles de servicio resultantes.

Una vez obtenida la tasa de abandono, la simulación permite flexibilizar el análisis mediante el uso de expresiones más complejas para representar el cambio en los patrones de llegada y servicio. Esto permite reflejar el funcionamiento real del sistema, reconociendo cada llamada con muchos de sus atributos.

Finalmente, en este capítulo se evalúa el rendimiento y nivel de servicio del Fono 133 comparándolo con los KPI de los estándares internacionales descritos en la sección 2.1.2.

### 4.1. Validez de la Simulación

Mediante la validación de la simulación, se descartan errores de modelamiento. En esta sección hacen varias pruebas aplicando modelos de teoría de colas como M/M/1 y M/M/c, para demostrar que se refleja la realidad. Se mejoran los modelos incrementalmente hasta llegar a aplicar un modelo con comportamiento de abandono, que permite estimar los tiempos de espera de los clientes. Se simula el funcionamiento del sistema bajo las mismas condiciones de cada modelo analítico aplicado, y los resultados se comparan para algunos horarios.

Las medidas de rendimiento, que se obtienen y comparan en este capítulo, son: promedio de personas en cola  $L_q$ , tiempo de espera en cola  $W_q$ , promedio de clientes en el sistema  $L$  y el tiempo total en el sistema  $W$ , obtenidas con teoría de colas, calculadora Erlang-X y simulación.

En la sección 4.3 se comparan los resultados obtenidos tras la aplicación de los distintos modelos.

Se observa cómo mejoran las estimaciones a través simulación, que es capaz de modelar sistemas más completos, que no necesariamente cumplen las condiciones y supuestos de la teoría de colas. El proceso principal de la simulación se modela en Arena, como muestra la figura 4.1.1.

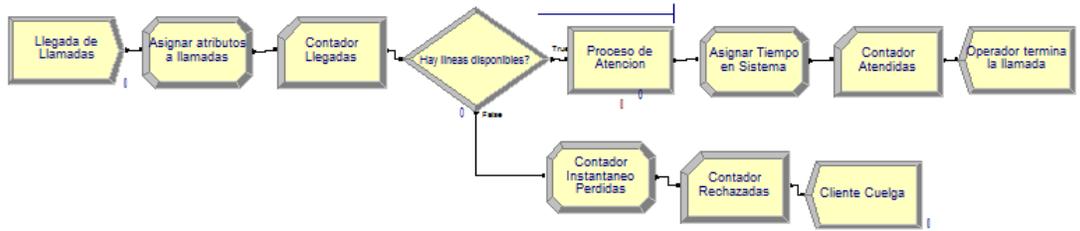


Figura 4.1.1: Modelo en Arena para simulación del sistema M/M/c y M/M/c/K.

#### 4.1.1. Aplicación del Modelo $M/M/c$

Se debe recordar, que la notación de Kendall asume los valores predeterminados de Capacidad del sistema, Población total de clientes y Disciplina de la Cola, cuando estos no se especifican (Infinita, Infinita y FIFO, respectivamente).

Dado que la capacidad del sistema es infinita, en este caso, ninguna llamada es rechazada por el sistema, todas pasan por el proceso de atención (ver figura 4.1.1).

La teoría de colas exige, para aplicar este modelo y conseguir estimar correctamente las medidas de rendimiento, que el nivel de utilización de los servidores cumpla  $\rho \leq 1$ . Además este tipo de modelo aplica para servidores que atienden un solo tipo de llamadas. Por lo tanto, Se utilizan las tasas generales de llegada y servicio, como muestra el resumen de la sección anterior, en la tabla 3.4.6.

**TEORÍA DE COLAS** Se usan las fórmulas descritas en la sección 2.3.2.6, acorde a las reglas del modelo, se asume que existe un solo tipo de llamadas. Éstas llegan de acuerdo a una distribución de Poisson de tasa  $\lambda$ , son atendidas a una cierta tasa exponencial  $\mu$  y los  $c$  operadores son idénticos.

##### HORARIO 1

Sólo en este horario es posible aplicar este modelo analítico, pues es el único en que se cumple la condición de estabilidad  $\rho \leq 1$ .

Sea  $\lambda = 0,112 \left[ \frac{\text{llamadas}}{\text{segundo}} \right]$ ,  $\mu = 0,0133 \left[ \frac{\text{llamadas}}{\text{segundo}} \right]$  y  $c = 11$  *servidores*. Primero se calcula  $\rho$  para determinar si el sistema, bajo estas condiciones, alcanza el estado estacionario.

$$\rho = \frac{\lambda}{\mu c} = 0,76 \text{ es el nivel de utilización de los servidores.}$$

Esto significa que los servidores están ocupados el 76 % del tiempo. Dado que  $\rho \leq 1$ , el sistema puede converger al estado estacionario y entonces se puede aplicar el modelo  $M/M/c$  (o Erlang-C) y ley de Little para calcular las medidas de rendimiento. La probabilidad de que una persona tenga que esperar en cola según el modelo, en estado estacionario,  $C\left(c, \frac{\lambda}{\mu}\right) = 0,311 = 31,1\%$ . Entonces, se calculan las medidas de rendimiento.

$L_q[\text{clientes}]$	$W_q[\text{segundos}]$	$L[\text{clientes}]$	$W[\text{segundos}]$
1.01	8.99	9.41	84.01

Tabla 4.1.1: Resultados de Aplicación del Modelo M/M/c en Horario 1.

La calculadora Erlang-X también da como resultado  $W_q = 8,99$  *segundos* (captura de pantalla en apéndice C.1).

## SIMULACIÓN

**HORARIO 1**

Se ejecuta simulación usando los mismos parámetros mostrados en la tabla 3.4.6, para el horario 1. Los resultados que se obtienen (detallados en apéndice D.1), en comparación con los obtenidos con el modelo M/M/c y calculadora Erlang-X, son prácticamente idénticos. Se resumen, y comparan en 4.3.

**4.1.2. Aplicación del Modelo M/M/c/K****TEORÍA DE COLAS**

Al igual que en el modelo anterior, las fórmulas para el cálculo de las medidas de rendimiento, en cuanto a tiempos de espera, tienen sentido sólo cuando  $\rho = \lambda/c\mu \leq 1$ , como se explica en la sección 2.3.2.7.

**HORARIO 1**

Sea  $\lambda = 0,112[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $\mu = 0,013[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $c = 11$  *servidores* y  $K = 60$  *líneas*.

Los resultados son idénticos a los del modelo Erlang-C. El número de líneas no afecta el comportamiento del sistema, ya que el sistema alcanza el estado estacionario ( $\rho \leq 1$ ). Por lo tanto los resultados de tiempos de espera son idénticos a los que se obtiene con el modelo M/M/c, mostrados en la tabla 4.1.1.

**HORARIO 2**

Sea  $\lambda = 430[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $\mu = 0,022[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $c = 11$  *servidores* y  $K = 60$  *líneas*. En este horario, la tasa de utilización  $\rho = 1,75 > 1$ . Un sistema que funciona con estos parámetros estaría sobreutilizado, y no alcanzaría estado estacionario, sino que se mantendría con personas en cola infinitamente.

A pesar de aquello, se puede calcular el promedio de personas en la cola y en el sistema, usando teoría de colas, y se obtiene:

$Lq[\text{clientes}]$	$L[\text{clientes}]$
47.67	58.67

Tabla 4.1.2: Resultados Modelo M/M/c/K en Horario 2.

No se ha desarrollado aún, un modelo que permita calcular tiempos de espera para sistemas inestables, como este. Sin embargo, si se tratara de un sistema estable, solo con el fin de comparar los resultados con los de otros modelos, los tiempos de espera serían:  $W_q = 110,80$  *segundos* y  $W = 155,54$  *segundos*.

La calculadora Erlang-X da como resultado  $W_q = 193,82$  *segundos* (captura de pantalla en apéndice C.2.1).

## SIMULACIÓN

**HORARIO 1**

Los resultados son los mismos que en teoría de colas. El número máximo de personas que se encuentran en el sistema en un determinado momento es 42 (detalles en apéndice D.1.3), por lo tanto no hay llamadas rechazadas por capacidad del sistema, en este horario. No hay momentos en que estén todas las líneas ocupadas. Luego, el número de líneas no afecta el comportamiento del sistema en el horario 1.

**HORARIO 2**

A modo de prueba, se simula el horario 2 con capacidad finita  $K$ . Las salidas de la simulación se detallan en el apéndice D.2.

$L_q[\text{clientes}]$	$W_q[\text{segundos}]$	$L[\text{clientes}]$	$W[\text{segundos}]$
47.58	193.21	58.57	237.92

Tabla 4.1.3: Resultados Simulación con Capacidad Finita en Horario 2.

Simulando el día completo, con los distintos parámetros y distribuciones de éstos, se obtiene un porcentaje de llamadas rechazadas de 47.42 %. Los resultados más importantes, de esta simulación completa, se encuentran en el apéndice D.3.

Se observa que las medidas de  $L_q$  y  $L$  son casi idénticas y se obtiene  $W_q$  similar al que entrega la calculadora Erlang-X, que usa supuestos Markovianos. La simulación cumple claramente la relación entre tiempos de espera y atención  $W_q + \text{Tiempo de servicio} = W$ .

## 4.2. Estimación de la Paciencia

El modelo Erlang-A se utiliza aquí como una herramienta para estimar los tiempos de paciencia de los clientes en cada horario. Luego, se simula considerando un modelo Markoviano con abandonos determinados por una distribución de la paciencia estimada. Esto permite, a partir de la simulación, estimar las medidas de rendimiento de un sistema con capacidad finita en cola y con abandonos, algo que no se puede hacer analíticamente con teoría de colas.

**TEORÍA DE COLAS** Se aplica el el Modelo  $M/M/c + M$  o Erlang-A, cuyas fórmulas se basan en los resultados del modelo Erlang-C mostrados en 4.1.2. A modo de ejemplo, se muestra el cálculo de la paciencia promedio en un horario, y luego se presentan los resultados en todos los horarios.

### HORARIO 2

Sea  $\lambda = 430[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $\mu = 0,022[\frac{\text{llamadas}}{\text{segundo}}]$ ,  $c = 11 \text{ servidores}$ ,  $K = 60 \text{ líneas}$  y un porcentaje de abandonos del 43,97 %.

Aplicando el modelo, se obtiene que la paciencia

$$R = \frac{110,80}{0,4397} = 251,99 \text{ segundos} = 4,17 \text{ minutos}$$

La calculadora Erlang-X, bajo los mismos parámetros, arroja  $R = 241,73 \text{ segundos} = 4,01 \text{ minutos}$  (captura en apéndice C.2).

### PACIENCIA EN CADA HORARIO

Haciendo lo mismo en todos los horarios, se obtiene una aproximación de los tiempos de paciencia de los clientes en cada horario.

Horario	Pacencia por Erlang-A [segundos]	Pacencia por Erlang-X [segundos]
1	40.90	-
2	252.00	241.73
3	123.64	156.81
4	13.70	143.52
5	3.93	149.85

Tabla 4.2.1: Tiempos de Paciencia Estimados usando Teoría de Colas

La paciencia estimada por el modelo Erlang-A difiere de los resultados que arroja la calculadora Erlang-X. Esto se asocia a que el modelo A no considera que existe capacidad finita en el sistema, es decir, no hay llamadas rechazadas. Además se basa en el modelo Erlang-C, que tiene error al predecir cuando  $\rho > 1$ , y no considera abandonos. Entonces los tiempos de paciencia estimados por el modelo Erlang-A, están generalmente subestimados.

En lo que sigue, se simulan los abandonos con una paciencia distribuida exponencialmente de media  $1/\theta_i$ , con  $i = 1, 2, 3, 4, 5$  los horarios, usando los resultados de la calculadora Erlang-X.

#### 4.2.1. Validez de la Paciencia Estimada

##### SIMULACIÓN

Esta vez se simula el sistema considerando el comportamiento de abandono en la cola de atención y capacidad finita. Los resultados de esta simulación permiten determinar si la paciencia estimada, en conjunto con los parámetros estimados del sistema, arrojan resultados coherentes con la realidad. De esta manera se puede apreciar si la paciencia está sobre estimada o subestimada en cada horario.

Este tipo de modelo, en teoría de colas, sería como un modelo  $M/M/c + M/K$ , sin embargo aún no se ha desarrollado este tipo de modelo analítico, y numéricamente el cálculo sería lento. La simulación, entonces, permite modelar sistemas más complejos de lo que la teoría de colas puede abordar. Por otro lado, la simulación considera parámetros que van cambiando durante el día, por lo tanto incluye el comportamiento del sistema en estado transiente.

De la simulación se extrae un reporte que permite analizar los abandonos en cada horario, luego se comparan los resultados de abandonos obtenidos en cada horario con los datos reales. Notar que la simulación, en este caso, considera los estados transientes, es decir, se simula el funcionamiento real, con todos los cambios de distribuciones que ocurren en cada horario.

Horario	Abandonos Simulación[%]	Abandonos Reales[%]
1	15.62	21.99
2	44.41	43.97
3	49.69	72.07
4	44.64	39.68
5	62.55	60.58

Tabla 4.2.2: Porcentaje de abandonos por horario.

Para ajustar los porcentajes de abandono resultantes a los reales, se prueban distintos tiempos de paciencia. Se requiere aumentar levemente los abandonos en promedio al día, concentrándose en los horarios 1 y 3.

Varias pruebas de tiempos de paciencia disminuida en aquellos horarios, no arrojan aumentos significativos en los abandonos. Al disminuir la paciencia, los tiempos de espera disminuyen y esto mismo provoca menos abandonos. Por ejemplo en el horario 3, se disminuye la paciencia a 78 segundos y aumentan a 50.5 % los abandonos, si se disminuye la paciencia a 10 segundos, los abandonos llegan al 49.76 % en el mismo horario, disminuyen a pesar de que la paciencia sea menor. Se determina que la paciencia no es la causa del error de estimación de los abandonos.

Algo que se observa en el Nivel 133, es que en la realidad hay menos operadores conectados atendiendo, que lo que muestran los datos, ya que ellos tienen la capacidad de ponerse en estado No Listo cuando no pueden contestar o salen de su módulo. Asumiendo que los operadores se ponen en estado No listo en los horarios 1 y 3, que son horarios más duros de trabajo (madrugada y almuerzo, respectivamente), se comprueba, que disminuyendo en 10 % la cantidad de operadores en dichos horario, mejora la estimación del volumen llamadas abandonadas. Luego, la siguiente tabla asume  $c_1=10$  y  $c_3=9$ , manteniendo la paciencia estimada. Los abandonos aumentan a niveles más parecidos a los reales.

Horario	Abandonos Simulación[%]	Abandonos Reales[%]
1	21.11	21.99
2	45.33	43.97
3	53.78	72.07
4	44.33	39.68
5	62.51	60.58

Tabla 4.2.3: Porcentaje de abandonos por horario. Resultado de la simulación con menos operadores en horarios 1 y 3.

Se comprueba que la paciencia está bien estimada. Ningún cambio razonable en la magnitud de la paciencia logra mejores estimaciones de los abandonos. En cambio, el número de operadores disponibles sí afecta dicha estimación. Simulando una cantidad de operadores más realista, los resultados se ajustan sorprendentemente.

Recordar que los parámetros que la simulación recibe como *input* son: las tasas de llegada, tasas de servicio, paciencia y número de operadores en cada horario del día, con todos los supuestos Markovianos que hace la teoría de colas. Dada la gran cantidad de variables aleatorias, se prueba los resultados de simulaciones de 30, 300 y 3000 repeticiones, y se obtienen estimaciones invariantes, es decir, convergen rápidamente y se confirma que la simulación de 30 repeticiones es válida en este caso.

### 4.3. Comparación de Métodos

Se comparan los resultados de los distintos modelos aplicados. La simulación que considera abandonos es lo que permite modelar el sistema de manera más realista.

#### MODELO $M/M/c$

Cuando se cumple la condición de estacionaridad  $\rho \leq 1$ , que es el caso del horario 1, la simulación y el modelo de teoría de colas coinciden en sus resultados, como se ve en el siguiente cuadro.

Medida de Servicio	M/M/c	Simulación	Diferencia
$L_q$	1.0071	1.0006	0.006 %
$W_q$	8.9932	8.9408	0.006 %
$L$	9.4077	9.3799	0.003 %
$W$	84.0110	83.9696	0.004 %

Tabla 4.3.1: Resultados Simulación vs. Teoría de colas para el modelo M/M/c en horario 1.

Esto además confirma que la simulación está modelada adecuadamente y se está siguiendo la lógica que plantea el modelo M/M/c de teoría de colas.

#### Modelo $M/M/c/K$

La complejidad del problema es que muchas veces en la realidad, no se cumple la condición de estabilidad  $\rho \leq 1$ , requisito del modelo  $M/M/c$ , entonces los resultados se obtienen mejor a través de una simulación. En en horario 2, el modelo analítico M/M/c/K y la simulación coinciden en cuanto a cantidad de personas en el sistema.

Medida de Servicio	M/M/c/K	Simulación	Diferencia
$L_q$	47.67	47.58	0.0019 %
$W_q$	110.80	193.21	-0.7437 %
$L$	58.67	58.57	0.0005 %
$W$	155.54	237.92	-1.4043 %

Tabla 4.3.2: Resultados Simulación vs. Teoría de colas para el modelo M/M/c/K en horario 2.

Entonces, al usar la ley de Little se subestima el tiempo de espera de los clientes.

Se simula el día completo, con los parámetros que varían en los distintos horarios. Se puede observar que existen llamadas rechazadas o bloqueadas por la capacidad del sistema (ver figura D.3.5) y estas llegan al 47%.

$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$
34.25	146.93	46.48	199.438

Tabla 4.3.3: Resultados Simulación con Capacidad Finita.

### Modelo con Abandonos

El modelo Markoviano  $M/M/c+M$  es la herramienta con que se obtienen estimaciones adecuadas de la paciencia. Incorporando el comportamiento de abandonos, dado por la paciencia de los clientes, en la simulación se obtienen medidas de rendimiento más realistas. En un día de semana corriente, las medidas de rendimiento obtenidas a través de simulación, son las siguientes.

Horario	$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$	Abandonos [%]
1	0.13	1.20	7.62	75.51	10.77
2	4.56	10.61	15.46	53.82	43.48
3	8.56	15.44	17.55	59.33	64.04
4	7.89	11.47	25.86	55.42	42.05
5	9.22	15.11	23.21	75.09	62.47

Tabla 4.3.4: Medidas de rendimiento obtenidas por simulación, en cada horario por separado.

Horario	$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$	Abandonos [%]
Todo el día	5.77	13.00	17.83	63.31	48.10

Tabla 4.3.5: Medidas de rendimiento obtenidas por simulación, considerando los estados transientes.

Entonces, dadas las estimaciones de paciencia de los clientes, la simulación estima abandonos de 48.10%, mientras los datos indican que realmente los abandonos alcanzan un 48.65%. Considerar que hay un porcentaje de llamadas “perdidas”, que es desconocido.

Recordar que en la realidad, no todos los supuestos Markovianos se cumplen necesariamente y además no se tienen datos completos, pues la tasa de llegadas está truncada. Esto se mostró en la sección 3.4.

## 4.4. Modelamiento del Sistema Real

Ya se tiene la mejor estimación de la paciencia de los clientes en cada horario. Ahora se simula el sistema con algunos cambios en los parámetros de entrada.

Primero, uno de los factores más importantes de estudio son los dos principales tipos de llamadas. Las llamadas de emergencia y llamadas inoficiosas tienen distintas tasas de llegada y servicio, como se muestra en la sección 3.4.8. El modelo simulado ahora es

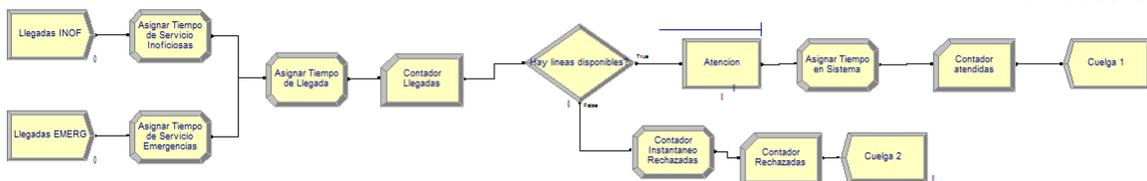


Figura 4.4.1: Modelo en Arena para simulación con dos tipos de llamadas.

Se comparan tres simulaciones con cambios relevantes, con respecto a los supuestos Markovianos que se asumían en la sección anterior. Se realiza un proceso de simulaciones incremental, incorporando mejoras en los parámetros de comportamiento. Se observan los resultados de la simulación con tasas de servicio distribuidas Lognormal, luego se agrega variabilidad en el número de servidores y finalmente se mejora la aproximación de las tasas de llegada de llamadas.

#### SIMULACIÓN CON TIEMPO DE SERVICIO DISTRIBUIDO LOGNORMAL

Esta vez, se asignan promedios de tasas de llegada, y sus desviaciones estándar, para simular cada tipo de llegada según una distribución Lognormal de los tiempos de servicio, que es la que se encontró en la sección 3.4.6.

Los abandonos por horario resultan un poco más parecidos a los reales, que lo obtenido con un solo tipo de llegadas.

Horario	Abandonos Simulación[%]	Abandonos Reales[%]
1	24.25	21.99
2	41.39	43.97
3	52.97	72.07
4	43.50	39.68
5	61.42	60.58

Tabla 4.4.1: Porcentaje de abandonos por horario simulando tiempos de servicio lognormales.

Se toman las medidas de servicio para un cliente cualquiera, que entrega la simulación.

$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$	Abandonos [%]
5.70	12.84	26.45	61.03	47.93

Tabla 4.4.2: Medidas de rendimiento simulando tiempos de servicio lognormales.

En este caso, los parámetros que nos interesan son las medidas de rendimiento para cada tipo de llamada, inoficiosas y emergencias, por separado. En promedio, un día cualquiera, éstas son las siguientes.

Tipo de Llamada	$Wq[segundos]$	$W[segundos]$	$L[clientes]$
Inoficiosa	12.91	28.20	10.98
Emergencia	12.35	118.51	6.47

Tabla 4.4.3: Parámetros de servicio para cada tipo de llamada simulando tiempos de servicio lognormales.

El tiempo total de un cliente en el sistema ( $W$ ) resulta bastante distinto, dado que cada uno tiene distintos tiempos de servicio. El tiempo de espera en cola es similar para lo dos tipos de clientes, ya que se asume que ambos tienen la misma distribución de paciencia.

**SIMULACIÓN ASUMIENDO VARIABILIDAD EN NÚMERO DE OPERADORES** Como siguiente paso, la simulación se ajusta a la realidad asumiendo un número promedio de operadores más parecido a la realidad, como se muestra en la sección 3.4.7.

El porcentaje de abandonos y medidas de servicio se muestran en las siguientes tablas, tal como se vieron en la sección anterior.

Horario	Abandonos Simulación[%]	Abandonos Reales[%]
1	29.31	21.99
2	46.48	43.97
3	59.91	72.07
4	46.49	39.68
5	65.02	60.58

Tabla 4.4.4: Porcentaje de abandonos por horario asumiendo variabilidad en el número de operadores.

$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$	Abandonos [%]
6.03	13.59	16.84	61.54	52.04

Tabla 4.4.5: Medidas de rendimiento con tiempos de servicio exponencial asumiendo variabilidad en el número de operadores.

Tipo de Llamada	$Wq[segundos]$	$W[segundos]$	$L[clientes]$
Inoficiosa	13.68	27.74	10.79
Emergencia	13.00	110.70	6.06

Tabla 4.4.6: Parámetros de servicio para cada tipo de llamada asumiendo variabilidad en el número de operadores.

**SIMULACIÓN ASUMIENDO VARIABILIDAD EN DISTRIBUCIÓN DE LLEGADAS** Esta simulación integra las tres mejoras posibles al modelamiento, para finalmente observar que se refleja la realidad mucho mejor que con meros modelos analíticos. Se considera tiempos de servicios distribuidos lognormal, mejor distribución del número de operadores conectados y variabilidad en la distribución de llegadas, sobre todo en los horarios 2 y 4, que como se observó en 3.4.8, claramente no distribuyen exponencial.

Horario	Abandonos Simulación[%]	Abandonos Reales[%]
1	29.41	21.99
2	45.29	43.97
3	58.62	72.07
4	45.82	39.68
5	64.92	60.58

Tabla 4.4.7: Porcentaje de abandonos por horario asumiendo variabilidad en distribución de llegadas.

$Lq[clientes]$	$Wq[segundos]$	$L[clientes]$	$W[segundos]$	Abandonos [%]
6.01	13.52	16.85	61.18	51.47

Tabla 4.4.8: Medidas de rendimiento con tiempos de servicio exponencial asumiendo variabilidad en distribución de llegadas.

Tipo de Llamada	$Wq[segundos]$	$W[segundos]$	$L[clientes]$
Inoficiosa	13.58	27.88	6.00
Emergencia	13.10	109.14	10.85

Tabla 4.4.9: Parámetros de servicio para cada tipo de llamada asumiendo variabilidad en distribución de llegadas.

Sabiendo que existen datos incompletos, que no se conoce el porcentaje de llamados rechazados, y que el muestreo indica que este porcentaje es cercano al 2%, entonces es esperable que los abandonos

simulados rodeen el 50 % .

Por lo tanto, se demuestra que la simulación, considerando patrones más realistas que los Markovianos que asume la teoría de colas, refleja muy bien la realidad del sistema y permite medir los niveles de servicio, según los indicadores definidos en la sección 2.1.2.

## 4.5. Nivel de Servicio

Los KPI a evaluar, en un servicio de emergencia como este, se relacionan con los tiempos de espera de las personas y los abandonos (como se explica en 2.1 y 2.1.2). Ya que tales medidas no se tienen en el 133, se estimó la tasa de abandono y se logró reflejar el funcionamiento real del sistema a través de la metodología propuesta, llegando a los resultados finales en 4.4.

La última simulación en la sección 4.4, es el modelo que refleja la realidad, y permite medir el nivel de servicio actual del fono de emergencias. El tiempo de espera en cola se considera la métrica de nivel de servicio a evaluar, ya que se trata de un servicio de emergencia y este tiempo se quiere minimizar.

En la sección 2.1 se describe cómo los principales KPI en los servicios telefónicos de emergencia internacionales se relacionan con el tiempo máximo de espera de los clientes en la cola.

La simulación permite observar los tiempos de espera de cada llamada, si es inoficiosa o emergencia y si es atendida, abandonada o rechazada, entre otros. El volumen de llamadas inoficiosas es muy grande, se mide el nivel de servicio para todas las llamadas, pero ¿Cuál es el nivel de servicio que están recibiendo los llamados de emergencia?. Son éstas las que nos deben preocupar y para las que se requiere otorgar un buen servicio. Por lo tanto, es necesario un análisis por separado de cada tipo de llamadas, se especifica entonces el nivel de servicio otorgado a las llamadas de emergencia.

### 4.5.1. Días hábiles

La simulación arroja que, en días hábiles, un 48.49 % de las llamadas son atendidas, y el resto, 51.51 %, abandonan la cola de espera, antes de la atención.

En la siguiente tabla se observa el porcentaje de llamadas que espera menos de 10 segundos, menos de 20 segundos y menos de 30 segundos, antes de ser atendidas.

Tiempo de espera $W_q$ [seg]	$\frac{\text{Llamadas Atendidas}}{\text{Total Llamadas}}$ [	Porcentaje acumulado [%]
< 10	19.60	19.60
< 20	26.05	45.66
< 30	2.83	48.49
> 30	0	48.49

Tabla 4.5.1: Nivel de Servicio Actual para Llamadas Atendidas. Días Hábiles.

La tabla 4.5.1 muestra que al menos el 48 % de las llamadas son atendidas antes de 30 segundos y nadie espera más de dicho tiempo, si se sobrepasa ese tiempo de espera, las personas cuelgan. Se debe recordar que más del 80 % de estas llamadas, atendidas o abandonadas, son inoficiosas. A continuación se mide el servicio para las emergencias.

El total de llamadas de emergencia entrantes al sistema, son 4700 llamadas por día, aproximadamente. Esto corresponde al 12.36 % de las llamadas entrantes totales. Es decir, más del 87 % son inoficiosas.

Sólo el 47.45 % de las emergencias son atendidas (5.86 % del total de llamadas entrantes). El 52.55 %, de las llamadas de emergencia, abandona el sistema antes de ser atendida. Esto significa que aproximadamente se atienden 2250 llamadas de emergencia en el 133, cada día.

Los detalles de los niveles de servicio para todas las llamadas se encuentran en el apéndice E.

Tiempo de espera $W_q[seg]$	$\frac{Emergencias\ Atendidas}{Total\ Emergencias}$ [%]	Porcentaje acumulado [%]
< 10	21.52	21.52
< 20	23.35	44.86
< 30	2.58	47.44
>30	0.00	47.45

Tabla 4.5.2: Nivel de Servicio en llamadas de Emergencia. Días Hábiles

#### 4.5.2. Fin de Semana

El comportamiento del sistema se estudia a partir de la base metodológica que se logra analizando el sistema en la semana. Se caracteriza el sistema según los parámetros relevantes, luego se estiman las tasas de abandono y se simula para obtener niveles de servicio. Sin la necesidad de pasar de nuevo por todo el proceso de validación.

Los resultados indican que llamadas atendidas, durante un fin de semana, corresponden al 32.4% y equivalen a 12600 llamadas atendidas por día, en promedio aproximadamente. La simulación no arroja llamadas rechazadas, pero un 67.6% abandona mientras espera atención.

Tiempo de espera $W_q[seg]$	$\frac{Llamadas\ Atendidas}{Total\ Llamadass}$ [%]	Porcentaje acumulado [%]
< 10	3.51	3.51
< 20	19.68	23.20
< 30	9.09	32.28
> 30	0.10	32.39

Tabla 4.5.3: Nivel de Servicio Actual. Fin de Semana.

En este caso, el mismo análisis indica que el 31.2% de las emergencias son atendidas (6.85% del total de llamadas entrantes). Entonces, un 68.8% de las personas que llaman para reportar emergencias, abandonan, sin llegar a ser atendidas.

Tiempo de espera $W_q[seg]$	$\frac{Emergencias\ Atendidas}{Total\ Emergencias}$ [%]	Porcentaje acumulado [%]
< 10	3.04	3.04
< 20	18.93	21.97
< 30	9.07	31.04
> 30	0.11	31.15

Tabla 4.5.4: Nivel de Servicio en llamadas de Emergencia. Fin de Semana.

En lo que sigue, se discuten estos resultados, finalmente se proponen y miden posibles soluciones para luego concluir.

## 4.6. Discusión

Este trabajo plantea una metodología de estimación de parámetros desconocidos de un sistema de colas para finalmente simular el funcionamiento real del servicio. Se valida la estimación al ver que la simulación logra reflejar el comportamiento actual. Aquello se consigue después del ajuste apropiado de las distribuciones de llegada, de servicio y de cantidad de operadores. Finalmente la simulación arroja cantidad de llamadas, porcentajes de llamadas inoficiosas y abandonos prácticamente idénticos a la realidad.

Los niveles de servicio resultan más bajos que los de que manejan internacionalmente. Un día de semana cualquiera, todas las llamadas emergencias atendidas esperan 30 segundos, pero esto no significa que se está atendiendo muy bien, pues se atiende solo un 47 %. Más del 52 % de estas llamadas cuelgan, porque la paciencia de los llamantes se agota. Esto indica que para una persona que necesita reportar una emergencia 30 segundos de espera es demasiado tiempo.

El mismo análisis se puede hacer para los fines de semana. En tal caso el resultado indica que casi el 70 % de las emergencias no son atendidas. Solo el 31 % son atendidas.

En los servicios telefónicos de emergencias de países desarrollados se tiene como meta atender al 90 %, agregando que esto debe ocurrir con un tiempo máximo de espera de 10 segundos. Cada día se trabaja, para lograr aquellas metas, y algunos la superan. Esto refleja que el nivel de servicio del Nivel 133 es bastante bajo, a pesar de que el número de llamadas rechazadas sea prácticamente despreciable. Las personas están esperando mucho tiempo en cola, el call center está congestionado.

Hay muchas medidas que se deben tomar para mejorar el servicio. Durante el trabajo de modelamiento del proceso de caracterización de los parámetros del sistema se encontró varios elementos sujetos a mejorar o cambiar. En el siguiente capítulo se discuten y evalúan varias propuestas de mejora.

¿Se podría decir que aumentar la cantidad de operadores mejoraría el servicio? ¿Cómo se podría reducir la carga del sistema? ¿Cómo mejoraría el servicio si se reducen las llamadas inoficiosas? Éstas y otras preguntas surgen y se responden a continuación.

## Capítulo 5

# Recomendaciones y Conclusión

Después de haber descrito el funcionamiento y medido el nivel de servicio del sistema actual de atención de emergencias, en la Región Metropolitana, se presentan algunas propuestas para mejorar los niveles de servicio. La solución definitiva, dependerá de la estrategia que se defina para el Nivel 133.

Primeramente, se presenta una medida que para muchos parece obvia. Más del 50 % de las personas que discan el 133, cuelgan mientras esperan en cola porque se agota su paciencia. El sistema está claramente congestionado. Esto implica que más de la mitad de las emergencias no son atendidas porque los operadores están ocupados, atendiendo otras llamadas. Entonces, se calcula cuántos es la dotación óptima de operadores y se discuten los resultados.

Otro punto importante a afrontar es el gran volumen de llamadas inoficiosas que se recibe. El 85 % de las llamadas consumen los escasos recursos de las operaciones del call center. Se plantean las primeras ideas sobre cómo disminuir la carga del call center y se evalúa el resultado que aquello podría conseguir.

Finalmente, se considera el hecho de que existe mucha información necesaria, que no está disponible. Por ejemplo, la demanda total y la paciencia de los clientes, son datos obligados para la evaluación de cualquier call center corporativo, la tecnología está disponible hace años. No existen datos de llamadas rechazadas ni abandonadas. No se conoce la tasa de rediscado, ni los tiempos de espera en cola. Toda esta información es muy relevante para medir niveles de servicio. Se explica la importancia y se presenta una idea de cómo mejorar este aspecto.

### 5.1. Dotación Óptima

La dotación óptima de operadores en el Nivel 133 es una gran preocupación en CENCO y Carabineros de Chile. A partir de este estudio, con simulaciones y teoría de colas no es difícil entregar una recomendación, tal que se contesten todas las llamada. Sin embargo, hay que tener muy presente que los recursos, actualmente están siendo utilizados en gran parte por las llamadas inoficiosas. Por otro lado, hay datos incompletos, parámetros importantes que han tenido que ser estimados, y esto podría originar un ligero nivel de error.

A priori, para equiparar tasa de llegada y tasa de servicio, usando teoría de colas basta con calcular  $a = \lambda / c\mu$ . En la semana, se obtiene que la cantidad de operadores debería ser 10, 20, 15, 30 y 30 operadores en promedio, en los 5 horarios respectivamente.

Una simulación que mantiene todos los parámetros actuales, y sólo aumenta el número de operadores a dichos niveles, muestra que efectivamente consigue mejoras en los niveles de servicio. Se obtiene que el 81 % de las emergencias se atienden. Un 69 % de las emergencias atendidas reciben servicio antes de 10 segundos, un 80 % antes de 20 segundos y un 81.26 % son atendidas antes de 30 segundos. En general, las llamadas atendidas, oficiosas o inoficiosas, alcanzan el 84 % (más de 26.000 llamadas), pero el 73 % del total de llamadas atendidas no son de emergencia (más de 23.000 llamadas). Un 16 % de las

llamadas, casi 6000 llamados al día, cuelgan, debido gran tiempo de espera.

Aumentar la cantidad de operadores a 30, después de las 17 horas (horarios 4 y 5), requiere contratar y capacitar operadores. Generalmente se contratan C.P.R., pero algunos de ellos no pueden hacer turnos largos, o de noche, como se explica en la sección 3.4.7. Poner más Carabineros atendiendo significaría sacar a algunos de otras labores, consideradas más importantes. Además, actualmente para aumentar la dotación a las 17 horas, se necesita aumentar la disponibilidad en el turno de las 14 horas y de las 21 horas. Lo que provocaría exceso de operadores en los horarios 1 y 3.

El fin de semana la utilización de los operadores es mucho mayor, por lo tanto es innegable que la situación empeora y se necesitaría aumentar mucho más el número de operadores.

El análisis demuestra que la dotación óptima debe ser calculada una vez que se tenga una estrategia para mejorar la gestión del Nivel 133. No vale la pena atender el 80 % de las llamadas, si la gran mayoría de éstas son bromas, llamadas equivocadas o solicitudes impertinentes de información, pues se estaría invirtiendo recursos valiosos, que podrían ayudar a la ciudadanía en otras circunstancias.

## 5.2. Disminuir la Demanda

Disminuir la demanda se refiere a lograr que las llamadas que no son emergencias, o que son tipificadas como emergencias, y no requieren presencia policial inmediata, dejen de llegar al 133.

Ya que se muestra que no es suficiente aumentar el número de operadores en el 133, se considera aumentarlos en el Nivel 139, de informaciones policiales. En esta unidad hay un solo operador, se sabe que muchas personas marcan el 139, pero como esta congestionado, llaman al 133.

Hay un gran problema de desinformación. Por ejemplo, en el audio de una llamada grabada en el Nivel 133, una persona indica que llama muchas veces al 139, durante más de 20 minutos, recibiendo tono ocupado. En el 133 reporta su problema, y se trata de un automóvil mal estacionado. Ese incidente, efectivamente está tipificado como emergencia y debe ser resuelto a través del 133. Por lo tanto el 139 podría estar sufriendo congestión también, por desinformación.

Se requiere una estrategia efectiva para educar a la población. Muchas campañas han traído malos resultados, pero para mejorar, se deben evaluar los errores y ejecutar medidas que hayan demostrado efectividad, aunque se hayan aplicado en otras industrias. Esto consistiría en otro trabajo complejo, pero muy importante.

Por ahora, se comenta una medida que se ha encontrado que funciona, en sociedades como la nuestra. Luego, se evalúa la efectividad de esta propuesta a través de análisis vía simulación y se observa cómo cambiarían los niveles de servicio.

### 5.2.1. Sancionar Llamados Inoficiosos

Usando los resultados de este estudio, Carabineros puede evaluar el modelo de respuesta oportuna. Si no se está conforme con el nivel de servicio, las medidas aquí presentadas servirían como argumento para proponer cambios de variada índole.

Existe gran interés por tipificar las llamadas inoficiosas como un delito o falta, por ejemplo como denegación de auxilio a ciudadanos en emergencias. El artículo 268 bis del Código Penal penaliza el delito de falsas alarmas de emergencia, desastres o calamidad. Actualmente, caen en esa condición, las llamadas de emergencia que generan un procedimiento, y al llegar los dispositivos al lugar, se dan cuenta que es una falsa alarma. El último año se discutió la posibilidad de establecer responsabilidad civil para las personas que hacen llamadas falsas, de otro tipo.

Según lo que dice el Código Penal, se podría proponer incluir en esa pena, a las llamadas inoficiosas o propiciar una modificación legal que permita sancionar de alguna manera. Países vecinos a Chile, han adoptado castigos en conjunto con las compañías telefónicas privadas, cortando el suministro telefónico de una persona que hace llamadas inoficiosas reiteradas.

### 5.2.2. Mejorar Nivel 139

Hay muchas formas de rediseñar el proceso para mejorar el nivel de servicio. Cada una debe ser evaluada. La primera propuesta de rediseño consiste en lograr que las llamadas de inoficiosas clasificadas como información o equivocado, se atiendan en el Nivel 139. Dichas clases de llamadas inoficiosas, suman en conjunto un 62.1 % del total de llamadas que no son emergencias.

Si se lograra concientizar a la población, a través de campañas educativas y de información adecuadas<sup>1</sup>, la carga del sistema disminuiría. Esto se reflejaría en la tasa de llegada de llamadas de inoficiosas. Asumiendo que el resto de los parámetros se mantiene constante, y se redujera en un 40 %<sup>2</sup> las llamadas inoficiosas, los niveles de servicio, manteniendo el resto de las condiciones actuales, mejoran.

Bajo los supuestos descritos llegarían proporcionalmente más llamadas de emergencia que inoficiosas. El resultado del modelamiento de este escenario, a través de la simulación, indica el 63 % son atendidas y cuelgan menos del 37 % de las llamadas de emergencia, mientras antes colgaba más del 50 %. Un 47 % de las emergencias que llaman, serían atendidas antes de 10 segundos de espera en cola.

Por lo tanto, se prueba que una disminución conservadora del 40 %, respecto al 62 % que realmente representan las llamadas inoficiosas, disminuye los abandonos, se descongestiona el sistema y se atiende proporcionalmente más emergencias.

Más adelante se indica que esta unidad también debería registrar los datos de las llamadas, para poder medir la demanda y cuantificar otros parámetros. Es la única manera de optimizar el servicio, como se propone aquí con el Nivel 133.

### 5.2.3. Rediseño para Atención de Alarmas Activadas

Existen llamadas de emergencia que también podrían estar congestionando el sistema. Las llamadas de Alarma Activada representan un 19.3 % de la demanda de llamadas de emergencia, durante una semana cualquiera. Existe una unidad, llamada Nivel Alpha II, que atiende Alarmas Corporativas durante las 24 horas. En esta unidad hay un solo operador, y las empresas pagan por el servicio. Ya que las llamadas de Alarma Activada, corresponden mayoritariamente a empresas de servicios de alarmas para inmuebles particulares, se podría pensar en atenderlas en Alpha II o de alguna forma, hacer que dejen de pasar por el 133.

El solo hecho de recibir estas llamadas fuera del Nivel 133, descongestiona el sistema y permite atender un 54 % de llamadas de emergencia, en porcentaje levemente mayor al 47 % atendido actualmente. Lo importante es que se observa que los abandonos de llamadas de emergencia disminuyen del 53 al 46 %. En general disminuyen los abandonos de 67 % a 40 %. Casi un 30 % de las llamadas de emergencia son atendidas en menos de 10 segundos. Más del 53 % reciben servicio antes de 30 segundos.

Dado que hay llamadas que consumen recursos valiosos, y no son de emergencia, o son de empresas de seguridad que siempre llaman y saben cómo funciona el proceso de atención, estas llamadas se podrían atender en otro proceso, o por otros servidores.

Se podría evaluar dejar a los operadores más eficientes contestando las llamadas de emergencia con mayor prioridad, de manera que se despachen recursos policiales rápidamente.

---

<sup>1</sup>Las medidas educativas y de información a generar no forman parte de este estudio, como se indica en 2.4.

<sup>2</sup>Se simula con 40 % de reducción de carga inoficiosa como forma de simular un escenario conservador de disminución, ya que en la realidad, sería complejo lograr reducir las al 100 % o al 60 % de inmediato. Este último porcentaje representa la demanda que se podría atender el 139.

Este tipo de soluciones se pueden simular, sin involucrar costos o errores prácticos. De manera de implementar un rediseño después de haber estudiado la respuesta del sistema completo.

Además se debe considerar que la carga de procedimientos en Alpha I podría variar de manera importante. Por eso esto sería otro estudio bastante complejo.

Se propone que Carabineros de Chile plantee una estrategia para disminuir la demanda de llamadas ineficaces, o de emergencia que no requieran presencia policial, como las denuncias civiles. Aquí se muestra que cada posible solución puede ser simulada para evaluar los costos y beneficios de implementarla.

### 5.3. Estudio de Tráfico

La compañía telefónica que administra las centrales que nutren el 133, menciona que es posible realizar un estudio de tráfico para entregar un reporte de la cantidad de llamadas que llega en total al call center.

Esta sería una medida a tomar de inmediato, para evaluar un poco más certeramente la gravedad de los problemas existentes. Sin embargo, es una medida provisoria, pues no otorga información completa, ni permite una evaluación periódica. A continuación se explica qué datos se necesitan registrar siempre y tener acceso ilimitado a ellos.

#### 5.3.1. Modelo de Datos

Para llevar a cabo una evaluación más precisa del nivel de servicio, poder medirlo periódicamente y optimizarlo, se necesita contar datos completos y consistentes<sup>3</sup> de las llamadas e indicadores actualizados que resuman la información relevante.

Trofimov et al. [14] proponen un Modelo de Datos para Callcenters que se podría tomar como base para construir el que se requiere en el Nivel 133.

Dado que próximamente se requiere licitar un sistema CAD, se debe estipular los requerimientos del nuevo sistema. Algunos de los datos que se requieren indispensablemente, se listan en seguida. Luego se explica cómo los atributos de las llamadas, guardados correctamente en los sistemas de Carabineros, podrían servir para mejorar los niveles de servicio de ahora en adelante.

**Información de llamados que no entran al sistema: Estos corresponden a llamados que recibieron el tono ocupado.** De estos llamados se necesita:

1. El número telefónico que realizó la llamada.
2. Fecha y hora del llamado.
3. Localización (Ubicación que registra la compañía telefónica)

**Para las llamadas que entran al sistema, se necesita la siguiente información:** 1. Número telefónico que realiza el llamado: esta información es relevante, entre otras razones, para identificar rediscadas y llamadas que cruzan otros servicios de emergencia (e.g. 139, ver más abajo).

2. Fecha y hora de Inicio y Término de un llamado: Esta información se requiere para determinar duración de un llamado, cantidad de llamadas por día, existencia de temporalidades y ciclos regulares además de otros estadísticos que puedan medir la carga de uso del sistema.

3. Indicador si el llamado fue atendido por un operador y la hora en que comenzó la atención: Esto se utiliza para estimar el número de llamados que colgaron antes de ser atendidos y el tiempo de espera de los usuarios.

Para las llamadas que efectivamente fueron atendidas por un operador, se necesita la siguiente información adicional:

---

<sup>3</sup>Se han encontrado inconsistencias en los datos, tales como: texto en lugar de números de teléfono, números en lugar de direcciones, número de teléfono del llamante distinto en distintas fuentes de datos, procedimientos creados antes de recibir una llamada, una misma llamada clasificada como ineficaz y emergencia a la vez, entre otras.

4. Ubicación geográfica del incidente: Para estudiar la procedencia geográfica de cada tipo de llamada.
5. Indicador si la llamada es inoficiosa: Llamadas que no generan procedimientos policiales, correspondientes a broma, inaudibles, insultos, información, niños o número equivocado.
6. Tipificación de la emergencia o llamada inoficiosa: Cada tipo de incidente que se reporta puede tener una relación con el tiempo total que usa un operador en atender la llamada. Según la información disponible, existen 103 distintos tipos de llamadas de emergencia y 6 tipos de llamadas inoficiosas.
7. Prioridad de la emergencia: Dado que no siempre se despacha el requerimiento con la prioridad pre-asignada, se debe estudiar la prioridad que asigna un operador.
8. Tiempos asociados a los siguientes eventos: a. Hora en que se inició la geo-referenciación y hora en la que finaliza. Hora en la que se dio notificación al despacho (en caso que se haya generado uno).
9. Identificación del operador que atiende y nivel de experiencia del mismo: Se estudiará si los distintos operadores tienen tiempos similares o distintos de atención para cada llamado, esto, contrastado con el nivel de experiencia que el operador tenga. El nivel de experiencia se refiere al tiempo que lleva el funcionario ejerciendo la actividad.

**Atención 139** También se requerirá datos como estos correspondientes al nivel 139 para estudiar sus niveles de uso, procedencia de llamados, tipos de llamadas recibidas y otros estadísticos que describan el sistema actual en profundidad. En particular, para los llamados al 139 se requiere:

1. Número telefónico de donde se inició la llamada.
2. Hora de inicio y término de cada llamado al 139.
3. Indicador si el llamado fue atendido por un operador, y la hora en que se inició la atención.
4. Clasificación de la llamada.

El contar con datos más completos y con menos errores, permitirá aplicar las simulaciones de manera más directa para medir el nivel de servicio. Incluso se podría automatizar el análisis. Además, lo más importante es que permitiría a Carabineros medir sus metas de servicio, evaluarlas y poder ir en contaste mejora, sin necesidad de pasar por todo el procesamiento de datos que se necesitó para llegar a obtener los parámetros del sistema en el capítulo 4.

## 5.4. Conclusiones

El nivel de servicio actual está por debajo del 52 %, considerando todas las llamadas atendidas antes de 30 segundos. Si se compara directamente con el nivel de servicios telefónicos de emergencia como el 911, menos del 20 % de las llamadas son atendidas antes de 10 segundos, un día de semana cualquiera. Los fines de semana, se atienden el 32 % de las llamadas antes de 30 segundos y solo un 3 % recibe servicio antes de 10 segundos, casi el 70 % de las llamadas de emergencia cuelgan debido a las largas esperas. Por lo tanto, está claro que se deben proponer estrategias de gestión del servicio para mejorar la calidad de la atención.

Según las metas que se formulen, se debe evaluar un proyecto de mejora que considere varias propuestas, y mida su impacto antes de ser implementado. Para medir las consecuencias de cambios en parámetros del sistema, tales como: operadores, tasas de llegadas, tasas de servicio o capacidad tecnológica, se podría usar la metodología y herramientas que se usaron en este trabajo.

La teoría de colas, se utilizó para estimar la tasa de abandono de los clientes, sin embargo, es limitada. Al mejorar los parámetros en los modelos, de acuerdo a la realidad, la simulación muestra el comportamiento del sistema tal como es en la realidad, con la ventaja de poder distinguir cada llamado y observar todas sus características, a través de bases de datos que se pueden generar.

Es imprescindible realizar un estudio de tráfico en las centrales del Nivel 133 y considerar el rediseño del sistema CAD. Más datos, no solo hacen el cálculo de los niveles de servicio más rápido, sino también permiten evaluar periódicamente el funcionamiento del sistema en cada región del país y optimizar el funcionamiento de los servicios telefónicos de emergencia.

# Bibliografía

- [1] George Antunes. Calling the cops: Police telephone operators and citizen calls for service. *Journal of Criminal Justice*, 9:165–179, April 2013. 2.1, 2.1.1
- [2] Athanassios N. Avramidis and Pierre L’Ecuyer. Modeling and simulation of call centers. *Proceedings of the 2005 Winter Simulation Conference*, pages 1–9, November 2005. 2.3.2.8
- [3] Lawrence Brown, Noah Gamb, Avishai Mandelbaum, Anat Sakov, Haipeng Shen, and Linda Zhao. Statistical analysis of a telephone call center: A queueing-science perspective. *Journal of the American Statistical Association*, 100(Applications and Case Studies):1–15, April 2005. 2.3.2.8, 2.3.2.8, 2.3.2.8, 3.4.6
- [4] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From data mining to knowledge discovery in databases. *American Association for Artificial Intelligence*, 17:37–54, October 1996. A.1.1
- [5] Donald Gross and Carl M. Harris. *Fundamentals of Queueing Theory*. February 1985. 2.3.2.2, 2.3.2.4
- [6] Hugo Hernández. Central de comunicaciones, cenco: El corazón de las operaciones policiales. *Revista Carabineros de Chile*, pages 1–80, September 2011. 1.1, 1.2, 3, 3.2
- [7] J. F. Kingman. The first erlang century—and the next. *Queueing Systems*, 63(1-4):3–12, 2009. 2.3.2, 2.3.2.8
- [8] Ger Koole. *Call Center Optimization*. 2012. 2.3.2.9
- [9] A. Mandelbaum, M. Reiman, and O. Garnett. Designing a call center with impatient customers. *Manufacturing & Service Operations Management*, 4:208–227, October 2002. 2.3.2.8
- [10] Avishai Mandelbaum. Call centers: Research bibliography with abstracts. pages 1–226, June 2006. 2.3.2.8, 2.3.2.8
- [11] Aysegul A. Yarpuzlu MD and Aysegul A. Yarpuzlu. The 112 emergency telephone service. *The Journal of Emergency Medicine*, 44(2):e289–e290, March 2013. doi: 10.1016/j.jemermed.2012.05.035. URL <http://linkinghub.elsevier.com/retrieve/pii/S0736467912011006>. 2.1.1
- [12] J. MacGregor Smith. M/g/c/k blocking probability models and system performance. *Performance Evaluation*, 52(4):237–267, June 2003. 2.3.2.7
- [13] William J. Stewart. *Probability, Markov chains, queues and simulation: the mathematical basis of performance modeling*. May 2009. (document), 2.3.2.1, 2.3.2.3, 2.3.3, 2.3.5
- [14] Valery Trofimov, Paul Feigin, Avishai Mandelbaum, Eva Ishay, and Ella Nadiharov. Data-mocca data model for call center analysis. pages 1–56, August 2006. 5.3.1
- [15] Mathias Weske. Business process management: Concepts llanguages architectures. pages 1–372, October 2007. 2.3.1
- [16] Stephen A. White. Microsoft word - bpmn introductory white paper.doc, June 2004. 2.3.1
- [17] Rüdiger Wirth and Jochen Hipp. Crisp-dm: Towards a standard process model for data mining. pages 1–11, December 2001. 2.2, A.2.1

Parte II  
Apéndices

# Apéndice A

## Procesamiento de Datos

### A.1. Proceso KDD

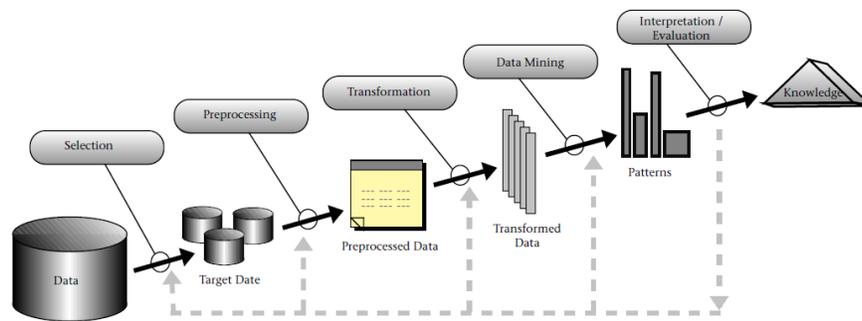


Figura A.1.1: Etapas de un proceso KDD  
Fuente:Fayyad et al. [4]

## A.2. Proceso CRISP-DM

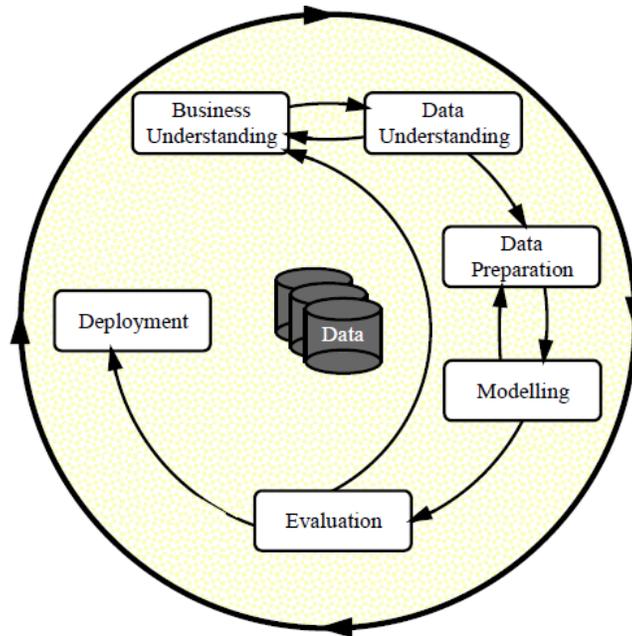


Figura A.2.1: Fases de el modelo CRISP-DM para Data Mining  
Fuente: Wirth and Hipp [17]

## A.3. Datos Disponibles

### A.3.1. Sistemas de Información

Actualmente, se usan los servicios de manejo de llamadas Power 911, Viper CAD y Power MIS, desarrollados por la empresa canadiense Intrado. El sistema Viper CAD se instaló en marzo del 2011, permitiendo a los operadores clasificar las llamadas y derivarlas a los despachadores con ciertas prioridades asignadas (ver detalle en sección 3.2). Los sistemas Power MIS y Contact Center permiten generar reportes como los que se usaron para el procesamiento de la información, como se describe en el apéndice A.

Los servidores se ubican en el departamento de Tecnologías de Información y Comunicación (TIC). Este departamento se encuentra separado de CENCO y se encarga de administrar otros tipos de datos de Carabineros.

## A.3.2. Cantidad de llamadas Abandonadas

Día	Mes	Entrantes	Atendidas	Abandonadas	Abandonadas [%]
Lunes	Agosto	36823	16938	19895	54
	Noviembre	37652	23256	14329	38
Martes	Agosto	37293	15633	21730	58
	Noviembre	38108	19561	18554	49
Miércoles	Agosto	39380	23454	15922	40
	Noviembre	38637	22452	16169	42
Jueves	Agosto	38949	22968	15984	41
	Noviembre	40009	24417	15605	39
Viernes	Agosto	39269	19460	19776	50
	Noviembre	40067	22264	17768	44
Sábado	Agosto	39415	14473	24964	63
	Noviembre	37744	12344	25396	67
Domingo	Agosto	39304	14009	25298	64
	Noviembre	37797	11587	26221	69

Tabla A.3.1: Llamadas entrantes atendidas y abandonadas por día

# Apéndice B

## Parámetros en Detalle

### B.1. Ajuste de Tasa de Servicio

La tasa de servicio se ajusta usando el software Stat Fit para comprobar la distribución Lognormal de éste parámetro. Se consideran todas las llamadas atendidas de la primera semana de Agosto de 2012, hasta una frecuencia acumulada del 80 %.

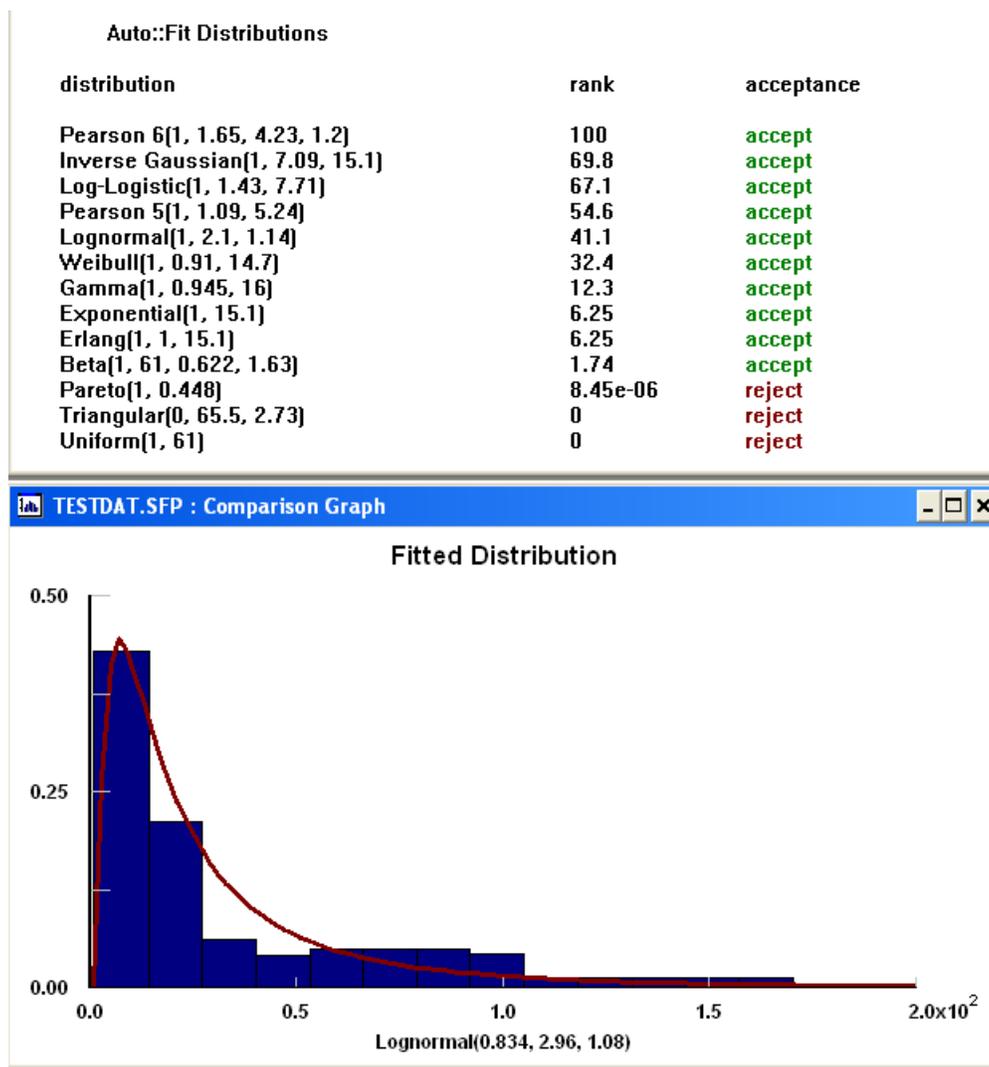


Figura B.1.1: Salida Stat Fit. Ajuste de tiempos de servicio.

Se observa que la distribución Lognormal es aceptable.

### B.1.1. Tiempos de Servicio de Emergencias Ponderados

Tipo de Emergencia	Total Llegadas	Horario 1 Llegadas [%]	Horario 2 Llegadas [%]	Horario 3 Llegadas [%]	Horario 4 Llegadas [%]	Horario 5 Llegadas [%]	Tiempo servicio Promedio	Horario 1 Tiempo [mm : ss.sss]	Horario 2 Tiempo [mm : ss.sss]	Horario 3 Tiempo [mm : ss.sss]	Horario 4 Tiempo [mm : ss.sss]	Horario 5 Tiempo [mm : ss.sss]	Tiempo Servicio Ponderado
ALARMA ACTIVADA	19.3	4.98	6.23	0.22	2.95	4.90	02:30.041	00:07.472	00:09.348	00:00.330	00:04.426	00:07.000	00:28.576
ROBO	7.4	1.03	1.73	0.16	1.67	2.82	03:36.001	00:02.225	00:03.737	00:00.346	00:03.607	00:06.091	00:16.006
ACC. TTO. SOLO DAÑOS	5.3	0.20	1.68	0.10	1.55	1.72	03:12.998	00:00.386	00:03.242	00:00.193	00:02.991	00:03.320	00:10.132
INFRACCION OTRAS LEYES	4.7	0.26	1.12	0.08	1.50	1.72	04:14.917	00:00.663	00:02.855	00:00.204	00:03.824	00:04.385	00:11.930
INDIVIDUO MERODEANDO (SOSP)	4.6	1.35	0.71	0.07	0.82	1.69	04:03.130	00:03.282	00:01.726	00:00.170	00:01.994	00:04.109	00:11.281
V.I.F.	4.6	0.92	0.80	0.13	0.89	1.83	03:42.960	00:02.051	00:01.784	00:00.290	00:01.984	00:04.080	00:10.189
INFORMACION	4.0	0.65	0.81	0.08	0.89	1.53	04:15.806	00:01.663	00:02.072	00:00.205	00:02.277	00:03.914	00:10.130
RUIDOS MOLESTOS	3.6	2.02	0.16	-	0.13	1.30	03:18.905	00:04.018	00:00.318	00:00.259	00:02.586	00:07.180	00:07.180
EBRIOS	3.4	0.94	0.33	0.10	0.46	1.60	03:09.806	00:01.784	00:00.626	00:00.190	00:00.873	00:03.037	00:06.510
ACC. TTO. LESIONADOS/MUERTOS	3.1	0.10	0.87	0.05	0.98	1.13	03:08.046	00:01.272	00:01.118	00:00.077	00:01.234	00:02.332	00:06.033
V.I.F. EFECTUANDOSE	3.1	0.66	0.58	0.04	0.64	1.21	03:12.741	00:00.188	00:01.636	00:00.094	00:01.843	00:02.125	00:05.886
IMPUTADO TIENDA COMERCIAL	2.9	-	0.32	0.12	1.18	1.26	02:19.891	-	00:00.448	00:00.168	00:01.651	00:01.763	00:04.029
AGRESION	2.5	0.36	0.41	0.08	0.71	0.95	03:34.164	00:00.771	00:00.878	00:00.171	00:01.521	00:02.035	00:05.376
VEHICULO ABANDONADO	2.4	0.20	0.83	0.06	0.60	0.75	03:52.659	00:00.465	00:01.931	00:00.140	00:01.396	00:01.745	00:05.677
DAÑOS	2.2	0.29	0.37	0.07	0.58	0.89	03:52.067	00:01.649	00:00.603	00:00.040	00:00.784	00:01.328	00:04.405
DESORDEN	2.2	0.82	0.30	0.02	0.39	0.66	03:21.146	00:00.673	00:00.859	00:00.162	00:01.346	00:02.065	00:05.105
VEHICULO MAL ESTACIONADO	2.0	0.04	0.75	0.12	0.50	0.61	03:39.639	00:00.088	00:01.647	00:00.264	00:01.098	00:01.340	00:04.437
LESIONADO	2.0	0.13	0.45	0.04	0.55	0.79	03:44.473	00:00.292	00:01.010	00:00.090	00:01.235	00:01.773	00:04.400
AMENAZA DE PALABRA	1.7	0.15	0.28	0.05	0.55	0.71	04:06.832	00:00.370	00:00.691	00:00.123	00:01.358	00:01.753	00:04.295
ROBO DE VEHICULO	1.4	0.27	0.26	0.01	0.30	0.59	04:25.030	00:00.716	00:00.689	00:00.027	00:00.795	00:01.564	00:03.790
DROGAS	1.4	0.21	0.18	0.04	0.34	0.59	04:02.450	00:00.509	00:00.436	00:00.097	00:00.824	00:01.430	00:03.297
VENTA/CONSUMO COOPERACION	1.3	0.23	0.30	0.04	0.39	0.36	05:14.336	00:00.723	00:00.943	00:00.126	00:01.226	00:01.132	00:04.149
INDIVIDUO INTER. DOMICILIO (SOSP)	1.3	0.53	0.18	0.03	0.13	0.41	03:40.453	00:01.168	00:00.397	00:00.066	00:00.287	00:00.904	00:02.822
...	...	...	...	...	...	...	...	...	...	...	...	...	...
Tiempo de Servicio de Emergencias	-	-	-	-	-	-	-	00:38.830	00:44.766	00:04.336	00:46.665	02:16.976	-

Tabla B.1.1: Ejemplo. Cálculo de tiempo de servicio ponderado.

## B.2. Resultados del Muestreo

Los resultados se separan en días de semana y fin de semana. En los gráficos, se puede ver que los tiempos de espera en cola, denotados por  $W_q$ , son mayores durante la noche. No es necesario ajustar los datos a una distribución de probabilidad, ya que esto más bien es un resultado dependiente de los parámetros de servicio del sistema (tasas de servicio y número de operadores) y del momento del día (hora y tasa de llegada de llamadas en ese horario).

Durante el fin de semana, un 1.89% de las llamadas son rechazadas. En la semana, un 15% de las llamadas fueron rechazadas, pero éstas se concentraron en tres horarios. De 0 a 3 hrs, de 13 a 14 hrs y de 17 a 22 hrs.

B.2.1. Tiempo de Espera en Cola

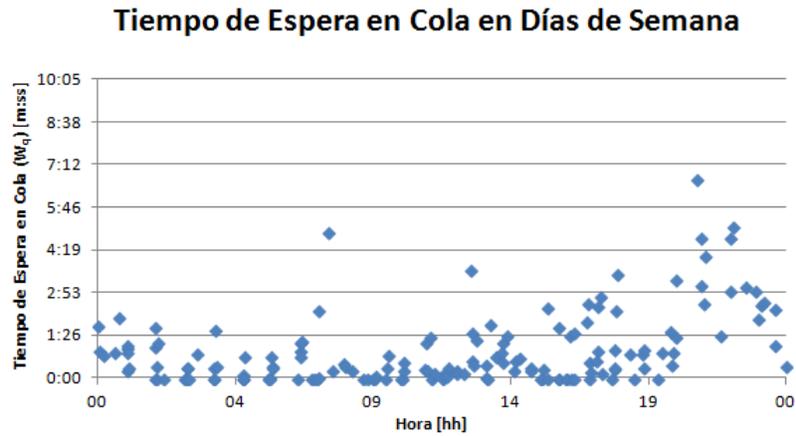


Figura B.2.1: Distribución de tiempos de espera en Cola en días de semana.

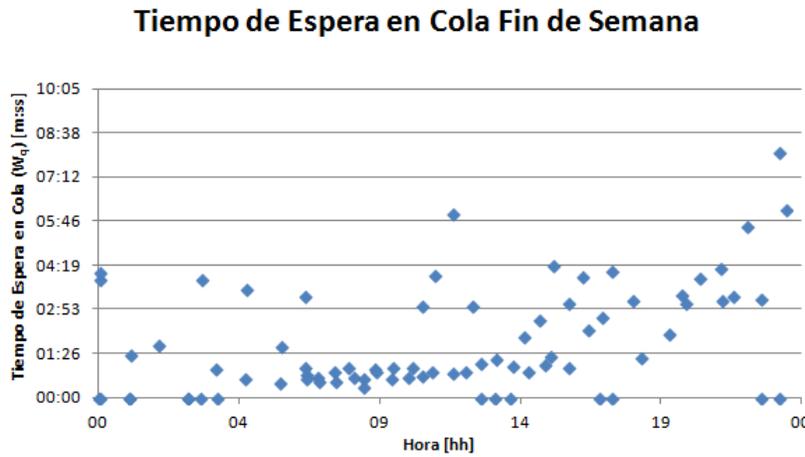


Figura B.2.2: Distribución de tiempos de espera en Cola en días de semana.

Día	Tiempo en de Espera en Cola Promedio por Hora [mm : ss]
Semana	01:03
Fin de Semana	02:29

Tabla B.2.1: Promedio Tiempo de Espera en Cola.

Horario	Días de Semana		Fin de Semana	
	Tiempo de Espera en Cola [mm : ss]	Tamaño Muestral	Tiempo de Espera en Cola [mm : ss]	Tamaño Muestral
1	00:33	37	02:14	16
2	00:33	45	01:20	28
3	00:47	9	01:12	4
4	00:47	34	02:20	16
5	02:14	34	04:26	16

Tabla B.2.2: Tiempo de espera en cola por horario.

# Apéndice C

## Salidas Calculadora Erlang-X

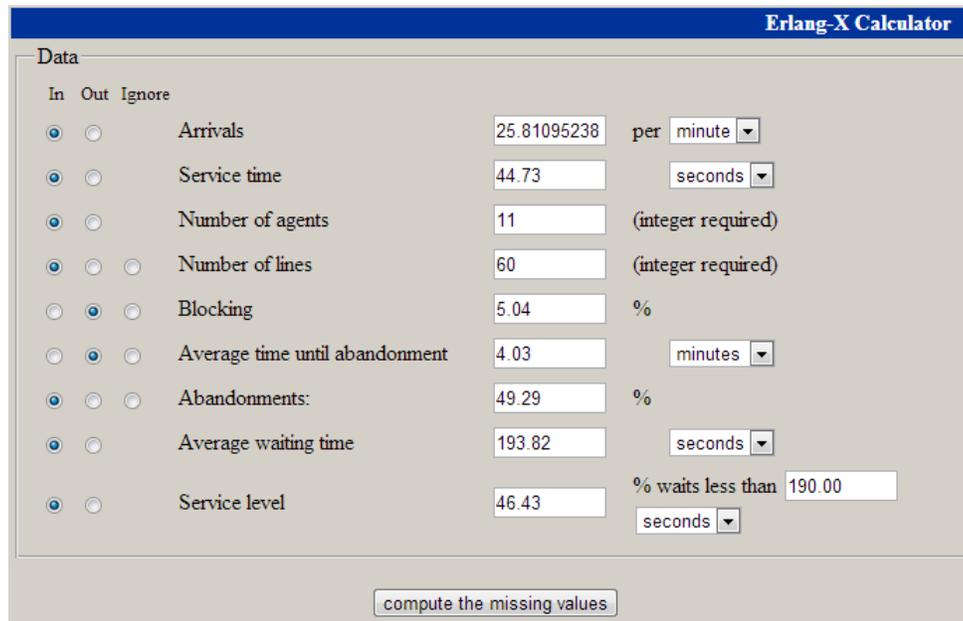
### C.1. Horario 1

The screenshot shows the 'Erlang-X Calculator' window. It has a title bar with the text 'Erlang-X Calculator'. Below the title bar is a section labeled 'Data' containing a table of input and output parameters. Each row has three radio buttons labeled 'In', 'Out', and 'Ignore'. The 'Arrivals' row has 'In' selected. The 'Service time' row has 'Out' selected. The 'Number of agents' row has 'In' selected. The 'Number of lines' row has 'Ignore' selected. The 'Blocking' row has 'Out' selected. The 'Average time until abandonment' row has 'Out' selected. The 'Abandonments:' row has 'Out' selected. The 'Average waiting time' row has 'Out' selected. The 'Service level' row has 'Out' selected. The 'Arrivals' value is 6.718888888 per minute. The 'Service time' value is 75.01776466 seconds. The 'Number of agents' value is 11. The 'Average waiting time' value is 8.99 seconds. The 'Service level' value is 99.97 % waits less than 200.00 seconds. A 'compute the missing values' button is located at the bottom of the window.

Parameter	In	Out	Ignore	Value	Unit
Arrivals	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	6.718888888	per minute
Service time	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	75.01776466	seconds
Number of agents	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	11	(integer required)
Number of lines	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>		(integer required)
Blocking	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>		%
Average time until abandonment	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>		seconds
Abandonments:	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>		%
Average waiting time	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	8.99	seconds
Service level	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	99.97	% waits less than 200.00 seconds

Figura C.1.1: Captura de pantalla. Resultados Erlang-X en Horario 1.

## C.2. Horario 2



The screenshot shows the 'Erlang-X Calculator' interface. It features a 'Data' section with a table of input and output parameters. Each row has three radio buttons labeled 'In', 'Out', and 'Ignore'. The 'In' radio button is selected for all rows. The parameters and their values are as follows:

Parameter	Value	Unit/Note
Arrivals	25.81095238	per minute
Service time	44.73	seconds
Number of agents	11	(integer required)
Number of lines	60	(integer required)
Blocking	5.04	%
Average time until abandonment	4.03	minutes
Abandonments:	49.29	%
Average waiting time	193.82	seconds
Service level	46.43	% waits less than 190.00 seconds

At the bottom of the interface, there is a button labeled 'compute the missing values'.

Figura C.2.1: Captura de pantalla. Resultados Erlang-X en Horario 2.

# Apéndice D

## Salidas Simulación

### D.1. M/M/c Horario 1

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	1.0006	0.08	0.5973	1.3979	0.00	31.0000

Figura D.1.1: Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c en horario 1.

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	8.9417	0.71	5.4548	12.3620	0.00	233.56

Figura D.1.2: Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c en horario 1.

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Llamada	9.3799	0.13	8.7257	10.1003	0.00	42.0000

Figura D.1.3: Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c en horario 1.

Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion	83.9696	0.93	79.6884	88.9806	0.00016035	1064.11

Figura D.1.4: Tiempo en el sistema ( $W$ ). Salida simulación M/M/c en horario 1.

### D.2. M/M/c/K Horario 2

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	47.5750	0.01	47.4983	47.6630	0.00	49.0000

Figura D.2.1: Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c/K en horario 2.

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	193.21	0.42	190.34	195.82	0.00	338.35

Figura D.2.2: Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c/K en horario 2.

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Llamada	58.5729	0.01	58.4962	58.6609	0.00	61.0000

Figura D.2.3: Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c/K en horario 2.

Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion	237.92	0.52	234.44	241.11	0.05047705	822.15

Figura D.2.4: Tiempo en el sistema ( $W$ ). Salida simulación M/M/c/K en horario 2.

### D.3. M/M/c/K 24 Horas

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	34.2502	0.04	34.0926	34.6236	0.00	50.0000

Figura D.3.1: Largo promedio de cola ( $L_q$ ). Salida simulación M/M/c/K

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	146.93	0.43	143.74	149.45	0.00	368.51

Figura D.3.2: Tiempo de espera en cola ( $W_q$ ). Salida simulación M/M/c/K.

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Llamada	46.4769	0.06	46.2089	46.9471	0.00	61.0000

Figura D.3.3: Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación M/M/c/K.

Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion	199.43	0.58	195.13	202.68	0.00016034	1064.08

Figura D.3.4: Tiempo en el sistema ( $W$ ). Salida simulación M/M/c/K.

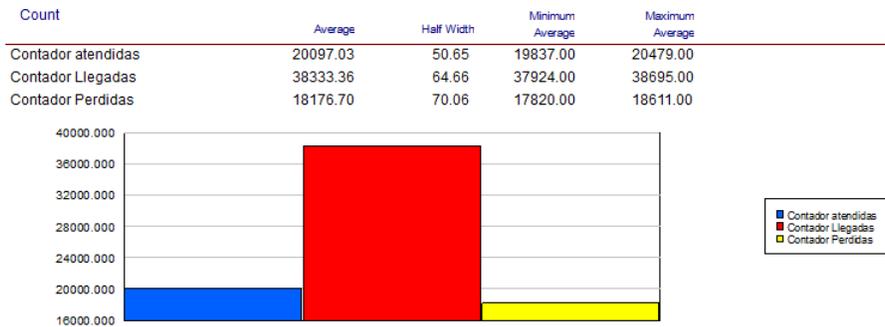


Figura D.3.5: Contador de llamadas M/M/c/K

### D.4. Simulación con Abandonos

Number Waiting	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	5.7675	0.02	5.6525	5.9102	0.00	31.0000

Figura D.4.1: Largo promedio de cola ( $L_q$ ). Salida simulación con Abandonos

Waiting Time	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion.Queue	12.9979	0.03	12.8129	13.1773	0.00	37.0710

Figura D.4.2: Tiempo de espera en cola ( $W_q$ ). Salida simulación con Abandonos.

WIP	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Llamada	17.8295	0.03	17.7049	18.0404	0.00	49.0000

Figura D.4.3: Cantidad de Clientes en el Sistema ( $L$ ). Salida simulación con Abandonos.

Total Time Per Entity	Average	Half Width	Minimum Average	Maximum Average	Minimum Value	Maximum Value
Atencion	63.3772	0.17	62.3025	64.5013	0.00004950	1064.08

Figura D.4.4: Tiempo en el sistema ( $W$ ). Salida simulación con Abandonos.



Figura D.4.5: Contador de llamadas. Simulación con Abandonos y Capacidad Finita.

## Apéndice E

# Resultados Nivel de Servicio

### E.1. Semana

	Tipo de llamada	Tiempo de Espera en Cola [seg]	Porcentaje [%]	Cantidad Promedio de Llamadas por día
Atendidas	Todas	-	48.49	18604
	Ineficaces	-	42.63	16354
		< 10	16.94	6500
		< 20	23.17	8888
		< 30	2.51	964
		> 30	0.01	2
	Emergencias		5.86	2249
		< 10	2.66	1020
		< 20	2.89	1107
		< 30	0.32	122
	> 30	0.00	0	
Abandonadas	Todas	-	51.51	19759
	Ineficaces	-	45.01	17267
		< 10	3.88	1488
		< 20	32.71	12547
		< 30	8.38	3216
		> 30	0.04	16
	Emergencias	-	6.49	2491
		< 10	0.57	219
		< 20	4.70	1803
		< 30	1.22	467
	> 30	0.01	2	
Total			100.00	38363

Tabla E.1.1: Nivel de servicio para todas las llamadas. Días hábiles.

## E.2. Fin de Semana

	Tipo de llamada	Tiempo de Espera en Cola [seg]	Porcentaje [%]	Cantidad de Llamadas
Atendidas	Todas	-	32.39	12577.30
	Inoficiosas	-	25.54	9916.89
		< 10	2.84	1104.67
		< 20	15.52	6027.04
		< 30	7.09	2754.22
		> 30	0.08	30.96
	Emergencias	-	6.85	2660.41
		< 10	0.67	259.85
		< 20	4.16	1616.78
		< 30	1.99	774.30
	> 30	0.02	9.48	
Abandonadas	Todas	-	67.61	26258.81
	Inoficiosas	-	52.47	20378.93
		< 10	0.86	332.67
		< 20	23.68	9196.22
		< 30	27.26	10585.22
		> 30	0.68	264.81
	Emergencias	-	15.14	5879.89
		< 10	0.23	87.81
		< 20	6.78	2632.74
		< 30	7.94	3082.48
	> 30	0.20	76.85	
Total			100.00	38836.11

Tabla E.2.1: Nivel de servicio para todas las llamadas. Fin de Semana

## Apéndice F

# Información Istitucional

### F.1. Fono 133

The screenshot displays the institutional portal of the Carabineros de Chile. At the top, it features the name of the Director, Gustavo González Jure, the Carabineros de Chile logo, the text 'CARABINEROS DE CHILE PORTAL INSTITUCIONAL', the Chilean flag, and the slogan 'Un Amigo Siempre!'. Below this is a banner image of a green police car at night with officers, and a navigation bar with 'INICIO' and 'FONOS DE EMERGENCIA'. The main content area includes a call to action: 'Llámenos, estamos a su disposición' with an image of a smiling operator and a hand holding a mobile phone. A video player shows a '133 Emergencia Policial' logo. To the right of the video, text explains that when dialing 133 with high traffic, a voice guide indicates that calls may not be answered immediately and advises staying on the line. It also provides statistics for 2011: 11,594,279 calls received, 2,026,274 answered (21.79%), 9,068,000 real emergencies, and 51,760 calls resulting in 6,921 police procedures. At the bottom, a section titled 'OTROS NUMEROS DE EMERGENCIA' lists: 133 Fono de Emergencias Policiales, 135 Fono Drogas, 139 Fono Informaciones Policiales, 147 Fono Niños, and 149 Fono Familia.

Figura F.1.1: Captura de Pantalla. Descripción del Fono 133

Fuente: <http://www.carabineros.cl>

## F.2. Sistemas de Atención de llamadas

A través del sistema Power 911, el operador puede contestar una llamada, registrar la información del usuario, localizar la dirección en un mapa, clasificar las llamadas inoficiosas y tipificar las emergencias. Al finalizar este proceso, el operador envía los requerimientos de procedimientos a Alpha I usando la interfaz del Viper CAD.

Para contestar una llamada, el operador presiona el botón "133", se cuelga con el botón "Liberar". El botón "No listo" permite una desconexión temporal cuando el operador esté ocupado en otras tareas y no esté listo para contestar una nueva llamada.



Figura F.2.1: Panel de recepción de llamada en Power 911

Al contestar una llamada se activa el panel para registrar los datos del llamante (ver figura F.2.2), donde se debe escribir la dirección (1), el sistema automáticamente muestra un mapa indicando el punto de ubicación, el operador lo puede manipular. Luego se ingresan notas o antecedentes (2). Se tipifica la emergencia eligiendo o digitando uno de los 103 tipos de emergencias que ofrece el sistema (3), el sistema sugiere preguntas para realizar al usuario y así ayudar al operador a seleccionar la tipificación. Finalmente se confirma la zona de despacho (4) y se envía el requerimiento al despachador correspondiente en Alpha I (5).

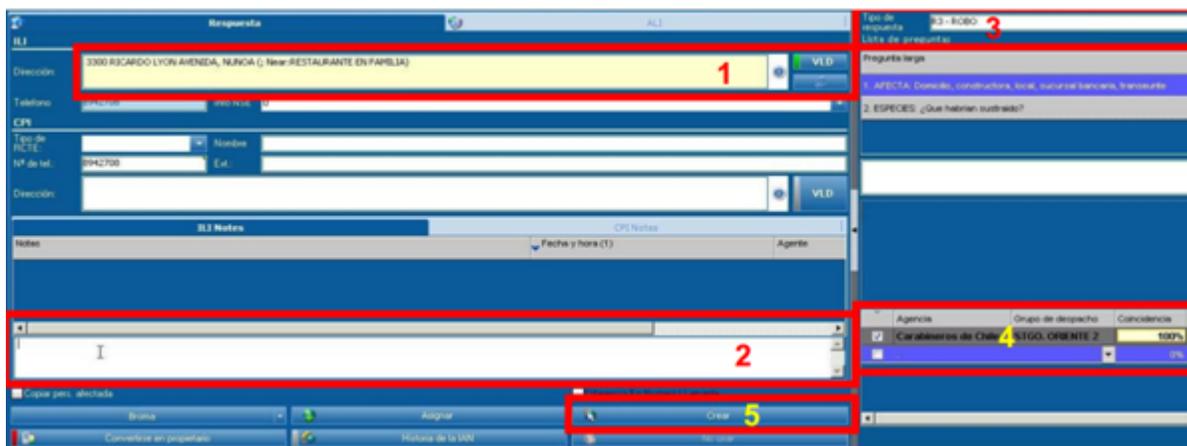


Figura F.2.2: Ingreso de datos de la llamada en Power 911

Los despachadores de la Unidad de Procedimientos Alpha I gestionan la acción policial según la prioridad que se le asignó a la emergencia.

Si la llamada no genera procedimiento, no se llenan aquellos campos. El operador procede a clasificar la llamada en una de las seis categorías de llamadas inoficiosas. En la imagen anterior aparece la opción "Broma" seleccionada. Luego se libera la llamada sin crear requerimiento.

### F.3. Tipificación de Emergencias

Códigos de tipificación, descripción y prioridad asignada por *default* a cada emergencia.

Nro.	Tipo	Prioridad	Nro.	Tipo	Prioridad
1	ABANDONO DE HOGAR	7	53	HURTO	5
2	ABIGEATO	5	54	INCENDIO	2
3	ABIGEATO EFECTUANDOSE	2	55	INDIVIDUO O GRUPO ARMADO	1
4	ABUSO SEXUAL	1	56	INDIVIDUO INTER. DOMICILIO (SOSP)	1
5	ACC. LABORAL	2	57	INDIVIDUO MERODEANDO (SOSP)	2
6	ACC. TTO. SOLO DAÑOS	5	58	INFORMACION	7
7	ACC. TTO. LESIONADOS/MUERTOS	2	59	INFRACCIÓN OTRAS LEYES	6
8	ACC. DE AVIACIÓN	1	60	INTENTO DE SUICIDIO	1
9	ACC. FERROVIARIO	1	61	INTOXICACIÓN	1
10	AGRESION	3	62	INUNDACION	2
11	AGRESION EFECTUANDOSE	1	63	IMPUTADO TIENDA COMERCIAL	5
12	ALARMA ACTIVADA	2	64	LESIONADO	3
13	ALARMA ALPHA II ACTIVADA	1	65	MALTRATO ANIMAL	5
14	ALARMA ALPHA II (P. COM.)	1	66	MALTRATO INFANTIL	1
15	ALARMA CAJERO AUTOMATICO	1	67	MATADERO CLANDESTINO	6
16	AMENAZA DE ARMAS	1	68	MEDIDAS CAUTELARES	1
17	AMENAZA DE PALABRA	6	69	MUDANZA SIN SALVO CONDUCTO	6
18	AMENAZA DE ATENTADOS	7	70	MUERTO	3
19	ANIMAL SUELTO VIA PUBLICA	6	71	OBTRUCCION VIA PUBLICA	7
20	BARRICADAS	2	72	OCUPACION O TOMA DE PROPIEDAD	5
21	CABLES CORTADOS	5	73	OFENSA A LA MORAL	6
22	CARAB. ATACADO/DIFICULTADES	1	74	PAQUETE SOSPECHOSO	1
23	CARAB. CON IMPUTADOS	1	75	PARTO	1
24	CATASTROFE NATURAL	1	76	PERSONA EXTRAVIADA	4
25	CIVILES CON IMPUTADOS	1	77	PRESUNTA DESGRACIA	4
26	COMERCiantES SIN PERMISO	6	78	PROSTITUCION VIA PUBLICA	6
27	CONDUCTOR EBRIO	6	79	PROTESTA/PARO Y/O MARCHA	2
28	CONGESTION DE TRANSITO	7	80	RAPTO	1
29	CONTAMNACION AMBIENTAL	6	81	RIÑA/PENDENCIA	1
30	CONSTANCIA DEL SERVICIO	7	82	ROBO	4
31	CONTROL DE VEHICULO	7	83	ROBO EFECTUANDOSE	1
32	COOPERACION	7	84	ROBO ARMADO	2
33	CORTE DE ENERGIA ELECTRICA	7	85	ROBO ARMADO EFECTUANDOSE	1
34	DAÑOS	5	86	ROBO DE VEHICULO	2
35	DAÑOS EFECTUANDOSE	2	87	ROBO DE VEHICULO EFECTUANDOSE	1
36	DERRUMBE	7	88	RUIDOS MOLESTOS	7
37	DESORDEN	6	89	SECUESTRO	1
38	DISPAROS	3	90	SEMAFORO MAL ESTADO	7
39	DROGAS VENTA/CONSUMO	4	91	SERVICIO ESPECIAL	7
40	EBRIOS	6	92	SUSTRACCION DE MENOR	1
41	ENCARGO (PERSONA/VEHICULO)	7	93	TRABAJOS DE CONSTRUCCION	6
42	ENFERMO/INSANO VIA PUBLICA	7	94	ULTRAJE A LA BANDERA	6
43	ESCAPE O FUGA DE GAS	1	95	USURPACION DE FUNCIONES Y/O IDENTIDAD	6
44	ESCOLTA A	7	96	VEHICULO ABANDONADO	6
45	ESTAFA	5	97	VEHICULO MAL ESTACIONADO	6
46	EVASION REOS/IMPUTADO	1	98	VEHICULO EN PANNE	7
47	EXPLOSIVO	1	99	VIGILANCIA ESPECIAL	7
48	FUEGOS ARTIFICIALES SIN PERMISO	3	100	VIOLACION	3
49	HALLAZGO ARMAMENTO	6	101	V.I.F.	3
50	HALLAZGO CADAVER	3	102	V.I.F. EFECTUANDOSE	1
51	HALLAZGO FETO	3	103	VULNERACION DE SUS DERECHOS	3
52	HOMICIDIO	3	104	PROC. DE PRUEBA	7

Tabla F.3.1: Tipos de Emergencias.