



**TESIS DE GRADO
EN INGENIERÍA INDUSTRIAL**

**DIFUSIÓN DE PRODUCTOS
- UN ENFOQUE EN REDES SOCIALES
Y SIMULACIÓN -**

Autor: Hernán Masaji Kina

Director de Tesis: Prof. Ing. Andrés Agres

2007

*A mis padres y hermana Gaviot,
por su apoyo y afecto.*

REFERENCIA BIBLIOGRÁFICA

Este trabajo propone una metodología práctica para el estudio de difusión de productos, incorporando el análisis de las redes sociales. De esta manera, se busca complementar las prácticas usuales de difusión de productos y a su vez acercar una metodología que integra diferentes herramientas permitiendo testear y diseñar nuevas estrategias de difusión. El trabajo esta acompañado por un caso práctico en el cual se aplica la metodología para el análisis de la difusión de productos en una compañía de telecomunicaciones.

Palabras Clave: Difusión de productos, Difusión en redes sociales, Difusión de innovaciones.

ABSTRACT

This work proposes a practical methodology for the study of product diffusion while considering the influence of the social network. Its intention is to complement present practices applied to product diffusion and at the same time to facilitate an integral methodology that allows to test and design new diffusion strategies. The work is accompanied by a case study that applies the methodology to the analysis of product diffusion in a telecommunications company.

Keywords: Product Diffusion, Diffusion in social networks, Innovations diffusion.

RESUMEN EJECUTIVO

La difusión del producto es un paso crucial e inevitable que enfrentan muchas empresas. En general, las estrategias de difusión de productos están enfocadas en que el potencial cliente directamente contactado se convierta en adoptador. Sin embargo, no se hace mucho énfasis en potenciar el efecto boca en boca ni tampoco se realizan estudios de la estructura de red social para diseñar las estrategias de difusión.

En base a lo expresado, este trabajo integra diferentes herramientas y teorías en una metodología práctica que permite considerar el impacto de las redes sociales en la difusión del producto. Dicha metodología se compone de seis etapas: La definición del problema a analizar, la representación de la topología de red social, el análisis de la red social, la teoría del comportamiento, el diseño de experimentos y las conclusiones.

En la primera sección del trabajo se introduce brevemente el estado de la cuestión. Posteriormente, se explica en detalle las seis etapas de la metodología propuesta y finalmente se realiza un caso práctico analizando la difusión de productos en una compañía de telecomunicaciones. En el caso práctico se estudia la red social de un pueblo en el cual la compañía brinda sus servicios.

El trabajo acerca un nuevo enfoque a la difusión de productos basado en la simulación y las redes sociales. El mismo permite considerar la estructura de las redes sociales en las cuales se insertará el producto, ensayar nuevas ideas y conocer su potencial impacto y definir la red social en estudio con el fin de complementarse con las técnicas actuales y mejorar la eficiencia de los procesos de difusión.

EXECUTIVE BRIEF

The product diffusion is a crucial and inevitable step that many companies must face. Generally, the product diffusion strategies are focused on converting the direct contacted person into an adopter. However, there is not much focus in leveraging the word of mouth effect and there are not many studies of the social network structure to design diffusion strategies.

As a way of tackling the expressed problems, this work integrates different tools

and theories into a practical methodology that allows considering the impact of the social networks in the product diffusion process. Such methodology is composed by six steps: The problem definition, the representation of the social network topology, the analysis of the social network, the theory of behaviour, the experiments design and the conclusions.

In the first section, the state of the art that investigates the problem is overviewed. Secondly, the six steps of the proposed methodology are explained in detail and finally a case study that analyses the product diffusion in a telecommunications company is presented.

This work presents a new point of view for the product diffusion based on simulation and social networks. This framework can be used to consider the social network structure in which the product will be marketed, to test new ideas while understanding their impact and to define the social network under study with the objective of complementing the present techniques and enhancing the efficiency of the product diffusion.

AGRADECIMIENTOS

Me gustaría agradecer a mi familia, Masaaki, Mitsuko y Naomi por su apoyo constante e incondicional en cada proyecto emprendido. También deseo agradecer a mi tutor, el Ing. Andrés Agres, quien además de guiarme en diferentes etapas de la tesis, fue buen consejero e hizo posible la realización del caso práctico con la compañía de telecomunicaciones.

Hernán Masaji Kina
Hernan_Kina@yahoo.com

ÍNDICE

1. INTRODUCCIÓN	1
2. ESTADO DE LA CUESTIÓN	3
2.1 Estrategias de Difusión Actuales	3
2.2 Difusión de Innovaciones	4
2.3 Análisis de Redes Sociales	11
2.4 Topología de Red	14
2.5 Simulación Basada en Agentes	20
3. DEFINICIÓN DEL PROBLEMA	25
4. METODOLOGÍA PROPUESTA	29
4.1 Primer Paso: Identificación del Problema	32
4.2 Segundo Paso: Representación de la Topología de Red Social	32
4.3 Tercer Paso: Análisis Estático de la Red Social	36
4.4 Cuarto Paso: Teoría del Comportamiento	38
4.5 Quinto Paso: Diseño de Experimentos	40
4.6 Sexto Paso: Conclusiones	42
5. CASO PRÁCTICO: COMPAÑÍA DE TELECOMUNICACIONES	43
5.1 Definición del Problema	47
5.2 Representación de la Topología de Red	49
5.3 Análisis Estático de la Red Social	69
5.4 Teoría del Comportamiento	71
5.5 Diseño de Experimentos	74
5.6 Resultados	84
5.7 Conclusiones	91
6. CONCLUSIONES GENERALES	95
6.1 Futuras Líneas de Investigación	96
7. BIBLIOGRAFÍA	99

LISTA DE FIGURAS

Figura N° 1: Áreas de estudio de la teoría de difusión	6
Figura N° 2: Tipos de Adopciones en el Modelo de Bass	8
Figura N° 3: Estructura analítica del Modelo de Bass	9
Figura N° 4: Esquema de una Red Libre de Escala	18
Figura N° 5: Esquema de la Metodología Propuesta	31
Figura N° 6: Orientación de la sociabilidad cara a cara y telefónica a través del ciclo de vida	34
Figura N° 7: Clasificación de Jager y Janssen	40
Figura N° 8: Paradigmas de simulación en Anylogic.....	41
Figura N° 9: Flujo de llamadas	52
Figura N° 10: Gráfico en Escala Logarítmica	55
Figura N° 11: Gráfico en Escala Logarítmica utilizando rangos logarítmicos ...	56
Figura N° 12: Método Base. Histograma.....	60
Figura N° 13: Método Base. Ajuste de Datos.....	62
Figura N° 14: Método Tiempo. Histograma	64
Figura N° 15: Método Tiempo. Ajuste de Datos	65
Figura N° 16: Método Llamada Recíproca. Histograma.....	67
Figura N° 17: Método Llamada Recíproca. Ajuste de Datos	68
Figura N° 18: Selección del modelo más adecuado.....	72
Figura N° 19: Panel de control de la simulación.....	75
Figura N° 20: Red social de base simulada. Histograma	77
Figura N° 21: Red social de base simulada. Ajuste de datos.....	77
Figura N° 22: Red social de base simulada. Enlaces de un nodo.....	80
Figura N° 23: Diagrama de Estados de nodos	81
Figura N° 24: Evolución de Adoptadores ADSL-Monte Equis	82
Figura N° 25: Evolución de Adoptadores	84
Figura N° 26: Etapas de Análisis de la Difusión	85
Figura N° 27: Sensibilidad Sobre el Grado Nodal de la Red	87
Figura N° 28: Sensibilidad Sobre el N° Adoptadores Iniciales	89
Figura N° 29: Variables del Sistema	92

LISTA DE TABLAS

Tabla N° 1: Método Base. Tabla de Frecuencias	59
Tabla N° 2: Método Base. Tabla en Logaritmos.....	61
Tabla N° 3: Método Tiempo. Tabla de Frecuencias	63
Tabla N° 4: Método Tiempo. Tabla en Logaritmos.....	64
Tabla N° 5: Método Llamada Recíproca. Tabla de Frecuencias	66
Tabla N° 6: Método Llamada Recíproca. Tabla en Logaritmos.....	67
Tabla N° 7: Descripción de la Red Social de Monte Equis.....	70
Tabla N° 8: Red Social de Base.....	74
Tabla N° 9: Modelo de la red social simulada. Tabla de Frecuencias	76
Tabla N° 10: Identificación de Hubs	79
Tabla N° 11: Tabla de Tiempos. Comparación de Estrategias.....	84
Tabla N° 12: Tabla de tiempos. Comparación por Grado Nodal	87
Tabla N° 13: Tabla de tiempos. Comparación por N° Adoptadores Iniciales ...	89

1. INTRODUCCIÓN

En general, las estrategias de difusión de productos están enfocadas en que el potencial cliente directamente contactado se convierta en adoptador. Sin embargo, no se hace mucho énfasis en potenciar el efecto boca en boca ni tampoco se realizan estudios de la estructura de red social para diseñar las estrategias de difusión.

Si bien existen diferentes herramientas para estudiar la difusión de innovaciones y las redes sociales, éstas se encuentran fragmentadas y no existe una metodología integradora que permita realizar análisis estáticos y dinámicos del problema.

En las últimas décadas se ha comenzado a descubrir la ubicuidad de las redes. Dentro de la gran diversidad de redes se encuentra la red social cuya esencia se basa en la identificación de las relaciones existentes entre los individuos, empresas, instituciones u otras organizaciones de la sociedad. El análisis de las redes sociales permite conocer su estructura e identificar a los diferentes actores de acuerdo a su grado de conexión o posición en la red social.

El aporte que este trabajo desea ofrecer es acercar un nuevo enfoque basado en la simulación y las redes sociales que permita complementar las prácticas actuales, haciendo más eficiente el proceso de difusión de innovaciones y sirviendo a su vez de fundamento para el desarrollo de planes estratégicos de difusión de productos.

En primer lugar, se presentan brevemente las estrategias actuales de difusión y el estado de las herramientas y teorías que se utilizan luego para el desarrollo del trabajo. En segundo lugar, se define en detalle el problema a abordar y las limitaciones que se han impuesto al trabajo. En tercer lugar, se presenta la metodología propuesta para el análisis de difusión de productos en redes sociales. La misma consta de seis fases: La identificación del problema, la representación de la topología de red social, el análisis estático de la red social, la teoría del comportamiento, el diseño de experimentos y las conclusiones. Posteriormente se aplica la metodología desarrollada en la resolución de un caso práctico con una empresa de telecomunicaciones. En el mismo se aborda el problema de la difusión de sus productos en Monte Equis¹ y se investiga el impacto del efecto boca en boca utilizando diferentes estrategias de difusión.

¹ Por cuestiones de confidencialidad se ha modificado el nombre del pueblo y de la compañía de telecomunicaciones sobre la cual se basa el caso práctico.

Finalmente, se establecen las conclusiones del caso práctico y del trabajo en general y se menciona brevemente algunos temas que sería de utilidad abordar en futuras líneas de investigación.

2. ESTADO DE LA CUESTIÓN

En la actualidad existen diferentes herramientas fragmentadas que estudian la difusión de innovaciones por un lado y las redes sociales por otro. Sin embargo, es raro encontrar metodologías integradoras con un enfoque práctico que permitan analizar la difusión de innovaciones aprovechando la información de las redes sociales mediante un enfoque estático y dinámico.

A continuación, se hará una breve mención sobre las estrategias de difusión actuales que están siendo empleadas. Posteriormente, se repasa el estado de algunas de las herramientas y teorías que estudian la difusión de innovaciones y las redes sociales. Al final de la sección, se introducirá un tipo de simulación denominado simulación basada en agentes que tiene un gran potencial de aplicación en el estudio de la difusión en redes sociales.

2.1 Estrategias de Difusión Actuales

En general, las estrategias de difusión de productos actuales se enfocan en que el potencial cliente directamente contactado se convierta en adoptador pero no analizan el alcance del efecto boca en boca ni tampoco hacen un estudio de la red social en la que el mercado objetivo esta inmerso.

Si bien existen algunas estrategias de difusión que contemplan de alguna manera el efecto boca en boca, en general no estudian la estructura de la red social y se basan en conceptos intuitivos cuya efectividad es difícil de medir. Por ejemplo las promociones por recomendación o por grupo (amigos, familiares, etc.) o el marketing viral. Dependiendo del producto que se pretende difundir estas herramientas son altamente efectivas, aunque poseen el inconveniente que al no analizar la red social las estrategias se plantean de manera general y no suelen ser personalizadas.

A continuación, se repasa brevemente la teoría del marketing viral.

Marketing Viral

El marketing viral puede ser definido como una estrategia que incentiva a que los individuos transmitan rápidamente un mensaje comercial a otros de manera tal de crear un crecimiento exponencial en la exposición de dicho mensaje.

“El marketing viral se utiliza generalmente para producir incrementos en conocimiento de marca y se logra mediante procesos de autorreplicación viral análogos a la expansión de un virus informático. Se suele basar en el boca a oreja mediante medios electrónicos; usa el efecto de red social creado por Internet y los modernos servicios de telefonía móvil para llegar a una gran cantidad de personas rápidamente.

También se usa el término marketing viral para describir campañas de marketing encubierto basadas en Internet, incluyendo el uso de blogs, de sitios aparentemente amateurs, y de otras formas diseñadas para crear el boca a oreja para un nuevo producto o servicio. Frecuentemente, el objetivo de las campañas de marketing viral es generar cobertura mediática mediante historias inusuales.

El término publicidad viral se refiere a la idea que la gente se pasará y compartirá contenidos divertidos e interesantes. Esta técnica a menudo está patrocinada por una marca, que busca generar conocimiento de un producto o servicio. Los anuncios virales toman a menudo la forma de divertidos video clips o juegos Flash interactivos, imágenes e incluso textos.

La popularidad creciente del marketing viral se debe a la facilidad de ejecución de la campaña, su coste relativamente bajo, (comparado con campañas de correo directo), buen targeting, y una tasa de respuesta alta y elevada. La principal ventaja de esta forma de marketing consiste en su capacidad de conseguir una gran cantidad de posibles clientes interesados, a un bajo costo”.²

2.2 Difusión de Innovaciones

La difusión de innovaciones fue formalizada por Everett Rogers en un libro editado en 1962, titulado “Difusión de Innovaciones”. Rogers estableció que los adoptadores de una innovación o idea pueden ser categorizados como innovadores (2,5%), adoptadores tempranos (13,5%), mayoría temprana (34%), mayoría tardía (34%) y laggards (17%), basados en una curva campana.

Algunas de las características de cada categoría de adoptadores incluyen:

- Innovadores: Aventureros, educados, poseen múltiples fuentes de información y una mayor propensión al riesgo.

² Miguel Aguirre; “Impactos del marketing y la tecnología” El marketing viral, efectividad en tiempos competitivos.

- Adoptadores tempranos: Líderes sociales, populares y educados.
- Mayoría temprana: Deliberados, poseen muchos contactos sociales informales.
- Mayoría tardía: Escépticos, tradicionales, bajo estatus socioeconómico.
- Laggards: Los vecinos y amigos son sus principales fuentes de información.

Los modelos de difusión de innovaciones fueron tradicionalmente utilizados en contextos de pronóstico de ventas. Sin embargo, así como señala Majan y Wind (1986) y Kalish y Lilien (1986), el pronóstico de ventas es sólo uno de los objetivos de los modelos de difusión. Además de pronósticos, los modelos de difusión pueden ser utilizados con objetivos de tipo descriptivo y normativo. Dado que los modelos de difusión son un acercamiento analítico a la descripción del fenómeno de difusión, pueden ser utilizados para testear hipótesis de difusión específicas. Adicionalmente, como los modelos de difusión están diseñados para capturar el ciclo de vida de un producto, pueden ser utilizados con objetivos normativos como base para conocer la manera en que un producto debe ser promocionado.

La difusión de un producto es el proceso mediante el cual el producto es comunicado a través de ciertos canales a través del tiempo entre los miembros de un sistema social. Al igual que la teoría de la comunicación, la teoría de la difusión se centra en el estudio de los diferentes canales de información, que son los medios a través de los cuales la información sobre una innovación es transmitida hacia o en el medio social. Estos medios se resumen en los medios masivos de comunicación y las comunicaciones interpersonales. Los miembros de un sistema social tienen diferentes actitudes y pueden ser más o menos propensos a confiar en los medios masivos o en los canales interpersonales cuando buscan información sobre una innovación. Las comunicaciones interpersonales, incluidas las observaciones no verbales, son influencias importantes que determinan la velocidad y la forma del proceso de difusión en un sistema social.

La teoría de difusión de innovaciones fue introducida en el marketing en la década del 60 (Arndt 1967; Bass 1969; Frank, Massy y Morrison 1964; otros) y a partir de ese momento fue analizado con el objetivo de nutrir áreas como el estudio de la conducta del consumidor, administración del marketing y la ciencia de la administración. Investigadores de la conducta del consumidor han evaluado la aplicabilidad del proceso general de difusión al estudio del consumidor. La literatura del marketing y administración se focalizó en las implicancias de estas hipótesis en la definición de los prospectos de nuevos productos y en el desarrollo de estrategias de marketing que tienen el objetivo

de atraer potenciales adoptadores. Otros investigadores han contribuido en el desarrollo de la teoría de la difusión sugiriendo modelos analíticos para describir y pronosticar la difusión de una innovación en un sistema social. A su vez, existe un interés creciente en desarrollar guías normativas sobre cómo las innovaciones deberían ser difundidas en un sistema social.

La teoría de la difusión puede ser dividida en cinco áreas de estudio:

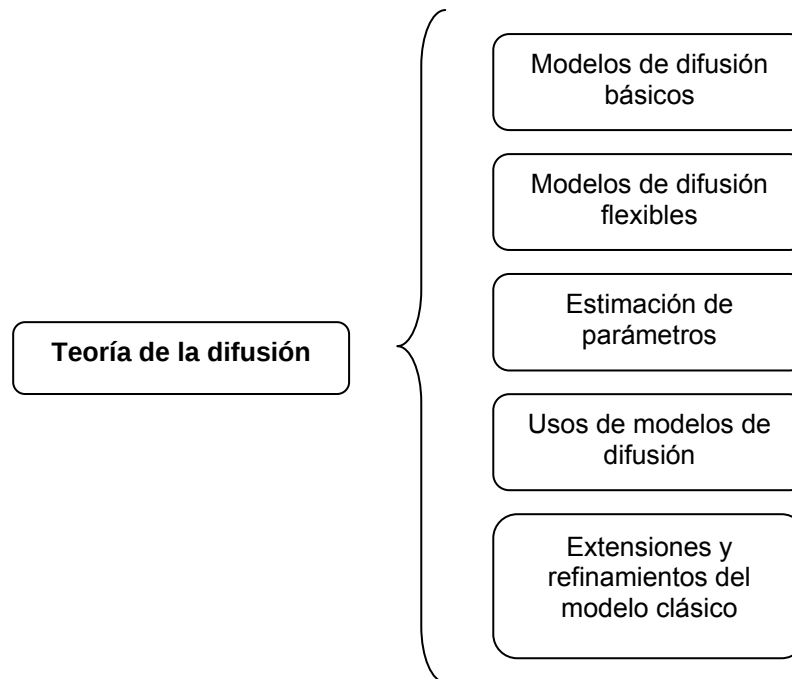


Figura N° 1: Áreas de estudio de la teoría de difusión

Proceso de difusión de innovaciones y proceso de compra

Rogers propuso un modelo de cinco estados para el proceso de difusión de innovaciones:

- Conocimiento: Aprender acerca de la existencia y función de la innovación.
- Persuasión: Convencerse acerca del valor de la innovación.
- Decisión: Comprometerse a la adopción de la innovación.
- Implementación: Ponerla en uso.
- Confirmación: La aceptación (o rechazo) final de la innovación.

En base a este modelo, dentro del área del marketing se ha esquematizado el proceso de decisión de compra típico. En este el consumidor pasa por cinco etapas: reconocimiento del problema, búsqueda de información, evaluación de alternativas, decisión de compra y conducta posterior a la compra. Es necesario destacar que en ocasiones los consumidores podrían saltarse algunas etapas o invertir su orden. No obstante, este modelo es útil dado que reúne la gama completa de consideraciones que surgen cuando un consumidor enfrenta una compra nueva de alta participación.

Modelos de difusión: “Aguja Hipodérmica” y “En Dos Pasos”

Al igual que la teoría de la comunicación, la teoría de la difusión se centra en el estudio de los diferentes canales de información, que son los medios a través de los cuales la información sobre una innovación es transmitida hacia o en el medio social. Estos medios se resumen en:

- Los medios masivos de comunicación.
- Las comunicaciones interpersonales (Efecto denominado Boca en Boca).

Surgen dentro de la teoría de difusión dos modelos principales que se diferencian en la temporalidad y preponderancia otorgada a los dos medios citados anteriormente.

En primer lugar, surge el “Modelo de la Aguja Hipodérmica” que postulaba que los medios masivos tenían efectos poderosos, directos e inmediatos en los miembros individuales de una audiencia masiva.

Posteriormente surge el “Modelo de difusión en Dos Pasos”, basado en la hipótesis que los mensajes fluyen desde una fuente, a través de canales de medios masivos, hacia los líderes de opinión, que luego transmiten los mensajes a los seguidores. Este modelo refuerza la importancia del Boca en Boca en el proceso de difusión de innovaciones.

Modelo De Bass

El modelo de Bass asume que los adoptadores potenciales de una innovación son influidos a través de dos medios de comunicación:

- Los medios masivos de comunicación (Radio, TV, revistas, etc.)
- Boca en Boca

En su desarrollo asume que los adoptadores de una innovación comprenden dos grupos. Un grupo influenciado solamente por los medios de comunicación masivos (influencias externas), y el segundo grupo influenciado solo por el Boca en Boca (influencias internas). Bass denominó al primer grupo “innovadores” mientras que al segundo lo llamó “imitadores”.

La siguiente figura es un gráfico conceptual y analítico de la estructura subyacente en el modelo de Bass.

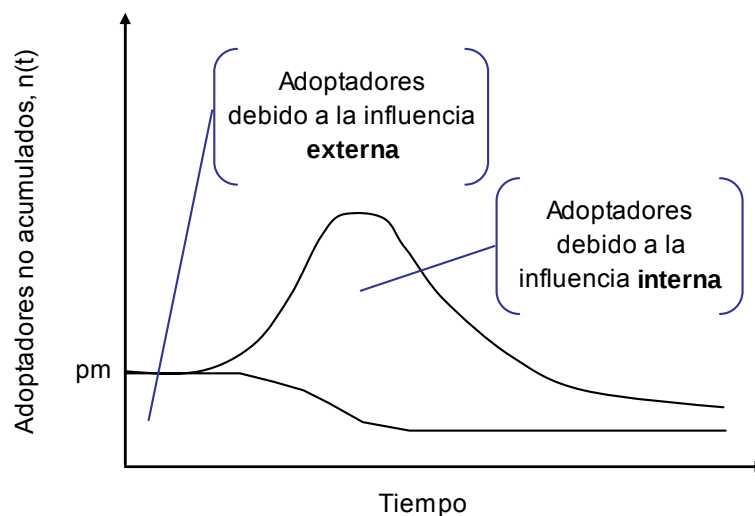


Figura N° 2: Tipos de Adopciones en el Modelo de Bass³

³ Trichy V. Krishnan; “The Bass Model”.

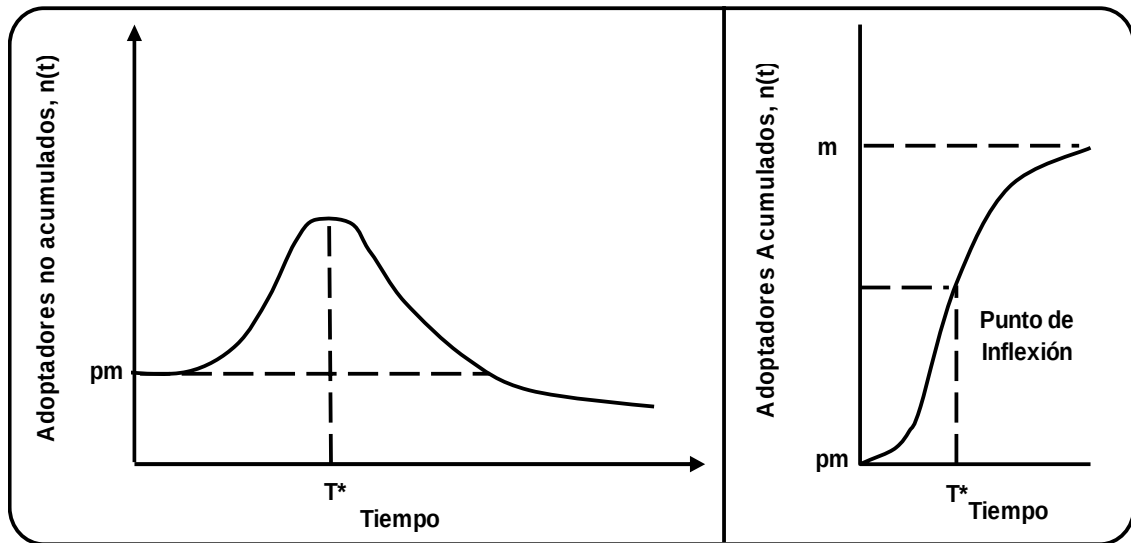


Figura N° 3: Estructura analítica del Modelo de Bass⁴

Como se observa en la *Figura N° 2*, el modelo de Bass asume conceptualmente que los “innovadores” están presentes en toda etapa del proceso de difusión.

La *Figura N° 3* muestra la estructura analítica subyacente en el modelo de Bass. Como se observa, la distribución no acumulada de adoptadores presenta un pico en el momento T^* , que representa el punto de inflexión de la curva acumulativa de adopción que posee forma de “S”.

Además la distribución del adoptador asume que un nivel pm (una constante) inicial de adoptadores compra el producto al inicio del proceso de difusión.

Una vez iniciado el proceso de adopción es simétrico con respecto al tiempo alrededor del pico T^* hasta $2T^*$. Es decir, la forma de la curva de adopción desde el tiempo T^* a $2T^*$ es la imagen reflejada de la curva de adopción desde el inicio del proceso de difusión hasta el tiempo T^* .

El modelo de Bass se desprende de la siguiente función:

$$f(t)/[1 - F(t)] = p + qF(t) \quad (1)$$

La función densidad de tiempo para la adopción esta dada por $f(t)$ y la fracción cumulativa de adoptadores al tiempo t esta dada por $F(t)$. Entonces, la premisa básica establece que parte de la influencia en la adopción depende de la imitación o del aprendizaje y parte de ella no. El parámetro q refleja esta

⁴ Trichy V. Krishnan; “The Bass Model”.

influencia y el parámetro p refleja una influencia que es independiente de la adopción previa. Si q es cero, $f(t)$ seguiría una distribución exponencial. Si m es el número potencial de adoptadores último, el número de adoptadores al tiempo t será:

$$mf(t) = n(t) \quad (2)$$

y el número acumulativo de adoptadores en el tiempo t será:

$$mF(t) = N(t) \quad (3)$$

Manipulando la formula inicial se obtiene:

$$n(t) = \frac{dN(t)}{dt} = p[m - N(t)] + \frac{q}{m} N(t)[m - N(t)] \quad (4)$$

El primer término $p[m - N(t)]$ en la ecuación (4) representa las adopciones debido a compradores que no son influenciados en el tiempo de su adopción por el número de personas que ya compró el producto. Bass se refiere al parámetro p como “Coeficiente de Innovación”. El segundo término en la

ecuación (4), $\frac{q}{m} N(t)[m - N(t)]$, representa las adopciones debido a compradores que fueron influenciados por compradores previos. Bass se refiere a q como “Coeficiente de Imitación”. Notar que en la ecuación (4) en $t = 0$, $n(0) = pm$.

La ecuación (4) es una ecuación diferencial de primer orden. Puede ser integrada para obtener la curva de distribución acumulada de adoptadores, $N(t)$.

Luego se pueden obtener las siguientes ecuaciones:

$$N(t) = m \left[\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}} \right] \quad (5)$$

$$n(t) = m \left[\frac{p(p+q)^2 e^{-(p+q)t}}{(p+qe^{-(p+q)t})^2} \right] \quad (6)$$

$$n(T^*) = \frac{1}{4q} (p+q)^2 \quad (7)$$

$$N(T^*) = m \left(\frac{1}{2} - \frac{p}{2q} \right) \quad (8)$$

$$T^* = -\frac{1}{(p+q)} \ln \left(\frac{p}{q} \right) \quad (9)$$

2.3 Análisis de Redes Sociales

El análisis de redes sociales consiste en analizar las relaciones existentes entre los actores del universo que definimos. En primera instancia nos remitimos a revisar una de las definiciones más comunes de sociedad que es justamente el sistema en estudio:

“La sociedad es el sistema o conjunto de relaciones que se establecen entre los individuos y grupos con la finalidad de constituir cierto tipo de colectividad, estructurada en campos definidos de actuación en los que se regulan los procesos de pertenencia, adaptación, participación, comportamiento, autoridad, burocracia, conflicto y otros.”

Es importante remarcar que sin relaciones entre las personas no existe sociedad alguna, ya que justamente la sociedad se conforma a partir de ellas.

El concepto empleado para definir sociedad establecerá de manera ineludible los límites en las propiedades que se estudiarán de tal objeto. En este trabajo se considera a la sociedad como un sistema inteligente que goza de las siguientes propiedades o componentes:

- Una Frontera
- Un Entorno
- Partes dentro de la sociedad
- Lazos de comunicación
- Una extensión limitada de espacio y tiempo
- Capacidad de aprender
- Objetivo compartido

Algunos ejemplos de sociedades son: estados, provincias, municipalidades, empresas industriales, partidos políticos, organizaciones religiosas y clubes. Una jauría de lobos o una manada de renos, también son sociedades, tal como definimos este término, porque están compuestos de sistemas inteligentes.

La Sociedad Como Sistema Complejo

Un sistema complejo es:

- “Un sistema compuesto por un número (usualmente elevado) de entidades interactivas, procesos, o agentes cuyo entendimiento requiere del desarrollo o uso de nuevas herramientas científicas, modelos no lineales, o simulaciones por computadora”.⁵
- “Un sistema que puede ser analizado a partir de varios componentes que poseen numerosas relaciones entre los mismos, de tal manera que el comportamiento de uno de los componentes depende del comportamiento del resto de ellos”.⁶
- “Un sistema que implica numerosos agentes interactuando cuyos comportamientos agregados deben ser comprendidos. Dicha agregación de comportamientos es no lineal, por lo cual no puede ser deducida a partir de la simple suma de los comportamientos individuales”.⁷

Un sistema complejo es cualquier sistema que presenta un número elevado de componentes interactuando (agentes, procesos, etc.) cuyas actividades agregadas es no lineal (no se deriva de la suma de la actividad individual de los componentes) y típicamente exhibe una auto-organización jerárquica bajo

⁵ Advances in Complex Systems Journal

⁶ Herbert Simon

⁷ Jerome Singer

presiones selectivas. Esta definición aplica a sistemas de un amplio rango de disciplinas científicas.

Casi todos los procesos interesantes en la naturaleza poseen relaciones cruzadas. Sin embargo, en muchos sistemas podemos distinguir un conjunto de bloques fundamentales que interactúan de manera no lineal para formar estructuras o funciones compuestas que no pueden ser explicados bajo los mismos paradigmas que el funcionamiento o estructura de los bloques individuales. Este proceso de emergencia de la necesidad de nuevos modos de descripción es conocido auto-organización jerárquica, y los sistemas que presentan esta característica son definidos como complejos.

Ejemplos de estos sistemas son las redes sociales que son el enfoque de este estudio, las colonias de hormigas, las redes neuronales en el cerebro, redes ecológicas, las economías y otros.

Redes Sociales

Una red social es un grafo en el cual los nodos representan individuos, instituciones, empresas, viviendas, u otros entes (a veces denominados actores) y los enlaces relaciones entre ellos. El contenido de esas relaciones debe ser definido por el investigador y puede ser cualquier relación, desde amistad a número de llamadas de teléfono o asistencia a las mismas películas.

Según la naturaleza de las relaciones, se pueden dividir entre:

- *Diádicas* (sólo indican ausencia o existencia de la relación) o *Valoradas* (en la que la cantidad de la relación pueda medirse en términos de orden o de peso como, por ejemplo, número de encuentros sexuales entre dos personas), o bien
- *Transitivas* (la relación en realidad es una unión entre actores que siempre es recíproca. Ejemplo: leemos el mismo blog habitualmente) o *Dirigidas* (que el individuo A tenga relación con el individuo B no implica que B tenga esa misma relación con B).

El análisis de redes sociales ha irrumpido en muchas ciencias sociales en los últimos veinte años como una nueva herramienta de análisis de realidad social.

Al centrarse en las relaciones de los individuos (o grupos de individuos) y no en las características de los mismos (raza, edad, ingresos, educación, etc.) ha sido capaz de abordar algunos temas con un éxito insospechado. La difusión

de información o el contagio de enfermedades son dos ejemplos de asuntos en los que la estructura de las relaciones puede llegar a ser más relevantes que las características de los individuos, o por lo menos, información clave para conocer los procesos. El uso de las ideas y herramientas de la rama de las matemáticas conocida como "teoría de grafos" ha ayudado a desarrollar una gran cantidad de herramientas y software de análisis.

2.4 Topología de Red

Como se vio anteriormente, las relaciones sociales pueden representarse como redes en las cuales los nodos pueden representar diferentes entes y los arcos las relaciones entre los mismos. De esta manera, de poder disponer de la información suficiente, se podría construir un mapa de la red y obtener conclusiones acerca del estado de la red y las acciones que se deberían tomar para alcanzar una difusión del producto que sea óptima.

Sin embargo, en la mayoría de los casos es muy costosa o complicada la compilación de dichos datos. Es por eso que el estudio de las topologías de red puede otorgar un conocimiento de base para entender las ventajas y desventajas de cada tipo de estructura de asociación de manera teórica.

Topologías Simples

El concepto de topología de red es muy utilizado en el campo de la informática. En este campo se estudia las diferentes disposiciones físicas en las que se conectan los nodos de una red de ordenadores o servidores. Sin embargo, desde el nacimiento del análisis de redes sociales el alcance del concepto de topología de red también se extendió para referirse al estudio de la estructura de una red de relaciones en una sociedad.

La diferencia principal que se encuentra al analizar una red de computadores o una red social es que la última supera ampliamente en complejidad a la primera, no sólo por el mayor número de componentes que puede conformar la red sino también porque las relaciones sociales no están enmarcadas dentro de un estándar de construcción por lo cual simplemente son generadas por la suma de comportamientos individuales influenciados por el contexto social.

A pesar de esta diferencia puede resultar interesante analizar brevemente las topologías básicas utilizadas principalmente en el ámbito informático. Entre las topologías simples se encuentran la topología en estrella, en árbol, en malla, anillo o las topologías híbridas (Ver Anexo: "Topologías Simples").

Topologías Complejas

El término redes complejas se refiere a redes que poseen ciertas características no triviales que no ocurren en las redes simples. Estas características no triviales incluyen:

- Una cola extensa en el grado de distribución,
- Un alto coeficiente de clustering o
- Una evidencia de estructuras jerárquicas entre otras.

Si bien las topologías simples son de utilidad para identificar algunas ventajas y desventajas que presentan los patrones de redes, la realidad es que las estructuras sociales no presentan estas formas tan simples. Es por esto que al analizar sistemas como la sociedad se hace uso de topologías complejas. Entre las topologías complejas se puede hallar a las redes mundo pequeño y las redes libre de escala, que han demostrado poseer usos en diversos campos.

Redes Mundo Pequeño

El concepto redes “mundo pequeño” ha recibido mucha atención últimamente, a pesar de que sus orígenes datan de trabajos realizados hace más de 40 años en sistemas complejos. Erdős y Renyi (ER) fueron pioneros en esta área aplicando métodos probabilísticos para resolver problemas de teoría de grafos, en los que estaban implicados un gran número de nodos. Bajo este supuesto modelaron extensos gráficos que utilizaban algoritmos en los que N nodos eran conectados aleatoriamente de acuerdo a una probabilidad p , y encontraron que cuando los nodos eran conectados de esta manera seguían una distribución de Poisson.

Un análisis más exhaustivo sobre los gráficos aleatorios puede ser encontrado en los trabajos de Bollobás. Continuando con los descubrimientos de ER, sus modelos de formación de redes aleatorias fueron utilizados ampliamente en diferentes disciplinas que examinaban las redes. La ausencia de datos topológicos para las redes complejas hizo que las redes aleatorias fueran ampliamente utilizadas en simulaciones de redes. A medida que el poder de cómputo aumentó y datos sobre las redes reales comenzaron a estar disponibles, emergieron nuevos descubrimientos. De los análisis de redes complejas aparecieron tres características de la red:

- Caminos de distancia promedio corta
- Alto nivel de clustering
- Distribuciones de tipo exponencial o ley de potencias

Caminos de distancia promedio corta indican que la distancia entre dos nodos cualesquiera es corta. El clustering ocurre cuando nodos localizados a corta distancia se conectan mutuamente formando grupos. Por último, la distribución de la frecuencia de densidad de nodos, denominado grados, habitualmente sigue la ley potencial.

Watts y Strogatz (WS) encontraron que las redes del mundo real no eran totalmente aleatorias, sino que presentaban un nivel de clustering significativo a nivel local. De acuerdo a WS, las redes “mundo pequeño” se caracterizan por su longitud de camino $L(p)$ y el grado de conectividad local en la red, medido a través del coeficiente de clustering $C(p)$.

$L(p)$ es el menor número de enlaces que se requieren para conectar un nodo con otro, promediado en la red entera. $C(p)$ es la fracción de nodos adyacentes conectados unos con otros. Se puede mirar a $L(p)$ como una propiedad global de la red y a $C(p)$ como una propiedad local de la misma.

WS demostraron que la red “mundo pequeño” esta entre una red “Regular Lattice” y una red aleatoria. Para demostrar esto, comenzaron con una red “Regular Lattice” con n nodos y k enlaces y la recablearon de tal manera de que se asemejara a una red aleatoria. Los atajos a través del gráfico entre diferentes clusters introdujeron un nivel de eficiencia no predicho por el modelo ER. La distribución no era Poisson, como en el modelo ER, sino sesgada y cayendo exponencialmente para grandes cantidades de nodos.

El trabajo de WS difundió la aparición de numerosos trabajos que se enfocaban en entender los atributos de las redes complejas, mientras que numerosos descubrimientos aparecieron. Dos estudios paralelos realizados por Albert, Jeong y Barabasi de Notredame y Adamic y Huberman de Xerox Parc encontraron que cuando uno observa la red de Internet como si fuese un grafo (las páginas de Internet son nodos y los hipervínculos enlaces) esta se ajusta no a una distribución de Poisson sino a una distribución de ley potencial.

Enfoques Metodológicos

En una distribución potencial una gran cantidad de nodos poseen pocos enlaces, y una pequeña pero significativa minoría posee una gran cantidad de enlaces. Se debería destacar que esto es ligeramente diferente de los modelos

ER y WS. La probabilidad de encontrar nodos altamente conectados en los modelos ER y WS decrece exponencialmente de manera tal que los nodos con alta conectividad son prácticamente inexistentes.

La razón de acuerdo a Barabasi y Albert era que su modelo agregaba otra perspectiva a las redes complejas, incorporando el crecimiento de la red, el número de nodos ya no permanecía constante como en los modelos de ER y WS.

Los modelos de Barabasi-Albert (BA) agregaban crecimiento a través del tiempo y la idea de que nuevos nodos se anexaban preferencialmente a los nodos altamente conectados ya existentes en la red. BA formalizaron esta idea en “Emergence of Scaling in Random Networks”. Determinaron que en redes complejas como Internet la probabilidad $P(k)$ de que un nodo en la red interactúe con otros k nodos decae como una ley potencial.

Al estudiar redes reales libre de escala, el exponente de ley potencial se encontró dentro de un rango de 2.1 a 4. El modelo de BA esta basado en tres mecanismos que determinan la evolución de la estructura del grafo a través del tiempo para producir relaciones de ley potencial:

- **Crecimiento Incremental** – Proviene de la observación que la mayoría de las redes se desarrollan a través del tiempo mediante la adición de nuevos nodos y nuevos enlaces a la red existente.
- **Conexión Preferencial** – Expresa el fenómeno frecuente que muestra la existencia de una mayor probabilidad de que un nuevo nodo se conecte o reconecte con un nodo de alto grado nodal que con uno que posee pocos enlaces.
- **Reconexión** – La reconexión adiciona flexibilidad en la formación de redes mediante la remoción de enlaces conectados a determinados nodos y reemplazándolos con nuevos enlaces.

AB ven las redes “mundo pequeño” y las redes libre de escala como las explicaciones de dos fenómenos diferentes que ocurren en redes complejas. El modelo de “mundo pequeño” de WS explica el clustering y la red de escala libre explica la distribución de red potencial. Sin embargo ha habido opiniones distintos en cuanto a la manera en que las redes mundo pequeño y las libre de escala deberían ser clasificadas. Amaral argumenta que las redes libre de escala son subclases de redes de mundo pequeño. Estableciendo la existencia de tres tipo de redes mundo pequeño:

- Redes libre de escala, caracterizadas por una distribución de la conectividad de nodos que decae como una ley potencial.
- Redes de ancha escala, caracterizadas por una distribución de la conectividad que es de tipo ley potencial seguido de una caída abrupta.
- Redes de escala única, caracterizadas por una distribución de la conectividad con una cola descendente.

Aún no existe una definición exacta sobre la divergencia entre las redes mundo pequeño y las redes libre de escala. Pero si esta claro que ambos están relacionados y que en general las redes libre de escala muestran el clustering y una corta distancia promedio como las redes mundo pequeño, pero no todas las redes mundo pequeño muestran la distribución de ley potencial de las redes de escala libre.

Redes Libre De Escala

En 1999, Lászlo Barabási y sus colaboradores de la Universidad de Notre Dame en Indiana (EEUU) hicieron un mapa de la Web, observando la manera en que las diferentes páginas web se enlazaban. Para su sorpresa, la web no presentaba una distribución del grado de conectividad usual. En lugar de esto, algunos pocos nodos, a los que llamaron "hubs", estaban mucho más conectados que el resto. Al mismo tiempo, una observación similar sobre Internet fue obtenida por los hermanos Faloutsos (1999).

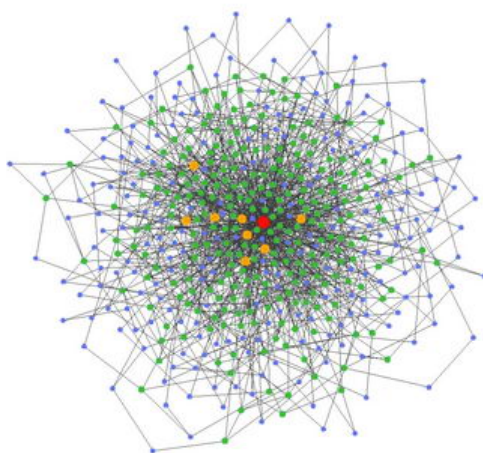


Figura N° 4: Esquema de una Red Libre de Escala

En general, encontraron que la probabilidad $p(x)$ de que un nodo de la red estuviera conectado con x nodos era proporcional a $k^{-\gamma}$, es decir, sigue una ley de potencias. El exponente γ no es universal, sino que depende del tipo específico de red. Para la mayor parte de los sistemas se encontró que dicho parámetro se encontraba en el rango $2 < \gamma \leq 3$ aunque hay excepciones.

Posteriormente, se ha verificado que una gran variedad de sistemas complejos comparten esta propiedad importante, algunos nodos poseen una enorme cantidad de conexiones, mientras que el resto de los nodos, que componen la mayoría, sólo poseen algunas conexiones. Los nodos populares, denominados hubs, pueden tener cientos, miles o incluso millones de enlaces. En este sentido la red parece no tener escala, es por eso que se denominó red libre de escala.

Además del ejemplo de las redes de páginas Web, existen muchas otras redes libres de escala. Entre estas encontramos:

- La red de contactos sexuales entre personas. Hay unas pocas personas que tienen muchas parejas a lo largo de su vida, mientras que la mayoría de las personas tiene unas pocas parejas.
- Las redes del crimen organizado, en los cuales unos cuantos "peces gordos" ordenan la actuación de muchos "peces chicos".
- La red de distribución eléctrica, en que existen estaciones enormes que abastecen a zonas enormes, y al mismo tiempo una miríada de transformadores pequeños.
- Las redes de comercio internacional, dado que los países desarrollados, que son la minoría, concentran la mayor cantidad de intercambio de bienes, mientras que en los países no desarrollados, que son la mayoría, el intercambio comercial es relativamente menor. Esto se aplica también a las redes de comercio entre empresas dentro de cada país.
- Las redes de citas bibliográficas incluyen unos pocos libros o escritos muy citados, mientras que la mayoría de los libros reciben pocas o incluso ninguna citación.
- Las redes de neuronas en los organismos dotados de sistema nervioso, lo que significa que permanentemente usamos mucho una fracción de las neuronas, mientras que la mayoría de las neuronas las ocupamos muy poco.
- Las redes de interacción de proteínas en el metabolismo celular, con unas pocas proteínas que aparecen en la mayoría de las reacciones, mientras que la mayoría de las proteínas aparecen sólo en situaciones muy específicas.

Las redes libres de escala poseen ciertas características relevantes. Son robustas frente a accidentes aleatorios, pero vulnerables frente a ataques coordinados. La comprensión de estas características podría llevar a nuevas

aplicaciones en muchas áreas. Por ejemplo, los científicos de la computación podrían identificar estrategias más efectivas para prevenir que los virus informáticos se difundan por redes como Internet.

2.5 Simulación Basada en Agentes

El mundo en el que vivimos es una estructura estratificada con muchos niveles. Entidades de cualquiera de los niveles forman parte de las entidades de niveles superiores. Partículas elementales conforman átomos. Átomos forman moléculas, simples y hasta las más complejas, y las moléculas forman todos los objetos del universo. Entre otros, moléculas orgánicas complejas forman células vivas. Las células forman tejidos y organismos. Organismos vivos forman ecosistemas. Neuronas, un tipo especial de células, forman el cerebro humano capaz de ser consciente. Los seres humanos forman grupos, sociedades y civilizaciones.

A través de la identificación de entidades clave y sus propiedades y modos de interacción, en cada uno de estos niveles un contexto de una disciplina científica particular fue establecido explicando el fenómeno del nivel correspondiente. Sin embargo, las entidades en cada uno de estos niveles no son simplemente clusters o conjuntos de los niveles inferiores, no son simplemente la suma de las partes, sino un “todo” más complejo que posee nuevas interrelaciones estructuras interactivas y nuevas características y regularidades. Reconociendo que las disciplinas científicas tradicionales son poco exitosas en describir y explicar formas en que partes relativamente simples se organizan o se auto organizan en un todo más complejo la relativamente nueva disciplina de la ciencia de la complejidad o teoría de la complejidad tiene el objetivo de abordar estos temas que fueron insuficientemente explorados.

De manera general se puede afirmar que la complejidad estudia sistemas complejos adaptativos. Es decir, sistemas dinámicos que consisten de muchas partes, generalmente no lineales, que interactúan y se modifican constantemente adaptándose al ambiente. La tarea principal de la ciencia de la complejidad resulta el explicar como patrones agregados, relativamente estables y macroscópicos son inducidos por interacciones de multitudes en entidades de bajo nivel.

¿Que es la simulación basada en agentes?

Como estas interacciones no lineales, son en su mayoría demasiado complejas para ser capturadas por expresiones analíticas, es usual utilizar simulaciones por computadora. La idea básica de estas simulaciones es especificar las reglas de comportamiento de entidades individuales, al igual que las reglas de interacción, para simular una multitud de entidades individuales usando una computadora y luego explorar las consecuencias de los comportamientos individuales a nivel macro. Las entidades simuladas son usualmente llamadas agentes y las simulaciones de su comportamiento e interacciones son conocidas como simulaciones basadas en agentes. Las propiedades individuales de los agentes que describen su comportamiento e interacciones son conocidas como propiedades elementales, y las propiedades que emergen en niveles superiores o colectivos se conocen como propiedades emergentes.

Propiedades básicas de las simulaciones basadas en agentes

En muchas ocasiones los modelos basados en agentes arrojan resultados que no son ni obvios ni esperables, aún en casos en los que las suposiciones sobre las propiedades individuales del agente son muy simples.

Esta capacidad de generar propiedades emergentes intrigantes y complejas no resultan tanto de las reglas internas de comportamiento de los agentes individuales sino de la complejidad de la red de interacciones entre los agentes. Precisamente esta multitud de agentes, al igual que la multitud y complejidad de las interacciones, son las razones principales por las cuales en muchos casos deducciones formales matemáticas de los resultados de modelos basados en agente no es posible.

La comunidad científica regularmente decide prestar atención a problemas definidos por conceptos y técnicas instrumentales que se encuentren a mano y con un nivel avanzado de desarrollo. Es por esto que sólo luego de los recientes avances de la tecnología de procesamiento de las computadoras que permitió los experimentos de simulación masivas, los temas de complejidad emergente se acercaron al foco de la investigación científica.

Además del modelado de efectos “botton-up” (por ejemplo los efectos que se originan a nivel individual influenciando el nivel colectivo), otros modelos más complejos basados en agentes son capaces de modelar efectos “top-down”, que surgen en niveles colectivos y que afectan el nivel individual de los agentes.

Ventajas de la simulación por agentes

- Los lenguajes de programación son más expresivos y menos abstractos que la mayoría de las técnicas matemáticas.
- Los programas de computación lidian más fácilmente con procesos en paralelo o con un orden no definido de acciones que sistemas de ecuaciones matemáticas.
- Los programas diseñados de acuerdo a los principios de la ingeniería del software son modulares, lo cual facilita su modificación; los sistemas matemáticos frecuentemente carecen de este sistema modular.
- Es más sencillo construir sistemas de simulación que incluyan agentes heterogéneos, mientras que esto resulta más difícil utilizando ecuaciones matemáticas.
- Posibilidad de modelar condiciones más “fluidas” o “turbulentas” cuando los agentes modelados y sus identidades no están dadas, sino que son susceptibles a cambios que pueden incluir la muerte o nacimiento de agentes individuales al igual que la adaptación de su comportamiento.
- Posibilidad de modelar agentes con niveles de racionalidad limitada, que toman decisiones y actúan en condiciones de información o conocimiento incompleto.
- Posibilidad de modelar procesos fuera del equilibrio.

El verdadero alcance de la simulación por agentes

Los modelos de simulación basados en agentes tienen un fin explicativo más que predictivo. Nos proveen de los medios para realizar simulaciones por computadora a través de experimentos que tienen el objetivo de mejorar nuestra intuición acerca del fenómeno modelado. Esta característica se torna más importante aún en casos en los cuales la experimentación en el mundo real es o bien muy costosa o simplemente imposible. Los resultados de los experimentos son contrastados con la teoría que fue utilizada al diseñar el experimento. Cuando un experimento pensado genera disonancias debemos cuestionarnos tanto la integridad de las actuales teorías, como la validez de las intuiciones que guiaron nuestros pensamientos a través del experimento.

Limitaciones y puntos débiles

El poder y versatilidad que nos ofrece la simulación tiene su contrapartida. Esta es la “opacidad explicativa”. Esto significa que el comportamiento de la simulación no es entendible por simple inspección. Por el contrario, es

necesario cierto grado de esfuerzo para explicar los resultados. Las simulaciones deben ser observadas y exploradas sistemáticamente antes de ser comprendidas, y este entendimiento puede realimentar teorías existentes.

Por supuesto, esto es más sencillo decirlo que hacerlo. Debido a un enorme número de parámetros del modelo y una cantidad masiva de datos generados por la simulación para la configuración de los parámetros, los resultados son frágiles. Esto significa que no siempre es sencillo descubrir si los resultados del modelo son un simple artefacto de una configuración específica de los parámetros o si son resultados realmente significativos.

La teoría subyacente en el diseño del modelo puede proveer una guía acerca de los rangos críticos de los parámetros a testear lo cual permite reducir la inspección.

A su vez, el método de diseño “paso a paso” provee otra manera de reducir la “opacidad explicativa”. A partir de comenzar el análisis de cero de vez en cuando y agregar características teóricas de manera incremental, este diseño de la estrategia hace mas sencillo el manejo de la complejidad.

Para facilitar la comunicación del modelo y la reproducibilidad de los resultados es recomendable el uso de paquetes de software relativamente estandarizados. Por el otro lado la utilización de paquetes de software son inevitablemente limitados en las funcionalidades que ofrecen, es por eso que también deberá considerarse la utilización de algún lenguaje de programación apropiado.

3. DEFINICIÓN DEL PROBLEMA

El problema que se abordará es un paso crucial e inevitable al que se enfrenta la mayoría de las empresas. Luego de superar diversos pasos como la concepción del producto, estudio del mercado, identificación de las propiedades de diseño relevantes, desarrollo de ingeniería, producción y definición de la plaza, la empresa se enfrenta a un paso decisivo: la difusión de su producto.

Más precisamente, el problema se enmarca dentro de la difusión de productos a través del efecto boca en boca. Se observa que si bien este efecto de difusión de productos de persona en persona es de gran relevancia, la falta de una metodología práctica e integradora provoca que en general las empresas no hagan foco en estudiar la manera de potenciar dicho efecto, ni establezcan estrategias de difusión que consideren el impacto de la estructura de red social en el proceso de difusión.

El objetivo que busca este trabajo y sobre el cual se define el criterio de éxito del mismo, es acercar un nuevo enfoque basado en la simulación y las redes sociales que permita hacer un aporte en la solución del problema citado anteriormente. De esta manera, se pretende hacer más eficiente el proceso de difusión de innovaciones y desarrollar algunos fundamentos para el desarrollo de planes estratégicos de difusión de productos y promociones.

Limites del trabajo

- Sobre el tipo de difusión al que se refiere el trabajo:

Dentro de la diversidad de temas que aborda la difusión de innovaciones, este trabajo hace foco en la difusión de productos en particular.

Cualquier otro tema que no sea el de la difusión de productos no será analizado en profundidad en este trabajo debido a que queda fuera del alcance del mismo. Sin embargo, esto no implica que los resultados que se obtengan o los conceptos desarrollados no puedan ser empleados en la difusión de innovaciones en general.

- Sobre el medio utilizado para la difusión del producto:

El ejemplo que se analiza en el caso práctico será particularmente útil para la difusión de productos que utilicen o puedan utilizar, entre otros medios de difusión, la publicidad dirigida (tele marketing, correo, etc.). Sin embargo, la metodología planteada puede ser aplicada con diferentes objetivos y de manera más general.

- Sobre el aspecto teórico-práctico del trabajo:

El trabajo de investigación propone la incorporación de un nuevo enfoque que complemente la teoría actualmente utilizada para el diseño de procesos de difusión de innovaciones.

Como ejemplo de como llevar a la práctica alguno de los conceptos desarrollados, se trabajará sobre un caso práctico. En este se abordará el problema de difusión de uno de los productos que posee una compañía de telecomunicaciones.

- Sobre el aspecto económico:

Este trabajo es fundamentalmente de investigación y por ende no pretende hacer un análisis del aspecto económico del proceso de difusión de innovaciones propuesto. Sin embargo, como se verá más adelante, la solución propuesta en el caso práctico es de simple aplicación y bajo costo para la empresa.

4. METODOLOGÍA PROPUESTA

La metodología propuesta combina diferentes herramientas de análisis de redes sociales, simulación basada en agentes, teorías de marketing y teorías de comunicación. De esta manera, se propone una serie de pasos generales que permiten realizar un análisis integral de la red social y diseñar una estrategia de difusión que potencie los efectos del boca en boca.

La estrategia se compone de diferentes etapas. En primer lugar, se define el problema que posee la empresa o mejora que se pretende alcanzar con el análisis. En segundo lugar, una vez definido el problema y habiendo considerado de qué manera la metodología puede aportar en la solución del problema, se realiza la representación de la topología de red. En tercer lugar, se realiza un análisis estático de la red social, midiendo los indicadores que se consideren necesarios de acuerdo al problema definido. En cuarto lugar, se define una teoría del comportamiento que servirá de base para modelar el comportamiento de los agentes al realizar la simulación de la difusión. En quinto lugar, se diseñan los experimentos de simulación adecuados para responder al problema planteado. Finalmente, se analizan los resultados y se establecen las conclusiones derivadas del estudio.

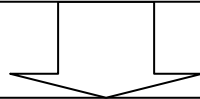
Esquema de la Metodología Empleada

En esta sección se describirá la metodología general para diseñar un modelo basado en agentes que permita enfocar el problema de difusión de productos considerando el efecto Boca en Boca y las topologías de red social.

I. Identificación del problema

Inicialmente, dentro del marco de la difusión de productos, se debe definir el problema o la mejora que se desea incorporar.

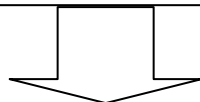
- ¿Cuál es el problema o mejora que se desea incorporar?
- ¿El problema puede ser cubierto con la metodología planteada?
- ¿Se cuentan con los recursos necesarios? (datos, tiempo, etc.)



II. Representación de la Topología de Red Social

Es necesario obtener una representación lo más fiel posible de la red social en estudio. De esta manera, una vez comprendida la estructura de la red social e identificados los aspectos relevantes de la misma, se podrá utilizar como base para la topología sobre la cual corra el modelo de simulación.

- ¿Qué representan los nodos? (Individuos, empresas, instituciones, etc.)
- ¿Qué relación representarán los enlaces entre nodos?
- ¿Qué datos están disponibles para modelar la topología de red en estudio?



III. Análisis de la Red Social

Se realiza un análisis de la red social de interés.

- ¿Qué índices son de interés para su medición?
- ¿Qué estructura de red presenta?

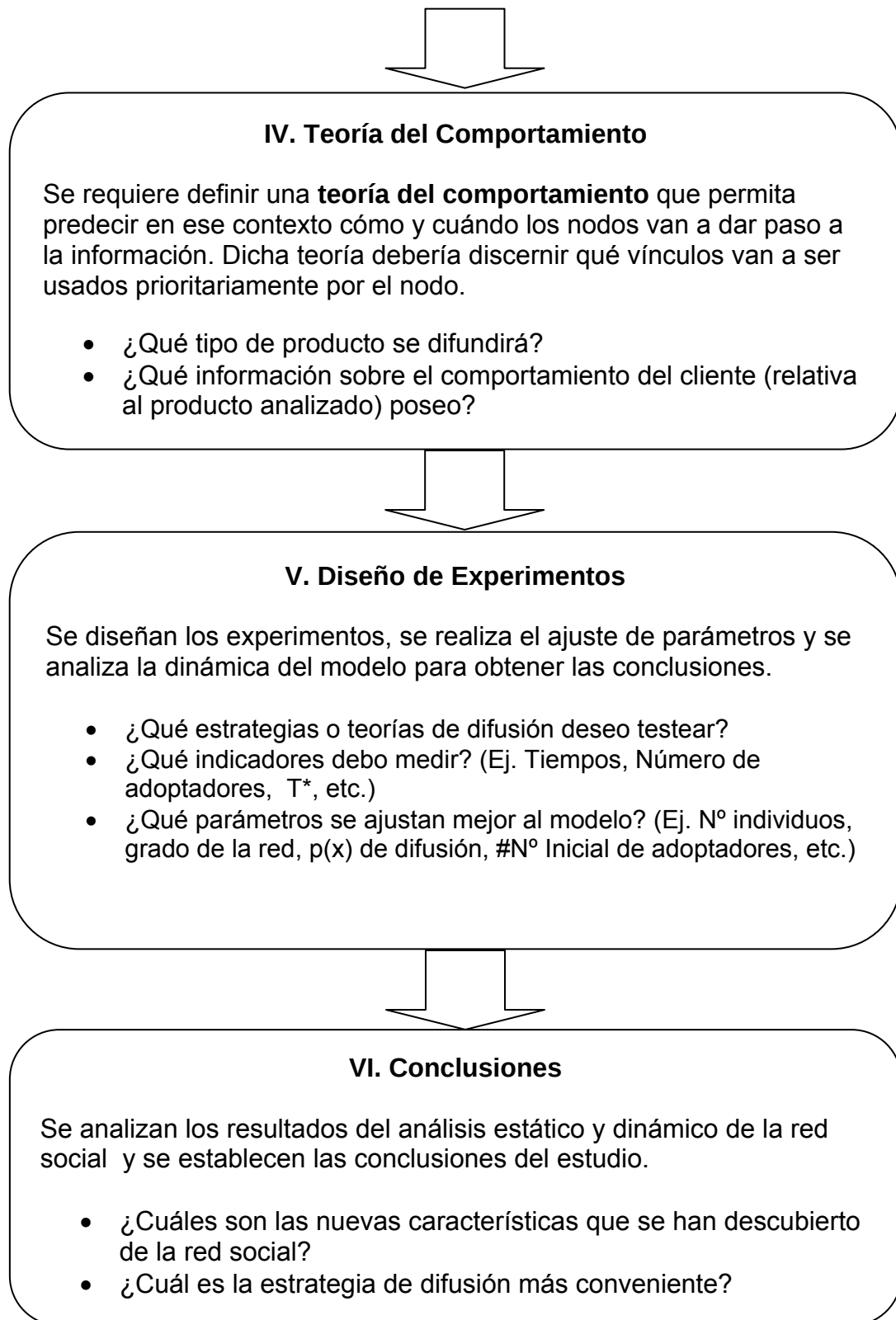


Figura N° 5: Esquema de la Metodología Propuesta

4.1 Primer Paso: Identificación del Problema

La primera tarea consiste en la definición del problema o mejora que se desea obtener a través del estudio de la red social. Como se mencionó anteriormente el estudio de la difusión de innovaciones posee una amplia diversidad de usos. Sin embargo, esta metodología se focaliza en la difusión de productos.

La metodología será de utilidad para estudiar la estructura de la red social del mercado objetivo al cual la compañía desee apuntar, y para diseñar una estrategia que permita potenciar el efecto boca en boca.

A su vez, se deberá considerar el tipo de empresa y hacer un análisis minucioso del tipo de producto que comercializa para establecer las ventajas potenciales del uso de la metodología.

Una vez definido el problema se aconseja especificar los recursos a los que se tendrá acceso tales como bases de datos, información comercial o de marketing, personas dedicadas al proyecto y tiempo. En función de esto, se procede a definir un cronograma de trabajo estableciendo las diferentes fases del proyecto y su duración estimada.

4.2 Segundo Paso: Representación de la Topología de Red Social

Con el objetivo de utilizar la teoría de redes sociales de manera tal de complementar la teoría actual de difusión de innovaciones se deberá realizar un análisis que permita modelar la red social en la cual se desea difundir el producto.

En esta instancia, es importante responder las siguientes preguntas:

- ¿Qué representarán los nodos?

Los nodos pueden representar diferentes entes dependiendo del problema planteado y los datos disponibles. Es común que representen personas (potenciales adoptadores, adoptadores, perfiles u otros), instituciones, empresas u hogares.

- ¿Qué tipo de relación representarán los enlaces?

A nivel general, los enlaces representan las relaciones entre nodos. Sin embargo, este término es ambiguo y debe ser definido con más detalle considerando el motivo del estudio.

Al investigar el proceso de difusión de determinado producto los enlaces entre nodos deberían representar idealmente las relaciones que potencialmente impliquen un contacto que arrastre el efecto boca en boca. En otras palabras, se debe definir la existencia de un enlace cuando se considera que entre dos nodos existe una relación tal que el flujo de información entre estos puede acarrear la difusión del producto.

- ¿Qué datos se poseen para la representación de la topología?

La información disponible definirá y limitará la elección de metodologías para la representación de la topología de red. Se deberá pensar en la manera más adecuada para la definición de la topología considerando los datos disponibles.

Dependiendo de los datos disponibles existen dos alternativas posibles para la representación de la red social que se complementan. Una de ellas es basarse en datos empíricos para poder aproximarnos a la red social más representativa de la sociedad. La otra alternativa consiste en basarse en modelos teóricos ya desarrollados estimando que la sociedad debería aproximarse al modelo planteado.

Método Basado En Datos Empíricos

El objetivo de este método es analizar cual es la estructura de la red social de la sociedad en estudio basándonos en datos de la misma. Al utilizar este método se requieren una gran cantidad de datos y detalles de la población.

En esta instancia deberíamos preguntarnos qué datos son de utilidad para la definición de la topología de la sociedad en estudio. El objetivo es poder conocer todas las relaciones existentes entre las personas bajo estudio, por lo cual deberíamos conocer en primer lugar los medios a través de los cuales se comunican.

Las relaciones interpersonales suceden a través de diferentes medios como el contacto cara a cara, teléfono fijo, teléfono celular, correo electrónico,

correspondencia, mensajería por Internet (MSN, Yahoo Messenger, etc.), telefonía IP, sistemas de radio personal, etc.

El gran inconveniente de este método es que, habitualmente, acceder a la información de millones de personas y de todos los medios citados no es una tarea sencilla ni económicamente viable.

Método Basado En Datos Empíricos De Telefonía Fija

Las relaciones interpersonales se manifiestan a través de diferentes canales de comunicación. Sin embargo se observa una correlación entre las llamadas realizadas a través de telefonía fija y los contactos cara a cara entre las personas.

Un estudio realizado en Francia en 1996 comparó la proporción de llamadas realizadas a la familia con las realizadas a amigos y analizó la variación de los contactos telefónicos y cara a cara de una persona en diferentes periodos de su vida.

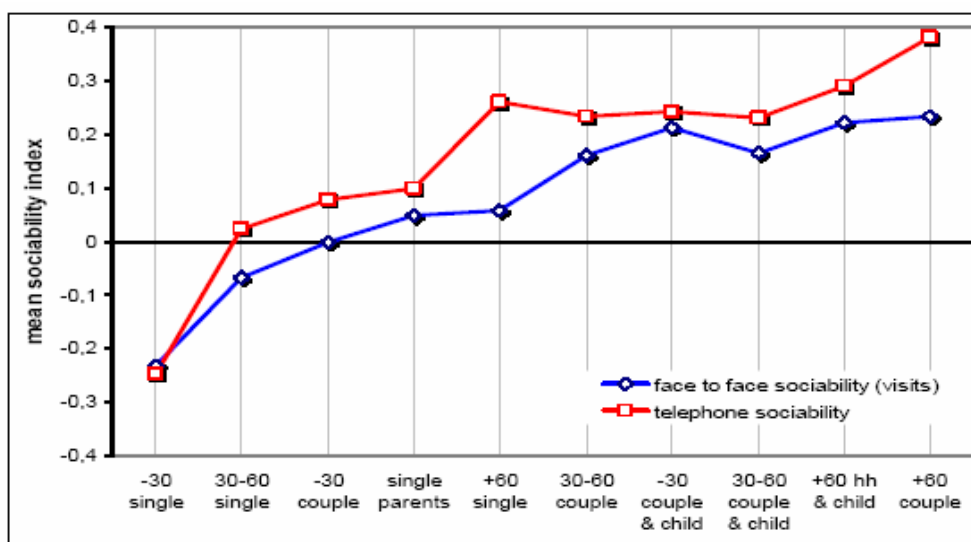


Figura Nº 6: Orientación de la sociabilidad cara a cara y telefónica a través del ciclo de vida⁸

Un índice de sociabilidad fue creado para reflejar la orientación a la familia de la red. Este índice fue calculado mediante la diferencia entre el número de miembros familiares llamados o vistos y el número de amigos llamados o

⁸ Zbigniew Smoreda, Frank Thomas. 2001. Social Networks and Residential ICT Adoption and Use. France Telecom R&D. Paper.

vistos, ponderado por el tamaño de la red familiar y amigos llamados o vistos. Como se puede observar en el gráfico el índice de sociabilidad de contactos telefónicos sigue de cerca el índice de contactos cara a cara. Se observa también un cambio en el comportamiento comunicacional cuando la persona se pone en pareja, la pareja tiene un hijo o entra en el período de retiro laboral.

Este resultado muestra claramente que los patrones de comunicación acompañan la orientación social a lo largo de diferentes etapas en la vida. A su vez, esta investigación indica que las comunicaciones electrónicas están correlacionadas con los contactos cara a cara.

Investigaciones sobre telefonía fija y móvil muestran que en las relaciones familiares o entre amigos la regla es “Mientras más de te veo, más te llamo”. El estudio demuestra que existe una fuerte correlación entre el número de personas llamadas y el número de personas contactadas cara a cara en el mismo período de tiempo.

Es por esto que al analizar las redes de llamadas a través de telefonía fija es posible inferir el grado de sociabilidad general de la persona. Esto significa que ya no será necesario el análisis de todos los medios de comunicación utilizados por una persona para inferir su grado de sociabilidad lo cual implicaría mucho tiempo y altos costos.

Método Basado En Topologías Teóricas

Este método consiste en trabajar directamente con estructuras de red teóricas ya conocidas. A través del tiempo, se ha descubierto que ciertas sociedades se agrupan mediante estructuras de redes complejas diversas como, las redes de escala libre, las redes aleatorias, las redes “mundo-pequeño” y otras.

Este método se basa en estudios realizados sobre el tipo de topología que presenta la sociedad en general y supone que la sociedad analizada también se adaptaría a determinado tipo de topología.

De esta manera, no se requieren datos sobre la población, pero se corre el riesgo de que la población no se corresponda con la topología de red que se le ha asignado.

Método Híbrido

Este método busca un equilibrio entre los dos métodos comentados anteriormente. Por un lado, se basa en modelos teóricos ya desarrollados, pero busca validar de alguna manera que la sociedad en estudio realmente se ajuste a dicho modelo. De esta manera, no se requiere una cantidad de datos tal que

imposibilite la ejecución real del ajuste y a su vez se valida la utilización de determinado modelo teórico de red social.

4.3 Tercer Paso: Análisis Estático de la Red Social

Una vez representada la red social, se realiza el análisis estático de la misma. Los índices que se miden dependerán del objetivo planteado. Sin embargo, es conveniente cubrir el análisis de los índices básicos de la red como ser, el número de nodos, el grado nodal promedio y la identificación de los nodos mejor conectados.

Para realizar el análisis de la estructura social se puede utilizar algún software especializado como el *Pajek*, *Ucinet* u otros existentes en el mercado. Entre otras cosas, estos programas permiten calcular los índices de manera automática y visualizar la red social de diferentes maneras.

Indicadores Característicos De La Red Social

Existen diferentes indicadores que permiten caracterizar a las redes sociales y comprender como influyen las diferentes características estructurales en la dinámica de la red. A continuación, se mencionan los términos e indicadores más relevantes.

Términos Generales

- Nodos de la Red: También llamados puntos o vértices, representan a los actores de la red.
- Líneas o enlaces: Representan la relación o relaciones que los unen.
- Red: Se define como un conjunto de nodos unidos por líneas.
- Influencia: Probabilidad asociada a un nodo de transmitir o impedir la transmisión de nuevas ideas o pautas de comportamiento en la red.
- Componentes: Un componente se definiría como un subconjunto de nodos en el cual cada nodo puede alcanzar cualquier otro a través de algún camino sin importar cuán largo sea. En una red conectada sólo habría un componente.
- Adyacencia: Decimos que dos nodos son adyacentes si existe una línea que les conecte.

Indicadores Generales

- Grado Nodal: Es el número de nodos con el que se relaciona un nodo.
- Grado Nodal de una red: Describe la densidad de la red. Es la media de los grados nodales de cada uno de los nodos (la sumatoria de todos los grados nodales dividido por el número de nodos).
- Camino: Un camino es una secuencia de nodos en el que cada nodo es adyacente al siguiente.
- Longitud (De un Camino): Es el número de enlaces que recorre para ir de un nodo a otro.
- Camino Geodésico: Es el camino de longitud menor entre dos nodos.
- Distancia (Entre dos nodos): Es la longitud del camino geodésico que los une.
- Diámetro (De una Red): Es la distancia máxima existente entre dos nodos.

Indicadores De Centralidad

- Punto de Corte: Es un nodo que al ser quitado genera la división de la red en dos componentes.
- Puente: Es un enlace que al ser quitado genera la división de la red en dos componentes.
- Puente Local de Grado n: Si al quitar el lazo no quedara separada la red en componentes pero los dos nodos que se unían quedarán separados por n nodos de distancia, el enlace sería un puente local de grado n.
- Cercanía ("Closeness"): Es la suma de las distancias que separan a un nodo del resto de nodos en la red; aproxima su "peso", su capacidad para llegar en pocos pasos a cualquiera.
- Intermediación: Es una medida del número de veces que un nodo aparece en el camino más corto entre otros dos nodos. Es la suma de los cocientes entre el número de caminos geodésicos que unen dos nodos y el número de ellos que pasan por el nodo en cuestión. La intermediación nos da una aproximación al peso como conector (como "hub") del nodo, su importancia cara a que la red se mantenga unida.
- Clustering: Es la tendencia a que dos nodos conectados a través de un tercero se conecten directamente entre sí. (Es la tendencia que tienen dos conocidos comunes a un tercero a conocerse entre sí).
- Conectores ("hubs"): Son nodos de fácil acceso desde distintas subredes (clusters).

- Densidad de la Red: Numero de lazos en el sistema dividido por el numero máximo de lazos posibles.

4.4 Cuarto Paso: Teoría del Comportamiento

En este apartado, se formulará una teoría del comportamiento que permita predecir en ese contexto cómo y cuando los nodos van a dar paso a la información. Dicha teoría determinará que vínculos van a ser usados prioritariamente por el nodo.

Diseñando el comportamiento del agente

El diseño del comportamiento de los agentes no es tarea sencilla, pues muchas veces implica comportamientos para los cuales la psicología no presenta un frente unificado.

Existe una gran cantidad de teorías del comportamiento que ayudan a explicar partes de los procesos que determinan el comportamiento humano. Hace hasta poco, no existían intentos convincentes de aunar los diversos modelos y teorías psicológicas. Sin embargo, recientemente Jager y Janssen (Jager 2000; Jager et al., 2000, Janssen y Jager, 1999) han provisto un marco integrador.

En su teoría integradora de procesamiento cognitivo Janssen y Jager identifican dos dimensiones. Estas son el procesamiento automático en oposición al racional, y el procesamiento determinado socialmente en oposición al determinado individualmente.

Para el procesamiento automático en oposición al racional, un factor clave es si el consumo de cierto producto o servicio es altamente satisfactorio (considerando los recursos que gastamos en él). Si fácilmente cubre nuestras necesidades, entonces será menos probable que el proceso implique una intensa búsqueda de otras posibilidades. En cuanto a si se determina individualmente o socialmente, el factor clave es el nivel de certeza que lleva el consumo. Si no estamos seguros de lo que vamos a obtener con el producto, lo más probable será que busquemos alrededor nuestro para observar cómo otras personas están actuando.

Dadas estas dos dimensiones Jager y Janssen clasifican en cuatro los posibles modos de procesamiento cognitivo:

1. **Hábito o Repetición** (determinado individualmente y mediante un procesamiento automatizado). Este modo ocurre con aquellos bienes que

brindan un alto grado de satisfacción (no hay una necesidad urgente de buscar otros bienes que brinden mayor o igual satisfacción) y bajo nivel de incertidumbre (sabemos lo que obtendremos al comprar el producto). Un ejemplo sería comprar pasta de dientes.

2. **Deliberación** (determinado individualmente y mediante un razonamiento explícito). Este modo es usual en situaciones en las que existe un bajo nivel de satisfacción con el bien, comparado con las cosas que resignamos para obtenerlo, y en consecuencia surge una necesidad de pensar conscientemente acerca de lo que se recibe realmente y de cuales son las alternativas. Sin embargo, con un alto nivel de certeza sobre las cualidades del producto no existe una necesidad de observar lo que otras personas están haciendo.

3. **Imitación** (determinado socialmente y mediante un procesamiento automatizado) Este modo ocurre cuando existe un alto grado de satisfacción de un producto o servicio pero al existir incerteza sobre sus cualidades se sigue automáticamente lo que otras personas hacen.

4. **Comparación Social** (razonamiento explícito y determinado socialmente) Al igual que la imitación, las personas observan lo que otras personas están haciendo, pero al existir bajos niveles de satisfacción, la gente esta predispuesta a razonar conscientemente sobre la decisión de consumo. En este proceso de razonamiento los agentes intentan inferir información sobre el producto mediante el nivel de interés que tienen otros agentes por el producto.

Razonamiento	Determinado	
	Individualmente	Socialmente
Explícito	Deliberación	Comparación Social
Automatizado	Hábito o Repetición	Imitación

Figura N° 7: Clasificación de Jager y Janssen

4.5 Quinto Paso: Diseño de Experimentos

Como herramienta de análisis del problema se realizará un modelo basado en agentes. La metodología general para la construcción de este modelo implica la construcción de la población o agentes. Estos agentes interactúan en un contexto determinado o espacio social.

Los agentes son diseñados para emular sus partes en la vida real y usualmente consisten de:

- Atributos (Ej. edad, sexo, preferencias).
- Comportamientos (Ej. Acciones basadas en la optimización de utilidades o en el estado de sus contactos directos).

Luego de definir cuales serán los atributos o comportamientos de los agentes, será necesario fijar el valor de sus parámetros. A tal fin, existen dos opciones. Uno requiere usar datos de encuestas y tarea de campo, mientras que el otro consiste en realizar la mejor aproximación para luego calibrar estos valores comparando el resultado del modelo con el sistema del mundo real.

Luego de que los agentes fueron creados y parametrizados el modelo puede ser corrido. En general, el tiempo se maneja como una variable discreta y en cada paso los agentes son activados para ejecutar alguna acción determinada,

sujeta a ciertas condiciones. La simulación puede correr por un número definido de iteraciones o hacerlo indefinidamente.

Uno de los objetivos que se suele perseguir con la simulación es la de proveer al usuario de una idea acerca de cómo el sistema real se comportaría bajo las condiciones reflejadas en el modelo. Las simulaciones del sistema proveen al operador del modelo una ventana a través de la cual observar estos sistemas bajo diferentes parámetros o configuraciones.

Anylogic

Para el diseño de experimentos se utilizará Anylogic, un software de simulación que soporta diferentes paradigmas: Sistemas dinámicos, eventos discretos, dinámica de sistemas e incluso simulación basada en agentes.

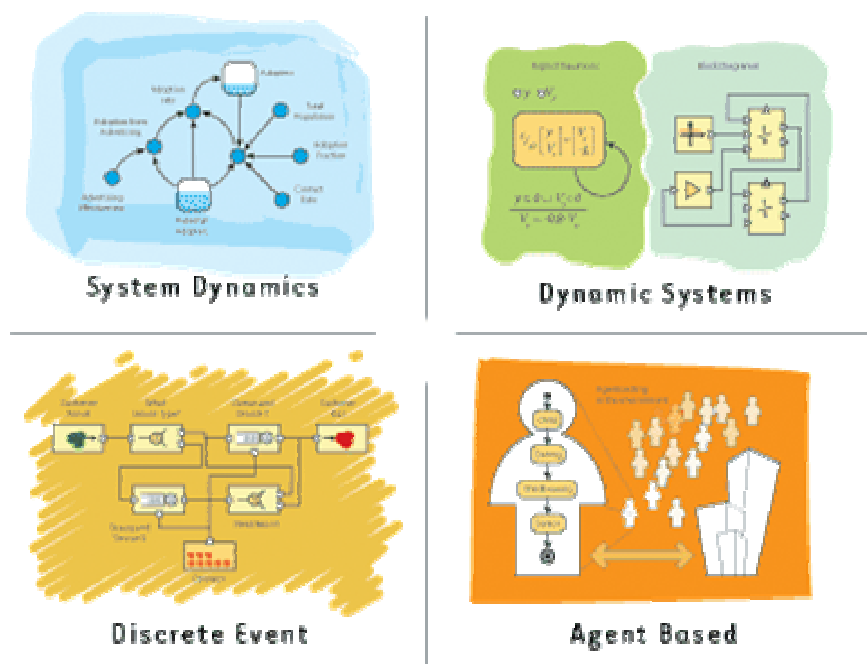


Figura N° 8: Paradigmas de simulación en Anylogic⁹

Sus principales características son su capacidad de:

- Crear modelos más eficientemente con objetos activos visuales, flexibles, extensibles y reutilizables, y mediante java.

⁹ Página Web de Anylogic

- Modelar más acertadamente y capturar más fenómenos mediante la utilización de múltiples paradigmas de simulación.
- Utilizar herramientas de análisis y optimización desde el panel de modelado.
- Integrarse fácilmente con planillas de cálculo, bases de dato, ERP y sistemas de CRM.
- Incrementar el ciclo de vida del modelo mediante el mantenimiento efectivo del mismo a medida que el sistema real evoluciona.

4.6 Sexto Paso: Conclusiones

La última etapa consiste en el análisis de los resultados de la simulación y de la información extraída del análisis estático de la red y el establecimiento de las conclusiones del estudio realizado.

5. CASO PRÁCTICO: COMPAÑÍA DE TELECOMUNICACIONES¹⁰

En esta sección se plantea un caso práctico aplicado a una empresa de telecomunicaciones. Se realiza una breve presentación de la empresa, se describe la situación actual y posteriormente se analiza el problema utilizando la metodología propuesta anteriormente.

Sobre la empresa

La compañía de telecomunicaciones es una empresa destinada a operar servicios de telefonía fija (nacional e internacional), telefonía pública y acceso a Internet Dial up y Banda Ancha. Pertenece junto a otras empresas al grupo TelCom.

El Grupo TelCom

El Grupo TelCom en la Argentina suministra actualmente una gama completa de servicios de telecomunicaciones, que incluyen telefonía fija; telefonía móvil; transmisión de datos y servicios de valor agregado; soluciones corporativas; acceso a Internet; servicios de “Contact Centers”; broker de seguros, y cables submarinos. Además, provee una variada gama de contenidos y servicios de soporte en áreas de recursos humanos, comercio exterior, seguros, logística, contabilidad, gestión inmobiliaria, seguridad, servicios generales y consultoría.

Desde el inicio de sus operaciones en el país, el mismo ha invertido más de 10.000 millones de dólares. Estos son algunos números relevantes del Grupo TelCom:

- Más de 13 millones de clientes totales en el país, el doble que en 2003.
- Más de 4,7 millones de líneas fijas instaladas.
- 9,7 millones de líneas de telefonía móvil.
- 120.600 teléfonos públicos en funcionamiento.
- Más de 117.000 clientes residenciales de Internet.
- 18.511 kilómetros de cable de fibra óptica instalados.
- Más de 440.000 accesos de banda ancha (ADSL) instalados.

En la actualidad, el Grupo TelCom emplea a casi 20.000 personas y realiza compras a sus empresas proveedoras por 3.000 millones de pesos al año, lo que equivale a casi el 1% del Producto Bruto Interno (PBI) nacional.

¹⁰ Por razones de confidencialidad se ha modificado el nombre de la compañía de telecomunicaciones y del grupo al cual pertenece.

Estrategia de Difusión Actual

La empresa posee una estrategia de difusión que combina publicidad en medios masivos (TV, radio, revistas, vía publica, Internet, etc.) con campañas de telemarketing.

Si bien las pautas generales de las campañas de publicidad son fijadas con ayuda de la casa matriz en el extranjero basándose en la experiencia de éxito de la misma, la empresa a nivel local se encarga de adaptar dicha estrategia al contexto local y de emprender numerosas campañas de telemarketing para explotar nichos no descubiertos.

La publicidad en medios masivos se utiliza generalmente para aumentar y/o mantener el conocimiento de marca, o para dar a conocer nuevos productos de la compañía como por ejemplo Internet por dial-up en su momento, o banda ancha en la actualidad. Mientras que el telemarketing presenta una manera personalizada de acercarse a un grupo target más específico y motivar la compra.

Para conocer el grado de efectividad de las campañas realizadas a través de medios masivos la empresa realiza diversas encuestas de percepción de sus spots publicitarios y medición del grado de influencia en la motivación de compra.

Con respecto a las campañas de telemarketing, la medición de la efectividad es más simple. Se generan listados de prospectos, se los llama y luego se mide el porcentaje de adoptadores sobre el total de individuos listados o total de contactados realmente.

Para la generación de dichos listados se aplican técnicas de minería de datos sobre la base de datos para la elección de prospectos con alta probabilidad de adopción del producto. Para la realización de dicho filtro se consideran variables referentes a información personal del individuo (edad, sexo, NSE, etc.), otros productos de la empresa que consume, nivel de consumo y otros datos.

Minería de Datos en Compañías de Telecomunicaciones

Las compañías de telecomunicaciones generan una gran cantidad de datos. Estos datos incluyen información detallada de las llamadas, que describen las llamadas que son transmitidas a través de las redes de telecomunicaciones; datos de la red, que describen el estado de los componentes de hardware y

software dentro de la red, e información del cliente; que describen a los usuarios de las redes de telefonía.

La cantidad de información es tan abundante que su análisis manual resultaría difícil, sino imposible. La necesidad de manipular tal cantidad de datos llevó al desarrollo de sistemas expertos basados en el conocimiento. Estos sistemas automatizados realizan funciones importantes tales como la identificación de llamadas fraudulentas o fallos en la red. El problema de este enfoque es la cantidad de tiempo que consume obtener el conocimiento de expertos humanos sin contar que en algunos casos los humanos pueden no tener el conocimiento suficiente.

El advenimiento de la tecnología de minería de datos prometió soluciones a estos problemas y por esta razón la industria de telecomunicaciones fue un adoptador temprano de dicha tecnología.

Tipos de Información Recopilada

A. Detalles de llamada

Cada vez que una llamada es realizada en una red de telecomunicaciones, información descriptiva sobre la llamada es guardada como un fichero de “detalle de llamada”. El número de ficheros de detalles de llamada es inmenso. (Por ejemplo, en los Estados Unidos de Norteamérica, los usuarios de larga distancia de AT&T generan más de 300 millones de ficheros por día¹¹). Dado que varios meses son mantenidos en línea, esto implica que decenas de miles de millones de ficheros de detalles de llamada deben estar almacenados en cualquier momento.

La información de las llamadas no es utilizada directamente dado que el objetivo de la minería de datos es extraer información al nivel del cliente y no al nivel de llamadas individuales. En consecuencia, los ficheros de detalle de llamadas asociados a un cliente deben ser resumidos en un solo fichero que describa el comportamiento del cliente. Debajo se presenta una lista de características que uno podría utilizar para generar la descripción del usuario basados en las llamadas que originan o reciben en un período de tiempo determinado (Delta T):

1. Duración promedio de la llamada
2. % de llamadas no atendidas

¹¹ Cortes & Pregibon, 2001.

3. % de llamadas desde/hacia un código de área diferente
4. % de llamadas en días de semana
5. Porcentaje de llamadas en horario diurno (9 AM- 5 PM)
6. Promedio del N° de llamadas recibidas durante el día
7. Promedio del N° de llamadas originadas durante el día
8. N° de códigos de área únicos llamados durante Delta T

Estas ocho características pueden ser utilizadas para crear un perfil del usuario. Este perfil tiene muchos usos potenciales. Por ejemplo podría ser utilizado para distinguir entre clientes comerciales o residenciales en base al porcentaje de llamadas en días de semana y horarios diurnos.

Citando otro uso menos común, la última característica enumerada, podría ser utilizada para identificar telemarketers dado que estos llaman a una cantidad diversa de códigos de área en un periodo de tiempo corto.

B. Datos de la Red

Las redes de telecomunicaciones son redes extremadamente complejas, compuestas por miles de componentes interconectados. Cada uno de los elementos de la red es capaz de generar mensajes de error y estado, que producen una gran cantidad de datos de red. Estos datos deben ser almacenados y analizados de manera tal de brindar soporte a las funciones de administración de la red, como por ejemplo aislamiento de fallas. Esta información incluye al menos el horario, un código que identifique de manera inequívoca el hardware o software generando el mensaje y un código que explique la razón del mensaje.

Debido a la enorme cantidad de datos, los técnicos no pueden manipular cada mensaje. Por esta razón, se han desarrollado sistemas expertos para analizar automáticamente estos mensajes y tomar acciones debidas, sólo requiriendo de un técnico en caso de que el problema no pueda ser solucionado automáticamente.

C. Datos del Cliente

Las compañías de telecomunicaciones poseen millones de clientes. Esto implica el mantenimiento de la información de estos clientes. Esta información incluye nombre, dirección, nivel crediticio, plan de servicio, información del contrato actual, ingreso familiar e historial de pago. Esta información puede ser complementada con datos de fuentes externas y suele ser utilizada junto con

los otros tipos de datos descriptos anteriormente para mejorar los resultados del análisis.

Usos de la información

Los usos son diversos y van desde la detección de llamadas fraudulentas, el aislamiento de fallas en la red hasta el soporte de acciones de marketing o comerciales.

Las compañías de telecomunicación mantienen una gran cantidad de datos de los clientes. Además de la información del cliente que mayoría de los negocios recolectan, las compañías de telecomunicaciones también almacenan datos con el detalle de las llamadas, que describen de manera precisa los hábitos de llamado de los clientes.

Con propósitos de marketing, esta información puede ser utilizada para crear perfiles de los clientes, diseñar productos específicos de acuerdo a los perfiles definidos, crear campañas de telemarketing, crear paquetes de productos que sean bien recibidos por los clientes, diseñar nuevas promociones, realizar ventas cruzadas de productos y muchas otras funciones.

Es importante destacar que muchas veces la manera en que se pueden utilizar las bases de datos esta restringida por leyes de competencia o privacidad de la información.

5.1 Definición del Problema

El problema es más bien una oportunidad. Una oportunidad para incorporar el análisis de las redes sociales en el diseño de estrategias de difusión. Actualmente, la compañía presenta una estructura formal que se encarga de explotar la información de la base de datos que almacena la empresa con el objetivo de diseñar nuevas campañas de telemarketing, diseñar nuevos productos en base a la necesidad que perciben y retroalimentar a diferentes áreas de la empresa con la información de reclamos de clientes.

El contexto de mercado se vio modificado de manera significativa. La desregularización del mercado de telefonía fija y la entrada de la compañía en nuevos segmentos como Internet de banda ancha generan el desafío de diseñar nuevas estrategias de difusión que se adapten a un mercado de competencia y altamente dinámico.

No es sólo el mercado de telecomunicaciones el que se ha modificado. También evolucionan con los años los medios de difusión para llegar al cliente

objetivo. De esta manera, han surgido nuevas estrategias, técnicas o métodos para transmitir un mensaje al cliente potencial. Dentro de las mismas se observan desde técnicas de impacto en medios masivos (Ej. repetición de cortes publicitarios en una misma tanda de comerciales) hasta el surgimiento de nuevas tendencias como el marketing viral o nuevos medios como la publicidad en Internet (banners, ad-sensing de google, pop-ups, etc).

Actualmente, la empresa no suele emplear un foco basado en las redes sociales. En consecuencia, no es posible diseñar una estrategia de difusión que contemple la posición de los potenciales clientes dentro de la red social y su impacto en la difusión del producto mediante el efecto boca en boca.

Recientemente la empresa ha difundido el producto de Internet de banda ancha (ADSL) en Monte Equis. Se utiliza este caso para estudiar de qué manera se aplicaría la metodología planteada. Si bien la cantidad de experimentos de simulación que se pueden realizar es tan extensa como objetivos y nuevas teorías se pretenda investigar, en este caso particular se diseña un modelo de simulación para analizar cómo afectaría una estrategia que comience la difusión a través de los hubs de la red social en el desempeño final de la difusión.

Monte Equis

Más específicamente la difusión se investiga en Monte Equis, un pueblo que posee una población estimada de 15.000 habitantes. La cantidad de líneas fijas es de aproximadamente 3.600. En los últimos años, la empresa de telecomunicaciones ha comenzado a brindar servicios de ADSL en el pueblo.

La elección del pueblo se debe a que:

- Es un pueblo en el cual no hay competencia para ADSL por lo cual se filtra el efecto que la misma pueda ocasionar en el análisis permitiéndonos concentrarnos en la difusión sin competencia.
- Es un pueblo relativamente pequeño y aislado que permite que la cantidad de datos de llamadas de telefonía fija sea más manejable en computadoras personales.

5.2 Representación de la Topología de Red

Para poder conocer la topología de red sobre la cual se difundirá el producto de la empresa se utilizará un modelo híbrido, en el cual haciendo uso de datos disponibles sobre las llamadas a través de telefonía fija de la compañía de telecomunicaciones se determinará el tipo de topología que mejor represente la red social bajo estudio.

Datos Disponibles

Como mencionamos anteriormente, es muy complejo determinar la topología exacta que represente la sociedad. Las razones de esta dificultad radican en la cantidad de datos de diferentes canales de comunicación que debería de recolectarse y en la dinámica de la red que hace que se modifiquen constantemente las relaciones.

A pesar de esto, se considera que la utilización de los datos de las comunicaciones a través de la telefonía fija permite obtener una idea acabada acerca de la estructura de red que soslaya en la sociedad debido a que:

- Existen estudios demostrando que el hecho de que dos personas se comuniquen frecuentemente a través de telefonía fija es un indicador de que probablemente lo hagan en forma personal (cara a cara) o a través de telefonía móvil, correo electrónico u otros medios.
- Dentro de la diversidad de canales para la comunicación interpersonal la telefonía fija es aún una de las más utilizadas.
- Los nodos representarán los diferentes hogares que son adoptadores potenciales del servicio de ADSL que provee la compañía de telecomunicaciones.
- Los enlaces representarán las relaciones entre los hogares y se modelarán en función de las comunicaciones que realizan a través de la red de telefonía fija.

Descripción de los datos

Se analizarán los datos de las llamadas telefónicas realizadas a través de líneas fijas en el barrio de Monte Equis.

Tabla “en Crudo”

La tabla en crudo presenta información sin procesar. La información correspondiente a las llamadas telefónicas de Monte Equis se recibió en formato separado con comas (“CSV”). Estos datos fueron exportados a la base de datos MySQL y luego se realizaron las consultas correspondientes para procesar los datos.

Los campos correspondientes a la tabla “en crudo” son:

- DDN (Discado Directo Nacional): También denominado código interurbano o código de área. Es el servicio que permite acceder a los códigos telefónicos correspondientes a las diferentes provincias y localidades del territorio argentino. La provincia en la que se ubica el pueblo bajo estudio posee 5 localidades diferentes a las cuales le corresponden los siguientes DDN:
 - Localidad A, Localidad B, Localidad C (03901)
 - Localidad D y Localidad E (03964)
- Código Urbano: Es el código correspondiente al barrio de Monte Equis. Posee las siguientes características: 540, 543, 544 o 545¹².
- Código de línea: Es el código que identifica el número de línea. Toma valores de 000 al 999.
- Fecha de llamada: Indica la fecha en que se realizó la llamada.
- Hora de llamada: Indica la hora en que la llamada fue realizada.
- Código de banda horaria: Distingue entre los dos códigos de banda horaria (1 o 2).
- Código destino: Es el número destino de la llamada y se compone del código interurbano, el código urbano y el código de línea.
- Duración de llamada: Indica la duración de la llamada en segundos.
- Duración tarifada: Indica la cantidad de pulsos consumidos por llamada.
- Código de producto: Es el código del producto de ADSL.

¹² Los códigos han sido modificados.

- Fecha de alta: Indica la fecha en que se adoptó el producto. Si el usuario no adoptó el producto entonces este campo estará vacío.

El total de llamadas realizadas por el pueblo en un lapso de 3 meses es de aproximadamente unos 3,2 millones. A su vez se observa que existen 3,600 líneas diferentes que realizaron los llamados (denominado de aquí en adelante como “códigos salientes”). Estos códigos salientes llaman a un total de 57,200 líneas diferentes (o “códigos entrantes”).

Cabe destacar que los códigos telefónicos corresponden a números residenciales. Los números comerciales se han filtrado dado que poseen un uso diferente de la línea telefónica y deberían de ser analizados de manera separada.

Procesamiento de datos

La cantidad de datos (3,2 millones de llamadas) que se deben analizar hace prohibitivo el uso de una planilla de cálculos para su procesamiento. Es por eso que se optó por el uso de la base de datos *MySQL 5.0* y el programa *TOAD* que permite realizar consultas y exportar los resultados al *MS Excel*.

Consulta 1: Obtención de la “Tabla Estandarizada”

Con el objetivo de estandarizar los datos, luego de exportada a la base de datos, se procede a crear una nueva tabla con los campos: `cod_out`, `cod_in` y `dur_llamada`.

- `Cod_out`: Corresponde a la línea desde la cual se realiza la llamada. Y se compone del Código urbano y el código de línea, concatenados en una misma línea de información. (Por Ej. 540 123).
- `Cod_in`: Corresponde a la línea que recibe la llamada. Al igual que el `cod_out` se compone del código urbano y el código de línea.
- `Dur_llamada`: Indica la duración de la llamada en segundos.

Dado que todos los códigos telefónicos están precedidos por el código interurbano 3901, correspondiente a la localidad en la que se ubica el pueblo, el dato es omitido dado que no agrega ninguna información.

Definición del Área Bajo Estudio

El siguiente esquema muestra el flujo de llamadas desde el punto de vista del área urbana:

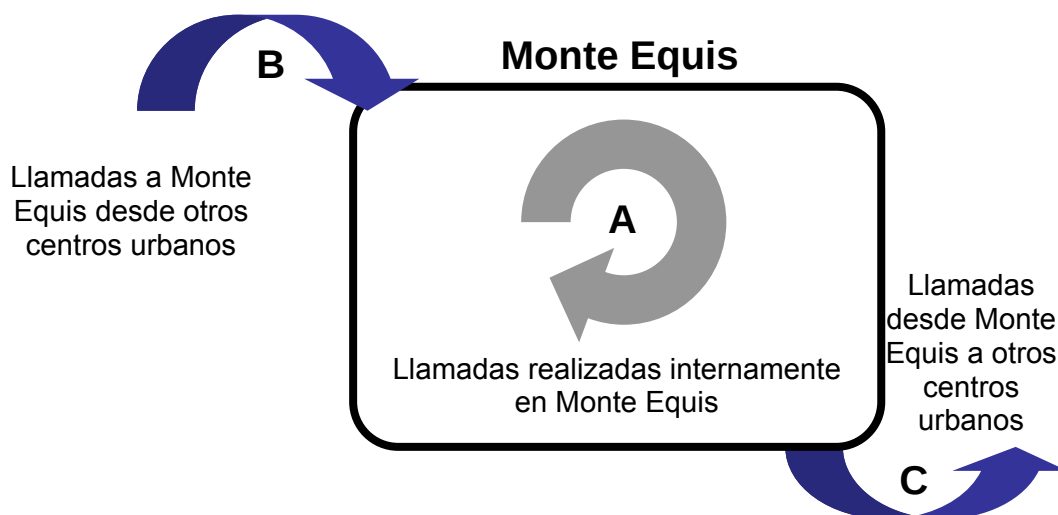


Figura N° 9: Flujo de Llamadas

A cada área geográfica le corresponde uno o varios códigos urbanos. Por ejemplo, a Monte Equis le corresponden los códigos urbanos 540, 543, 544 o 545.

El flujo A muestra las llamadas que son realizadas internamente en Monte Equis. El flujo B representa las llamadas que son realizadas desde otros centros urbanos hacia Monte Equis. Finalmente, el flujo C muestra las llamadas salientes desde Monte Equis hacia otros centros urbanos.

Este estudio hará foco en analizar la red social de Monte Equis y la influencia que sus habitantes tienen en otros habitantes del mismo centro urbano. El hecho de querer representar sólo la red social de Monte Equis, implica que los nodos a representar deberán corresponder solamente a esta región. Por esta razón se aislarán y estudiarán los datos correspondientes al flujo A.

Si bien no es el foco de este estudio, se podría realizar un análisis de influencia interurbano para el cual sería necesario incluir el estudio de los flujos B y C.

Consulta 2: Obtención de la “Tabla Estandarizada con llamadas dentro de Monte Equis”

Si bien todas los *cod_out* poseen un código urbano correspondiente a pueblo de Monte Equis, se observa que dentro de los *cod_in* existen códigos que no corresponden a códigos dentro del pueblo.

Debido a la naturaleza de este estudio que busca analizar la red social de Monte Equis de manera aislada, se analizan las llamadas que son realizadas dentro del pueblo. En consecuencia, se procede a filtrar todas las llamadas que tienen como código destino un código urbano correspondiente a Monte Equis (540, 543, 544 o 545). Dentro de los códigos destino también se filtraron otros números como los 0810, 0800, *124 o códigos alfanuméricos que no corresponden a códigos telefónicos residenciales.

El resultado de este filtro es una tabla con las llamadas que fueron realizadas desde Monte Equis y que tuvieron como destino otro número telefónico localizado en el mismo pueblo.

Esta tabla contiene alrededor de 800.000 llamadas. El número de códigos diferentes que realizaron los llamados fue de 3.600 y el número de códigos diferentes que recibieron llamados fue de 3.700.

Interpretación De Los Datos

En esta instancia, es importante responder las siguientes preguntas:

- ¿Qué representarán los nodos?

Los nodos serán representados por las diferentes líneas telefónicas. Es decir, que los nodos representarán a su vez las diferentes viviendas con línea telefónica de Monte Equis.

- ¿Qué relación representarán los enlaces?

Al investigar el proceso de difusión de determinado producto los enlaces entre nodos deberían representar idealmente las relaciones que potencialmente impliquen un contacto que arrastre el efecto boca en boca.

Se definirán tres métodos diferentes para representar los enlaces y se los comparará posteriormente:

Método Base: Se indicará un enlace entre dos nodos cuando una persona realiza al menos un llamado al nodo destino. Por ejemplo si una persona realiza cinco llamadas a cinco códigos de línea diferentes entonces se representará mediante cinco enlaces diferentes a cada uno de los nodos destino. Si en cambio la persona realiza cinco llamadas a tres códigos de línea diferentes esto se representará mediante tres enlaces a cada uno de los tres códigos-destino diferentes.

Método Tiempo: Se indicará como un enlace entre dos nodos a cada llamada con una duración mayor a tres minutos (o 180 segundos).

Método Llamada Recíproca: Se indicará como un enlace entre dos nodos al caso en que la llamada sea recíproca. Es decir que si una persona llama a otra y esta a su vez llama a la primera persona en un período de 3 meses, entonces esta situación representará un enlace o relación entre las dos personas. Si en cambio una persona llama a otra y ésta no llama a la primera entonces se considera que esta situación no representará un enlace entre los dos nodos.

Metodologías Para El Ajuste De Datos

Se ha encontrado que diversas redes sociales se ajustan a la estructura de red de escala libre. Sin embargo, dichos estudios no se han realizado hasta el momento con datos de comunicaciones a través de telefonía fija en argentina. Es por eso que se procederá a realizar un ajuste de los datos para determinar si la red social de Monte Equis también presenta una estructura de red de escala libre.

Las redes de escala libre presentan algunos nodos con un elevado grado nodal mientras que la mayoría de los nodos presentan un grado nodal bajo.

$$p(x) = C \cdot x^{-\gamma} \quad (1)$$

Las distribuciones de la forma (1) siguen la ley potencial. La constante γ es denominada el exponente de la ley potencial. Mientras que C corresponde a un valor constante y X representa el grado nodal (cantidad de conexiones) del nodo.

Para revelar la forma de la ley potencial, es más conveniente graficar el histograma en una escala logarítmica. Al hacerlo, si los datos efectivamente

siguen una ley potencial se observará que presentan una forma de línea recta (Ver ecuación 4)

$$\text{Log}(p(x)) = \text{Log}(C.x^{-\gamma}) \quad (2)$$

$$\text{Log}(p(x)) = \text{Log}(C) + \text{Log}(x^{-\gamma}) \quad (3)$$

$$\text{Log}(p(x)) = \text{Log}(C) + (-\gamma)\text{Log}(x) \quad (4)$$

Sin embargo, el gráfico de escala logarítmica presenta un inconveniente. La cola del gráfico presenta un gran ruido debido a que en esa sección la cantidad de datos por rango (“bin”) es baja.

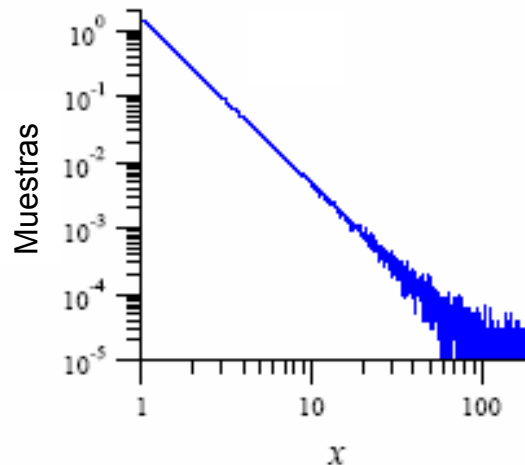


Figura N° 10: Gráfico en Escala Logarítmica

Una manera de lidiar con este inconveniente es dejar fuera del análisis los datos en la cola de la curva. Sin embargo, esos datos pueden presentar información útil y a su vez muchas distribuciones presentan un comportamiento de ley potencial sólo en la cola.

Una solución alternativa es la de utilizar rangos logarítmicos. En este caso, los rangos en la cola de la distribución poseen más muestras que si los rangos fuesen fijos, y esto reduce los errores estadísticos en la cola. En ese caso el gráfico disminuirá el ruido.

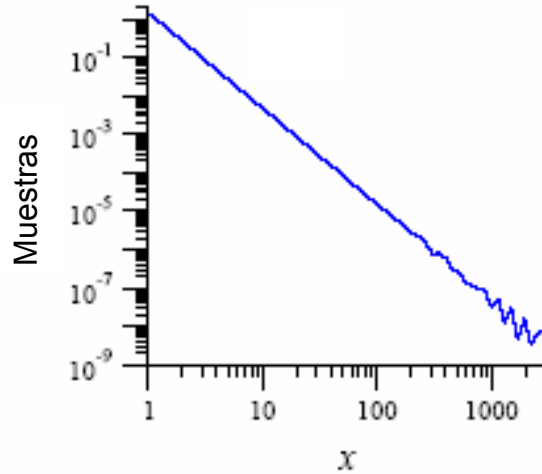


Figura N° 11: Gráfico en Escala Logarítmica utilizando rangos logarítmicos

Aún utilizando rangos logarítmicos existe algún ruido en la cola, a pesar de que el mismo es reducido drásticamente. Supongamos que el límite inferior del rango más bajo este a $x_{\min}$ y que el ratio del ancho de los rangos sucesivos sea a . De esta manera, el rango k se extendería desde $x_{k-1} = x_{\min} a^{k-1}$ hasta $x_k = x_{\min} a^k$ y el número de muestras esperado en ese rango es

$$\begin{aligned} \int_{x_{k-1}}^{x_k} p(x) dx &= C \int_{x_{k-1}}^{x_k} x^{-\gamma} dx \\ &= C \frac{a^{\gamma-1} - 1}{\gamma - 1} (x_{\min} a^k)^{-\gamma+1} \end{aligned}$$

Entonces, siempre que $\gamma > 1$, el número de muestras por rango disminuirá a medida que aumenta k y los rangos en la cola tendrán más ruido estadístico que aquellos rangos que lo preceden. Como vimos anteriormente la mayoría de las distribuciones de ley potencial que ocurren en la naturaleza tienen un $2 \leq \gamma \leq 3$, por lo cual las colas ruidosas son la norma.

Otro método superior para graficar los datos es calcular la función de distribución acumulada. En lugar de graficar un histograma de los datos, se realiza un gráfico de la probabilidad $P(x)$ de que x tenga un valor superior o igual a x :

$$P(x) = \int_x^{\infty} p(x') dx'$$

Si la distribución sigue la ley potencial $p(x) = C \cdot x^{-\gamma}$, entonces

$$P(x) = C \int_x^{\infty} x'^{-\gamma} dx' = \frac{C}{\gamma-1} x^{-(\gamma-1)}$$

Por ende la función de distribución acumulada $P(x)$ también sigue una ley potencial pero con un exponente diferente $(\gamma-1)$, que es menor por una unidad que el exponente original. De esta manera, si graficásemos $P(x)$ en escala logarítmica deberíamos obtener nuevamente una recta pero con una pendiente menor.

En este caso no existe necesidad de utilizar rangos para agrupar los datos y calcular $P(x)$. Por definición, $P(x)$ está definido para todo valor de x , por lo cual puede ser graficado normalmente sin necesidad de utilizar rangos. Esto evita cualquier interrogante sobre el tamaño que deberían tener los rangos. A su vez, hace un uso más efectivo de los datos dado que al no agrupar los datos en rangos, no se desecha la información contenida en los valores individuales.

Para estimar el valor de γ se ajustan los datos mediante una recta y se obtiene la ecuación de la misma. Un método alternativo, simple y confiable es extraer el valor del exponente utilizando la fórmula

$$\gamma = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \right]^{-1}$$

Aquí las cantidades x_i , $i = 1 \dots n$ son los valores medidos de x y $x_{\min}$ es el valor mínimo de x . En situaciones prácticas $x_{\min}$ corresponde usualmente no al mínimo valor de x sino al mínimo valor para el cual la ley potencial se mantiene. Una estimación de error estadístico esperado está dado por:

$$\sigma = \sqrt{n} \left[\sum_{i=n}^n \ln \frac{x_i}{x_{\min}} \right]^{-1} = \frac{\gamma - 1}{\sqrt{n}}$$

Ajuste De Datos

Para el ajuste de datos se utilizará el método de ajuste lineal de la función de distribución de probabilidad acumulada. Las condiciones de aceptación del ajuste de la red a una red del tipo escala libre son: 1) Coeficiente de Correlación: $R^2 > 0,75$. Un coeficiente de correlación mayor a 0,75 es indicador del buen ajuste de la red considerando que el ajuste es de tipo social. 2) Exponente de la ley potencial: En el caso de que se ajuste $p(X)$ el exponente deberá estar entre $2 \leq \gamma \leq 3$. Sin embargo, por las ventajas antes detalladas se procederá a graficar y ajustar la distribución de probabilidad acumulada $P(X)$. En este caso se espera que el exponente este entre $1 \leq \beta \leq 2$.

Método Base

A continuación se observa la tabla de frecuencias en donde se indica la cantidad de nodos que poseen k cantidad de contactos. Es decir, se establece la frecuencia por grado nodal:

K	Frecuencia	K	Frecuencia	K	Frecuencia	K	Frecuencia
1	44	44	43	88	4	148	1
2	44	45	30	89	5	152	1
3	41	46	30	90	1	155	1
4	50	47	31	91	1	156	1
5	53	48	17	92	2	161	1
6	52	49	25	93	4	162	2
7	56	50	28	94	1	168	1
8	80	51	27	95	3	169	1
9	92	52	18	96	1	172	1
10	83	53	27	98	3	173	1
11	73	54	15	99	1	181	1
12	101	55	25	100	1	189	1
13	98	56	21	101	4	197	2
14	100	57	15	102	1	201	1
15	82	58	14	103	2	216	1
16	84	59	13	104	3	223	1
17	101	60	10	107	1	227	1
18	92	61	18	110	2	231	1
19	95	62	9	111	2	234	1
20	94	63	8	112	1	235	1
21	97	64	6	114	2	246	1
22	98	65	6	116	2	256	1
23	72	66	6	118	1	260	1
24	88	67	6	119	2	262	1
25	76	68	7	121	1	263	1
26	78	70	10	124	1	264	1
27	87	71	3	125	1	286	1
28	73	72	2	126	1	310	1
29	59	73	9	127	1	327	1
30	70	74	2	128	1	336	1
31	61	75	9	129	1	340	1
32	54	76	8	131	2	351	1
33	59	77	8	134	1	354	1
34	56	78	4	135	2	374	1
35	55	79	6	136	1	381	1
36	44	80	3	137	1	388	1
37	37	81	3	139	1	390	1
38	52	82	4	140	1	422	1
39	33	83	2	141	2	451	1
40	48	84	6	142	1	502	1
41	34	85	6	143	1	510	1
42	29	86	1	144	1	513	1
43	45	87	4	145	4	663	1

Tabla N° 1: Método Base. Tabla de Frecuencias

Para observar la tendencia se procederá a realizar el gráfico de la tabla anterior:

Método Base - Histograma

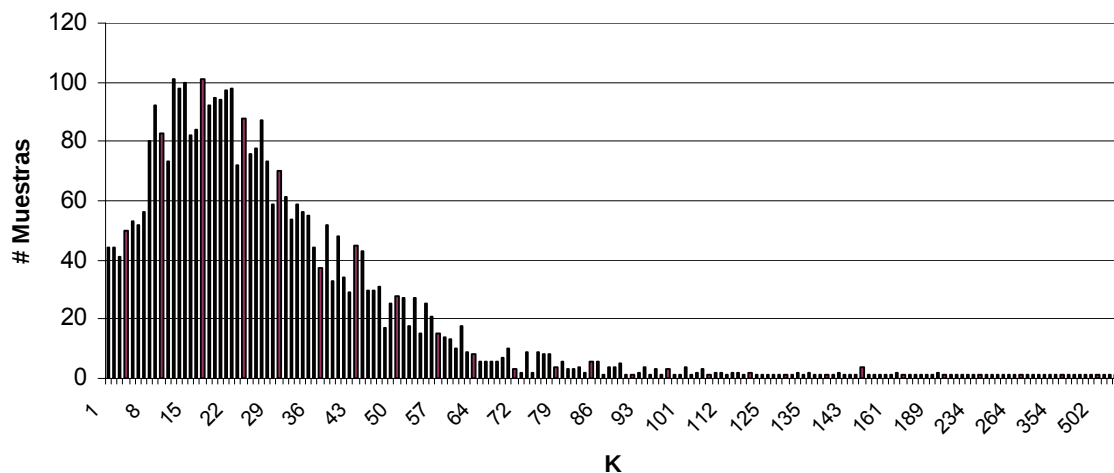


Figura N° 12: Método Base. Histograma

Al analizar el histograma se observa que existen muchos nodos con pocas conexiones mientras que existen una minoría de nodos muy bien conectados. A su vez, se observa que la curva de frecuencias tiene un pico entre $K = 12$ y $K = 17$, lo cual estaría indicando que si bien se sigue la tendencia de que existe una minoría mejor conectada, lo más frecuente es que se contacten entre 12 a 17 personas diferentes por vivienda.

Proceso de Ajuste

A partir de la tabla anterior se obtiene otra tabla que tiene como campos Log (P (K)) y Log (K):

Log(K)	Log(P(K))	Log(K)	Log(P(K))	Log(K)	Log(P(K))
0,000	0,000	1,763	-1,059	2,104	-1,735
0,301	-0,005	1,771	-1,079	2,107	-1,742
0,477	-0,011	1,778	-1,098	2,111	-1,748
0,602	-0,016	1,785	-1,114	2,117	-1,755
0,699	-0,022	1,792	-1,143	2,127	-1,769
0,778	-0,029	1,799	-1,158	2,130	-1,776
0,845	-0,036	1,806	-1,172	2,134	-1,791
0,903	-0,043	1,813	-1,183	2,137	-1,799
0,954	-0,054	1,820	-1,195	2,143	-1,806
1,000	-0,067	1,826	-1,206	2,146	-1,814
1,041	-0,079	1,833	-1,218	2,149	-1,822
1,079	-0,090	1,845	-1,232	2,152	-1,838
1,114	-0,105	1,851	-1,253	2,155	-1,847
1,146	-0,120	1,857	-1,260	2,158	-1,856
1,176	-0,137	1,863	-1,264	2,161	-1,864
1,204	-0,150	1,869	-1,285	2,170	-1,901
1,230	-0,165	1,875	-1,290	2,182	-1,911
1,255	-0,183	1,881	-1,311	2,190	-1,921
1,279	-0,201	1,886	-1,332	2,193	-1,931
1,301	-0,219	1,892	-1,353	2,207	-1,942
1,322	-0,239	1,898	-1,364	2,210	-1,952
1,342	-0,259	1,903	-1,381	2,225	-1,975
1,362	-0,281	1,908	-1,390	2,228	-1,986
1,380	-0,298	1,914	-1,399	2,236	-1,998
1,398	-0,320	1,919	-1,411	2,238	-2,010
1,415	-0,340	1,924	-1,418	2,258	-2,023
1,431	-0,361	1,929	-1,437	2,276	-2,036
1,447	-0,386	1,934	-1,458	2,294	-2,049
1,462	-0,408	1,940	-1,461	2,303	-2,077
1,477	-0,427	1,944	-1,475	2,334	-2,092
1,491	-0,450	1,949	-1,490	2,348	-2,107
1,505	-0,471	1,954	-1,509	2,356	-2,123
1,519	-0,491	1,959	-1,513	2,364	-2,140
1,531	-0,514	1,964	-1,517	2,369	-2,157
1,544	-0,537	1,968	-1,525	2,371	-2,174
1,556	-0,560	1,973	-1,542	2,391	-2,193
1,568	-0,580	1,978	-1,546	2,408	-2,212
1,580	-0,597	1,982	-1,559	2,415	-2,232
1,591	-0,623	1,991	-1,563	2,418	-2,253
1,602	-0,640	1,996	-1,577	2,420	-2,276
1,613	-0,666	2,000	-1,581	2,422	-2,299
1,623	-0,686	2,004	-1,586	2,456	-2,324
1,633	-0,703	2,009	-1,605	2,491	-2,350
1,643	-0,732	2,013	-1,610	2,515	-2,378
1,653	-0,761	2,017	-1,620	2,526	-2,408
1,663	-0,782	2,029	-1,635	2,531	-2,441
1,672	-0,805	2,041	-1,641	2,545	-2,475
1,681	-0,829	2,045	-1,651	2,549	-2,513
1,690	-0,844	2,049	-1,662	2,573	-2,554
1,699	-0,865	2,057	-1,668	2,581	-2,600
1,708	-0,891	2,064	-1,679	2,589	-2,651
1,716	-0,917	2,072	-1,691	2,591	-2,709
1,724	-0,935	2,076	-1,697	2,625	-2,776
1,732	-0,965	2,083	-1,709	2,654	-2,856
1,740	-0,982	2,093	-1,716	2,701	-2,952
1,748	-1,012	2,097	-1,722	2,708	-3,077
1,756	-1,039	2,100	-1,728	2,710	-3,253
				2,822	-3,554

Tabla Nº 2: Método Base. Tabla en Logaritmos

A partir de la tabla anterior se realiza un gráfico y se realiza el ajuste lineal:

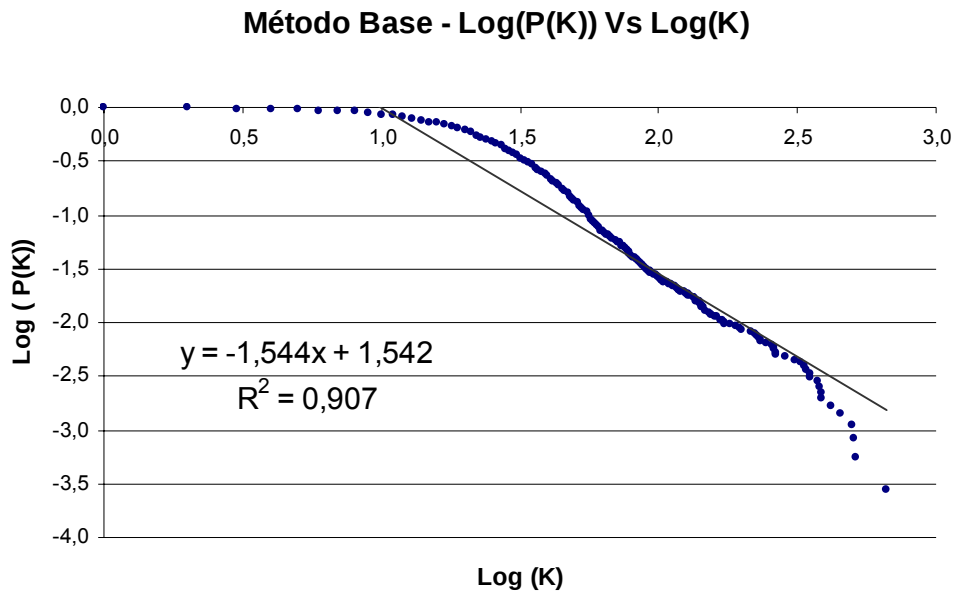


Figura N° 13: Método Base. Ajuste de Datos

Del análisis de los resultados del ajuste se observa que el R^2 es de 0.91, y el valor del exponente de la red de escala libre $\beta = 1,5$ ($\gamma = 2.5$).

Método Tiempo

En primer lugar se obtiene una tabla que muestra la frecuencia por cantidad de contactos (k):

K	Frecuencia	K	Frecuencia
1	283	29	6
2	344	30	4
3	328	31	5
4	327	32	6
5	316	33	5
6	271	34	1
7	225	35	4
8	192	36	1
9	175	37	1
10	139	38	1
11	118	39	3
12	110	40	1
13	91	41	2
14	75	42	1
15	51	45	2
16	47	46	2
17	39	47	2
18	35	49	1
19	21	53	1
20	14	54	2
21	17	55	1
22	13	64	1
23	16	74	1
24	6	75	1
25	6	81	1
26	8	89	1
27	5	94	2
28	10	121	1

Tabla N° 3: Método Tiempo. Tabla de Frecuencias

A grandes rasgos se puede observar que existen muchos nodos con pocas conexiones mientras que existen una minoría de nodos muy bien conectados.

Método Tiempo - Histograma

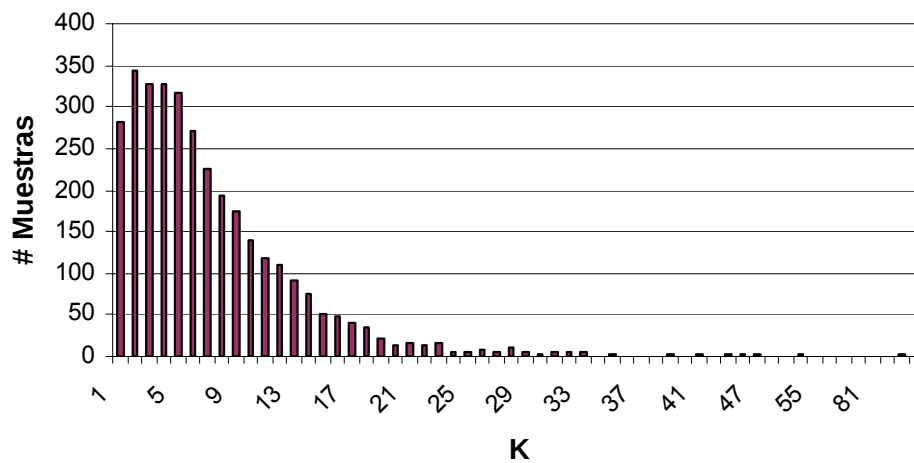


Figura N° 14: Método Tiempo. Histograma

A partir de la tabla anterior se obtiene una tabla que tiene como campos $\log(k)$ y $\log(P(k))$ y posteriormente se procede a realizar un ajuste lineal de los datos.

Log(K)	Log(P(K))	Log(K)	Log(P(K))
0,000	0,000	1,462	-1,746
0,301	-0,038	1,477	-1,792
0,477	-0,090	1,491	-1,825
0,602	-0,146	1,505	-1,871
0,699	-0,210	1,519	-1,933
0,778	-0,282	1,531	-1,993
0,845	-0,356	1,544	-2,005
0,903	-0,428	1,556	-2,062
0,954	-0,500	1,568	-2,077
1,000	-0,579	1,580	-2,093
1,041	-0,654	1,591	-2,109
1,079	-0,729	1,602	-2,162
1,114	-0,813	1,613	-2,182
1,146	-0,898	1,623	-2,223
1,176	-0,982	1,653	-2,245
1,204	-1,051	1,663	-2,294
1,230	-1,126	1,672	-2,348
1,255	-1,200	1,690	-2,410
1,279	-1,278	1,724	-2,445
1,301	-1,334	1,732	-2,483
1,322	-1,375	1,740	-2,570
1,342	-1,431	1,806	-2,621
1,362	-1,479	1,869	-2,679
1,380	-1,546	1,875	-2,746
1,398	-1,575	1,908	-2,825
1,415	-1,605	1,949	-2,922
1,431	-1,649	1,973	-3,047
1,447	-1,679	2,083	-3,524

Tabla N° 4: Método Tiempo. Tabla en Logaritmos

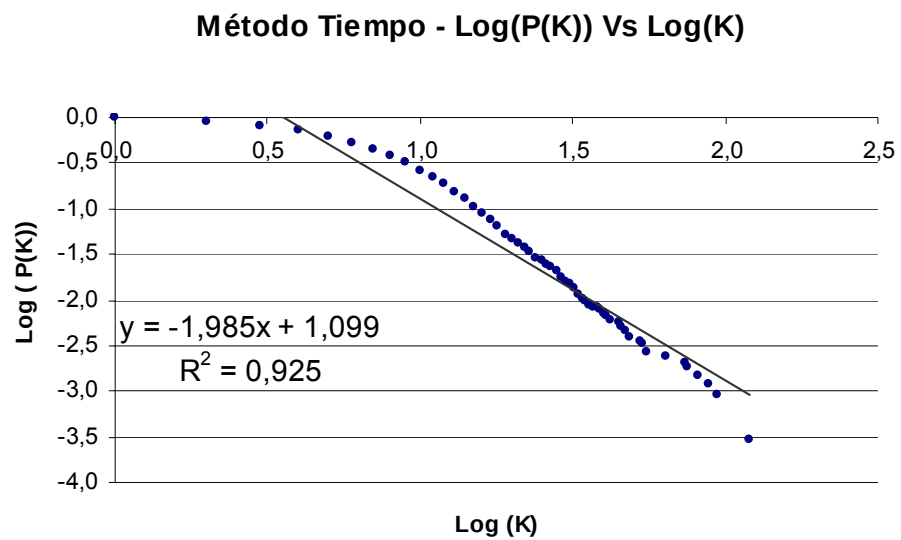


Figura N° 15: Método Tiempo. Ajuste de Datos

Del análisis del ajuste lineal se observa que en este caso los datos se ajustan con un R^2 de 0.93 a un tipo de red de escala libre con exponente $\beta = 2,0$ ($\gamma = 3,0$).

Método Llamada Recíproca

Al igual que en los métodos anteriores se obtiene una tabla que muestra la frecuencia por cantidad de contactos (k):

K	Frecuencia	K	Frecuencia
1	198	34	6
2	221	35	11
3	253	36	3
4	215	37	6
5	249	38	5
6	213	39	2
7	222	40	1
8	194	41	1
9	197	42	3
10	175	43	2
11	135	44	1
12	121	45	1
13	122	47	2
14	120	49	1
15	92	50	3
16	80	53	1
17	77	54	1
18	75	55	1
19	65	56	1
20	57	57	2
21	44	58	1
22	28	59	1
23	30	65	1
24	25	69	1
25	25	70	1
26	13	75	1
27	29	77	1
28	17	79	1
29	13	84	2
30	17	92	1
31	10	96	1
32	4	105	1
33	8	120	1

Tabla N° 5: Método Llamada Recíproca. Tabla de Frecuencias

Posteriormente, se muestra la tabla en un gráfico de frecuencias para observar la tendencia:

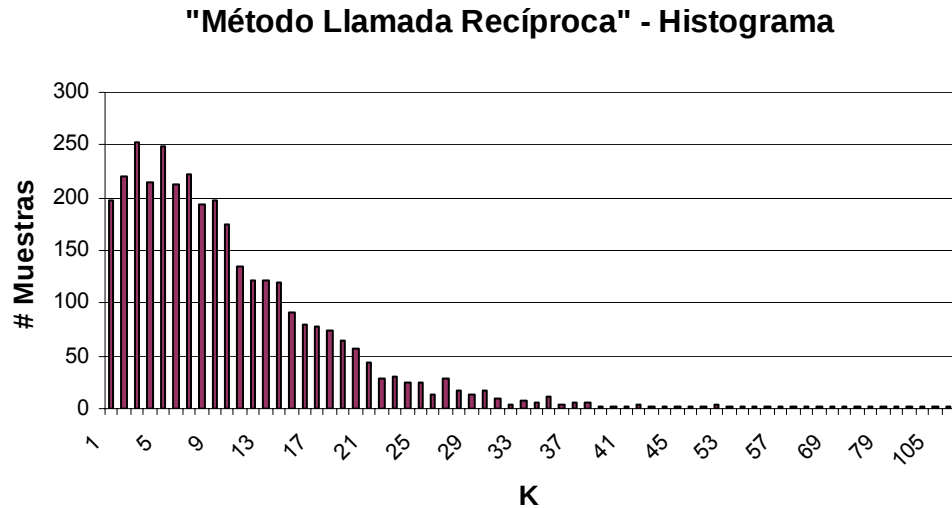


Figura N° 16: Método Llamada Recíproca. Histograma

A partir de los datos se obtiene la siguiente tabla y se realizan el ajuste correspondiente:

Log(K)	Log(P(K))	Log(K)	Log(P(K))	Log(K)	Log(P(K))
0,0000	0,0000	1,3617	-1,1197	1,6532	-2,1016
0,3010	-0,0260	1,3802	-1,1732	1,6721	-2,1180
0,4771	-0,0569	1,3979	-1,2234	1,6902	-2,1528
0,6021	-0,0953	1,4150	-1,2802	1,6990	-2,1713
0,6990	-0,1307	1,4314	-1,3129	1,7243	-2,2320
0,7782	-0,1758	1,4472	-1,3963	1,7324	-2,2543
0,8451	-0,2185	1,4624	-1,4538	1,7404	-2,2777
0,9031	-0,2680	1,4771	-1,5036	1,7482	-2,3026
0,9542	-0,3163	1,4914	-1,5788	1,7559	-2,3289
1,0000	-0,3716	1,5051	-1,6299	1,7634	-2,3869
1,0414	-0,4275	1,5185	-1,6522	1,7709	-2,4191
1,0792	-0,4761	1,5315	-1,7005	1,8129	-2,4538
1,1139	-0,5248	1,5441	-1,7406	1,8388	-2,4916
1,1461	-0,5802	1,5563	-1,8254	1,8451	-2,5330
1,1761	-0,6426	1,5682	-1,8518	1,8751	-2,5788
1,2041	-0,6973	1,5798	-1,9098	1,8865	-2,6299
1,2304	-0,7513	1,5911	-1,9648	1,8976	-2,6879
1,2553	-0,8104	1,6021	-1,9889	1,9243	-2,7549
1,2788	-0,8769	1,6128	-2,0015	1,9638	-2,9309
1,3010	-0,9442	1,6232	-2,0145	1,9823	-3,0559
1,3222	-1,0132	1,6335	-2,0559	2,0212	-3,2320
1,3424	-1,0751	1,6435	-2,0859	2,0792	-3,5330

Tabla N° 6: Método Llamada Recíproca. Tabla en Logaritmos

"Método Llamada Recíproca" - $\text{Log}(P(K))$ Vs $\text{Log}(K)$

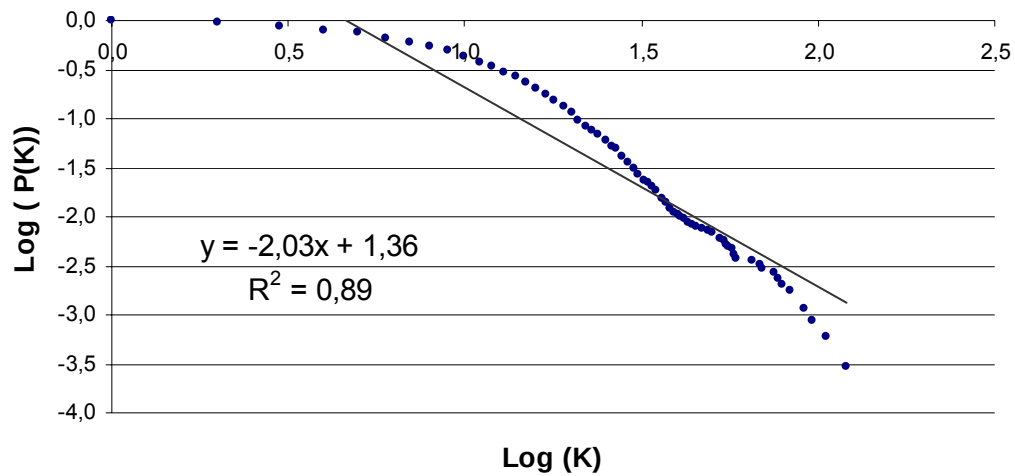


Figura N° 17: Método Llamada Recíproca. Ajuste de Datos

El R^2 es de 0.89 por lo cual el ajuste lineal se puede considerar adecuado. A su vez el valor del exponente de la red de escala libre β es de 2,0 ($\gamma = 3,0$) por lo cual también se encuentra dentro de los parámetros esperados de una red de escala libre.

Resultados

En la discusión de resultados resulta importante mencionar las diferencias que derivan de la consideración de los diferentes métodos para establecer qué condiciones representan una conexión con otra persona.

Utilizamos tres metodologías diferentes. El primer método (Método Base) consistió en considerar que dos personas se encuentran conectadas siempre y cuando haya ocurrido al menos un llamado entre estas dos personas en cualquiera de los dos sentidos. En el segundo método (Método Tiempo) se agregó la condición de que la llamada que se realizase fuese mayor a los 3 minutos (duración de una llamada promedio). El último método (Método Llamada Recíproca) consistió en considerar un enlace entre dos personas cuando estas se llamaban mutuamente al menos una vez.

El estudio de la conducta de los usuarios de teléfono y de sus perfiles muestra que:

- La duración de la llamada es un indicativo de la fuerza de la relación entre dos personas.
- A su vez, ocurre que en relaciones fuertes, las personas se llaman mutuamente.

En consecuencia, los métodos considerados más adecuados para representar relaciones fuertes entre dos personas resultan los métodos que consideren tanto la duración de la llamada como el hecho de que la llamada sea recíproca.

En el Método Tiempo se observa que la red social se ajusta con un R^2 de 0,93 a una red de escala libre con exponente $\lambda = 2,0$ ($\gamma = 3,0$). Mientras que con el Método Llamada Recíproca se observa que la red se ajusta con un R^2 de 0,89 a una red de escala libre con exponente $\lambda = 2,0$ ($\gamma = 3,0$).

Finalmente, es posible afirmar que efectivamente las comunicaciones a través de telefonía fija entre nodos que representan hogares en Monte Equis presentan una estructura de red de escala libre con un exponente $\lambda = 2,0$ (más precisamente dentro de un rango de $1,98 < \lambda < 2,03$). Esta observación servirá para definir la red social que se utilizará de base para el modelo de simulación.

5.3 Análisis Estático de la Red Social

Una vez representada la red social, se realiza el análisis estático de la misma. En este caso se realiza el análisis de los índices básicos de la red como ser, el

número de nodos, el grado nodal promedio y la identificación de los nodos mejor conectados.

Como vimos anteriormente los métodos más adecuados para la determinación de la red social son aquellos que consideren las relaciones más fuertes, las cuales pueden ser capturadas mediante el análisis de las llamadas recíprocas o considerando la duración de las llamadas.

Red Social de Monte Equis

Método	Nº Nodos	Nº Nodos c/grado nodal ≥ 1	Grado Nodal de la Red	Ajuste a Red de Escala Libre	Coeficiente de Red de Escala Libre (γ)
Llamada Recíproca	3.585	3.412	10,3	Si	3,03
Tiempo	3.585	3.342	7,6	Si	2,98

Tabla Nº 7: Descripción de la Red Social de Monte Equis

Monte Equis es un pueblo pequeño ubicado en Argentina. Posee en total aproximadamente 3.600 líneas de telefonía fija de tipo residencial. Al estimar el grado nodal de su red social mediante los dos métodos citados, se encuentra que el mismo varía entre 7.6 y 10.3. Esto implica que en promedio cada nodo se relaciona con aproximadamente otros 9 nodos. Se encontró que la red social de Monte Equis se ajusta a un tipo de red de escala libre con un coeficiente de red de aproximadamente $\gamma = 3,0$.

5.4 Teoría del Comportamiento

En esta instancia, se debe analizar el tipo de producto y definir que teoría se adapta mejor este. Para ello se utiliza como base de análisis la teoría de Jager y Janssen.

El producto de ADSL tiene un nivel de incertidumbre medio-alto. Esto implica que a diferencia de otros productos como una golosina o un shampoo en los cuales se tiene una idea acabada de lo que realmente se obtendrá, en el producto ADSL existe un cuestionamiento e incertidumbre mayor acerca de las cualidades del servicio (servicio técnico, problemas de facturación, calidad del servicio, etc.).

Generalmente, el nivel de incertidumbre es alto en productos nuevos, de tipo tecnológico y con un costo significativo, debido a que el conocimiento general que se posee del mismo es bajo y la satisfacción que brinda el producto en relación a su costo es incierta.

Es por esto que el proceso de adopción de ADSL, al igual que los de telefonía móvil (mayormente en planes de abono), está caracterizado por una previa investigación y acumulación de información por parte del potencial cliente, previa a la suscripción del servicio. En general, una persona que esta pensando en adoptar ADSL analiza las compañías que brindan este servicio en su zona y consulta con sus contactos que ya poseen el servicio su opinión sobre el mismo.

El proceso de compra típico se puede representar en los siguientes pasos:

En primer lugar, las campañas publicitarias, la misma difusión boca en boca y el contacto cotidiano con la tecnología que va ganando ubicuidad, provoca que los potenciales adoptadores reconozcan la necesidad (o deseo) y se dispara un interés por la búsqueda de información sobre dicho producto.

Posteriormente, la persona interesada obtiene información acerca de los productos de ADSL promocionados por cada una de las empresas y realiza un primer análisis. En esta etapa se consideran aspectos visibles de los productos como ser velocidad de conexión, precio mensual, precio de instalación, etc.

Antes de evaluar las diferentes alternativas, el potencial adoptador consulta a sus contactos que poseen el servicio de determinada compañía su experiencia con el mismo. De esta manera, el potencial adoptador busca conocer las características que no son fáciles de percibir. En esta etapa la persona busca obtener información fidedigna acerca del producto. Pregunta sobre las fallas del

servicio, los problemas de cobro que puede haber, la facilidad de contactar al servicio técnico, la existencia de costos “ocultos”, etc. Estas características u opiniones influyen enormemente en la decisión final de compra y la fuente prioritaria de información son los usuarios reales del servicio que no deben tener un vínculo de interés con la compañía que los provee.

Utilizando las categorías de Jager y Janssen, se define que el procesamiento cognitivo que mejor se ajusta a este producto es el de comparación social, dado que la decisión de compra requiere un razonamiento explícito y es determinado solicitando opiniones de sus contactos.

Modos de Procesamiento Cognitivo

	Determinado:	
Razonamiento	Individualmente	Socialmente
Explícito	Deliberación	Comparación Social
Automatizado	Hábito o Repetición (Ej. Caramelos, pasta de dientes, etc.)	Imitación

Figura N° 18: Selección del modelo más adecuado

Comparación Social (razonamiento explícito y determinado socialmente) Al igual que la imitación, las personas observan lo que otras personas están haciendo, pero al existir bajos niveles de satisfacción, la gente esta predispuesta a razonar conscientemente sobre la decisión de consumo. En este proceso de razonamiento los agentes intentan inferir información sobre el producto mediante el nivel de interés que tienen otros agentes por el producto.

Representación del Comportamiento en el Modelo de Simulación

Una vez definido el modo de procesamiento cognitivo, se debe analizar la mejor manera de representar este comportamiento en el modelo de simulación.

El comportamiento se definirá de la siguiente manera:

- En primer lugar, se representa la red social de potenciales clientes. La misma esta conformada por nodos que representan las diferentes viviendas que se consideran parte del target group y por los enlaces que los unen.
- Al igual que en el modelo de Bass, para comenzar el proceso de difusión se “contagia” (se transforma en adoptador) a un porcentaje definido de la red social. Estas personas adoptan el producto debido a la influencia de la publicidad en medios masivos o de la publicidad dirigida.
- Posteriormente, comienza el proceso de difusión boca en boca. Se definen períodos de tiempo dentro de los cuales cada uno de los adoptadores, se contacta con los nodos adyacentes. A su vez, se define una efectividad de contagio (nivel de influencia) que determina la probabilidad con la que una persona contagia a otra persona contactada.

En la próxima sección se observa en detalle como se modela el comportamiento del consumidor en la simulación.

5.5 Diseño de Experimentos

Red Social de Base

Los experimentos se realizan sobre una red social que se modela en base a los análisis previamente realizados. El programa Anylogic posee un módulo dedicado a la simulación basada en agentes dentro de la cual se puede establecer el formato de red, la cantidad de nodos a modelar y los parámetros de la red.

Variable	Valor
Tipo de Red	Red de Escala Libre
Nº de Nodos	1.000
“Scale M” ¹²	5
Grado Nodal de la Red	10
Nº de Adoptadores iniciales	10

Tabla Nº 8: Red Social de Base

Cabe recordar que los nodos de la red representan líneas de telefonía fija domésticas. Es decir que cada nodo corresponderá a una vivienda con diferentes miembros familiares dentro de ella. El hecho de que los nodos representen viviendas es conveniente dado que para la difusión de servicios como Internet o telefonía fija, los clientes se identifican por vivienda.

Otro tema relevante a destacar es que la red que se obtuvo del análisis de los llamados telefónicos en Monte Equis representa una red de relaciones que refleja el grado de contacto social general y no sólo telefónico. Como se mencionó anteriormente las relaciones a través de contactos telefónicos son un buen indicador del nivel de contacto social de la persona o vivienda. Es decir que es de esperar que una vivienda que posea una gran cantidad de contactos telefónicos frecuentes posea además un alto nivel de contacto con otras personas a través de otros medios de comunicación incluidos los contactos personales (Ver sección “Método Basado En Datos Empíricos De Telefonía Fija”).

¹³ “Scale M” define el número mínimo y más frecuente de nodos adyacentes.

En la figura siguiente se puede observar el modelo de la red social que se utilizará como base para las simulaciones:

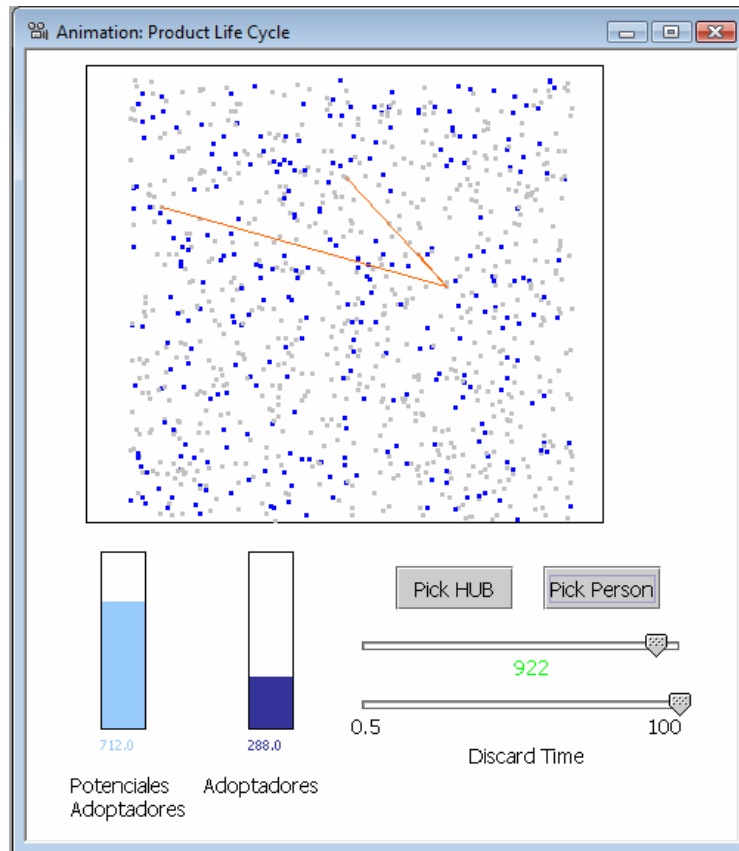


Figura N° 19: Panel de control de la simulación

En el cuadro enmarcado se observan los diferentes nodos, que representan las líneas fijas de Monte Equis. El color del nodo indica si es un potencial adoptador (nodos que aún no adoptaron) o si el nodo ha adoptado el producto. Para el primer caso el nodo es color gris, mientras que los adoptadores se representan con un nodo color azul.

Los enlaces entre nodos se indican con líneas color naranja. Si bien todos los nodos poseen al menos un enlace, estos no se representan simultáneamente debido a que provocaría una saturación en el gráfico. Sin embargo, es posible visualizar los enlaces de cada uno de los nodos de manera individual utilizando el botón "Pick Person" (Por ejemplo, la línea naranja que se observa conectando diferentes nodos muestra las conexiones de un nodo que en este caso particular posee sólo dos contactos).

Para asegurarnos que la red generada por Anylogic se ajusta a la red que deseamos representar se utiliza el siguiente código que corre al iniciarse la simulación:

```
for ( int i = 0; i < people.size(); i++ ) {

    Person p = people.item (i);
    int s = p.agentBase.getContacts().size();

    dataset1.add(i,s); }
```

Estas líneas de código almacenan el número de contactos de cada uno de los nodos en una tabla que luego puede ser exportada al MS Excel para su análisis:

K	Frecuencia	K	Frecuencia
5	286	32	1
6	165	33	3
7	137	34	2
8	75	35	1
9	68	36	2
10	37	42	1
11	44	43	2
12	19	44	1
13	16	45	1
14	19	46	2
15	14	47	1
16	14	53	1
17	16	54	1
18	7	56	1
19	9	57	1
20	9	58	1
21	5	60	1
22	5	61	1
23	2	69	1
24	2	73	1
25	6	76	1
26	4	82	1
27	1	96	1
28	3	105	1
29	1	111	1
30	3	123	1
31	1		

Tabla N° 9: Modelo de la red social simulada. Tabla de Frecuencias

Posteriormente, se muestra la tabla en un gráfico de frecuencias para observar la tendencia:

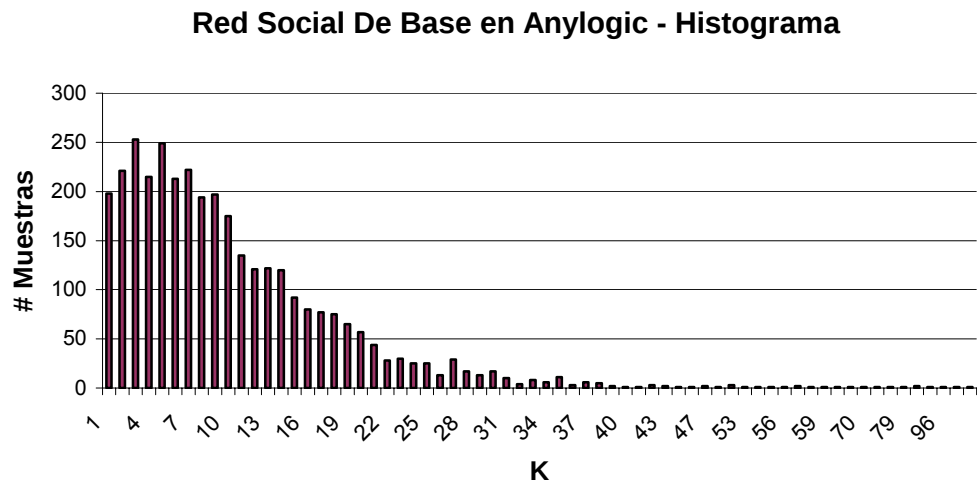


Figura N° 20: Red social de base simulada. Histograma

Luego se realiza el ajuste de los datos para observar su ajuste a una red de escala libre:

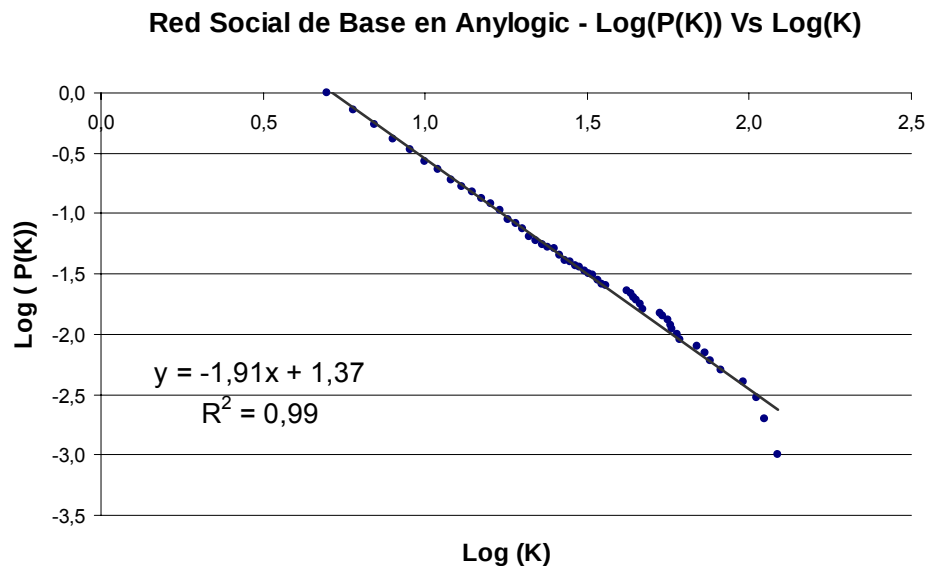


Figura N° 21: Red social de base simulada. Ajuste de datos

El resultado es un ajuste con un R^2 de 0.99 y un exponente de la red de escala libre $\lambda = 1,91$ ($\gamma = 2.91$). En síntesis, se puede afirmar que la red social de base modelada en el programa Anylogic se asemeja a la red social de Monte Equis en estudio.

Modelos de Simulación

Utilizando la teoría de comportamiento definida, se hace una serie de ensayos para determinar la estrategia óptima que potencie el efecto boca en boca en la difusión.

Para ello, se comparará una estrategia de difusión aleatoria (Estrategia Aleatoria), con otra que se base en comenzar la difusión a través de los hubs de la red social (Estrategia Hub). Para ambas estrategias se utiliza una red social de base definida en el apartado anterior.

Sobre el modelo base se corren las diferentes estrategias y los correspondientes análisis de sensibilidad.

En primer lugar, manteniendo el modelo base se estudia la evolución de adoptadores bajo la Estrategia Aleatoria variando el grado nodal de la red (10, 20, 30, 40 y 50). Este estudio implica 5 casos diferentes, realizando para cada uno de ellos 50 corridas. Esto significa que se realizan un total de 250 corridas en la simulación.

Posteriormente, nuevamente sobre el modelo base, se corre la Estrategia Aleatoria variando el número de adoptadores iniciales (10, 20, 30, 40 y 50). Este estudio presenta otros 5 casos particulares, realizando para cada uno de ellos 50 corridas. Es decir, un total de 250 corridas adicionales en la simulación.

Finalmente, se extraen y analizan los resultados para verificar las potenciales ventajas que se obtienen al considerar el contexto de la red social.

Estrategia De Difusión Aleatoria

En este modelo se refleja la estrategia actual de la empresa. En la misma se generan listados con potenciales clientes en base a diferentes atributos (NSE, productos adoptados, edad, etc.) y luego se realiza publicidad dirigida a los mismos sin considerar su grado de conexión.

Se utiliza en la simulación la red social de base (Ver “Red Social de Base”) que fue generada anteriormente considerando las características del pueblo bajo estudio.

Estrategia De Difusión Hub

En este modelo se refleja una estrategia de difusión hub en la cual una vez seleccionado el grupo de potenciales clientes en base a diversos atributos, la publicidad dirigida se realiza comenzando por los hubs (individuos mejor conectados) del grupo.

Este modelo posee las mismas características de red y parámetros de difusión que el modelo anterior con la única diferencia que en este caso la elección de los adoptadores iniciales se realiza en función de su grado de conexión, eligiendo a las 10 viviendas mejor conectadas.

Para la identificación de los hubs se genera una tabla con la cantidad de conexiones que posee cada una de las viviendas y luego se eligen las viviendas mejor conectadas.

La siguiente tabla muestra las viviendas mejor conectadas y el número de contactos que cada uno posee:

Código de Entidad	Grado Nodal
1	123
7	111
10	105
9	96
8	82
5	76
11	73
41	69
25	61
15	60

Tabla N° 10: Identificación de Hubs

Dinámica del Modelo

Al comienzo de la simulación, todas las entidades serán potenciales clientes (es decir todavía no han adoptado el producto) identificados con un color gris claro. Al convertirse en adoptadores su color cambiará a azul.

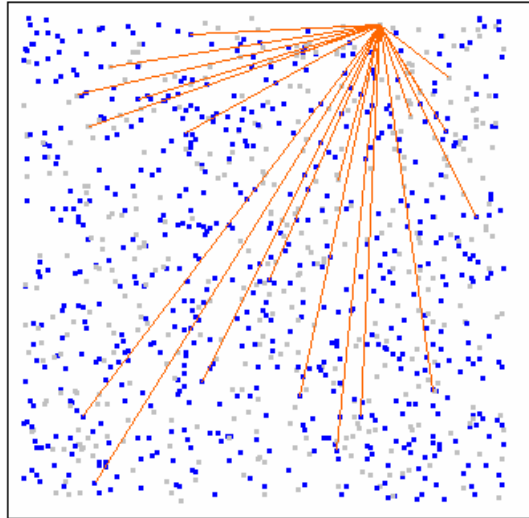


Figura N° 22: Red social de base simulada. Enlaces de un nodo

El objetivo de la simulación es observar el efecto boca en boca, pero para ello debe haber entidades que ya hayan adoptado y que comiencen el proceso de difusión. Para ello se define un disparador con el objetivo que al comienzo del modelo 10 nodos se conviertan en adoptadores. Dichas nodos serán seleccionadas aleatoriamente o en función de su grado nodal dependiendo de la estrategia que sea planteada.

Posteriormente, los adoptadores difundirán el producto mediante un proceso boca en boca. Las variables que definen la velocidad de difusión son la tasa de contacto y la probabilidad de adopción una vez que una persona es contactada.

En cada período de tiempo definido los adoptadores toman contacto con sus relaciones y difunden su producto.

Diagrama de Estados de Nodos

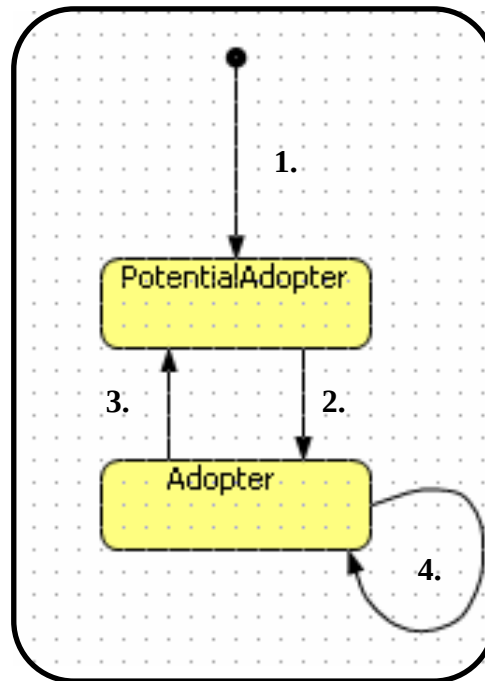


Figura Nº 23: Diagrama de Estados de nodos

La Figura Nº 24 muestra los diferentes estados por los que cada entidad puede pasar:

1. Al comenzar la simulación todas las entidades pasan al estado de adoptador potencial.
2. Se define un disparador para que diez nodos (elegidos aleatoriamente o en función de su grado nodal) se conviertan en adoptadores y comiencen la difusión boca en boca.
3. Este lazo de transición no se utilizará, pero permitiría modelar los casos en que un adoptador deja de serlo debido a la duración finita del producto o al cambio en la elección de marca de producto por parte del cliente.
4. Permite modelar el efecto en boca. Dentro de esta flecha de cambio de estado se definen las circunstancias bajo las cuales se desea que se produzca un contagio por efecto boca en boca:

Cada un período de tiempo definido por una distribución exponencial con parámetro λ , se ejecuta unas líneas de código. Estas líneas de código ordenan que las entidades tomen contacto con su círculo de relaciones (nodos adyacentes) y disparen en estos la compra enviando la señal “buy”. Sin embargo, no todas las entidades que son contactadas se convierten en adoptadores sino que sólo lo hacen con una probabilidad determinada.

```
for ( int i = 0; i < agentBase.getContacts().size(); i++ ) {

    Person buddy = (Person)agentBase.getContacts().get( i );

    buddy.adoption.fireEvent( "buy" ); }
```

Parametrización del Modelo

Para estimar el parámetro λ se analiza la evolución de adoptadores de ADSL en Monte Equis y en otros dos pueblos similares con el objetivo de estudiar la curva de adopción del producto.

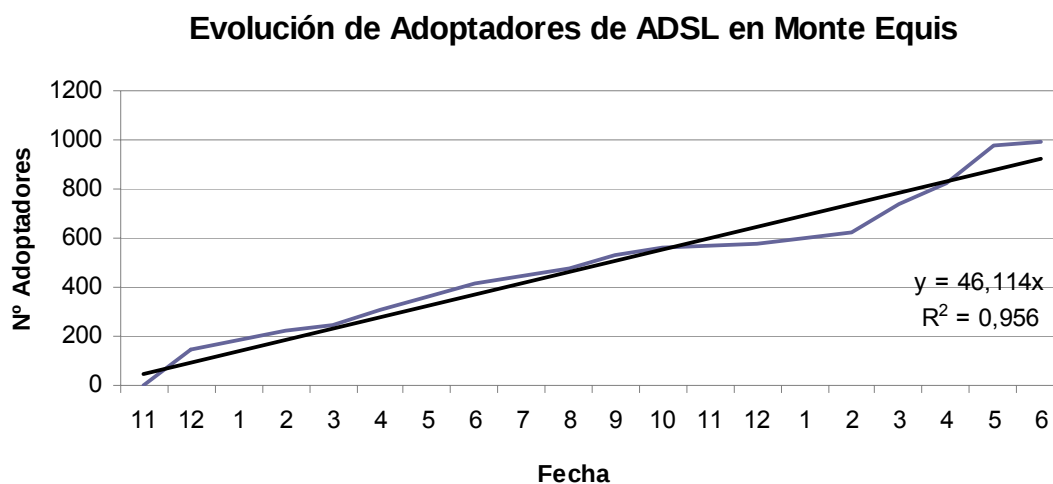


Figura Nº 24: Evolución de Adoptadores ADSL-Monte Equis

En Monte Equis, en un período de veinte meses hubo aproximadamente 1.000 adoptadores de ADSL. Si bien la curva de adopción total presentaría la forma de Bass, se disponen de datos de la etapa inicial de la curva de adopción. Esta

sección puede ser ajustada linealmente con un R^2 de 0.95. La cantidad de adopciones de ADSL por mes es de aproximadamente 46 durante este período.

En otros pueblos similares (Pehuajó y Rada Tilly) se observa que la cantidad de adopciones es de aproximadamente 25 por mes. Cabe destacar que en estos casos, se utilizan datos de otra sección de la curva de adopción y que obviamente debido a las características propias de cada pueblo la cantidad de adopciones por mes varía.

Finalmente, se corre el modelo con diferentes parámetros de λ y se estima el valor que mejor refleja el comportamiento observado en Monte Equis. El resultado obtenido es un parámetro $\lambda = 0.23$.

5.6 Resultados

Comparación de Estrategias de Difusión

Utilizando el modelo de red base, se corren 50 simulaciones de cada una de las estrategias de difusión. Para cada uno de los modelos se analiza la evolución de los adoptadores en función del tiempo obteniendo el siguiente gráfico:

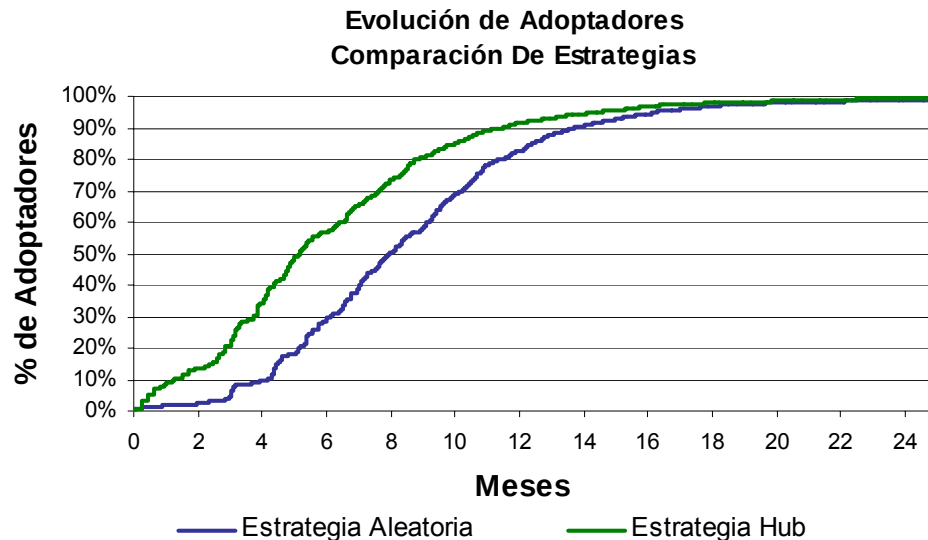


Figura Nº 25: Evolución de Adoptadores

A su vez, se obtiene una tabla con los tiempos que toma alcanzar diferentes niveles de penetración en el mercado objetivo en las diferentes estrategias:

% Adoptadores	Estrategia Aleatoria	Estrategia Hub	Delta Tiempo
	Meses	Meses	%
10%	4,2	1,2	70%
20%	5,2	2,8	46%
30%	6,1	3,8	39%
40%	7,1	4,3	39%
50%	8,0	5,1	36%
75%	10,7	8,3	22%
90%	13,7	11,4	16%
95%	16,2	14,3	12%

Tabla Nº 11: Tabla de Tiempos. Comparación de Estrategias

El análisis del gráfico se puede dividir en tres etapas, de acuerdo al comportamiento de la curva:

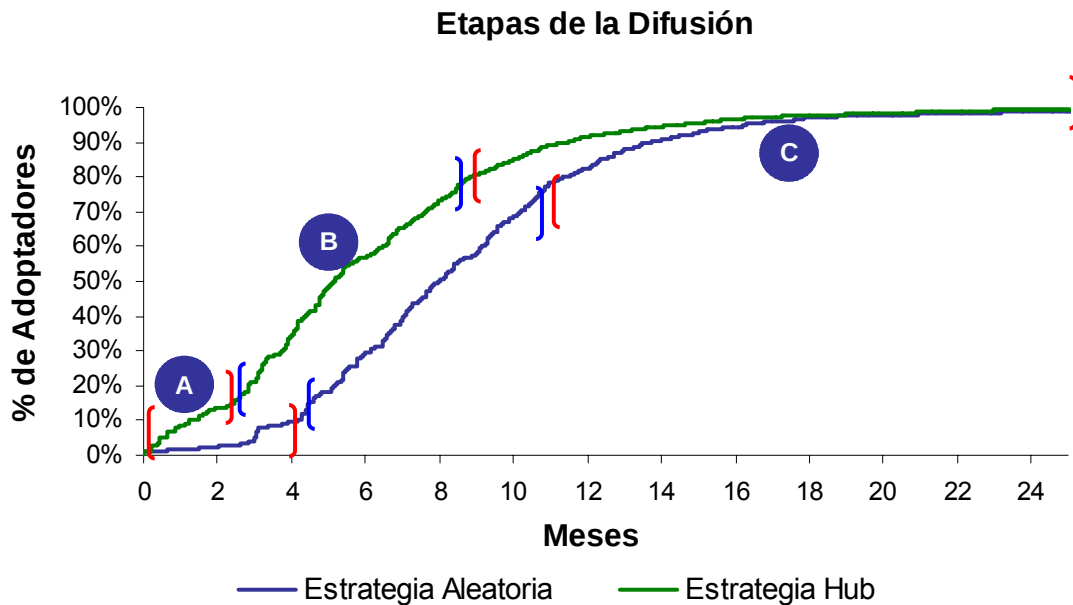


Figura N° 26: Etapas de Análisis de la Difusión

En la etapa A se observa que al cabo de dos meses la estrategia Hub alcanza un 10% del market share mientras que la estrategia aleatoria apenas alcanza el 1%. Este hecho se debe a que al iniciar la difusión por los hubs el contagio del producto en la etapa inicial es mayor que utilizando la estrategia aleatoria.

La etapa B presenta una curva bastante similar para ambas estrategias pero en tiempos desfasados. Al observar el final de la etapa A, es posible notar que la pendiente de la curva correspondiente a la estrategia aleatoria va aumentando debido a que el producto ha comenzado a difundirse por una parte de la población que incluye alguno de los hubs. En la etapa B, las pendientes de las curvas de las dos estrategias son similares dado que en esa instancia los nodos con un grado nodal medio o alto ya han adoptado el producto y contagian a un mismo ritmo.

Por otro lado, en la etapa C la mayoría de los nodos ya son adoptadores. En consecuencia, la cantidad de nuevas adopciones decrece considerablemente. Esto implica que la velocidad de difusión y la pendiente de la curva disminuyen. En contraposición a lo que ocurre con los hubs que son contagiados en las etapas iniciales, los nodos menos conectados o aislados son los últimos en adoptar el producto.

Comparando las dos estrategias de difusión se observa que la estrategia hub se desempeña de mejor manera que la estrategia aleatoria. A su vez, se observa que la diferencia entre los tiempos en que se alcanzan diferentes porcentajes de penetración de mercado para las dos estrategias, va disminuyendo (de 70% a 12%) a medida que aumenta la penetración de mercado (Ver *Tabla N° 11*).

Para explicar esto, en primer lugar, es preciso identificar el rol de las viviendas hubs y su comportamiento. Estas viviendas son las más conectadas siendo más probable que se conviertan en adoptadores en una fase temprana de la difusión debido a que poseen una gran variedad de contactos que al contagiarse los influenciarán para convertirse en adoptadores. Es por esto que, sin importar la estrategia empleada, los hubs tenderán a contagiarse en las fases tempranas (Etapa A o B) de la difusión.

A su vez, se observa otro comportamiento que está relacionado con la cantidad de viviendas que se han convertido en adoptadores. Si consideráramos una cantidad de viviendas infinita, a medida que hay más adoptadores la velocidad de difusión (N° contagios/tiempo) aumentará. Sin embargo, lo que ocurre en poblaciones finitas es que al saturarse de adoptadores la velocidad de difusión disminuye dado que hay menos viviendas para contagiar.

La combinación de estos dos comportamientos permite comprender la forma de la curva de difusión que alcanza su pendiente máxima cuando una mayoría de hubs han adoptado y la cantidad de adoptadores es grande pero aún quedan suficientes individuos para contagiar.

En la última etapa de la difusión se observa que la velocidad de difusión se hace mínima. Esto se debe a que la cantidad de viviendas que aún no han adoptado es mínima y a su vez se trata de nodos poco conectados.

Sensibilidad sobre el Grado Nodal de la Red

Con el objetivo de conocer el impacto de la variación del grado nodal de la red en la difusión del producto se realiza un análisis de sensibilidad. En este análisis, realizado bajo una estrategia aleatoria, se realizan 50 corridas para cada caso de grado nodal de red diferente (10, 12, 14, 16, 18 y 20). Los resultados fueron:

	Grado Nodal De La Red					
	10	12	14	16	18	20
% Adoptadores	Meses	Meses	Meses	Meses	Meses	Meses
10%	4,2	4,2	3,9	3,3	3,8	2,9
20%	5,2	5,1	4,5	4,4	4,3	3,5
30%	6,1	5,9	5,4	5,1	5,0	4,3
40%	7,1	6,6	6,3	6,0	5,7	5,0
50%	8,0	7,2	7,1	6,7	6,2	5,4
75%	10,7	9,7	9,4	8,7	7,9	7,0
90%	13,7	12,8	12,1	10,8	10,0	8,9
95%	16,2	14,4	13,8	12,0	11,7	10,5

Tabla Nº 12: Tabla de tiempos. Comparación por Grado Nodal

Al graficar la evolución de adoptadores en función del tiempo se obtiene el siguiente gráfico:

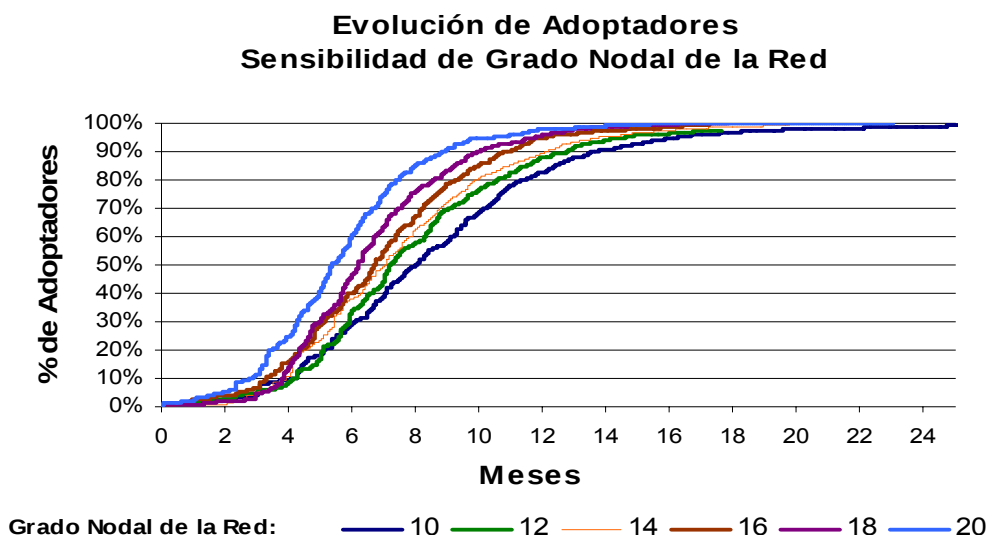


Figura Nº 27: Sensibilidad Sobre el Grado Nodal de la Red

Los resultados permiten comprobar la idea intuitiva que a mayor grado nodal de la red la difusión del producto se realiza más rápidamente. Esto se debe a que un mayor grado nodal implica que los nodos de la red social poseen más conexiones en promedio. En consecuencia, la difusión boca en boca posee una mayor efectividad.

Considerando una corrida cualquiera, al seleccionar los 10 adoptadores iniciales de manera aleatoria puede ocurrir que a pesar de que el grado nodal promedio de la red sea por ejemplo de 20, el grado nodal de los 10 adoptadores elegidos sea menor o mayor que el grado nodal de la red. Sin embargo, al realizar una gran cantidad de corridas el grado nodal promedio de los 10 adoptadores elegidos a través de las diferentes corridas tiende al grado nodal de la red.

En las columnas de grado nodal de red 16 y 18 de la Tabla N° 11 se puede observar que el 10% de penetración de mercado se alcanza primero en la red de Grado Nodal 16. Esto se puede deber a la aleatoriedad de la selección de los 10 adoptadores iniciales que pudo producir que los adoptadores iniciales del ensayo sobre la red de Grado Nodal 16 estuvieran mejor conectados que los de la red de Grado Nodal 18. Sin embargo, aún en este caso se observa que las penetraciones de mercado posteriores evolucionan más rápidamente en la red de Grado Nodal 18 que en la red de Grado Nodal 16.

Finalmente, resulta interesante observar que la pendiente de la Etapa B se incrementa en redes de grado nodal más alto. Sin embargo, al comparar estrategias bajo un mismo grado nodal de red (Ver *Figura N° 25*), se observa que la pendiente en la etapa B es similar para las ambas estrategias.

Sensibilidad sobre el Número de Adoptadores Iniciales

Con el objetivo de conocer el impacto de la variación del número de adoptadores iniciales en la difusión del producto se realiza un análisis de sensibilidad. En este análisis, realizado bajo una estrategia aleatoria y utilizando la Red Social Base (Ver *Tabla N° 8: Red Social De Base*) se realizan 50 corridas para diferentes casos en los que el número de adoptadores iniciales varía (10, 20, 30, 40 y 50).

Los resultados fueron:

%	N° Adoptadores Iniciales				
	10	20	30	40	50
Adoptadores	Meses	Meses	Meses	Meses	Meses
10%	4,2	3,4	2,1	1,4	0,7
20%	5,2	4,4	2,9	2,4	1,7
30%	6,1	5,4	4,0	3,3	2,6
40%	7,1	6,3	4,9	4,4	3,5
50%	8,0	7,0	6,0	5,4	4,5
75%	10,7	10,0	9,4	8,9	8,2
90%	13,7	13,7	12,5	13,0	11,9
95%	16,2	16,2	14,8	15,6	14,5

Tabla N° 13: Tabla de tiempos. Comparación por N° Adoptadores Iniciales

Al graficar la evolución de adoptadores en función del tiempo se obtiene el siguiente gráfico:

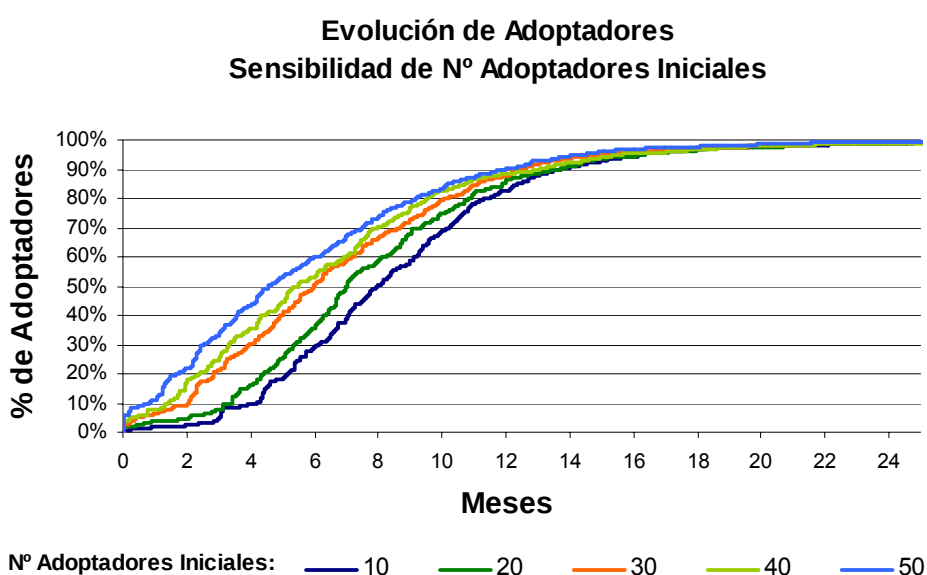


Figura N° 28: Sensibilidad Sobre el N° Adoptadores Iniciales

Se observa que a medida que se incrementa el número de adoptadores iniciales la penetración de mercado en las etapas tempranas se alcanza más rápido. Sin embargo, es notable el hecho que a partir de un grado de penetración de mercado de aproximadamente 65%, la brecha entre los diferentes casos disminuye.

5.7 Conclusiones

En el estudio del caso práctico se analiza la red social del pueblo y su dinámica. El mismo permite avanzar en la comprensión del efecto boca en boca y ensayar teorías que apalanquen dicho efecto para mejorar la difusión del producto ADSL en el pueblo estudiado.

En primer lugar, se ha identificado que la red social del pueblo bajo estudio posee una estructura de red de escala libre. Este descubrimiento es de gran relevancia dado que abre las puertas a que se extiendan los análisis a otras regiones, brindando una base empírica y metodológica para estudiar las redes sociales. En general, uno de los principales retos es contar con los datos para definir la estructura social del mercado en el cual deseamos ingresar. Sin embargo, la utilización de las llamadas a través de telefonía fija para el establecimiento de la estructura de red social es una solución simple y de poco costo para una compañía telefónica.

El resultado de los experimentos realizados permite confirmar que, bajo los supuestos utilizados, la estrategia hub es superior que la estrategia aleatoria dado que permite penetrar en la base de potenciales clientes de una manera más efectiva y en una menor cantidad de tiempo.

Si bien en las etapas finales de la difusión la brecha de tiempos entre ambas estrategias disminuye, en la etapa inicial la diferencia de tiempos es significativa.

La ventaja en la etapa inicial de la difusión resulta de gran importancia sobre todo cuando nos encontramos en ambientes de competencia. En estos casos la expansión más veloz de la marca propia con respecto a la de la competencia es vital dado que brinda la posibilidad de posicionar el producto con antelación y ganar un mayor porcentaje de participación del mercado. A su vez, utilizando la estrategia hub la empresa se asegura el contagio de los nodos más influyentes al inicio de la difusión y de esta manera los mismos colaborarán en la difusión del producto.

De manera general, se identifican diferentes variables que influyen en la curva de adopción. Algunas de estas variables son más bien fijas (variables estructurales) mientras que otras pueden ser modificadas (variables dinámicas):

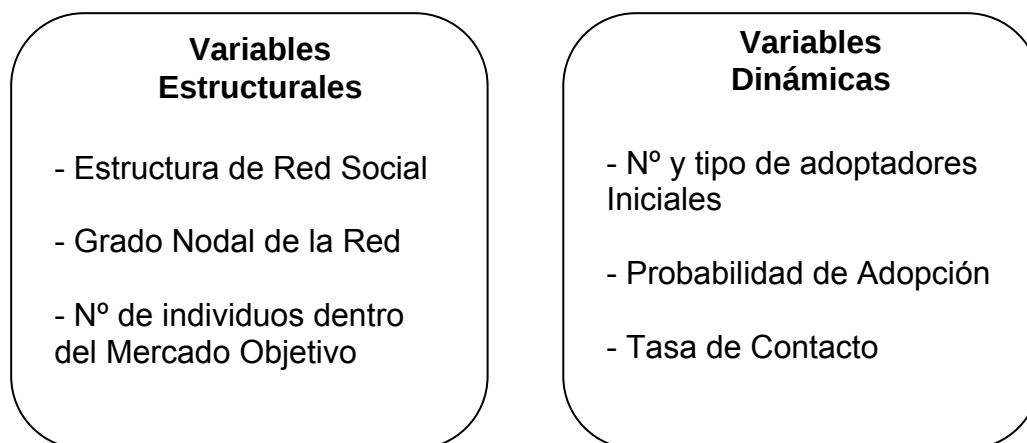


Figura N° 29: Variables del Sistema

Las variables estructurales (Estructura de la Red Social, Grado Nodal de la Red y N° individuos que conforman el mercado objetivo) representan el contexto. Es por esto que para conocer la situación en la que se difundirá el producto es necesario el estudio de estas variables. En general, estas variables son estructurales por lo cual su modificación es menos factible.

Por otro lado, se presentan las variables dinámicas (N° y tipo de adoptadores iniciales, probabilidad de adopción y tasa de contacto). Las mismas son más flexibles y pueden ser influenciadas en buena medida por la estrategia de difusión que se emprenda. Por lo cual al diseñar cualquier estrategia de difusión es necesario hacer foco en el apalancamiento de estas variables para lograr un efecto boca en boca que sea óptimo.

A continuación, se mencionan algunos factores que pueden influenciar las variables dinámicas:

- **Número de Adoptadores Iniciales:** En un ambiente de competencia resultará clave contar con una masa crítica de adoptadores iniciales que produzca un boca en boca positivo del producto. Para esto, será necesario combinar una campaña en medios masivos que aumente el conocimiento de marca, lanzar promociones que alienten el consumo del producto y establecer campañas de telemarketing dirigidas a personas con alta probabilidad de adopción.
- **Tipo de Adoptadores Iniciales:** Luego del estudio realizado se puede afirmar que la red social del pueblo bajo estudio presenta una estructura de red de escala libre. En esta red existen una minoría de nodos bien conectados, denominados hubs, a los cuales es conveniente apuntar.

Un medio adecuado para asegurarnos la calidad de difusión del adoptador es el telemarketing. De esta manera, es posible realizar los llamados a potenciales clientes con alto grado nodal. A su vez las campañas de recomendación podrían ayudar a captar clientes dispuestos a recomendar el producto a cambio de abonos mensuales gratis u otros beneficios.

- **Probabilidad de Adopción:** Las propiedades del producto, la adaptación de las mismas al potencial consumidor y el timing con el cual el mensaje es enviado, son claves en la probabilidad de adopción.
- **Tasa de Transmisión del Mensaje:** La frecuencia con la cual una persona recomienda un producto o transmite determinado mensaje esta ligada a la importancia o novedad del mismo. En general, una persona recomienda un producto cuando alguna de sus características es altamente satisfactoria o cuando existe algo sorpresivo o nuevo en este. Las campañas de recomendación pueden ser útiles para aumentar la tasa de transmisión del mensaje. De alguna manera, estas campañas funcionan con un concepto similar al del marketing directo, permitiendo que personas convencionales se conviertan en “vendedores” del producto.
La tasa de transmisión del mensaje también depende de las herramientas brindadas al usuario para que recomiende el producto. Un ejemplo de este hecho fue Gmail. Además de que en sus comienzos la suscripción al servicio era sólo a través de una invitación (creando una sensación de exclusividad), la facilidad con la que se podía y aún puede invitar a un amigo a suscribirse al servicio ha influido en la gran difusión del servicio.

6. CONCLUSIONES GENERALES

El presente trabajo propone una metodología integradora para el estudio de difusión de productos mediante un enfoque en redes sociales y simulación basada en agentes. Dicha metodología práctica, compuesta por seis etapas fundamentales (Ver “Metodología Propuesta”), integra diferentes teorías y herramientas otorgando un marco general que permite la consideración del impacto de las redes sociales en la difusión de productos.

El trabajo permite cubrir el problema planteado (Ver “Definición del Problema”) brindando una metodología de base para el estudio de problemas de difusión de productos en redes sociales. A su vez, al integrar y clarificar las etapas fundamentales del análisis, dicho trabajo aporta en la disminución de la brecha existente entre la teoría y la práctica. Como muestra de ello se analiza un caso real en una compañía de telecomunicaciones que busca desarrollar nuevas estrategias de difusión con el objetivo de mejorar la difusión del producto ADSL en un pueblo de Argentina.

Respecto al caso práctico, se han presentado diferentes desafíos y soluciones. En primer lugar, al realizar la representación de la red social del pueblo bajo estudio se utiliza un método híbrido que no requiere de una excesiva cantidad de datos pero que posee una alta fidelidad a la red social real. Dicho método consiste en la utilización de datos de llamadas de telefonía fija del pueblo para fijar la estructura de la red, y de un posterior ajuste a un tipo de red teórica. En segundo lugar, se ha comprobado que la red social del pueblo estudiado se ajusta a una estructura de red de escala libre. Este hecho resulta de gran relevancia dado que brinda nueva información acerca de la manera en que se organiza la sociedad, ofrece un modelo de base para realizar simulaciones o experimentos y abre las puertas para continuar el análisis con otros pueblos y ciudades de la Argentina. Finalmente, se ha demostrado que la estrategia hub aventaja a la estrategia aleatoria (particularmente en las etapas iniciales de la difusión) y se han definido las variables dinámicas del sistema sobre las cuales es imprescindible actuar al diseñar una estrategia que potencie el efecto boca en boca. A su vez, se muestra la sensibilidad de la difusión frente al cambio del número de adoptadores iniciales y al cambio del grado nodal de la red.

El alcance del trabajo está definido por el tipo de difusión estudiado, que se refiere a la difusión de productos, por lo cual el caso práctico elegido y las conclusiones también se realizaron siguiendo esa premisa. A su vez, al modelar el proceso de difusión ha sido necesario respetar el criterio de

simplicidad y los límites del trabajo, por lo cual se incluyen en el análisis sólo las variables más relevantes que afectan el fenómeno que se desea estudiar. De esta manera, el lector encontrará que factores como la influencia de la competencia o el efecto del boca en boca negativo no han sido estudiados porque exceden los límites del presente trabajo.

La metodología propuesta no pretende reemplazar sino complementar las actuales prácticas de difusión de productos. De esta manera, además de aplicar las técnicas de minería de datos y otros métodos actuales se propone la utilización de esta nueva herramienta con el objeto de generar nuevos experimentos y diseñar nuevas estrategias de difusión que contemplen la red social.

Como reflexión final, se considera que dicho trabajo de investigación representa un primer avance en aunar diferentes teorías con el objetivo de diseñar una metodología práctica para el estudio de difusión de productos y sirve como base para la realización de posteriores análisis (Ver “Futuras Líneas de Investigación”).

6.1 Futuras Líneas de Investigación

Al ser un tema de reciente difusión y considerando que las herramientas están en constante desarrollo, surgen muchos temas sobre los cuales parecería ser interesante realizar una ampliación de la investigación. Por ejemplo podríamos citar:

1) El análisis de la red social a partir de otras fuentes de información como detalles de llamadas a través de telefonía móvil u otros medios de comunicación actuales como el correo electrónico o la mensajería instantánea. Realizar comparaciones entre los comportamientos y características de las redes sociales detectadas.

2) La introducción del factor competencia en estrategias de difusión de productos y su impacto. Análisis de las diferentes etapas del ciclo de vida del producto y el análisis de estrategias de difusión diferenciales para cada una de estas.

3) Las redes sociales virtuales y el fenómeno Web 2.0. En lugar de intentar descubrir las redes sociales subyacentes algunos sitios otorgan las herramientas y reglas para establecer redes sociales virtuales con diferentes

fines. Por ejemplo Linked In otorga la posibilidad de crear una red social con objetivos profesionales.

4) En algunas compañías telefónicas se han comenzado a realizar promociones de recomendación. El cruce de esta información con el presente estudio puede dar luz a nuevos descubrimientos.

7. BIBLIOGRAFÍA

- Jager, W. 2000. Modelling Consumer Behaviour. Centre for Environmental and Traffic Psychology. University of Groningen, Groningen. Doctoral thesis.
- Jager Wander, Janssen Marco. 2003. Simulating Market Dynamics: Interactions between Consumer Psychology and Social Networks. Artificial Life, Volumen 9. Págs 346-356.
- Zbigniew Smoreda, Frank Thomas. 2001. Social Networks and Residential ICT Adoption and Use. France Telecom R&D. Paper.
- Philip Kotler. 2001. Dirección de Marketing. 792 Páginas. Editorial Prentice Hall. ISBN 968-444-422-2.
- Robert Axelrod. 1997. The Complexity of Cooperation. 279 Páginas. Editorial Princeton University Press. ISBN 069-101-567-8.
- Mitchell Resnick. 1994. Turtles, Termites and Traffic Jams. 222 Páginas. Editorial The Mit Press. ISBN 847-432-834-9.
- Gary M. Weiss. 1999. Data Mining in Telecommunications. Department of Computer and Information Science. Fordham University.
- Bass, Frank M. 1969. A New Product Growth Model for Consumer Durables, Management Science. Paper.
- Rogers, M. Everett. 1995. Diffusion of Innovations. Editorial Free Press.
- Rosen, Emanuel. 2000. The Anatomy of Buzz. New York, Doubleday.