



UNIVERSIDAD DE LAMBAYEQUE

FACULTAD DE CIENCIAS DE INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

**EVALUACIÓN DE LA CAPACIDAD DE LAS LÍNEAS DE
COMUNICACIÓN APLICANDO LA TEORÍA DE LA INFORMACIÓN
DE SHANNON EN LA RED DE DATOS DE LA UNIVERSIDAD DE
LAMBAYEQUE**

**PRESENTADA PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS**

AUTOR:

Bach. Sánchez Carrasco Edward Félix

ASESOR:

Mg. Carlos Antonio Rojas Ortiz

Líneas de Investigación:

Gestión de Infraestructura de TI

Chiclayo – Perú

2019

FIRMA DEL ASESOR Y JURADO DE TESIS

Mg. Carlos Antonio Rojas Ortiz
ASESOR

Mg. Enrique Santos Nauca Torres
PRESIDENTE

Dr. Juan Luis Rodríguez Vega
SECRETARIO

Mg. Carlos Antonio Rojas Ortiz
VOCAL

DEDICATORIA

A mi esposa Kely compañera y sobre todo amiga y a mi hijo Manuel Fermín que fueron mi motor y apoyo para culminar esta carrera profesional, quienes dieron sentido a mi vida y siguen apoyándome incondicionalmente.

AGRADECIMIENTO

Agradezco a Dios, por haberme permitido llegar hasta este momento y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A mi esposa e hijo por su aliento y ayuda diaria para cumplir esta realización de esta carrera profesional.

A mis padres Fermín y Victoria, Manuel y Anita mis queridos suegros, a mis hermanos y cuñados, por sus buenos deseos y consejos que sin pedir nada a cambio, me brindaron su apoyo incondicional.

Nuestro agradecimiento a todos mis amigos que me ayudaron con sus conocimientos y apoyo moral para culminar mi carrera profesional.

A mi asesor de tesis el Ingeniero Carlos Rojas Ortiz por la orientación y ayuda que nos brindó para realizar esta tesis, su paciencia, amistad y consejos que nos ayudaron a entender y mejorar el tema expuesto en esta tesis.

Agradecer de forma especial a nuestros colaboradores que nos brindaron su apoyo en la realización de este trabajo.

CONTENIDO

DEDICATORIA	3
AGRADECIMIENTO	4
CONTENIDO	5
ÍNDICE DE FIGURAS.....	7
RESUMEN.....	9
ABSTRACT	10
I.INTRODUCCIÓN.....	11
1.1.DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA	11
1.2. FORMULACIÓN DEL PROBLEMA.....	14
1.2.1. PROBLEMA GENERAL	14
1.3. OBJETIVOS	14
1.3.1. OBJETIVO GENERAL	14
1.3.2. OBJETIVOS ESPECÍFICOS	14
1.4. JUSTIFICACIÓN	14
1.4.1. JUSTIFICACIÓN E IMPORTANCIA DE LA INVESTIGACIÓN.....	14
II. MARCO TEORICO	17
2.1. ANTECEDENTES BIBLIOGRÁFICOS DEL PROBLEMA	17
2.2. BASES TEÓRICO – CIENTÍFICAS.....	21
2.2.1. REDES DE COMPUTADORAS	21
2.2.2. COMUNICACIÓN:.....	22
2.2.3. INFORMACIÓN:.....	23
2.2.4. PRINCIPIOS DE LA MEDICIÓN DE LA INFORMACIÓN.....	25
2.2.5. UNIDAD DE INFORMACIÓN:	26
2.2.6. REDUNDANCIA:.....	29
2.2.7. TEORÍA MATEMÁTICA DE LA COMUNICACIÓN.....	31
2.2.8. CAPACIDAD DEL CANAL DE COMUNICACIÓN	43
2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS.....	53
2.4. HIPÓTESIS	53
2.4.1. FORMULACIÓN DE LA HIPÓTESIS	53
III. MATERIALES Y MÉTODOS	54
3.1. VARIABLES Y OPERACIONALIZACIÓN DE VARIABLES	54
3.2. TIPO DE ESTUDIO Y DISEÑO DE INVESTIGACIÓN.....	54
3.2.1. TIPO DE INVESTIGACIÓN.....	54

3.3. POBLACIÓN Y MUESTRA EN ESTUDIO	54
3.3.1. POBLACIÓN	54
3.3.2. MUESTRA.....	55
3.4. MÉTODOS TÉCNICA E INSTRUMENTO PARA LA RECOLECCIÓN DE DATOS	55
3.4.1. <i>Análisis documental</i>	55
3.4.2. <i>Observación</i>	55
3.4.3. <i>Trabajo en escritorio</i>	55
3.4.4. <i>Entrevista</i>	55
IV RESULTADOS Y CONCLUSIONES	56
4.1. ANÁLISIS DE LA INFRAESTRUCTURA DE LA RED DE DATOS EN LA UDL.....	56
4.1.1. <i>Análisis de la infraestructura del servicio de telefonía</i>	56
4.1.2. <i>Análisis de la infraestructura de la red de comunicación de datos</i>	56
4.1.2.1. <i>Backbone de Fibra Óptica</i>	59
4.1.2.2. <i>Subsistema de cableado horizontal</i>	62
4.1.2.3. <i>Infraestructura de equipos activos de comunicaciones</i>	63
4.1.3. <i>Dispositivos de enrutamiento e Internet de Banda Ancha</i>	65
4.2. ANÁLISIS DE LA CAPACIDAD DEL CANAL CON LA INFRAESTRUCTURA EXISTENTE	68
4.2.1. <i>Identificación de los canales de comunicación</i>	68
4.2.2. <i>Análisis del tráfico y disponibilidad de recursos</i>	69
4.2.3. <i>Análisis del Tráfico de datos en enlace Internet dedicado de 10 Mbps y 2 Mbps</i>	74
4.3. ANÁLISIS DE LA CAPACIDAD DEL CANAL CON LOS VALORES TEÓRICOS OBTENIDOS APLICANDO EL TEOREMA DE SHANNON.	77
4.3.1. <i>Cálculo de los valores de las variables de la investigación aplicando el Teorema de Shannon</i>	77
V. DISCUSION	82
VI. CONCLUSIONES.....	83
VII. RECOMENDACIONES	85
VIII. REFERENCIAS BIBLIOGRÁFICAS	86
IX. ANEXOS	88
ANEXO: HOJA DE ENCUESTAS.	88
ANEXO: FOTO DE ENTREVISTA AL ADMINISTRADOR DE LA RED	90

ÍNDICE DE FIGURAS

<i>Figura 1: Red de computadoras</i>	<i>21</i>
<i>Figura 2: Representación de un bit de información</i>	<i>26</i>
<i>Figura 3. Cantidad de información que se puede obtener con 3 bits.....</i>	<i>27</i>
<i>Figura 4. Elementos de un sistema de comunicación general.....</i>	<i>33</i>
<i>Figura 5. Relación entre las entropías de la fuente $H(S)$ y del destinatario $H(D)$</i>	<i>38</i>
<i>Figura 6. Esquema de codificación de un canal con ruido.....</i>	<i>41</i>
<i>Figura 7. Sistema de comunicación, según Shannon.....</i>	<i>43</i>
<i>Figura 8. Sistema de comunicación con ruido, según Shannon- Hartley.....</i>	<i>49</i>
<i>Figura 9. Nivel de ruido en relación a la velocidad de transmisión</i>	<i>50</i>
<i>Figura 10. https://www.3cx.es/voip-sip/fxs-fxo/.....</i>	<i>56</i>
<i>Figura 11. Estructura de red UDL.....</i>	<i>59</i>
<i>Figura 12. Marco de conexión principal de fibra optica.....</i>	<i>61</i>
<i>Figura 13. Marco de conexión principal de fibra óptica.....</i>	<i>63</i>
<i>Figura 14. Router Cisco 2821.....</i>	<i>68</i>
<i>Figura 15. Máximo trafico soportado desde equipos activos entre nodos</i>	<i>69</i>
<i>Figura 16. Utilización de ancho de banda desde el nodo A.....</i>	<i>70</i>
<i>Figura 17. Utilización mensual de ancho de banda- línea dedicada internet 2Mbps</i>	<i>74</i>
<i>Figura 18. Utilización anual de ancho de banda- lineal dedicada internet 2Mbps</i>	<i>74</i>
<i>Figura 19. Utilización mensual de ancho de banda – lineal dedicada internet 10 Mbps</i>	<i>75</i>
<i>Figura 20. Utilización anual de ancho de banda – lineal dedicada internet 10 Mbps</i>	<i>75</i>

TABLAS

<i>Tabla N° 1. Representación de las posibles elecciones de signos con tres bits.....</i>	<i>28</i>
<i>Tabla N° 2. Variables e indicadores</i>	<i>54</i>
<i>Tabla N° 3. Utilización de ancho de banda, latencia y confiabilidad de la red en la transmisión de paquetes.....</i>	<i>73</i>
<i>Tabla N° 4. Utilización de ancho de banda de internet en líneas dedicadas en la UDL</i>	<i>76</i>
<i>Tabla N° 5. Parámetros para el análisis de la velocidad de transmisión en un canal sin ruido.....</i>	<i>77</i>
<i>Tabla N° 6. Parámetros para el análisis de la velocidad de transmisor en un canal con ruido.....</i>	<i>78</i>
<i>Tabla N° 7. Comparativa de las velocidades de transmisión real y teórica (Shannon)..</i>	<i>79</i>
<i>Tabla N° 8. Calculo del ancho de banda de los tramos de los canales de transmisión analizados</i>	<i>80</i>

RESUMEN

El propósito de esta investigación fue realizar una evaluación de la capacidad de las líneas de comunicación aplicando la teoría de la información de Shannon; complementando con los aportes Nyquist y Hartley de en la red de Datos de la Universidad de Lambayeque, con este teorema matemático de la comunicación; obtendremos la medida de la capacidad de datos que permitirá encontrar los límites fundamentales en las operaciones de transmisión y procesamiento de datos esto ayudará a prevenir la saturación en momentos adecuados como en de los diferentes canales de información cuando haya mayor acceso de los usuarios. Con el propósito de facilitar mayor información como herramientas tecnológicas que permita determinar el comportamiento bajo ciertos parámetros de la red datos. Velocidad de conexión, ancho de banda y capacidad de transmisión, a fin de proporcionar recomendaciones que permitan evaluar la red de la Universidad de Lambayeque. El estudio fue descriptivo; se utilizó una población con un número de 210 computadoras.

Para evaluar la capacidad de las líneas de comunicación de la red de datos

Se aplicó toda la población y se utilizó un software de especialidad para el análisis de redes.

La base teórica manejada para el desarrollo de la tesis fue obtenida por documentación y publicaciones de tesis.

En la recomendación se menciona que se debe evaluar constantemente la capacidad de las líneas de comunicación de la red de Datos de La Universidad de Lambayeque, adquirir software analizador de red de última generación, se recomienda la evaluación de otros parámetros para la medición del tráfico, valorar constantemente la plataforma tecnológica, en cuanto a su configuración, modernización y aplicación para mejorar el funcionamiento de la red.

Palabras claves: Teoría de Shannon, velocidad de transmisión, banda ancha, ruido, capacidad de un canal.

ABSTRACT

The purpose of this investigation was to carry out an evaluation of the capacity of the communication lines applying Shannon's information theory; complementing with the contributions Nyquist and Hartley of in the data network of the University of Lambayeque, with this mathematical theorem of communication; we will obtain the measure of the data capacity that will allow to find the fundamental limits in the operations of transmission and data processing, this will help to prevent the saturation at appropriate moments as in the different information channels when there is greater access of the users. With the purpose of providing more information as technological tools that allow determining the behavior under certain parameters of the data network. Connection speed, bandwidth and transmission capacity, in order to provide recommendations to evaluate the network of the University of Lambayeque. The study was descriptive; a population with a number of 210 computers was used.

To evaluate the capacity of the communication lines of the data network

The entire population was applied and specialized software was used for network analysis.

The theoretical base managed for the development of the thesis was obtained by documentation and thesis publications.

The recommendation mentions that the capacity of the communication lines of the data network of the University of Lambayeque should be constantly evaluated, to acquire last generation network analyzer software, the evaluation of other parameters for the measurement of traffic is recommended, constantly assess the technological platform, in terms of its configuration, modernization and application to improve the functioning of the network.

Keywords: Shannon theory, transmission speed, bandwidth, noise, capacity of a communication channel

I.INTRODUCCIÓN

1.1. DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA

Algunos autores, han llegado a afirmar que vivimos en la 'era de la información'. Independientemente del debate acerca de si esta caracterización histórica es adecuada o no, lo cierto es que el estudio de los problemas técnicos asociados a la transmisión, medición y procesamiento de la información ha cobrado vital importancia para las relaciones sociales y comerciales de la sociedad globalizada. Sin ir más lejos, los mercados actuales dependen fuertemente de que los canales informacionales entre los distintos agentes económicos funcionen en forma fluida y eficiente. Pero la importancia del estudio de los flujos informacionales y su correcto funcionamiento va más allá de la esfera estrictamente económica y abarca fines militares, científicos y recreativos.

En este contexto, el estudio de la Teoría de la Información ha tenido un creciente desarrollo, en el que cada vez más investigadores aúnan esfuerzos para resolver los desafíos planteados por la existencia de unos volúmenes de flujo de información muy difíciles de manejar y administrar. Y es especialmente importante que estos desarrollos contemplen el aspecto matemático formal de la noción de información, de forma tal que ésta se pueda medir y cuantificar para ser aplicada con fines tecnológicos (Holik, 2018).

Físicos, matemáticos e ingenieros electrónicos (como C. Shannon, N. Wiener, Von Neuman y N. Morgenster) abordaron los problemas de la comunicación desde la perspectiva de determinar formalmente cuáles son las condiciones generales para la transmisión de mensajes con independencia de cuál pueda ser el «contenido» de los mensajes transmitidos. Desde esta perspectiva informativa se analiza cuál es la cantidad de señales que pueden transportarse por un determinado canal de modo que puedan transmitirse mayor cantidad de mensajes distintos y con la menor ambigüedad y de qué forma esto es posible.

La teoría que formaliza una respuesta a esta pregunta se denomina Teoría de la Información, o también en su origen, Teoría matemática de la Comunicación. Ambas acepciones sirven para distinguir esta teoría de cualquier otra. Los postulados de la Teoría de la Información consideran el transporte de señales y la medida de la complejidad como criterios últimos de los que depende la posibilidad de transmitir mensajes (Serrano, Piñuel, Gracia, & Arias, 2012).

Los aspectos de los procesos comunicativos que interesan a la Teoría de la Información son (Serrano, Piñuel, Gracia, & Arias, 2012):

Determinar la cantidad de información (que se simboliza por la letra H) que puede contener un flujo de mensajes. Este cálculo se establece a partir de la relación entre el número de señales distintas efectivamente transportadas en el mensaje, y el número total de señales distintas de que se dispone para transmitirlo en la fuente. Dicho de otra manera, se trata de calcular la «previsibilidad» de las señales. Es decir, la cantidad de información, H, se funda en el cálculo de probabilidades. El cálculo de esa previsibilidad permite conocer en qué medida los mensajes son formalmente distintos y diferentes. Cuanto mayor es la complejidad del sistema de señales, mayor es la capacidad que tiene el agente de informarla (organizarla) para la transmisión de mensajes distintos.

La Red de Datos de la Universidad de Lambayeque se encarga de planificar, administrar y mantener la operatividad de la infraestructura de telecomunicaciones y los servicios informáticos de esta Casa de Estudios, estableciendo para ello las políticas y normativas necesarias para el desarrollo informático y de telecomunicaciones.

La Red de la Universidad de Lambayeque, tiene como misión brindar un servicio ágil y eficiente, elaborando políticas en el rubro informático basado en los lineamientos establecidos por el Rectorado, desarrollando soluciones

integrales a los problemas comunes de la universidad, utilizando para ello los recursos humanos y logísticos con los que cuenta esta organización.

A medida que la demanda de acceso a la red en la Universidad de Lambayeque comienza a crecer, ésta se vuelve cada vez menos eficiente. Esto debido a que el tráfico de la red incrementa provocándole congestión, y, por ende, el rendimiento de la red disminuye, además que, al aumentar el número de dispositivos y aplicaciones, los tiempos de respuesta se degradan, provocando lentitud.

Esto trae como consecuencia, un incremento en el costo de los servicios de soporte; además, limita la implantación de nuevas soluciones y aplicaciones para la institución.

El retardo de propagación es constante, según las condiciones de la red en cada instante en cuanto al tráfico y tamaño de los paquetes principalmente, los retardos desde el origen al destino varían con cada paquete, esto es un inconveniente para algunos tipos de servicios.

El control de flujo de la información se realiza entre nodos de la red y entre nodo y terminal de usuario por lo que pierde tiempo y capacidad de transferencia de información.

Por lo tanto, la motivación para el desarrollo de esta tesis ha sido aplicar los fundamentos teóricos de la Teoría matemática de la comunicación para evaluar el sistema de comunicación que utiliza la Universidad de Lambayeque para brindar sus diferentes servicios de información y comunicación, que sobre ella se ejecutan. Esta plataforma de comunicación es conocida como Red de Datos.

1.2. FORMULACIÓN DEL PROBLEMA

1.2.1. PROBLEMA GENERAL

¿Cuál es la capacidad de las líneas de comunicación en la Red de Datos de la Universidad de Lambayeque para transmitir el volumen de información de los diferentes servicios de Tecnología de información que se brindan?

1.3. OBJETIVOS

1.3.1. OBJETIVO GENERAL

Evaluar la red de datos de la Universidad de Lambayeque para comparar las líneas de comunicación de la información a través de la Teoría de la Información de Shannon.

1.3.2. OBJETIVOS ESPECÍFICOS

- a) Comprender la Teoría de Shannon para evaluar los distintos canales de información en la Universidad de Lambayeque.
- b) Analizar la cantidad de datos que pueden transportarse por un determinado canal de modo que puedan transmitirse mayor cantidad de mensajes distintos y con la menor confusión.

1.4. JUSTIFICACIÓN

1.4.1. JUSTIFICACIÓN E IMPORTANCIA DE LA INVESTIGACIÓN

La presente investigación se justifica y es importante por las siguientes razones:

a. Pertinencia de la investigación

Esta es una investigación pertinente porque es oportuna y conveniente.

Es oportuna, porque de la revisión de antecedentes realizadas encontramos que existe muy poca documentación relacionada con la aplicación de la Teoría de la Información en contextos de sistemas de comunicación con tecnologías de información actuales, como el Internet o redes de computadoras conectadas a través de sistemas inalámbricos o por cableado de fibra óptica, como es el caso específico del caso de estudio de la Red Telemática de la Universidad de Lambayeque. Ante este escenario, la oportunidad de la investigación se expresa en el aporte que esta investigación brindará a la confirmación de la aplicabilidad de la teoría estudiada.

Es conveniente, porque al concluir la investigación permitirán a los responsables de la gestión de las infraestructuras de comunicación y redes, contar con un modelo que les permita evaluar la eficiencia de sus sistemas de comunicación en el proceso de transmisión de la información, específicamente para determinar la capacidad de sus canales de comunicación y el nivel de ruido o distorsión de la información que en éstos existe.

b. Relevancia de la investigación

La presente investigación realizará un análisis comparativo de los resultados que se obtengan con la aplicación del modelo propuesto basado en la Teoría de la Información de Shannon, con los que se obtienen en la realidad, en las mediciones de las capacidades de canal en la transmisión de la información en la Red Telemática; lo cual permitirá identificar problemas en las comunicaciones y toma de decisiones para su mejora.

Seguridad perimetral. Proporcionando una protección más óptima de su información y podrá gestionar el acceso de los usuarios.

c. Acotación de la investigación

Este trabajo de tesis está debidamente delimitado para ser aplicado a un caso de estudio específico, como son los servicios de comunicación que se han implementado en la Red de Datos de la Universidad de Lambayeque, de los cuales se cuenta con la información suficiente para poder realizar los cálculos necesarios en la aplicación de la propuesta.

d. Viabilidad de la investigación

La investigación es viable operativamente, porque el modelo desarrollado es fácil de comprender y aplicar y no requiere de muchos conocimientos matemáticos, más que los conceptos básicos de logaritmos y probabilidades.

También es viable tecnológicamente, porque la tecnología requerida para la realización de las pruebas es la que se tiene en la Red de Datos de la Universidad de Lambayeque, a la cual se tiene autorización de acceso.

II. MARCO TEORICO

2.1. ANTECEDENTES BIBLIOGRÁFICOS DEL PROBLEMA

Flores Heroy (2003), Optimización de una Red de Transporte de Telecomunicaciones, el propósito de esta investigación fue evaluar y determinar, mejor forma de optimizar la red de transporte de telecomunicaciones de CANTV en la ciudad de Maracaibo.

La investigación fue descriptiva de tipo campo, donde se concluyó que se pueden dar situaciones de diversas índoles y que van desde problemas de software o programaciones hasta averías de hardware tanto en el área de transmisión como conmutación, causando pérdidas en los porcentajes de competición.

Por último, recomendó implementar planes de ampliación de rutas y equilibrio o balanceo de los procesadores y las redes de las centrales, para lograr un mejor uso de los recursos debido a la presencia de tráfico tan variado como el de Internet y tráfico comercial. Esta investigación constituye un antecedente para el presente trabajo por cuanto destaca una serie de parámetros y procesos que logran elevar la calidad de servicio en una red de telecomunicaciones.

Rincón C. Carlos A. (2003), realizó una investigación titulada Modelo Matemático para la Estimación del Performance de una Red Ethernet, el propósito de esta investigación fue formular un modelo matemático que permita la estimación del rendimiento de una red Ethernet. La investigación fue de campo de carácter explicativo, concluyó que el comportamiento de la red Ethernet de la licenciatura en computación de la Universidad del Zulia, permitió definir las variables fundamentales del modelo matemático a formular.

Planteó algunas recomendaciones como el análisis de otros parámetros para la medición del rendimiento de redes Ethernet, como retardo promedio y

throughput, entre otros. Esta investigación constituye un antecedente para el presente trabajo por cuanto destaca los conceptos y parámetros utilizados para medir el rendimiento de una red Ethernet.

Armando Urdaneta (2005), realizó una investigación titulada Análisis de Tráfico en una Red LAN aplicando la Tecnología de Redes Neuronales, el propósito de esta investigación fue evaluar la aplicabilidad de la tecnología de redes neuronales para el análisis de tráfico en la red de área local del colegio universitario "Rafael Belloso Chacín". La investigación fue descriptiva de carácter explicativo.

El mencionado autor concluyó que se puede caracterizar y pronosticar la conducta de la tasa de bits erróneo como un parámetro de tráfico fundamental en cualquier canal de comunicación y como inciden otros parámetros como velocidad de transmisión, capacidad de canal, tiempo de transmisión, frecuencia de transmisión, nivel de ruido, entre otros.

Recomendó evaluar constantemente el nivel de tráfico existente en la red de área local mediante la medición de sus respectivos parámetros.

Esta investigación constituye un gran aporte para el presente trabajo por cuanto destaca la importancia de realizar un análisis del tráfico de red para identificar y solucionar problemas relacionados con el estudio de las redes de telecomunicaciones. Y resalta la importancia que tiene valorar constantemente la plataforma tecnológica, en cuanto a su configuración, instalación, modernización y aplicación para optimizar el funcionamiento de la red.

En el año 2010 Ana Cristina Guzmán, realizó su Tesis titulada "Diseño y Evaluación del Tráfico de una Red LAN para la empresa SERVTEC S.A", de la Universidad Politécnica Nacional – Quito, Este proyecto contempló una solución de Diseño e implementación de una red LAN en un sistema de

comunicaciones que garantice las aplicaciones de voz y datos de forma confiable y eficiente para el transporte de información.

Se concluyó que de acuerdo a cada una de las actividades o servicios que ejerce cada uno de los departamentos, se puede realizar el análisis del respectivo dimensionamiento del tráfico que circulará por toda la red LAN; considerando así la velocidad efectiva promedio de las aplicaciones con respecto al correo electrónico, acceso a Internet, tráfico del servidor de datos, local, para los plotters, impresora y entre los departamentos.

Por lo tanto, se evidencia que existe una corriente de investigación que utiliza el modelo matemático de Claude Shannon para evaluar las líneas de transmisión de datos.

Teorema de la máxima capacidad de Shannon

En 1928 Harry Nyquist, un investigador en el área de telegrafía, publicó una ecuación llamada la Razón Nyquist que media la razón de transmisión de la señal en bauds. La razón de Nyquist es igual a $2B$ símbolos (o señales) por segundo, donde B es el ancho de banda del canal de transmisión. Así, usando esta ecuación, el ancho de banda de un canal telefónico de 3,000 Hz puede transmitido hasta $2 \times 3,000$, o 6,000 bauds o Hz.

Claude Shannon después de la investigación de Nyquist estudio el como el ruido afecta a la transmisión de datos. Shannon tomo en cuenta la razón señal-a-ruido del canal de transmisión (medido en decibeles o dB) y derivo el teorema de Capacidad de Shannon.

$$C = B \log_2 (1+S/N) \text{ bps}$$

Un típico canal telefónico de voz tiene una razón de señal a ruido de 30 dB ($10^{(30/10)} = 1000$) y un ancho de banda de 3,000 Hz. Si sustituimos esos valores en el teorema de Shannon:

$$C = 3,000 \log_2 (1+1000) = 30,000 \text{ bps}$$

Debido a que $\log_2 (1001)$ es igual al logaritmo natural de $\ln (1001)/ \ln (2)$ y es igual a 9.97, el teorema nos demuestra que la capacidad máxima* de un canal telefónico es aproximadamente a 30,000 bps.

Debido a que los canales de comunicación no son perfectos, ya que están delimitados por el ruido y el ancho de banda. El teorema de Shannon-Hartley nos dice que es posible transmitir información libre de ruido siempre y cuando la tasa de información no exceda la Capacidad del Canal.

Así, si el nivel de S/N es menor, o sea la calidad de la señal es más cercana al ruido, la capacidad del canal disminuirá.

Esta capacidad máxima es inalcanzable, ya que la fórmula de Shannon supone unas condiciones que en la práctica no se dan. No tiene en cuenta el ruido impulsivo, ni la atenuación ni la distorsión. Representa el límite teórico máximo alcanzable.

¿Cuánto nivel de S/N requeriríamos para transmitir sobre la capacidad del canal telefónico, digamos a 56,000 bps?

De la fórmula de Shannon;

$$C = B \log_2(S/N + 1) = \text{bps} \quad \text{bps} = B \log_2 (10^{(dB/10)} + 1)$$

Despejando los dB

$$\text{bps}/B = \log_2 (10^{(dB/10)} + 1)$$

$$2^{(\text{bps}/B)} = 10^{(dB/10)} + 1$$

$$10^{(dB/10)} = 2^{(\text{bps}/B)} - 1$$

$$dB/10 = 1 \log_{10} (2^{(\text{bps}/B)} - 1)$$

$$dB = 10 * 1 \log_{10} (2^{(\text{bps}/B)} - 1)$$

Sustituyendo

$B = 3,000$ y $\text{bps} = 56,000$

$\text{dB} = 10 \cdot \log_{10} (2^{(56,000/3,000)} - 1)$

$\text{dB} = S/N = 56.2 \text{ dB}$

Lo que significa que si queremos rebasar el límite de Shannon debemos de aumentar el nivel de S/N.

2.2. BASES TEÓRICO – CIENTÍFICAS

2.2.1. REDES DE COMPUTADORAS

Conjunto de computadoras autónomas interconectadas. Se dice que dos computadoras están interconectadas si pueden intercambiar información. No es necesario que la conexión se realice mediante un cable de cobre; también se pueden utilizar las fibras ópticas, las microondas, los rayos infrarrojos y los satélites de comunicaciones (ANDREW 2009).



Figura 1: Red de computadoras

Fuente: (Salcedo, 2017).

2.2.2. COMUNICACIÓN:

El concepto de comunicación en el contexto de la Teoría de la Información es empleado en un sentido muy amplio en el que "quedan incluidos todos los procedimientos mediante los cuales una mente puede influir en otra". De esta manera, se consideran todas las formas que el hombre utiliza para transmitir sus ideas: la palabra hablada, escrita o transmitida (teléfono, radio, telégrafo, etc.), los gestos, la música, las imágenes, los movimientos, etc. (López, Parada, & Simonetti, 1995)

En su empeño de abordar teóricamente y con garantías esta manera laxa de entender la comunicación, Weaver propone la distinción de tres tipos o niveles de problemas comunicacionales: el problema técnico, el semántico y el pragmático (Pérez, 2015). En el nivel técnico se analizan aquellos problemas que surgen en torno a la fidelidad con que la información puede ser transmitida desde el emisor hasta el receptor. En el semántico se estudia todo aquello que se refiera al significado del mensaje y su interpretación. Por último, en el nivel pragmático se analizan los efectos conductuales de la comunicación, la influencia o efectividad del mensaje en tanto da lugar a una conducta. Es importante destacar que la Teoría de la Información se desarrolla como una respuesta a los problemas técnicos del proceso de comunicación, aun cuando sus principios puedan aplicarse en otros contextos (López, Parada, & Simonetti, 1995).

Estos problemas pueden ser enunciados de manera más precisa de la siguiente manera (Pérez, 2015):

Nivel a: ¿Con qué grado de exactitud se pueden transmitir los símbolos empleados en la comunicación? (El problema técnico).

Nivel b: ¿Con qué grado de precisión transmiten el significado deseado los símbolos empleados en la comunicación? (El problema semántico).

Nivel c: ¿Con qué grado de eficacia afectan los significados recibidos a las conductas? (El problema pragmático o de la efectividad).

2.2.3. INFORMACIÓN:

El concepto de información es definido en términos estrictamente estadísticos, bajo el supuesto que puede ser tratado de manera semejante a como son tratadas las cantidades físicas como la masa y la energía. La palabra "información" no está relacionada con lo que decimos, sino más bien, con lo que podríamos decir. El concepto de información se relaciona con la libertad de elección que tenemos para seleccionar un mensaje determinado de un conjunto de posibles mensajes. Si nos encontramos en una situación en la que tenemos que elegir entre dos únicos mensajes posibles, se dice, de un modo arbitrario, que la información correspondiente a esta situación es la unidad. La Teoría de la Información, entonces, conceptualiza el término información como el grado de libertad de una fuente para elegir un mensaje de un conjunto de posibles mensajes (López, Parada, & Simonetti, 1995).

El concepto de información supone la existencia de duda o incertidumbre. La incertidumbre implica que existen diferentes alternativas que deberán ser elegidas, seleccionadas o discriminadas. Las alternativas se refieren a cualquier conjunto de signos construidos para comunicarse, sean estas letras, palabras, números, ondas, etc. En este contexto, las señales contienen información en virtud de su potencial para hacer elecciones. Estas señales operan sobre las alternativas que conforman la incertidumbre del receptor y proporcionan el poder para seleccionar o discriminar entre algunas de estas alternativas (López, Parada, & Simonetti, 1995).

Se asume que en los dos extremos del canal de comunicación -fuente y receptor- se maneja el mismo código o conjunto de signos. La función de la fuente de información será seleccionar sucesivamente aquellas señales que constituyen el mensaje y luego transmitir las al receptor mediante un determinado canal (Chen, 2016).

Existen diversos tipos de situaciones de elección. Las más sencillas son aquellas en que la fuente escoge entre un número de mensajes

concretos. Por ejemplo, elegir una entre varias postales para enviarle a un amigo. Otras situaciones más complejas son aquellas en que la fuente realiza una serie de elecciones sucesivas de un conjunto de símbolos elementales tales como letras o palabras. En este caso, el mensaje estará constituido por la sucesión de símbolos elegidos. El ejemplo más típico aquí es el del lenguaje.

Al medir cuánta información proporciona la fuente al receptor al enviar un mensaje, se parte del supuesto que cada elección está asociada a cierta probabilidad, siendo algunos mensajes más probables que otros. Uno de los objetivos de esta teoría es determinar la cantidad de información que proporciona un mensaje, la cual puede ser calculada a partir de su probabilidad de ser enviada (Chen, 2016).

El tipo de elección más simple es el que existe entre dos posibilidades, en que cada una tiene una probabilidad de $1/2$ (0,5). Por ejemplo, al tirar una moneda al aire ambas posibilidades -cara y sello- tienen la misma probabilidad de salir. El caso del lenguaje e idioma es diferente. En éstos la elección de los símbolos que formaran el mensaje dependerá de las elecciones anteriores. Por ejemplo, si en el idioma español el último símbolo elegido es "un", la probabilidad que la siguiente palabra sea un verbo es bastante menor que la probabilidad que sea un sustantivo o un adjetivo. Asimismo, la probabilidad que a continuación de las siguientes tres palabras "el esquema siguiente" aparezca el verbo "representa" es bastante mayor que la probabilidad que aparezca "pera". Incluso se ha comprobado que, en el caso del lenguaje, es posible seleccionar aleatoriamente letras que luego son ordenadas según sus probabilidades de ocurrencia y éstas tienden a originar palabras dotadas de sentido.

2.2.4. PRINCIPIOS DE LA MEDICIÓN DE LA INFORMACIÓN

Primer principio

Un primer principio de la medición de información, establece que mientras más probable sea un mensaje menos información proporcionará. Esto puede expresarse de la siguiente manera (López, Parada, & Simonetti, 1995):

$$I(x_i) > I(x_k) \text{ si y solo si } p(x_i) < p(x_k)$$

(Fórmula N° 1)

Donde:

$I(x_i)$: cantidad de información proporcionada por x_i

$p(x_i)$: probabilidad de x_i

De acuerdo a este principio, es la probabilidad que tiene un mensaje de ser enviado y no su contenido, lo que determina su valor informativo. El contenido sólo es importante en la medida que afecta la probabilidad. La cantidad de información que proporciona un mensaje varía de un contexto a otro, porque la probabilidad de enviar un mensaje varía de un contexto a otro.

Segundo principio

Un segundo principio que guarda relación con las elecciones sucesivas establece que, si son seleccionados los mensajes X e Y, la cantidad de información proporcionada por ambos mensajes será igual a la cantidad de información proporcionada por X más la cantidad de información proporcionada por Y, dado que X ya ha sido seleccionada. Esto puede ser expresado así (López, Parada, & Simonetti, 1995):

$$I(x_i \text{ e } y_j) > f p(x_i) + f p\left(\frac{y_j}{x_i}\right)$$

(Fórmula N° 2)

Donde:

$I(x_i \text{ e } y_j)$: cantidad de información proporcionada por los mensajes x_i e y_j

f : función

$p(x_i)$: probabilidad de x_i

$p(y_j/x_i)$: probabilidad de y_j dado que x_i ha sido seleccionado.

2.2.5. UNIDAD DE INFORMACIÓN:

Una vez que hemos seleccionado el mensaje expresado en un lenguaje determinado es posible transcribirlo a un código de tipo binario. Este consta de sólo dos tipos de señales que indican Si o No, y que generalmente se codifican como 1 o 0. La cantidad de información proporcionada por cada elección entre dos alternativas posibles constituye la unidad básica de información, y se denomina dígito binario, o abreviadamente bit.

La elección existente al tener un bit de información puede ser esquematizada de la siguiente manera:

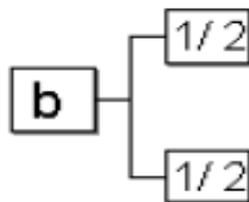


Figura 2: Representación de un bit de información

Fuente: (López, Parada, & Simonetti, 1995)

En la elección (b) tanto la línea superior como la inferior, es decir ambas posibilidades, pueden ser elegidas con la misma probabilidad de $r/2$.

Si existen N posibilidades, todas igualmente probables, la cantidad de información será igual a Log_2N . Es, entonces, el Log_2N la función matemática que nos indicará la cantidad de bits de información de una situación determinada. Esto puede esquematizarse de la siguiente manera:

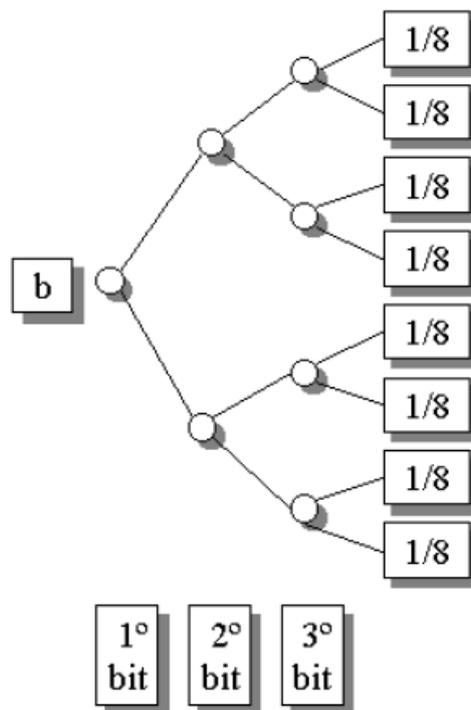


Figura 3. Cantidad de información que se puede obtener con 3 bits

Fuente: (López, Parada, & Simonetti, 1995)

La figura nos muestra una situación con 8 posibilidades, cada una con una misma probabilidad de $1/8$. Para poder determinar una posibilidad específica de estas 8, la elección requiere como mínimo 3 etapas, cada una de las cuales arroja un bit de información. El primer bit corresponde a la elección entre las primeras cuatro o segundas cuatro elegidas. El último bit determina el primer o segundo miembro del par y especifica la posibilidad elegida. Como vemos, el primero de bits que se requieren en esta situación para determinar una posibilidad específica es de 3, lo que corresponde al $\text{Log}_2 8$.

Esto se puede ejemplificar en la siguiente tabla.

Tabla N° 1. Representación de las posibles elecciones de signos con tres bits

signo	Elecciones		
	1°	2°	3°
A	1	1	1
B	1	1	0
C	1	0	1
D	1	0	0
E	0	1	1
F	0	1	0
G	0	0	1
H	0	0	0

Esta figura nos muestra un alfabeto compuesto por sólo 8 signos. Pensemos que una fuente de información selecciona un signo y de alguna manera se lo señala al receptor. La pregunta sería entonces, ¿cuánta Información deberá conocer el receptor para identificar correctamente el signo escogido?

Asumamos que a partir de elecciones anteriores sabemos que cada uno de los 8 signos tiene la misma probabilidad de ser seleccionado. La incertidumbre, entonces, se ha repartido uniformemente sobre nuestro "alfabeto", o lo que es lo mismo, las probabilidades a priori de los signos son iguales; en este caso $1/8$.

Las señales que llegan al receptor representan instrucciones para seleccionar alternativas. La primera instrucción responde a la pregunta ¿está en la primera mitad del alfabeto, sí o no? (en la figura, si = 1 y no = 0). La respuesta nos proporciona un bit de información y reduce el rango de incertidumbre exactamente a la mitad. Luego, una segunda instrucción divide cada mitad nuevamente en la mitad y, una tercera instrucción, otra vez en la mitad. En este caso, bastan tres simples instrucciones Si-No (1-

0) para identificar un signo cualquiera de un total de ocho. La letra F, por ejemplo, podría ser identificada de la siguiente manera: 010. La respuesta a nuestra pregunta es entonces, el receptor deberá obtener tres bits de información para identificar correctamente el signo escogido

2.2.6. REDUNDANCIA:

No obstante, lo anterior, la mayoría de las fuentes de información producen mensajes que no consisten en una única elección entre posibilidades de igual probabilidad, sino en elecciones sucesivas entre posibilidades de probabilidad variable y dependiente. A este tipo de secuencias se les denomina procesos estocásticos. El caso más típico son las letras y palabras que conforman el lenguaje. El escribir en español constituye un proceso de elecciones dependientes. Por ejemplo, al formar una palabra se elige una primera letra de todas las posibles primeras letras con diferentes probabilidades; luego, se elige la segunda letra cuya probabilidad depende de la primera letra seleccionada, y así sucesivamente hasta formar la palabra deseada. Lo mismo ocurre en el caso de las palabras para formar oraciones.

Lo importante aquí es señalar el hecho de que, en la medida que se avanza en la formación de una palabra u oración, el rango de posibles letras o palabras a ser seleccionadas va disminuyendo y la probabilidad de que ciertas letras o palabras específicas sean seleccionadas va aumentando. Dicho de otra forma, tanto la incertidumbre como la información de las últimas letras de una palabra o de las últimas palabras de una oración es menor comparada con las primeras.

La mayoría de los mensajes se constituyen a partir de un número limitado de posibilidades, por ejemplo, sólo 29 letras en el caso de nuestro idioma. Como vimos, la probabilidad de ocurrencia de una de estas posibilidades dentro de un mensaje depende de las posibilidades seleccionadas previamente; por ejemplo, la probabilidad de que ocurra la letra "q" luego de una "p" es 0. Son estos dos hechos los que en conjunto determinan que todo mensaje contenga cierto grado de redundancia. En

otras palabras, la redundancia se refiere a que las posibilidades dentro de un mensaje se repiten, y se repiten de una cierta manera predecible. Mientras mayor sea, entonces, la redundancia de un mensaje, menor será su incertidumbre y menor la información que contenga.

El inglés escrito es un tipo de fuente de información que ha sido ampliamente estudiado. Se ha llegado a determinar que la redundancia de la lengua inglesa está muy próxima al 50%. Es decir, al escribir inglés aproximadamente la mitad de las letras y palabras que se emplean dependen de la libre elección de quien escribe, mientras que la otra mitad está determinada por la estructura probabilística del idioma.

La redundancia de los idiomas permite que si se pierde una fracción de un mensaje sea posible completarlo en forma muy aproximada al original. Este hecho se puede observar al eliminar varias letras de una oración sin que ello impida al lector completar las omisiones y rehacer la oración. Por ejemplo, en la siguiente frase han sido omitidas las vocales: CMPLT ST FRS.

Otra función importante de la redundancia es que nos permite ahorrar tiempo en la decodificación de los mensajes. Generalmente, no leemos cada una de las letras y palabras que conforman un texto, sino que vamos adivinando lo que viene. En el caso del telégrafo, por ejemplo, podríamos ahorrar tiempo ideando un código poco redundante y transmitiendo el mensaje a través de un canal sin ruido. Sin embargo, cuando el canal utilizado tiene ruido es conveniente no emplear un proceso de codificación que elimine toda la redundancia, pues la redundancia nos ayuda a combatir el ruido. Si se pierde parte del mensaje por el ruido que afecta al canal, la redundancia nos permite rehacer en forma aproximada el mensaje.

Por el contrario, la fracción de un mensaje no redundante que se pierde por el ruido es imposible de ser recuperada. La redundancia de los mensajes nos permite, entonces, corregir con facilidad los errores u omisiones que hayan podido ocurrir durante la transmisión.

2.2.7. TEORÍA MATEMÁTICA DE LA COMUNICACIÓN

2.2.7.1. El propósito de la Teoría matemática de la comunicación

El artículo de Claude Shannon "Una teoría matemática de la comunicación", publicado en julio y octubre de 1948, es la Carta Magna de la era de la información. El descubrimiento de Shannon de las leyes fundamentales de la compresión y transmisión de datos marca el nacimiento de la teoría de la información (Chen, 2016). Es sin duda uno de los avances científicos más importantes del siglo XX. El principal objetivo de esta teoría es el de proporcionar una definición rigurosa de la noción de información que permita cuantificarla. Fue desarrollada con el objetivo de encontrar límites fundamentales en las operaciones de procesamiento de señales tales como compresión de datos, almacenamiento y comunicación. Sus aplicaciones se extienden a campos diversos, entre ellos la física, la química, la biología, la inferencia estadística, la robótica, la criptografía, la computación, la lingüística, el reconocimiento de patrones y la teoría de la comunicación (Holik, 2018).

La teoría de Shannon se basó en teorías previas, como el mismo Shannon lo reconoce en su artículo original:

“El desarrollo de varios métodos de modulación, tales como PCM y PPM que intercambian ancho de banda para una tasa de señal a ruido ha intensificado el interés en una teoría general de la comunicación. Una base para esta teoría está contenida en los importantes artículos de Nyquist y Hartley en este tema. En el presente artículo extenderemos la teoría para incluir un número de nuevos factores, en particular el efecto del ruido en el canal, y los cambios posibles debido a la estructura estadística del mensaje original y debido a la naturaleza final de la destinación de la información.” (Shannon & Weaver, 1949)

Uno de los principales aportes de Shannon fue el especificar cuál es el problema informacional, y separar los aspectos formales de los aspectos semánticos asociados con el contenido del mensaje.

“El problema fundamental de la comunicación es el de reproducir en un punto exacto o aproximadamente, un mensaje seleccionado en otro punto. Frecuentemente los mensajes tienen significado; esto es, ellos refieren o están correlacionados de acuerdo a algún sistema con ciertas entidades físicas o conceptuales. Estos aspectos semánticos de la comunicación son irrelevantes para el problema de ingeniería. El aspecto significativo es que el mensaje de hecho es uno seleccionado a partir de un conjunto de posibles mensajes. El sistema debe ser diseñado para operar para cada posible selección, no sólo para aquella que será seleccionada de hecho, dado que es desconocida al momento del diseño.” (Shannon & Weaver, 1949)

De este modo, Shannon define el problema en términos técnicos precisos, separando el problema de ingeniería concreto asociado al proceso de comunicación con los aspectos semánticos de la noción de información vinculada a su uso generalizado en el lenguaje cotidiano. Esto no significa que estos últimos desaparezcan o que no sean importantes; de hecho, los aspectos filosóficos de la Teoría de la Información han dado lugar a importantes debates en la literatura. Lo que señala Shannon es simplemente que los aspectos semánticos son irrelevantes para el problema planteado en la ingeniería del proceso comunicacional.

Su trabajo se centró en algunos de los siguientes problemas que surgen en los sistemas destinados a manipular información: cómo hablar, los mejores métodos para utilizar los diversos sistemas de comunicación; cómo establecer el mejor método para separar las

señales del ruido y cómo determinar los límites posibles de un canal (Cherry, 1956).

2.2.7.2. El problema de la codificación

Shannon estableció resultados matemáticos acerca de los recursos que se necesitan para la codificación óptima y para la comunicación libre de errores. Estos resultados tuvieron aplicaciones en diversos campos ligados a la teoría de las comunicaciones, como en la radio, la televisión y la telefonía. Hoy en día se siguen aplicando en diversas disciplinas (Holik, 2018) como el Internet y las redes de computadoras.

De acuerdo a Shannon, un sistema de comunicación general consta de varias partes. Una fuente, la cual genera un mensaje a ser recibido en el destinatario. Un transmisor, que transforma el mensaje generado en la fuente en una señal a ser transmitida. En los casos en los que la información es codificada, el proceso de codificación también es implementado por el transmisor. Un **canal** es cualquier medio que sirva para que la señal se transmita desde el transmisor al receptor. Este puede ser, por ejemplo, un cable, una fibra óptica o una señal inalámbrica. Un **receptor**, que reconstruye el mensaje a partir de la señal, y finalmente, un destinatario, que es quien recibe el mensaje. En el siguiente dibujo se representan estos elementos en forma esquemática:

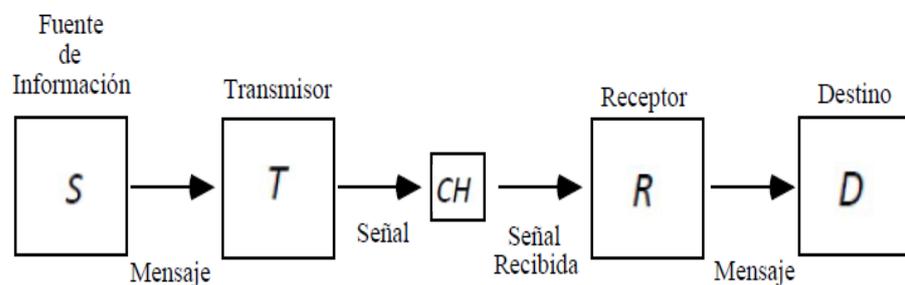


Figura 4. Elementos de un sistema de comunicación general

Fuente: (Pérez, 2015)

La fuente **S** es un sistema que contiene un conjunto de estados diferentes S_1, \dots, S_n , llamados usualmente letras. Un aspecto central de la teoría de Shannon es que es posible asignar probabilidades de ocurrencia para los distintos estados de la fuente. Es decir, los estados S_1, \dots, S_n son producidos con probabilidades $p(S_1), \dots, p(S_n)$. La cantidad de información generada por la fuente debido a la ocurrencia del estado S_i se define como:

$$I(S_i) = \log\left(\frac{1}{p(S_i)}\right) = -\log p(S_i)$$

(Fórmula N° 3)

Dado que **S** produce sucesiones de estados (estas sucesiones son usualmente llamadas mensajes), la entropía de la fuente **S** se define como la cantidad promedio de información producida por la fuente:

$$H(S) = \sum_{i=1}^n p(S_i) \log\left(\frac{1}{p(S_i)}\right) = -\sum_{i=1}^n p(S_i) \log p(S_i)$$

(Fórmula N° 4)

En forma análoga, el destinatario **D** es un sistema con un rango de estados posibles D_1, \dots, D_m , a los cuales se le asignan probabilidades $p(D_1), \dots, p(D_m)$. La cantidad de información $I(D_j)$ recibida en el destinatario debido a la ocurrencia de d_j se define como:

$$I(D_j) = \log\left(\frac{1}{p(D_j)}\right) = -\log p(D_j)$$

(Fórmula N° 5)

Y la entropía del destinatario se define como la cantidad promedio de información recibida:

$$H(D) = \sum_{j=1}^m p(D_j) \log\left(\frac{1}{p(D_j)}\right) = -\sum_{j=1}^m p(D_j) \log p(D_j)$$

(Fórmula N° 6)

Este sistema de comunicación es lo suficientemente amplio como para incluir los diferentes contextos en que se da la comunicación (conversación, T.V., danza, etc.). Tomemos como ejemplo lo que ocurre en el caso de la radio. La fuente de información corresponde a la persona que habla por el micrófono. El mensaje son las palabras y sonidos que esta persona emite. El micrófono y el resto del equipo electrónico constituyen el transmisor que transforma este mensaje en ondas electromagnéticas, las cuales corresponden a la señal. El espacio que existe entre las antenas transmisoras y receptoras es el canal, mientras que lo que altera la señal original constituye la fuente de ruido. El aparato de radio de cada hogar es el receptor y el sonido que éste emite corresponde al mensaje recobrado. Las personas que escuchan este mensaje radial son los destinatarios (López, Parada, & Simonetti, 1995).

También podemos ejemplificar esto mediante la lectura de un artículo que lee. En este caso, nuestros cerebros son la fuente de información y nuestros pensamientos, el mensaje. La máquina de escribir constituye el transmisor que transforma nuestros pensamientos en lenguaje escrito, el cual corresponde a la señal. El papel es el canal y cualquier error de tipeo o puntuación, manchas, espacios en blanco, etc., constituyen la fuente de ruido. Por último, usted que está leyendo este ejemplo es a la vez el receptor y destinatario, que a través de la lectura recobra el mensaje por nosotros enviado (López, Parada, & Simonetti, 1995).

Es importante considerar que el problema del significado del mensaje no es relevante en este contexto. El interés principal de la Teoría de la Información lo constituye todo aquello relacionado con

la capacidad y fidelidad para transmitir información de los diferentes sistemas de comunicación. En el ejemplo anterior, el mensaje podría haber consistido en una secuencia de letras carentes de todo significado e igualmente el problema de cuánta información es transmitida estaría presente. En un sentido amplio, la Teoría de la Información trata acerca de la cantidad de información que es transmitida por la fuente al receptor al enviar un determinado mensaje, sin considerar el significado o propósito de dicho mensaje. No interesa tanto la pregunta: "¿Qué tipo de información?" sino más bien, "¿Cuánta información?" es la que transmite la fuente.

2.2.7.3. Medidas de información

En su artículo original, Shannon discute la conveniencia del uso de una función logarítmica en la definición de entropías, dado que muchos parámetros importantes en ingeniería varían linealmente con el logaritmo del número de posibilidades. Señala que es al mismo tiempo una medida intuitiva, porque es usual medir magnitudes comparando linealmente con unidades de medida. Además de estas ventajas, es matemáticamente adecuada dado que resulta ser que muchas operaciones llevadas a cabo con el logaritmo se vuelven mucho más simples que en términos de otras formas funcionales (Holik, 2018).

La elección de una base logarítmica se corresponderá con la elección de una unidad de medida de la información. Si la base usada es 2, la unidad resultante se llamará 'bit'. Este término proviene de la contracción del término en inglés 'binary digit'. Con estas definiciones, un bit terminará siendo la cantidad de información que se obtiene cuando una de dos alternativas igualmente probables es especificada; en otras palabras, una fuente S con dos estados que puede adoptar los valores 1 y 2 con la misma probabilidad, tendrá una cantidad de información de un bit.

Otras bases dan lugar a diferentes unidades de la información. Por ejemplo, si se usan logaritmos naturales, la unidad de medida en este caso será el nat (que es la contracción del término en inglés natural unit). Cuando se usa la base 10 de los logaritmos, la unidad se llama Hartley. La posibilidad de usar diferentes unidades para cuantificar la información muestra la diferencia entre la cantidad de información asociada a un evento y el número de símbolos binarios necesarios para codificarlo. Es importante mencionar que el “bit” es una unidad de medida, y el hecho de que distintas unidades de medida puedan ser usadas, no afecta a la naturaleza de la información.

De las definiciones es bastante claro que $H(S)$ y $H(D)$ son cantidades de información promedio. Sin embargo, en la literatura son llamadas usualmente **entropías**. Efectivamente, la forma funcional que poseen estas medidas de la información, es completamente análoga a las formas funcionales que la entropía adquiere en distintas teorías físicas. Esta terminología refleja una problemática profunda en los fundamentos de la teoría de la información. Esto es así a tal punto que el carácter intrigante de esta magnitud y sus propiedades se manifestó al comienzo de la formulación de la teoría. En efecto, de acuerdo a las palabras mismas de Shannon refiriéndose a cómo llamar a su nueva medida.

“Mi preocupación más grande era cómo llamarla. Pensé en llamarla ‘información’, pero esa palabra estaba muy usada, de forma tal que decidí llamarla ‘incerteza’. Cuando discutí el asunto con John von Neumann, él tuvo una idea mejor. Von Neumann me dijo: ‘Deberías llamarla entropía, por dos motivos. En primer lugar, tu función de incerteza ha sido usada en la mecánica estadística bajo ese nombre, y por ello, ya tiene un nombre. En segundo lugar, y lo que es más importante, nadie sabe lo que es la entropía realmente, por ello, en un debate, siempre llevarás ventaja.’”

La relación entre las entropías de la fuente $H(S)$ y del destinatario $H(D)$ se puede representar intuitivamente por el siguiente diagrama:

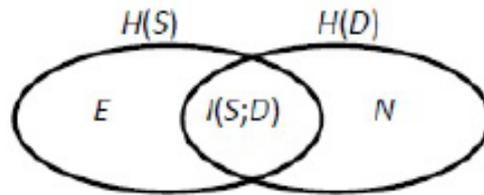


Figura 5. Relación entre las entropías de la fuente $H(S)$ y del destinatario $H(D)$

Fuente: (Pérez, 2015)

Donde $I(S;D)$ es la información mutua: la cantidad promedio de información generada en la fuente S y recibida en el destinatario D . E es la equivocación: la cantidad promedio de información generada en S pero no recibida en D . N es el ruido: la cantidad promedio de información recibida en D pero no generada en S . Como el diagrama sugiere desde un punto de vista intuitivo, la información mutua se define como:

$$I(S;D) = H(S) - E = I(X;Y) = H(D) - N$$

(Fórmula N° 7)

La equivocación E y el ruido N son medidas de la independencia entre la fuente y el destinatario, dado que si S y D son completamente independientes, los valores de E y N son máximos ($E = H(S)$ y $N = H(D)$), y el valor de $I(S;D)$ es mínimo ($I(S;D)=0$). Por otro lado, si la dependencia entre S y D es máxima, los valores de E y N son mínimos ($E = N = 0$), y el valor de $I(S;D)$ es máximo ($I(S;D)=H(S)=H(D)$).

Los valores de E y N son funciones no solo de la fuente y del destinatario, sino también del canal de comunicación CH . La

introducción del canal de comunicación lleva directamente a la posibilidad de que se produzcan errores en el proceso de transmisión. El canal CH se puede definir como una matriz $[p(d_j|s_i)]$, donde $p(d_j|s_i)$ es la probabilidad condicional de ocurrencia de d_j en el destinatario dado que s_i ocurrió en la fuente S. Los elementos de cada columna de $[p(d_j|s_i)]$ suman uno. De este modo, E y N se pueden expresar como:

$$N = \sum_{i=1}^n p(s_i) \sum_{j=1}^m p(d_j|s_i) \log \left(\frac{1}{p(d_j|s_i)} \right)$$

(Fórmula N° 8)

$$N = \sum_{j=1}^m p(d_j) \sum_{i=1}^n p(s_i|d_j) \log \left(\frac{1}{p(s_i|d_j)} \right)$$

(Fórmula N° 9)

$$\text{Donde: } p(s_i|d_j) = \frac{p(d_j|s_i) p(s_i)}{p(d_j)}$$

(Fórmula N° 10)

La capacidad del canal se define como:

$$C = \max_{p(s_i)} I(S; D)$$

(Fórmula N° 11)

Donde el máximo se toma sobre todas las posibles distribuciones $p(s_i)$ en la fuente. La magnitud C se interpreta como la cantidad de información promedio más grande que puede ser transmitida por el canal de comunicación CH.

2.2.7.4. Teoremas de Shannon

Shannon planteó dos teoremas:

El primer teorema, o teorema de codificación sin ruido, plantea que el valor de la entropía $H(S)$ de la fuente es igual al número de símbolos necesarios en promedio para codificar una letra de la fuente usando un código ideal: $H(S)$ mide la compresión óptima de los mensajes de la fuente. La prueba del teorema se basa en el hecho de que los mensajes de longitud N que produce S pueden ser clasificados en dos tipos. Una de estas clases tiene aproximadamente $2^{NH(S)}$ mensajes típicos. La otra clase, contiene los mensajes atípicos. Cuando $N \rightarrow \infty$, la probabilidad de un mensaje atípico se vuelve despreciable. Por ello, se puede concebir a la fuente como produciendo sólo $2^{NH(S)}$ mensajes posibles. Esto sugiere una estrategia natural para codificar: cada mensaje típico es codificado en una secuencia binaria de longitud $NH(S)$, que es en general más corta que la longitud N del mensaje original. Se pensaba entonces, que el incremento de la tasa de transmisión de información sobre un canal de comunicación siempre aumentaría la probabilidad de error.

En resumen, Desde esta perspectiva, el significado de la entropía $H(S)$ queda definido por el primer teorema de Shannon: “preguntar cuánta información produce una fuente es preguntar hasta qué grado es compresible la salida de la fuente?” ... También: “La información de Shannon $H(X)$ [...] mide en qué medida se pueden comprimir los mensajes de la fuente.” (Timpson, 2008)

El segundo teorema, conocido como **teorema de codificación en un canal con ruido,** sorprendió a la comunidad de la teoría de la comunicación probando que esa hipótesis no era cierta si se mantenía la tasa de comunicación por debajo de la capacidad del

canal. La capacidad del canal es igual a la tasa máxima con la cual la información puede ser enviada por el canal y recuperada en el destinatario con una probabilidad de error despreciable.

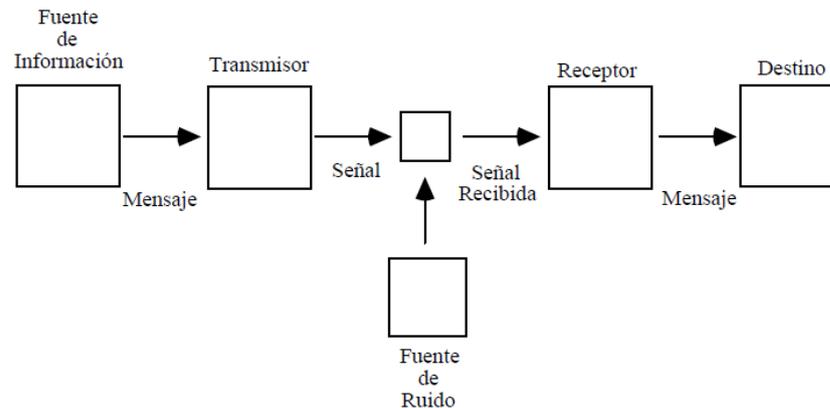


Figura 6. Esquema de codificación de un canal con ruido

Fuente: (Pérez, 2015)

La fuente consiste en un sistema S de n estados S_i , que se pueden pensar como letras de un alfabeto $AS = \{s_1, \dots, s_n\}$, cada una con su propia probabilidad $p(s_i)$. Las secuencias de letras son llamadas mensajes. La entropía de la fuente $H(S)$ se puede calcular exclusivamente en términos de estos elementos, y se mide en bits cuando el logaritmo tiene base 2. A su vez, el transmisor codifica el mensaje de la fuente y esto equivale a hacer una conversión entre el alfabeto de la fuente $AS = \{s_1, \dots, s_n\}$, y el código del alfabeto del transmisor T , que viene dado por $AC = \{c_1, \dots, c_q\}$. Los elementos C_i son:

Llamados símbolos. La secuencia de símbolos producidos por el transmisor y que entran al canal se llama señal. El alfabeto de n símbolos AS puede variar mucho dependiendo de los distintos dispositivos empleados

En resumen, el Segundo teorema de Shannon es usado para explicar la información mutua $H(S;D)$: “La interpretación más importante de la información mutua deriva del teorema de codificación ruidoso.” (Timpson, 2008). Es importante remarcar que, si bien la estrategia de definir la información de Shannon usando los teoremas de codificación lleva ciertamente a una perspectiva interesante, algunos autores la han criticado.

Por otro lado, en muchos ejemplos de interés, conviene elegir un AC binario, es decir, con $q = 2$. En este caso, los símbolos son directamente dígitos binarios. Pero en el caso más general, el alfabeto del código se puede implementar físicamente por medio de sistemas que tengan una cantidad q de estados disponibles. Para el caso particular en que $q = 2$, los sistemas de dos niveles se pueden llamar c bits.

En el contexto de la teoría de la información de Shannon, codificar implica establecer un mapa entre las letras S_i del alfabeto de la fuente AS el conjunto de cadenas de longitud finita de símbolos del alfabeto del código AC . Estas suelen llamarse palabras-código. En general, las palabras-código no tienen la misma longitud. Cada palabra-código w_i que corresponde a la letra S_i , va a tener una longitud l_i . Pero las longitudes l_i de las distintas palabras-código pueden variar. Es entonces útil definir una longitud de palabra-código promedio como:

$$L = \sum_{i=1}^n p(S_i) l_i$$

(Fórmula N° 12)

L es entonces una medida de cuán compacto es el código. En otras palabras, un código con un valor de L más pequeño será más eficiente, dado que, en promedio, economiza más recursos informacionales para codificar los mensajes. El Teorema del Canal Sin ruido afirma que existe un proceso de codificación óptimo tal

que la longitud de palabra-código promedio L está tan cerca como se quiera del límite inferior L_{\min} para L :

$$L_{\min} = \frac{H(S)}{\log q}$$

(Fórmula N° 13)

En la fórmula de arriba, si $H(S)$ se mide en bits, \log es el logaritmo en base 2. Es importante enfatizar la diferencia entre el proceso de generación de la información en la fuente y la codificación de la información en el transmisor. La información generada en la fuente se mide con $H(S)$, y sólo depende de las características de la fuente.

2.2.8. CAPACIDAD DEL CANAL DE COMUNICACIÓN

2.2.8.1. Canal de comunicación

Al igual que todos los tipos de comunicación, mantener un diálogo con otro sistema necesita un medio físico para transmitir los datos. En el caso de la Arquitectura de Redes, a ese medio se le denomina canal.



Figura 7. Sistema de comunicación, según Shannon

Fuente: (Heras, 2013)

La X representa el espacio entre las señales que puede ser transmitida y la Y el espacio de señales recibidas, durante un bloque de tiempo sobre el canal (Heras, 2013).

$$p_{Y|X}(y|x)$$

Con esta fórmula obtenemos la distribución condicional de la función de Y sobre X . Tratando el canal como un sistema estático conocido,

$p_{Y|X}(y|x)$, esto es una propiedad inherente del canal de

comunicación (que representa la naturaleza del ruido en él). Entonces la distribución conjunta $p_{Y|X}(y|x)$ de X e Y esta completamente determinada por el canal y por la elección de la distribución marginal de las señales que nosotros elegimos enviar sobre el canal:

$$p_X(x) = \int_y^x p_{X,Y}(x,y) dy$$

La distribución conjunta puede ser recuperada usando la identidad:

$$p_{Y,X}(y,x) = p_{Y|X}(y|x) p_X(x)$$

Sobre estas limitaciones, se debe después maximizar la cantidad de información, o de mensaje, que uno puede comunicar sobre el canal. La medida apropiada para esto es la información mutua $I(X;Y)$, y esta máxima información mutua es el llamado canal de capacidad y viene dado por:

$$C = \sup_{p_X} I(X;Y)$$

(Fórmula N° 14)

2.2.8.2. Elementos que intervienen en la definición de la capacidad de un canal de transmisión

Ahora que ya hemos precisado el concepto -de información y los conceptos relacionados con él (incertidumbre, bit, redundancia) podemos volver a plantearnos el problema inicial de definir la capacidad de un canal determinado para transmitir información. Dado un canal con una capacidad de C unidades por segundo que recibe señales de una fuente de información de H unidades por segundo, la pregunta es ¿cuánto es el máximo número de bits por segundo que puede ser transmitido a través de este canal? Por ejemplo, un teletipo consta de 32 símbolos posibles que supondremos son empleados con igual frecuencia. Cada símbolo

representa entonces 5 bits ($\log_2 32$) de información. De esta forma, si en ausencia total de ruido podemos enviar N símbolos por segundo a través de este canal, entonces podremos enviar $5N$ bits de información por segundo a través de dicho canal.

Son estas dos cantidades, la tasa de transmisión H por la fuente de información y la capacidad C del canal, las que determinan la efectividad del sistema para transmitir información. Si $H > C$ será ciertamente imposible transmitir toda la información de la fuente, no habrá suficiente espacio disponible. Si $H \leq C$ será posible transmitir la información con eficiencia. La información, entonces, puede ser transmitida por el canal solamente si H no es mayor que C .

El teorema fundamental para un canal sin ruido que transmite símbolos discretos afirma que si se emplea un procedimiento adecuado de codificación para el transmisor es posible conseguir que el ritmo medio de transmisión de símbolos por el canal sea muy próximo a C/H . Por muy perfecto que sea el procedimiento de codificación, dicho ritmo nunca podrá ser mayor de C/H .

Sin embargo, el problema de calcular la capacidad del canal se complica por la presencia de ruido. La presencia de ruido durante la transmisión provocará que el mensaje recibido contenga ciertos errores que contribuirán a aumentar la incertidumbre. Cuanto mayor sea la libertad de elección, mayor será la falta de seguridad en el hecho de que el mensaje enviado sea uno determinado. La incertidumbre será mayor y mayor la cantidad de información posible. De esta forma, si el ruido aumenta la incertidumbre, aumentará la información. Esto parecería indicar que el ruido es beneficioso, puesto que cuando hay ruido, la señal recibida es seleccionada a partir de un mayor conjunto de señales que las deseadas por el emisor. Sin embargo, la incertidumbre originada por la libertad de elección del emisor es una incertidumbre

deseable; la incertidumbre debida a errores por la influencia del ruido es una incertidumbre no deseable.

Para extraer la información útil de la señal recibida es necesario suprimir la ambigüedad introducida por el ruido. Para ello se recurre a un factor de corrección matemático que no entraremos a analizar. El teorema para la capacidad de un canal con ruido se define como el ritmo máximo a que la información útil (incertidumbre total menos la incertidumbre debida al ruido) puede ser transmitida a través del canal.

2.2.8.3. Definición de capacidad de un canal de transmisión

Se llama capacidad de un canal a la velocidad, expresada en bps (bits por segundo), a la que se pueden transmitir los datos en un canal o ruta de comunicación. La capacidad del canal depende de su ancho de banda y de la relación S/N (Relación señal/ruido). La capacidad del canal limita la cantidad de información (se denomina régimen binario y se mide en bits por segundo, bps) que puede transmitir la señal que se envía a través de él. La capacidad máxima de un canal viene dada por la fórmula:

$$C = B \log_2(1+S/R) \text{ (bps)}$$

(Fórmula N° 15)

El régimen binario de una señal que se propaga por un canal no puede ser mayor que la capacidad del canal y depende del número de niveles o estados que se usan para codificar la información.

$$\begin{aligned} \text{Régimen binario de la señal} &= nVt(\text{baudios}) = 2Bn = 2B \log_2 M \\ &= C \text{ (bps)} \end{aligned}$$

(Fórmula N° 16)

donde:

N = es el número de bits por cada elemento de la señal.

M = es el conjunto de elementos diferentes que puede adoptar la señal.

C = es la capacidad del canal según el teorema de Nyquist.

La capacidad del canal depende de la naturaleza del soporte, es decir, de los portadores canales de gran ancho de banda, como la fibra óptica, su capacidad siempre tiene un límite. Nyquist demostró la existencia de ese límite cuando se envían señales digitales por canales analógicos.

Las limitaciones en el ancho de banda surgen de las propiedades físicas de los medios de transmisión o por limitaciones que se imponen deliberadamente en el transmisor para prevenir interferencia con otras fuentes que comparten el mismo medio. Cuanto mayor es el ancho de banda mayor el costo del canal.

Lo deseable es conseguir la mayor velocidad posible dado un ancho de banda limitado, no superando la tasa de errores permitida; el mayor inconveniente para conseguir esto es el RUIDO.

2.2.8.4. Teorema de NYQUIST para calcular el ancho de banda de un canal sin ruido

Nyquist supuso en su teorema un canal exento de ruido (ideal). Por lo tanto, la limitación de la velocidad de transmisión permitida en el canal, es la impuesta exclusivamente por el ancho de banda del canal.

El teorema de Nyquist establece que la velocidad máxima de transmisión en bits por segundo para un canal (sin ruido) con ancho de banda B (Hz) es:

$$C = 2B \log M$$

(Fórmula N° 17)

Dónde:

M= niveles de la señal

Si M=2 entonces $\log(2)=1$, por lo tanto:

$$C=2B$$

(Fórmula N° 18)

Por ejemplo: En un medio de transmisión de cobre tiene un ancho de banda de 10Mhz ¿Cuántos bits/s se puede enviar si se utilizan señales digitales de cuatro niveles? Supongamos que se trata de un canal ideal exento de ruido.

Datos:

$$B= 10 \text{ Mhz} = 10^7 \text{Hz}$$

$$M = 4$$

APLICAMOS EL TEOREMA DE NYQUIST

$$C= 2 \cdot B \log_2 M$$

$$C = 10 \text{ Mhz} = 10^7 \text{Hz} \log_2 4 = 40 \text{ Mbps}$$

Por ejemplo, si suponemos que un canal de voz con un ancho de banda de 3100 Hz se utiliza con un modem para transmitir datos digitales (2 niveles), la capacidad C del canal es:

$$2B= 6200 \text{ bps.}$$

Si se usan señales de más de 2 niveles; es decir, cada elemento de señal puede representar a más de 2 bits, por ejemplo, si se usa una señal con cuatro niveles de tensión, cada elemento de dicha señal podrá representar dos bits (dibits). Aplicando la fórmula de Nyquist tendremos:

$$C=2 B \log(4)$$

$$C= 2(3100) (2)$$

$$C=12,400 \text{ bps}$$

Para señales múltiples Nyquist formuló que

$$C = 2B \log_2 M$$

(Fórmula N° 19)

Por ejemplo, para calcular la capacidad del canal si la velocidad de transmisión es igual a 1200 bps con 5 niveles de señal, entonces tenemos que:

$$C=2B \log_2 M$$

$$C=2(1200) \log_2(5)$$

$$C= 2(1200) \ln 5 / \ln 2$$

$$C= 5572.62 \text{ bps}$$

El teorema establece la máxima cantidad de datos digitales que pueden ser transmitidos sin error sobre dicho enlace de comunicaciones con un ancho de banda específico y que está sometido a la presencia de la interferencia del ruido.

2.2.8.5. Con ruido Teorema de Shannon – Hartley para calcular el ancho de banda de un canal.

Como se mencionó anteriormente, Shannon estudió el caso general de un sistema de comunicación, compuesto por un emisor, un receptor, un canal de transmisión y una fuente de ruido, como se muestra en la gráfica siguiente:

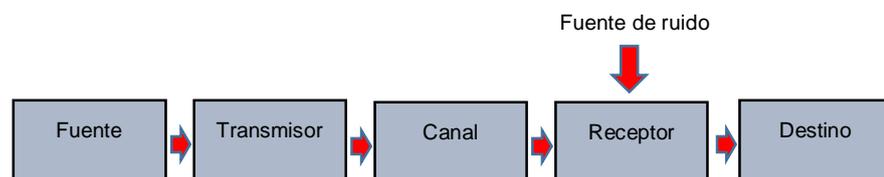


Figura 8. Sistema de comunicación con ruido, según Shannon-Hartley

Fuente: (Pérez, 2015)

En teoría de la información, el teorema de Shannon-Hartley (debe su nombre a Claude Shannon y Ralph Hartley) es una aplicación del teorema de codificación para canales con ruido. Un caso muy frecuente es el de un canal de comunicación analógico continuo en el tiempo que presenta un ruido gaussiano, que se encuentra asociado con la radiación electromagnética. Ya que no podemos tener comunicación eléctrica sin electrones, es imposible evitar el ruido.

El teorema establece la capacidad del canal de Shannon, una cota superior que establece la máxima cantidad de datos digitales que pueden ser transmitidos sin error (esto es, información) sobre dicho enlace de comunicaciones con un ancho de banda específico y que está sometido a la presencia de la interferencia del ruido.

En las hipótesis de partida, para la correcta aplicación del teorema, se asume una limitación en la potencia de la señal y, además, que el proceso del ruido gaussiano es caracterizado por una potencia conocida o una densidad espectral de potencia.

Dado un nivel de ruido, cuanto mayor es la velocidad de transmisión mayor es la tasa de errores

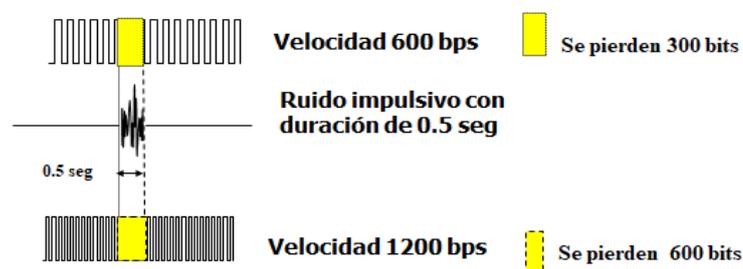


Figura 9. Nivel de ruido en relación a la velocidad de transmisión

Fuente: (Pérez, 2015)

El teorema de Shannon establece que:

$$C = B \log (1+S/N) \text{ (bps)}$$

(Fórmula N° 20)

Dónde:

C=capacidad teórica máxima en bps

B=ancho de banda del canal Hz.

S/N=relación de señal a ruido, S y N dados en watts.

Por ejemplo, supóngase que el espectro de un canal está situado entre 3Mhz y 4 Mhz y que la SNR es de 24 dB.

$$B = 4\text{Mhz} - 3\text{Mhz} = 1\text{Mhz}$$

$$\text{SNR} = 24 \text{ dB} = 10 \log(\text{SNR}) = 251$$

Usando la fórmula de Shannon se tiene que:

$$C = 10^6 \log (1+251) = 8 \text{ Mbps}$$

Este es un límite teórico difícil de alcanzar.

Según Nyquist para alcanzar este límite ¿Cuántos niveles serán requeridos?

$$C = 2 B \log_2 M = 8 \times 10^6 = 2 \times 10^6 \times \log_2 M$$

$$4 = \log_2 M$$

$$M = 16 \text{ niveles}$$

Para un nivel de ruido dado, podría parecer que la velocidad de transmisión se puede aumentar incrementando tanto la energía de la señal como el ancho de banda.

Sin embargo, al aumentar la energía de la señal, también lo hacen las no linealidades del sistema dando lugar a un aumento en el ruido de intermodulación.

Ya que el ruido se ha supuesto blanco, cuanto mayor sea el ancho de banda, más ruido se introducirá al sistema. Por lo tanto, cuando aumenta B, disminuye SNR.

2.2.8.6. Eficiencia de un canal de comunicación

La eficiencia E de un canal de comunicación, es la relación entre su capacidad y su ancho de banda e indica el número de bits por segundo de información que se puede transmitir por cada herzio de su ancho de banda:

$$E = C/B \text{ (bps/Hz)}$$

(Fórmula N° 21)

La eficiencia de un canal está establecida por el número máximo de estados distintos que puede adoptar la señal que se transmite por él para codificar la información. La capacidad de un canal de ancho de banda determinado podría elevarse a infinito utilizando señales con un número infinito de niveles. Pero esto solo se puede obtener con un canal ideal, sin ruidos ni distorsión alguna, los cuales no existen en el mundo real. Por tanto, el número de estados posibles está limitado por la sensibilidad y la capacidad de resolución del propio receptor para diferenciar entre niveles contiguos de la señal que recibe junto con el ruido.

2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS

Fuente: Componente de naturaleza humana o mecánica que determina el tipo de mensaje que se transmitirá y su grado de complejidad.

Transmisor: Recurso técnico que transforma el mensaje originado por la fuente de información en señales apropiadas.

Canal: Medio generalmente físico que transporta las señales en el espacio (cumple funciones de mediación y transporte)

Receptor: Recurso técnico que transforma las señales recibidas.

Destino: Componente terminal del proceso de comunicación al cual está dirigido el mensaje.

2.4. HIPÓTESIS

2.4.1. FORMULACIÓN DE LA HIPÓTESIS

Se puede determinar la capacidad de los canales de comunicación aplicando la Teoría de la Información de Shannon.

III. MATERIALES Y MÉTODOS

3.1. VARIABLES Y OPERACIONALIZACIÓN DE VARIABLES

Tabla N° 2. Variables e indicadores

Variables	Dimensiones	Indicador	Instrumentos
Transmisión de las líneas de comunicación aplicando la teoría de Shannon.	Diagnostico. Eficiencia en la transmisión	Velocidad de conexión Ancho de banda Tasa de transmisión Tiempo de respuesta Disponibilidad del servicio Rendimiento	Entrevista Software

Fuente: Elaboración Propia

3.2. TIPO DE ESTUDIO Y DISEÑO DE INVESTIGACIÓN

3.2.1. TIPO DE INVESTIGACIÓN

La siguiente investigación es de tipo de estudio descriptiva, porque tiene como finalidad evaluar la capacidad de las líneas de comunicación de la red de datos de la Universidad de Lambayeque aplicando la teoría de Shannon. La finalidad de la siguiente investigación es descubrir nuevas técnicas con estudios ya realizados para poder solucionar los problemas mencionados como lo que se exponen en la presente tesis.

3.3. POBLACIÓN Y MUESTRA EN ESTUDIO

3.3.1. POBLACIÓN

Para la presente investigación las unidades de análisis objeto de estudio están ubicadas en la Universidad de Lambayeque. La población está conformada por todos los equipos de cómputo, 130 equipos de

cómputo académicos y 80 administrativos sumando un total de 210; que tienen acceso a muchos servicios, siendo estos suficientes a objeto de evaluar la capacidad de las líneas de transmisión de la red de datos en la presente investigación.

3.3.2. MUESTRA

La muestra que se presenta es total de la población que constan de 210 computadoras.

3.4. MÉTODOS TÉCNICA E INSTRUMENTO PARA LA RECOLECCIÓN DE DATOS

Para el análisis descriptivo de la investigación se aplicó:

3.4.1. Análisis documental

La finalidad de la aplicación de esta técnica fue para recolectar la información relacionada con la descripción de las especificaciones y características técnicas de los equipos y de los canales de transmisión evaluados.

3.4.2. Observación

Esta técnica se aplicó para recopilar los datos de la ejecución del software especializado de análisis de la red de datos.

3.4.3. Trabajo en escritorio

Esta técnica se aplicó para recopilar los datos obtenidos de los cálculos utilizando las fórmulas de la Teoría de Shannon.

3.4.4. Entrevista

Entrevista realizada al encargado de la administración de la red de la universidad de Lambayeque para obtener respuestas que nos ayudaron a entender la realidad problemática de la actualidad.

IV RESULTADOS Y CONCLUSIONES

4.1. ANÁLISIS DE LA INFRAESTRUCTURA DE LA RED DE DATOS EN LA UDL

4.1.1. Análisis de la infraestructura del servicio de telefonía

La infraestructura del servicio de telefonía en la Universidad de Lambayeque no cuenta con telefonía IP; es por ese motivo que funciona con anexos de telefonía.

En la figura muestra como está diseñado la red de telefonía en la universidad de Lambayeque.

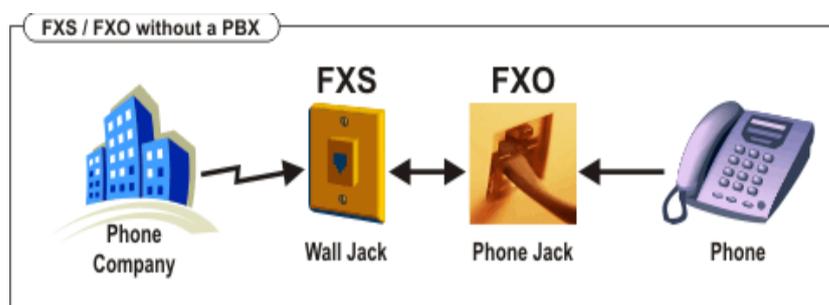


Figura 10. <https://www.3cx.es/voip-sip/fxs-fxo/>.

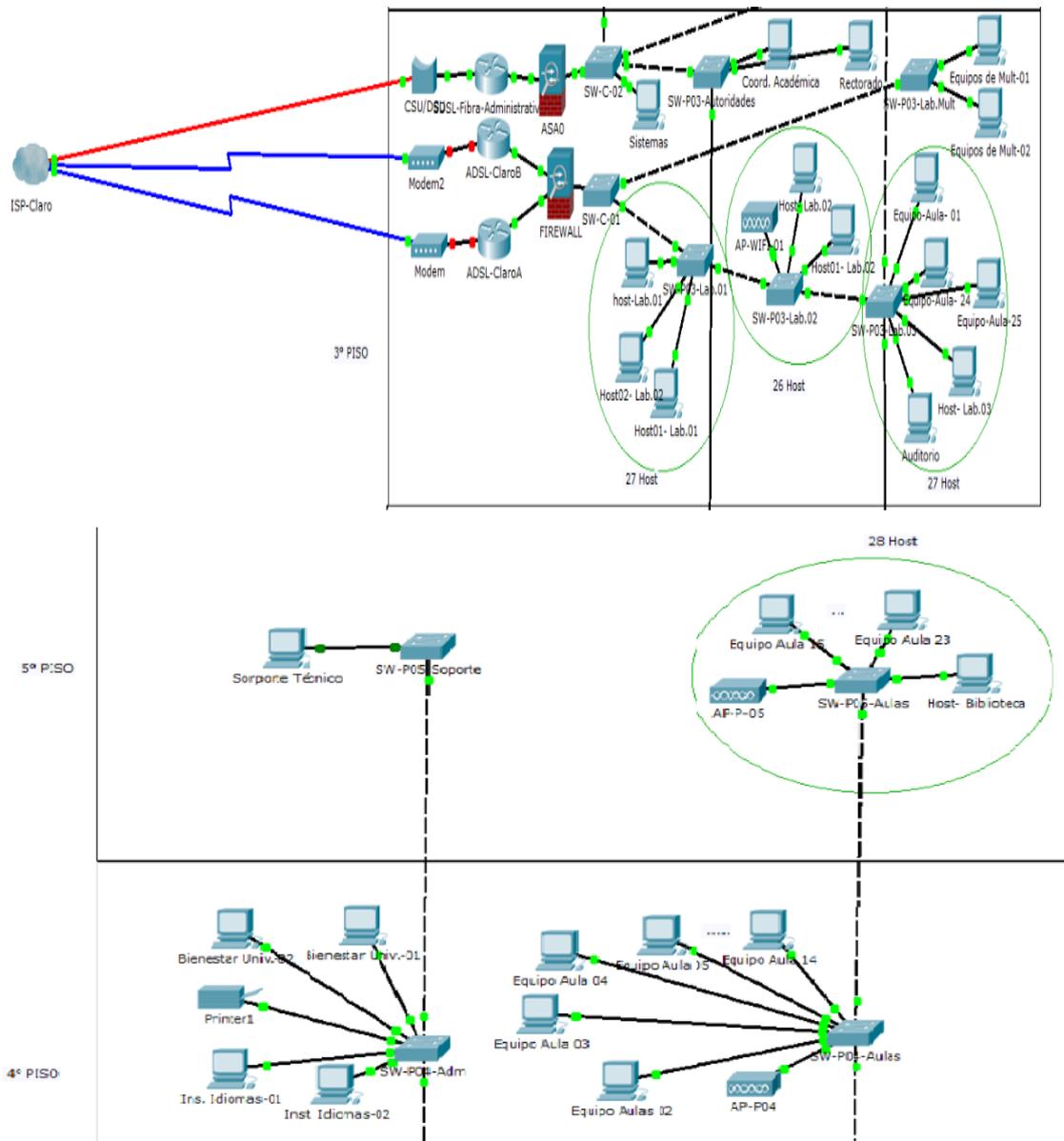
4.1.2. Análisis de la infraestructura de la red de comunicación de datos

La infraestructura de telecomunicaciones se soporta sobre los siguientes componentes:

- a. Sistema de cableado estructurado con backbone de fibra óptica bajo el estándar 802.3ae con un soporte de tasa de transferencia máxima de 10Gbps y subsistema de cableado horizontal con medio de transmisión que soporta 1 Gbps en la capa de acceso, cumpliendo los estándares internacionales para sistemas de cableado estructurado, lo que garantiza el uso de múltiples aplicaciones sobre la misma infraestructura de telecomunicaciones.
- b. Sistema Activo de comunicaciones, con tecnología de conmutación de datos distribuida en un sistema jerárquico en 2 capas preparada para actualizarse a 3 capas. Implementan una red escalable y con capacidades de redundancia, disponibilidad y administración. No utiliza el 100% de la capacidad de tráfico que puede soportar la red del Campus Universitario.

- c. Sistema Activo de comunicaciones con equipos de enrutamiento de alta confiabilidad, redundancia, disponibilidad y administración.
- d. Sistema de seguridad con hardware y software de protección en la frontera de la red pública y privada, así como la disponibilidad de herramientas de seguridad en todos los niveles del modelo de red TCP/IP de la red privada de la Universidad.
- e. Data Center con capacidad de soportar hardware de servidores adicional para el procesamiento y almacenamiento de nuevas aplicaciones.
- f. Sistemas de refrigeración y control de temperatura, protección y sistema de suministro eléctrico con backup de UPS's y generador de energía en condiciones operativas.
- g. Personal técnico capacitado y dedicado a la administración permanente de la red de comunicaciones IP.
- h. El sistema de cableado estructurado cumple con las normas ANSI/TIA/EIA-568 para cableado de telecomunicaciones, ANSI/TIA/EIA-569 rutas, espacios y canalizaciones para telecomunicaciones, ANSI/J-STD-607 aterramientos para los sistemas de telecomunicaciones.

Infraestructura de la red de datos de la universidad de Lambayeque



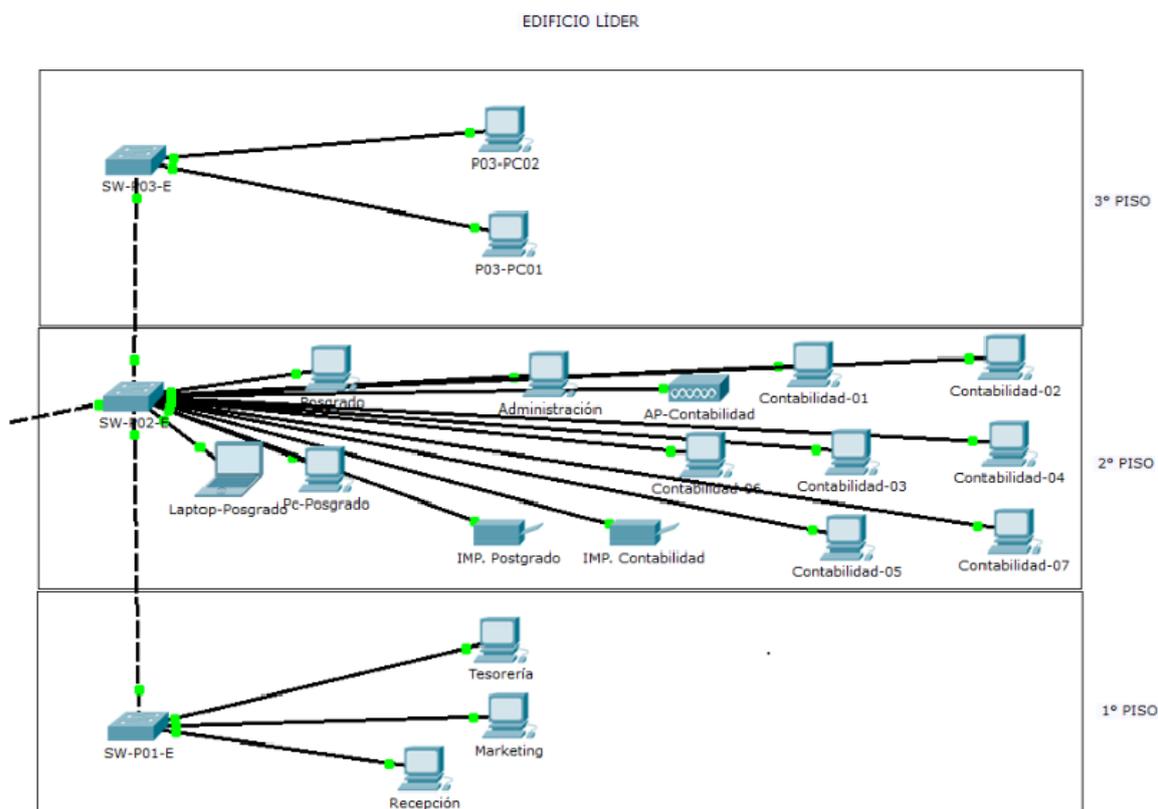


Figura 11. Estructura de red UDL

Fuente: propia

4.1.2.1. Backbone de Fibra Óptica

El backbone de la red de la Universidad está implementado bajo el estándar 802.3ae con Fibra Óptica multimodo y monomodo que conforman el cableado vertebral de campus y edificio.

El cableado vertebral de campus y edificios, se extiende desde el marco de conexión principal (MC) o distribuidor de campus (CD) (Nodo T – Sala de equipos) hacia los marcos de conexión horizontal (HC) o distribuidores de piso (FD) desde donde se extiende el subsistema de cableado horizontal con cable par trenzado UTP categoría 6, en cada dependencia principal de la Universidad.

El cableado de fibra óptica integra todos los laboratorios y dependencias principales de la Universidad dentro del Campus Universitario.

El medio de transmisión de fibra óptica tiene especificación propietaria del fabricante Siemon XGLO, el cual es un cable de fibra óptica Ethernet 10 Gigabit Ethernet mejorado sobre la especificación IEEE 802.3 ae para fibra óptica multimodo 50/125 μm y para fibra óptica monomodo (estándar especificado en la Subcapa de Control de Acceso al Medio – MAC, IEEE 802.3, por medio de Acceso Múltiple con Detección de Portadora y Detección de Colisiones - CSMA/CD).

El cable instalado considera cables de 6 hilos de fibra de tipo monomodo (número de parte: 9F5LB1-6B) y cables de 6 fibras de tipo multimodo (número de parte 9F8LB1-6B). Asimismo, el backbone está conformado por cables de fibra óptica con especificaciones 9F5LB1-24B (multimodo) y 9F8LB1-6B (monomodo) con 24 hilos de fibra óptica.

La instalación de fibra óptica en la Universidad cuenta con los test satisfactorios de certificación que garantizan el tráfico esperado definido por su especificación (tal como muestran los informes de certificación de Sistema de cableado estructurado: Test de conectorización y empalme de la fibra óptica instalada en la Red Telemática).

Instalación fibra óptica en Data Center de la Universidad.

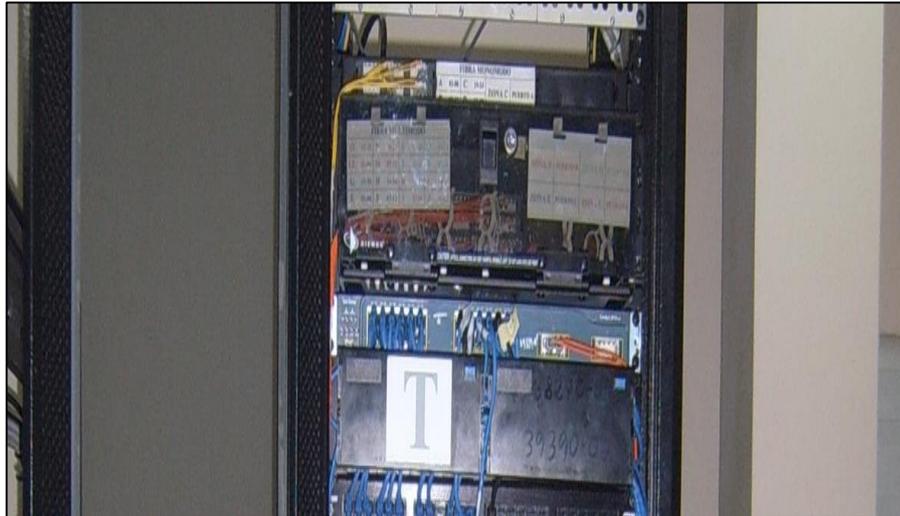


Figura 12. Marco de conexión principal de fibra optica

Fuente: Red UDL

4.1.2.2. Subsistema de cableado horizontal

El cableado horizontal de la Red Telemática, integra las dependencias de la Universidad y se extiende desde los gabinetes de telecomunicaciones, desde el marco de conexión horizontal hasta las salidas de telecomunicaciones en cada área de trabajo.

Este sistema está implementado con medio de transmisión cable UTP categoría 6 del fabricante Siemon Company en su especificación propietaria System 6, el cual asegura un ancho de banda de hasta 250 MHz permitiendo el soporte para tasas de transferencia mayores a 1 Gbps. El medio de transmisión y las canalizaciones que protegen el medio en el subsistema de cableado horizontal hasta el área de trabajo, están adecuadamente instaladas de acuerdo a las normas ANSI/TIA/EIA-568 y ANSI/TIA/EIA-569

En el área de trabajo, la Universidad cuenta con un total de 780 salidas de telecomunicaciones instalados con test satisfactorios de certificación, dentro de los cuales están considerados las oficinas que requieren del servicio de voz en un futuro (de acuerdo a los informes de certificación de Sistema de cableado estructurado).

Los test de certificación permiten garantizar el cumplimiento de las normas en las instalaciones del cableado de LAN a las propuestas de Categoría 6 de TIA e ISO. Estos test se hicieron con el equipo de certificación Agilente WireScope 350, el cual es una herramienta de certificación de cable de alta performance. Estas salidas de telecomunicación (780 puntos de red) están instaladas en las oficinas principales de todas las dependencias de la Universidad, quedando por integrar a la red de

comunicaciones a otras oficinas administrativas y laboratorios de cómputo cuya integración a la red de voz aún no es prescindible.

En consecuencia, luego de analizar el requerimiento de terminales de voz en la Universidad, todas las oficinas que requieren necesariamente del servicio de voz cuentan ya con la salida de telecomunicación estándar que los integra a la red de comunicaciones IP.



Figura 13. Marco de conexión principal de fibra óptica

Fuente: Red telemática

4.1.2.3. Infraestructura de equipos activos de comunicaciones

Los equipos activos de comunicaciones IP de la Red Telemática de la Universidad se distribuyen en un diseño de red modular Collapsed Core en 2 capas (Core y Acceso) con conexión e integración de los diferentes tipos de hosts (servidores, computadoras, impresoras) en la capa de acceso tratando de asegurar la performance y escalabilidad sin afectar el rendimiento de las diferentes redes conformadas.

Los equipos de seguridad permiten proteger la red privada de la red pública a través de software y hardware de supervisión, detección y filtrado de paquetes. La red

jerárquica tiene potencial para soportar la implementación de seguridad en todas las capas del modelo de red TCP/IP. El direccionamiento IP se distribuye por VLAN's y se asigna los diferentes recursos de la red a VLAN's específicas para una mejor administración.

La redundancia de los principales servicios y componentes hardware como el switch core (redundancia en componentes del mismo equipo), servidores de administración del dominio y el internet permiten mantener una disponibilidad del 99.900% en el año, de todos los servicios de red. Sin embargo, las redes de telefonía IP necesitan, adicionalmente, tener garantizada la disponibilidad del suministro eléctrico también en el área de trabajo, dado que un dispositivo hardware de voz requiere de por lo menos 15 watts de energía. La Universidad no cuenta con un generador de energía que garantice alta disponibilidad del suministro eléctrico.

Todos los equipos de acceso y Core (de conmutación, enrutamiento y seguridad) están preparados para soportar el tráfico de diferentes tipos de aplicaciones IP y tienen características que soportan la administración específica del tráfico de voz.

4.1.2.4. Capa Core: Switch Cisco Catalyst 4507R

El Switch Cisco Catalyst 4507 R de la Universidad de Lambayeque es un equipo de comunicación modular multicapa que concentra la conmutación central en el backbone de fibra óptica. En el modelo Jerárquico de 3 capas, este Switch estaría cumpliendo la función de CORE por ser el núcleo de la red, sin embargo, la funcionalidad actual del dispositivo en la red UDL, muestran su desempeño en las 2 capas jerárquicas: Core y Acceso lo

que la hace parte de una red modular Collapsed Core en 2 capas con capacidad de aprovechar las funcionalidades de esta capa en una red de voz y datos.

Este switch central soporta todas las conexiones de fibra óptica que vienen desde el marco de conexión principal del sistema de cableado estructurado administrando todo el tráfico de red de las diferentes dependencias de la Universidad. Cuenta con 2 fuentes de alimentación eléctrica, dos tarjetas supervisors (1 activa y 1 en standby) y disponibilidad de 1 tarjeta de 06 interfaces de fibra óptica lo cual le permiten operar con la redundancia que requiere ante problemas de hardware manteniendo la disponibilidad de los servicios de redes que soporta. Ante una falla mayor, en la tarjeta principal del switch por ejemplo, la redundancia de éste es prácticamente nula siendo un aspecto crítico a considerar dado su carácter nuclear en la red. Sin embargo, el registro de falla y de errores de hardware del switch es nulo.

Soporta la base de datos principal de VLAN's administradas en la red corporativa permitiendo hasta un máximo de 4096 VLAN's, Listas de control de acceso de las redes definidas en las VLAN's.

4.1.3. Dispositivos de enrutamiento e Internet de Banda Ancha

El acceso al servicio de Internet se implementa a través de 3 enlaces. 2 Routers Cisco modelo 2821 y 1 router Cisco modelo 2801 conforman los dispositivos front end hacia la red de internet con enlaces de 10, 2 y 2 Mbps respectivamente.

Un primer router principal Cisco 2821 es conectado a un dispositivo OptiX Metro1000 de la marca Huawei con tecnología de transporte óptico SDH el cual permite el enlace dedicado de banda ancha de 10 Mbps a internet con overbooking de 1:1

nacional e internacional a través de fibra óptica. Sin embargo, este equipo también permite multiplexar varias señales correspondientes a otros servicios sobre un mismo medio de transmisión (en este caso el enlace fibra óptica) identificando entre ellos a los enlaces E1/T1; este es el enlace principal utilizado para los principales servicios web de la Universidad publicados hacia internet incluyendo la publicación de los resultados de exámenes de admisión.

El router Cisco 2821 presenta las siguientes características físicas y lógicas totalmente compatibles con VoIP incluyen el software Cisco CallManager Express para la administración de llamadas VoIP, lo cual justifica aún más su soporte para una eficiente administración del tráfico de voz por parte del router.

HSRP, ACLs, multiple T1/E1/xDSL connections.

Digital Voice Support: Up to 192 calls, Voice and Video Enabled VPN (V3PN) for quality-of-service (QoS), Cisco CallManager Express (CME) embedded in Cisco IOS Software that provides call processing for Cisco IP phones, including wired and cordless WLAN phones, converged IP telephony solution for up to 96 IP phones, allows network managers to provide scalable analog and digital telephony

Ability to connect up to 12 T1/E1s trunks, 52 foreign-exchange-station (FXS) ports, or 36 foreign-exchange-office (FXO) ports. supports more than 90 modules, including WICs, VICs, network modules, PVDMs, and AIMS, Four integrated HWIC (High-Performance WIC) slots, Enhanced Network-Module (NME) Slots (The NME slots support existing network modules), HWICs slots can also support WICs, VICs, and VWICs, Extension-Voice-Module (EVM) Slot (The EVM supports additional voice services and density without consuming the network-module slot)

Multiprotocol Label Switching (MPLS) VPN Support, Packet Voice DSP Module (PVDM) Slots integrated on the Motherboard,

IP Phone Support: Optional support for Cisco in-line power distribution to Ethernet switch network modules and HWICs can be used to power Cisco IP phones; EVM Module Slots: provide support for the Cisco High-Density Analog and Digital Extension Module for Voice and Fax, providing support for up to 24 total voice and fax sessions without consuming a Network Module Slot.; PVDM (DSP) Slots on Motherboard: DSP (PVDM2) modules deliver support for analog and digital voice

Conferencing, transcoding, and secure Real-Time Transport Protocol (RTP) applications; Support for up to a 250 mailboxes using the Cisco Unity® Express voice messaging system is possible with the integration of an optional voice-mail AIM or network module

Amplia gama de interfaces de voz: Interfaces for public switched telephone network (PSTN), private branch exchange (PBX), and key system connections include FXS; FXO; analog direct inward dialing (DID); ear and mouth (E&M); Centralized Automated Message Accounting (CAMA); ISDN Basic Rate Interface (BRI); and T1, E1, and J1 with ISDN Primary Rate Interface (PRI); QSIG; E1 R2; and several additional channel-associated-signaling (CAS) signaling schemes; Survivable Remote Site Telephony (SRST).

El segundo enlace dedicado a internet se conecta a través de un router Cisco 2821 con conexión de fibra óptica a la red del proveedor con una tasa de transferencia máxima de 2 Mbps con overbooking de 1:1. Las características del Reuter que conecta a este enlace son las que se describieron anteriormente. No evaluamos este enlace debido a su uso restrictivo para aplicaciones específicas y a la transferencia de datos que requieren dedicación de ancho de banda.

El tercer enlace dedicado a internet se conecta a través de un router Cisco 2801 con conexión dedicada a la red del proveedor con una tasa de transferencia máxima de 2 Mbps con overbooking

de 1:2. Este enlace soporta un servicio de VPN el cual permite conexiones seguras en túneles con la red privada de la Universidad. Presenta las siguientes características compatibles con aplicaciones de voz incluyendo un software (Cisco CallManager Express) para la administración de llamadas VoIP.

Cisco Unified Communications Manager

02 puertos HWIC, 02 puertos Fast Ethernet, 01 puerto WIC, 01 puerto VWIC y VIC (sólo voz), 02 ranuras PVDM en placa para voz

02 módulos avanzados de integración

Soporte para Firewall de Cisco IOS, IPS, NAC y VPN SSL

VPN IPsec, hasta 1500 túneles, calidad de servicio (QoS) avanzada

802.1Q y 802.1X, SRST, correo de voz y PoE

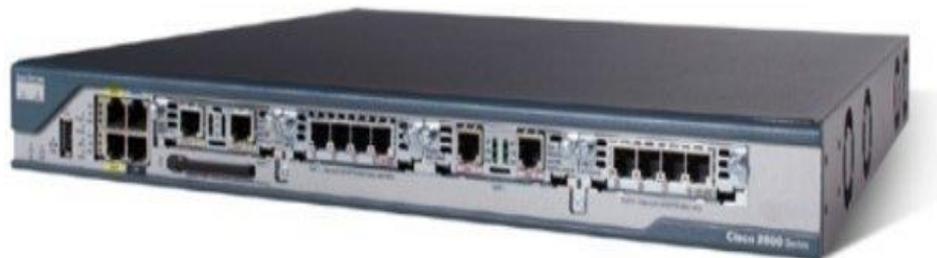


Figura 14. Router Cisco 2821

Fuente: Red UDL

4.2. Análisis de la capacidad del canal con la infraestructura existente

4.2.1. Identificación de los canales de comunicación

La red de datos en la Universidad de Lambayeque tiene una topología horizontal y estrella. En el Campus Universitario se encuentran ubicados nodos de borde. Cada nodo secundario se conecta al nodo principal mediante una F.O. de seis (06) hilos.

4.2.2. Análisis del tráfico y disponibilidad de recursos

Las herramientas software utilizadas para la recolección de datos durante varios períodos alternos de 2 semanas en horas pico de tráfico de red en los diferentes nodos de la Red Telemática de la Universidad son:

Cisco Network Assistant 5.3: Monitoreo y gestión de Redes Cisco.

Jperf 2.0.2 (Iperf): Medición del máximo rendimiento de ancho de banda

Mensajes de solicitud y respuesta de eco del Protocolo (ICMP).

A continuación, se muestra el resultado de mediciones realizadas para conocer el tráfico generado y el máximo tráfico soportado entre la capa de acceso correspondiente al Nodo A (edificio de Rectorado y Vicerrectorados) y la capa de acceso de otro nodo de referencia (nodo T).



Figura 15. Máximo tráfico soportado desde equipos activos entre nodos

Fuente: medición con software Jperf

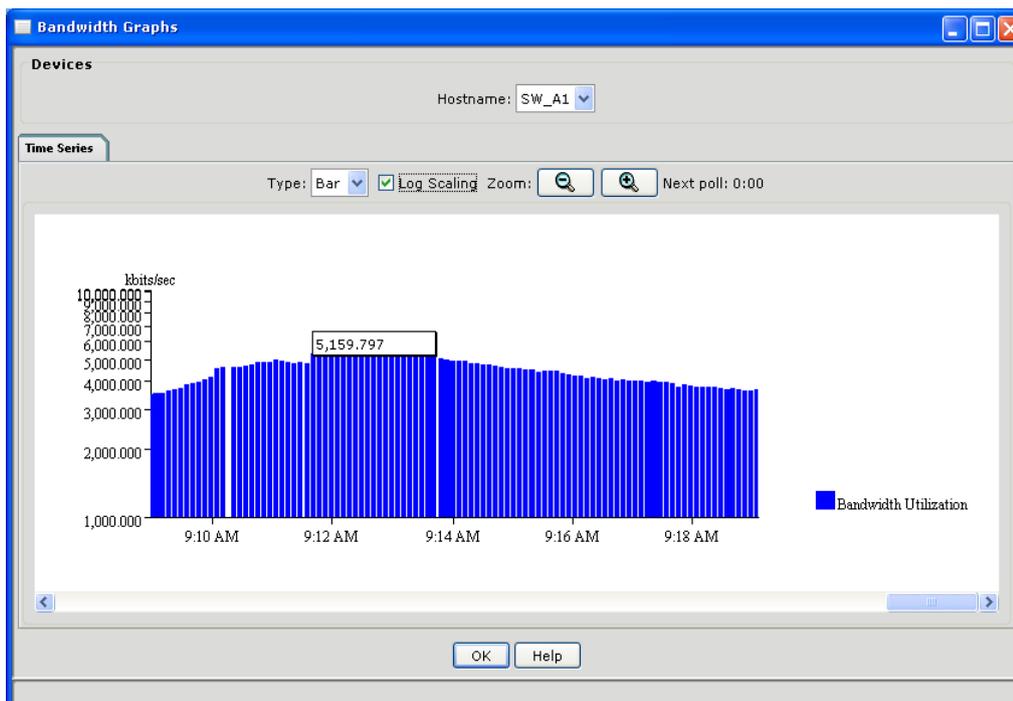


Figura 16. Utilización de ancho de banda desde el nodo A

Fuente: Medición con software Cisco Network Assistant

Obtenemos la latencia para las dependencias de cada Nodo de la Red Telemática con respecto a un nodo específico (Nodo T) enviando paquetes ICMP de solicitud y respuesta entre hosts de la capa de acceso de cada nodo.

El siguiente resultado muestra la latencia obtenida luego de enviar paquetes ICMP desde un host (pcAsecgnr01) del nodo A hacia un host (192.168.40.10) del nodo T:

```
administrador@pcAsecgnr01:~/Escritorio$ ping -c 5 192.168.40.10
```

```
PING 192.168.40.10(192.168.40.10) 56(84) bytes of data.
```

```
64 bytes from 192.168.40.10: icmp_req=1 ttl=63 time=0.359 ms
```

```
64 bytes from 192.168.40.10: icmp_req=2 ttl=63 time=0.278 ms
```

```
64 bytes from 192.168.40.10: icmp_req=3 ttl=63 time=0.276 ms
```

```
64 bytes from 192.168.40.10: icmp_req=4 ttl=63 time=0.261 ms
```

64 bytes from 192.168.40.10: icmp_req=5 ttl=63 time=0.265 ms

--- 192.168.40.10 ping statistics ---

5 packets transmitted, 5 received, 0% packet loss, time 4000ms

rtt min/avg/max/mdev = 0.261/0.287/0.359/0.042 ms

administrador@pcAsecgnr01:~/Escritorio\$

La capacidad disponible de procesamiento de los switches de acceso en cada nodo se obtiene de la información que muestra el software de gestión (IOS o Sistema Operativo de Internetworkin de Cisco) de los equipos en producción (en horas pico de tráfico) de la red de tal manera que se pueda garantizar la gestión de tráfico adicional para nuevas aplicaciones de red.

PROHIBIDO EL INGRESO A PERSONAL NO AUTORIZADO

ADMINISTRACION DE LA RED - UDL

SWITCH: NODO A ID:1

User Access Verification

Password:

SW_A1>en

Password:

SW_A1#

SW_A1#sh processes cpu

CPU utilization for five seconds: 23%/7%; one minute: 21%; five minutes:
21%

SW_A1#

```
SW_A1#sh processes memory
```

```
Total: 7675264, Used: 4916008, Free: 2759256
```

```
SW_A1#
```

```
PROHIBIDO EL INGRESO A PERSONAL NO AUTORIZADO
```

```
ADMINISTRACION DE LA RED - UDL
```

```
SWITCH: NODO B          ID:1
```

```
User Access Verification
```

```
Password:
```

```
sw_B1>en
```

```
Password:
```

```
sw_B1#sh processes cpu
```

```
CPU utilization for five seconds: 32%/13%; one minute: 30%; five minutes:  
30%
```

```
sw_B1#sh processes memory
```

```
Total: 7947968, Used: 4992584, Free: 2955384
```

```
PROHIBIDO EL INGRESO A PERSONAL NO AUTORIZADO
```

```
ADMINISTRACION DE LA RED - UDL
```

```
SWITCH: NODO TNC       ID:1
```

```
User Access Verification
```

```
Password:
```

```
sw_TNC1>en
```

```
Password:
```

```
sw_TNC1#sh processes cpu
```

CPU utilization for five seconds: 20%/8%; one minute: 17%; five minutes: 17%

sw_TNC1#sh processes memory

Total: 7669664, Used: 3590436, Free: 4079228

sw_TNC1#

De esta manera se detallan a continuación los resultados obtenidos para la capa de acceso de los nodos de la red de comunicaciones de la Universidad.

Tabla N° 3. Utilización de ancho de banda, latencia y confiabilidad de la red en la transmisión de paquetes

Nodo	Utilización de BW - kbps	Utilization de BW Máx (conec. all ports) - kbps	Tráfico máximo soportado entre Switches de acceso (kbps)	BW disponible (kbps)	% de Utilización de BW	RTT Round Trip Delay Time) máx. (ms)	Paquet esper didos	% CPU/ RAM disponible
1	4159.79	9,906.81	90,632	80,725	9	0.28	0%	77/3
2	9126.81	16,981.1	69,233	50,251	22	0.35	0%	68/3
3	2103.30	3,883.03	91,233	87,349	4	0.15	0%	91/6
4	6354.22	23,461.7	84,322	60,860	28	0.45	0%	73/3
5	57784.9	60,297.3	78,913	18,615	76	0.30	0%	58/3
6	3850.32	15,401.3	90,297	74,895	17	0.23	0%	79/5
7	1386.75	8,320.54	87,233	78,912	10	0.27	0%	75/4

Estos datos obtenidos, que procesamos y nos permiten afirmar que la infraestructura de Red puede soportar tráfico adicional generado por otros servicios de redes.

4.2.3. Análisis del Tráfico de datos en enlace Internet dedicado de 10 Mbps y 2 Mbps

Evaluamos la utilización de ancho de banda en cada enlace a internet en la Universidad. Los datos se obtienen de las mediciones realizadas al tráfico que gestionan los routers Cisco 2821 (10Mbps 1:1) y 2801 (2Mbps 1:2).



Figura 17. Utilización mensual de ancho de banda- línea dedicada internet 2Mbps

Fuente: Red UDL

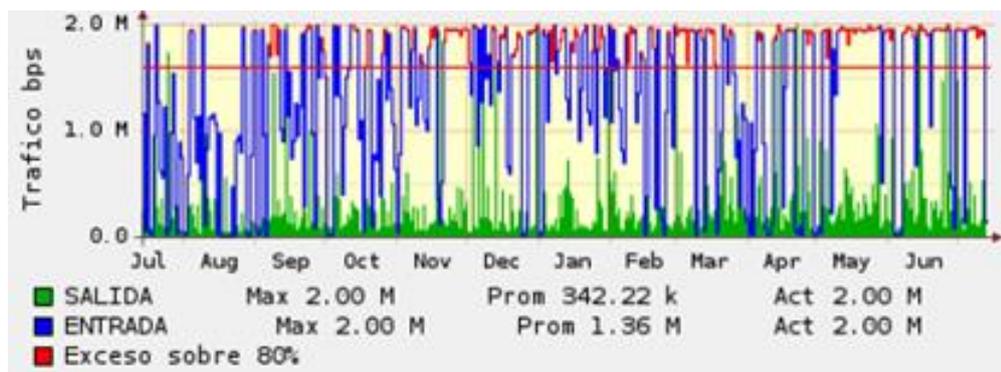


Figura 18. Utilización anual de ancho de banda- lineal dedicada internet 2Mbps

Fuente: Red UDL

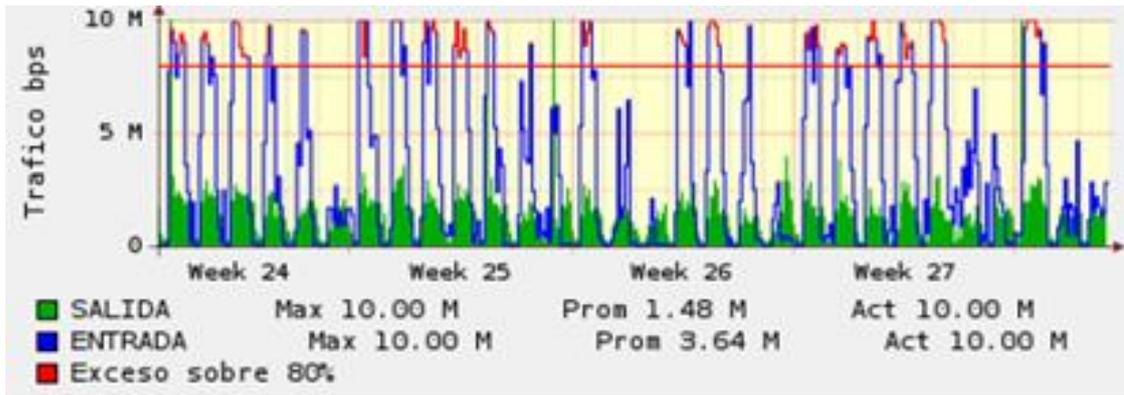


Figura 19. Utilización mensual de ancho de banda – lineal dedicada internet 10 Mbps

Fuente: Red UDL

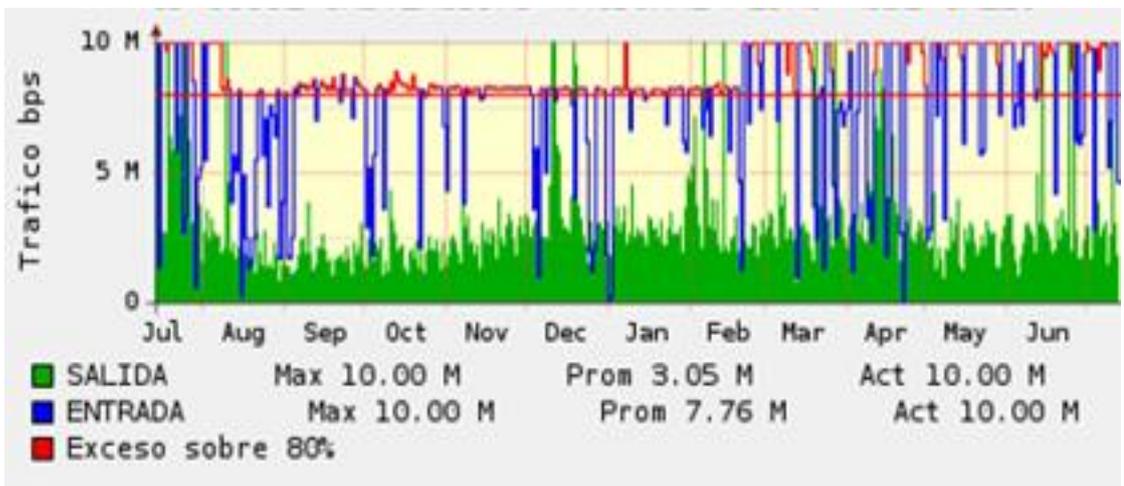


Figura 20. Utilización anual de ancho de banda – lineal dedicada internet 10 Mbps

Fuente: Red UDL

Tabla N° 4. Utilización de ancho de banda de internet en líneas dedicadas en la UDL

	Prom.Bandwidth In (uso en kbps)	Prom.Bandwidth Out (uso en kbps)	Prom. Total Bandwidth uso en kbps)	% utiliz.	Prom. Total Bandwidth (dispo.en kbps)	%dispon.
2Mbps mes	607.33	90.73	711.06	29.60	1136.94	50.40
2Mbps año	1392.64	342.22	1734.86	84.71	313.14	15.29
10Mbps mes	3527.36	1415.5	5242.88	51.20	4997.12	48.80
10Mbps año	7946.24	3123.2	11069.44	> 100	> 10240 (2Mbps)	0

De los datos registrados en cada enlace, podemos apreciar, que, en el análisis de un periodo mensual, ambos enlaces de internet (de 2 Mbps y 10Mbps), en promedio disponen de ancho de banda para la ejecución de otras aplicaciones. Siendo el enlace de 10 Mbps el enlace que mayor disponibilidad tiene (4997.12 kbps contra 1236.94).

Mientras que el enlace de 2 Mbps aún presenta ancho de banda disponible (17%) sin llegar a saturarse.

Esto se puede explicar considerando que el enlace de 10Mbps es el principal enlace para los procesos de Exámenes de Admisión.

Teniendo en cuenta esta información, podemos concluir en que la Universidad cuenta con un enlace dedicado de 2 Mbps, con un ancho de banda disponible mensual de 60.40%, lo que representa 1236.94 kbps, para la ejecución de otras aplicaciones sobre esta

línea de datos, la cual tomaremos en cuenta para la propuesta de tráfico a considerar sobre esta línea.

4.3. Análisis de la capacidad del canal con los valores teóricos obtenidos aplicando el Teorema de Shannon.

4.3.1. Cálculo de los valores de las variables de la investigación aplicando el Teorema de Shannon

a. Cálculo de la velocidad la transmisión

Para el cálculo de la capacidad teórica de transmisión se aplicará las fórmulas N° 11 y 13, es decir, el análisis de los canales de comunicación ideal y con ruido respectivamente.

Para análisis del canal sin ruido:

Se utiliza la fórmula: $C = 2Bn = 2B \log_2 M$ (bps)

Los parámetros a considerar son:

B: ancho de banda (Hz)

M: número de niveles posibles de la señal.

Tabla N° 5. Parámetros para el análisis de la velocidad de transmisión en un canal sin ruido

Parámetro	Valor	Características reales
Número de niveles posibles de la señal (M)	2	Los tramos con FO MM y FO SM son canales de datos. Se utilizan switchs para transmitir señales digitales. Por tanto, el número de niveles posibles de la señal es 2
Ancho de banda (B)	500 MHz	La FO MM de 1300 nm instalada tiene un ancho de banda de 500 a 600 MHz. En este caso se toma el valor de 500 MHz porque ninguno de los tramos supera el Km.

Aplicando la fórmula:

$$C = 2B \log_2 M \text{ (bps)}$$

$$C = 2 \cdot (500) \cdot \log_2(2)$$

$$C = 2 \cdot 500 \cdot 10^6$$

$$C = 1 \times 10^9$$

$$C = 1 \text{ Gbps}$$

La velocidad máxima de transmisión del canal sin ruido (ideal) es de 1 Gbps

Para el análisis del canal con ruido

Se utiliza la fórmula: $C = B \log_2 (1+S/N)$ (bps)

Para ello, se necesita conocer los siguientes parámetros:

Tabla N° 6. Parámetros para el análisis de la velocidad de transmisor en un canal con ruido

Parámetro	Valor	Observación
Ancho de banda (B)	500 MHz-Km	La FO MM de 1300 nm instalada tiene un ancho de banda de 500 MHz En este caso se toma el valor de 500 MHz porque ninguno de los tramos supera el Km
potencia de la señal (S)	5 dB (amplifican en 5 veces la señal)	En este caso la potencia de señal es de ganancia debido a la amplificación de la señal que generan los switch (transmisor) para transmitir la señal. Como todos los tramos utilizan los mismos switch, las potencias de señal son las mismas. La potencia de señal se calcula entre dos puntos. Se utiliza la fórmula $10 \log_{10} (p1/p2)$ dB.
potencia del ruido presente en el canal (N)	10 dB	La señal recibida en los switch es una señal modificada debido a las distorsiones introducidas en la transmisión, ya sea en el switch emisor o en el receptor.

Aplicando la fórmula:

$$C = B \log_2 (1+S/N) \text{ (bps)}$$

$$C = (500 * 10^6) * \log_2(1+10/3)$$

$$C = (500 * 10^6) * 1,4150$$

$$C = 0.7075 \text{ Gbps}$$

La velocidad máxima de transmisión del canal con ruido es de 0,70 Gbps

En la siguiente tabla se muestran las comparativas de las velocidades de transmisión obtenidas en la investigación, primero, con el software especializado de análisis de redes de datos y, segundo, con la aplicación de las fórmulas de la Teoría de la Información de Shannon, el cuál será nuestro modelo teórico de referencia.

Tabla N° 7. Comparativa de las velocidades de transmisión real y teórica (Shannon)

Nodo	Datos reales	Modelo Teórico de Referencia (Teoría de la Información de Shannon)		Comparativa	
	Velocidad de transmisión Gbps	Velocidad de transmisión Teórica en canal sin ruido Gbps	Velocidad de transmisión Teórica en canal con ruido Gbps	Diferencia en Velocidad de transmisión en canal sin ruido %	Diferencia en Velocidad de transmisión en canal con ruido %
1	0.425	1	0.70	34.00	60.71
2	0.932	1	0.40	83.56	129.1
3	0.210	1	0.70	17.53	24.74
4	0.6354	1	0.70	52.9	64.75

5	1.778	1	0.70	80.15	128.6
6	0.374	1	0.70	32.08	45.9
7	0.131	1	0.70	9.56	13.3

b. Cálculo del uso del ancho de banda

Para el cálculo del ancho de banda utilizado teóricamente, se debe tener en cuenta que es el ancho de banda óptico (se descarta el ancho de banda eléctrico).

Para ello, se considera que:

Ancho de banda óptico = B * longitud del cable instalado

B: es el ancho de banda referencial en las especificaciones técnicas del cable. En este caso es de 500 MHz-Km.

Tabla N° 8. Calculo del ancho de banda de los tramos de los canales de transmisión analizados

Tramo	Distancia del switch principal al switch de borde (tendido de FO en Km)	Ancho de banda óptico (en MHz)
Edificio líder – área adm.	0,180	108.00
Edificio líder-laboratorios	0,233	139.80

c. Cálculo del ruido generado en la transmisión

Dado que el switch funciona cumple una función de amplificación de la señal, el ruido que genera, es un ruido por amplificación, que depende de dos parámetros:

G: ganancia que genera el amplificador

Teq: Temperatura equivalente de ruido que modela el ruido que genera el amplificador. Viene dada por el fabricante conocido como Factor de ruido)

Por tanto, ruido (n) es:

$$n = g \cdot k \cdot T_{eq} \cdot B$$

g: ganancia que genera el amplificador

k: Constante de Boltzman ($1,38 \times 10^{-23} \text{ w/}^\circ\text{K Hz}$)

B: Ancho de banda

Aplicando la fórmula, se tiene:

$$n = 3 \text{ dB} * (1,38 \times 10^{-23} \text{ w/}^\circ\text{K Hz}) * 293^\circ\text{K} * (600 \text{ MHz})$$

$$n = 0,005 \text{ dB}$$

V. DISCUSION

La problemática siempre ocurre en cada centro de base de datos entre equipos informáticos ya que no se hace el estudio de datos de transmisión entre máquinas y siempre suele colgarse el sistema en momentos indicados y esto nos genera percances.

En relación a la tesis se están haciéndose mediciones en los equipos informáticos para luego ser comparados si la fórmula de Shannon es accesible para su aplicación.

Se tomó la teoría de Shannon porque a través de ello nos brinda la fórmula, donde se utilizaron datos de transmisión de comunicación entre equipos informáticos.

Al medir datos se tomarán para su aplicación y se obtendrá un resultado de los datos reales que esto nos da mejor alcance para solucionar problemas y llegar a los objetivos propuestos.

VI. CONCLUSIONES

En relación de la capacidad de las líneas de comunicación.

La teoría de la información da solución a la existencia de volúmenes de flujo de información muy difíciles de manejar y administrar. Y es especialmente importante que estos desarrollos contemplan al aspecto matemático formal de la noción de la información, de forma tal que estas se puedan medir y cuantificar para ser aplicadas con fines tecnológicos. Para evaluar la capacidad de las líneas de comunicación de la red de datos de la Universidad de Lambayeque se aplicó el modelo matemático de Shannon, debido a que los canales de comunicación analizados se encuentran asociados a la radiación electromagnética, generando ruido gaussiano. No podemos tener comunicación eléctrica sin electrones, y es imposible evitar el ruido. Es decir, se trata de canales de comunicación con ruido. $C = B \log (1+S/N)$ (bps)

Comparamos valores teóricos aplicando el teorema de Shannon donde el valor del canal de transmisor con ruido debe alcanzar es 0,70 Gbps

Mediante la apreciación de la ecuación llamada la Razón Nyquist que media la razón de transmisión de la señal en bauds. La razón de Nyquist es igual a $2B$ símbolos (o señales) por segundo, donde B es el ancho de banda del canal de transmisión. Así, usando esta ecuación, el ancho de banda de un canal telefónico de 3,000 Hz puede transmitido hasta $2 \times 3,000$, o 6,000 bauds o Hz.

En relación a la capacidad de un canal que puede transportar información.

Ancho de banda es una medida de la capacidad de un canal de comunicación que depende de las distancias del tendido del medio físico, en este caso la fibra óptica, los cálculos realizados para estimar

el ancho de banda, se realizó para los tramos donde la red tiene tendido FO.

El ancho de banda utilizado para las comunicaciones entre los puntos de los tramos analizados de la red de Datos en esta investigación, demuestran que el ancho de banda del tendido de FO está siendo subutilizado, debido a que en ninguno de los casos se llega a superar ni el 30% del ancho de banda del canal.

a. Para el caso del ruido generado por los switch, sobre la señal de entrada, se desprecia los siguientes casos:

Ruido térmico debido a la temperatura es despreciable, porque los switch no tienen resistencias que generen altas temperaturas.

Ruido de intermodulación. No ocurre porque la FO instalada tiene un hilo para envío y otro para recepción. Por tanto, no ocurren frecuencias distintas en un mismo canal.

Atenuación, debido a que la distancia en el tendido de la FO es despreciable

Se considera el ruido generado por la ganancia que el switch para amplificar la señal recibida. En este caso, el valor obtenido es $n = 0,005$ dB, el cual es un valor despreciable.

VII. RECOMENDACIONES

Se debe de evaluar constantemente la capacidad de líneas de transmisor de la red de datos de la universidad de Lambayeque; adquirir software analizador de red de última generación.

Aplicar el método matemático para determinar o hacer diagnósticos para red de datos.

Las variables que los teóricos de la información calculan sirven para medir y verificar las condiciones óptimas para la transmisión de mensajes.

Al número más reducido posible de señales se ajuste la mayor cantidad de mensajes diferentes posible.

Que se seleccione el canal más apropiado para hacer circular más información y al mayor número de usuarios.

Que se implemente equipos más rentables para el mayor número posible de mensajes y con la menor ambigüedad en su determinación.

Se deben limitarse al mínimo los «riesgos» de la distorsión y del ruido en la transmisión, provocados por aquellos fenómenos con los que forzosamente hay que contar al no poder ser eliminados de la naturaleza.

VIII. REFERENCIAS BIBLIOGRÁFICAS

- C.J. Date (2001), Introducción de los sistemas de base de datos, (7a. ed), Editorial Pearson Educación, México.
- Julio, I. de Vicente M. (2008) Medidas de Información, incertidumbre y entrelazamiento en Mecánica Cuántica, Tesis Doctoral en Matemáticas, Universidad Carlos III de Madrid España.
- Moreno, R. (2013). Análisis, diseño implementación dirigida para el área de Ventas y recursos humanos de una empresa exportadora e importadora de productos alimenticios, Tesis para optar el título profesional de Ingeniero informático, Póntica Universidad Católica del Perú, Lima, Perú.
- Br. Arroyo Raymundo, Héctor Esteban (2017), Propuesta de un Datamart para mejorar el proceso de toma de decisiones la empresa Cresko, 2016, Tesis para optar el título profesional de Ingeniero de Sistemas e Informática, Universidad Norbert Wiener Lima, Perú
- Ronald J Tocci, Neals. Widmer, Gregory L. Moss (2007), Sistemas Digitales principios y aplicaciones (10ma. ed.), Editorial Pearson Educación de México.
- Behrouz A. Foruzan (2002), Transmisión de datos y redes de comunicación (2a. ed), Editorial MCRAW HILL/ Interamericana de España.
- Shannon, C. (1948), 'A mathematical theory of communication', Bell System Technical Journal.
- Armando Urdaneta (2005), realizó una investigación titulada Análisis de Tráfico en una Red LAN aplicando la Tecnología de Redes Neuronales
- Ana Cristina Guzmán (2010), realizo su Tesis titulada "Diseño y Evaluación del Tráfico de una Red LAN para la empresa SERVTEC S.A"

Flores Heroy (2003), Optimización de una Red de Transporte de Telecomunicaciones, el propósito de esta investigación fue evaluar y determinar, mejor forma de optimizar la red de transporte de telecomunicaciones de CANTV en la ciudad de Maracaibo.

Rincón C. Carlos A. (2003), realizó una investigación titulada Modelo Matemático para la Estimación del Performance de una Red Ethernet.

IX. ANEXOS

ANEXO: Hoja de encuestas.

Entrevista para el Administrador de la red de Datos de la Universidad de Lambayeque	
-------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------

Nombre del entrevistado: _____	Fecha de la entrevista: _____
--------------------------------	-------------------------------

Objetivo: Capacidad de comunicación en la transmisión de información en la red de la Universidad de Lambayeque.

<ol style="list-style-type: none">1. ¿Cuáles son los protocolos de la red?2. ¿Con cuántos megas cuenta actualmente la red y quien es su proveedor?3. ¿en dicha red cuanta información como máximo se envían y reciben en un día hábil?4. ¿hay algún software que mide la transmisión de la cantidad de datos?5. ¿con cuantas computadoras instaladas a la Red Telemática cuenta la Universidad actualmente?6. ¿cuentan con distribución de telefonía IP?

7. ¿Cuál es el sistema de cableado estructurado, cuantos Gbps soporta y cuál es su categoría?
8. ¿La red cuenta con una infraestructura de equipos activos de comunicación? ¿Cuáles son y que función cumplen?
9. ¿en algún momento la red telemática a sufrido alguna caída por saturación de información?

ANEXO: Foto de entrevista al administrador de la red

