



**UNIVERSIDAD NACIONAL DE CHIMBORAZO
FACULTAD INGENIERÍA
CARRERA ELECTRÓNICA Y TELECOMUNICACIONES**

**SIMULACIÓN EN NS3 Y ANÁLISIS DE SEÑALES DE
INFORMACIÓN EN UN SISTEMA DE TRANSMISIÓN
MOLECULAR BASADO EN LA EXPRESIÓN GENÉTICA**

**Trabajo de Titulación para optar al título de Ingeniero en
Electrónica y Telecomunicaciones**

Autor:

Malán Chacaguasay Bryan Paúl
Pino Urquizo Yeimy Alexandra

Tutor:

Yesenia Elizabeth Cevallos Villacrés. PhD

Riobamba, Ecuador. 2024

DERECHOS DE AUTORÍA

Nosotros, **Yeimy Alexandra Pino Urquizo**, con cedula de ciudadanía **0605325869** y **Bryan Paúl Malán Chacaguasay**, con cédula de ciudadanía **0605997394**, autores del trabajo de investigación titulado: **Simulación en NS3 y análisis de señales de información en un sistema de transmisión molecular basado en la expresión genética**, certifico que la producción, ideas, opiniones, criterios, contenidos y conclusiones expuestas son de nuestra exclusiva responsabilidad.

Asimismo, cedo a la Universidad Nacional de Chimborazo, en forma no exclusiva, los derechos para su uso, comunicación pública, distribución, divulgación y/o reproducción total o parcial, por medio físico o digital; en esta cesión se entiende que el cesionario no podrá obtener beneficios económicos. La posible reclamación de terceros respecto de los derechos de autor (a) de la obra referida, será de mi entera responsabilidad; librando a la Universidad Nacional de Chimborazo de posibles obligaciones.

En Riobamba, 8 de abril de 2024.



Yeimy Alexandra Pino Urquizo

C.I: 0605325869



Bryan Paúl Malán Chacaguasay

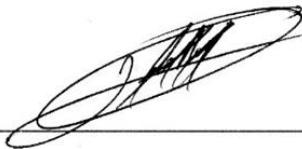
C.I: 0605997394

DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL;

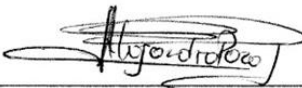
Quienes suscribimos, catedráticos designados Tutor y Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **Simulación en NS3 y análisis de señales de información en un sistema de transmisión molecular basado en la expresión genética**, presentado por **Bryan Paúl Malán Chacaguasay** con cedula de identidad número **0605997394** certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha asesorado durante el desarrollo, revisado y evaluado el trabajo de investigación escrito y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 8 de abril de 2024.

Ing. Leonardo Fabián Rentería Bustamante, PhD.
PRESIDENTE DEL TRIBUNAL DE GRADO




Ing. Alejandra del Pilar Pozo Jara
MIEMBRO DEL TRIBUNAL DE GRADO



Dr. Klever Hernán Torres Rodriguez
MIEMBRO DEL TRIBUNAL DE GRADO



Ing. Yesenia Elizabeth Villacrés, PhD.
TUTOR



DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL;

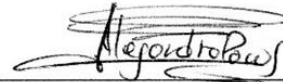
Quienes suscribimos, catedráticos designados Tutor y Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **Simulación en NS3 y análisis de señales de información en un sistema de transmisión molecular basado en la expresión genética**, presentado por **Yeimy Alexandra Pino Urquizo**, con cédula de identidad número **0605325869** certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha asesorado durante el desarrollo, revisado y evaluado el trabajo de investigación escrito y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 8 de abril de 2024.

Ing. Leonardo Fabián Rentería Bustamante, PhD.
PRESIDENTE DEL TRIBUNAL DE GRADO



Ing. Alejandra del Pilar Pozo Jara
MIEMBRO DEL TRIBUNAL DE GRADO



Dr. Klever Hernán Torres Rodriguez
MIEMBRO DEL TRIBUNAL DE GRADO



Ing. Yesenia Elizabeth Villacrés, PhD.
TUTOR



CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

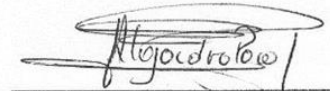
Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **Simulación en NS3 y análisis de señales de información en un sistema de transmisión molecular basado en la expresión genética** presentado por **Bryan Paúl Malán Chacaguasay**, con cedula de identidad número **0605997394**, bajo la tutoría de Ing. Yesenia Elizabeth Cevallos Villacrés. PhD; certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 8 de abril de 2024.

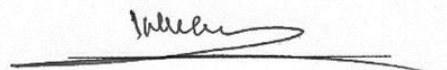
Presidente del Tribunal de Grado
Leonardo Fabián Rentería Bustamante,
PhD.



Miembro del Tribunal de Grado
Ing. Alejandra del Pilar Pozo Jara



Miembro del Tribunal de Grado
Dr. Klever Hernán Torres Rodruiguez



CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **Simulación en NS3 y análisis de señales de información en un sistema de transmisión molecular basado en la expresión genética** presentado por **Yeimy Alexandra Pino Urquizo**, con cédula de identidad número **0605325869**, bajo la tutoría de Ing. Yesenia Elizabeth Cevallos Villacrés. PhD; certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 8 de abril de 2024.

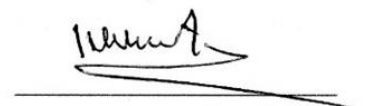
Presidente del Tribunal de Grado
Leonardo Fabián Rentería Bustamante,
PhD.



Miembro del Tribunal de Grado
Ing. Alejandra del Pilar Pozo Jara



Miembro del Tribunal de Grado
Dr. Klever Hernán Torres Rodruiguez





CERTIFICACIÓN

Que, **MALAN CHACAGUASAY BRYAN PAUL** con CC: **0605997394**, estudiante de la Carrera **ELECTRÓNICA Y TELECOMUNICACIONES, NO VIGENTE**, Facultad de **INGENIERÍA**; ha trabajado bajo mi tutoría el trabajo de investigación titulado "**SIMULACIÓN EN NS3 Y ANÁLISIS DE SEÑALES DE INFORMACIÓN EN UN SISTEMA DE TRANSMISIÓN MOLECULAR BASADO EN LA EXPRESIÓN GENÉTICA**", cumple con el 9 %, de acuerdo al reporte del sistema Anti plagio **TURNITIN**, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 21 de marzo de 2024



Firmado electrónicamente por:
**YESENIA ELIZABETH
CEVALLOS VILLACRES**

PhD. Yesenia Cevallos
TUTORA TRABAJO DE INVESTIGACIÓN



CERTIFICACIÓN

Que, **PINO URQUIZO YEIMY ALEXANDRA** con CC: **0605325869**, estudiante de la Carrera **ELECTRÓNICA Y TELECOMUNICACIONES, NO VIGENTE**, Facultad de **INGENIERÍA**; ha trabajado bajo mi tutoría el trabajo de investigación titulado "**SIMULACIÓN EN NS3 Y ANÁLISIS DE SEÑALES DE INFORMACIÓN EN UN SISTEMA DE TRANSMISIÓN MOLECULAR BASADO EN LA EXPRESIÓN GENÉTICA**", cumple con el 9 %, de acuerdo al reporte del sistema Anti plagio **TURNITIN**, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 21 de marzo de 2024



YVESNIA ELIZABETH
CEVALLOS VILLACRES

PhD. Yesenia Cevallos
TUTORA TRABAJO DE INVESTIGACIÓN

DEDICATORIA

A mi madre, que con amor y dedicación ha sido el pilar en mi formación como profesional íntegro. Le estoy profundamente agradecido por sus consejos y apoyo constante, que han iluminado cada paso de mi trayectoria.

A Gaby, quien reconoció en mí un potencial que yo mismo no veía, y que hasta el final de sus días me impulsó a ser mejor persona. Hoy concluye un proyecto que comencé con ella, y desde el cielo celebra nuestro éxito. Hasta después de muerto te quedará mi corazón despierto.

A mi familia y amigos, quienes estuvieron constantemente pendientes y me brindaron un apoyo inquebrantable, motivándome para que culminara con éxito esta etapa de mi vida.

Bryan Paúl Malán Chacaguasay

AGRADECIMIENTO

Agradezco a Dios por dotarme de sabiduría y salud, que me permiten luchar por mis metas y objetivos.

A mi madre, quien con su inquebrantable motivación me impulsó a dar lo mejor de mí. Sus consejos y dedicación han sido fuente de inspiración para convertirme en la mejor versión de mí.

A mi padre, por su apoyo a lo largo de mi trayectoria estudiantil.

A Gaby, agradezco profundamente el amor que me brindo y los valiosos consejos que guardaré siempre en mi corazón.

A mi hermano, quien me demostró que la madurez no está limitada por la edad y me ofreció su apoyo incondicional en todo momento.

A mi familia y amigos, cuyo apoyo moral y económico en los momentos más difíciles fue invaluable.

A los docentes de la carrera de Ingeniería en Electrónica y Telecomunicaciones, y en especial a la Ing. Yessenia Cevallos, PhD, gracias por su dedicación y guía constante. Su influencia ha sido fundamental en mi formación académica y profesional.

Bryan Paúl Malán Chacaguasay

DEDICATORIA

Este trabajo va dedicado a Corina Urquizo y Marcos Pino, quienes sacrificaron tanto para cumplir mis metas y que día a día me enseñaron a luchar por ello, el apoyo incondicional de ustedes es lo que me ayuda cada día a seguir creciendo y mejorando como persona.

A mis hermanos Marcia, Juan y Hernán por ser mis modelos, todo el apoyo y compañía ofrecidos en cada etapa de mi vida, han logrado que siempre tenga un motivo más para seguir adelante y convertirme en una gran ser humano.

A mis pequeños sobrinos sin su existencia hoy en día no sería ni la mitad de persona que soy, ustedes me han dado las mejores alegrías que he tenido y me han motivado siempre a ser un gran ejemplo y a luchar por conseguir mis objetivos.

Yeimy Alexandra Pino Urquizo

AGRADECIMIENTO

Agradezco a mis padres por todo el apoyo y las enseñanzas que me han brindado cada día de mi vida para poder cumplir todas mis metas.

A mis hermanos por siempre ayudarme en cada cosa que necesite y mostrarme un gran apoyo en cada decisión que he tomado.

A mis hermosos sobrinos por siempre hacerme reír en mis momentos tristes y hacer un poco más fácil seguir despertando cada día.

Agradezco infinitamente a mis maestros de la Universidad por enseñarme los conocimientos necesarios para ser una gran profesional y persona, en especial a nuestra tutora Ing. Yesenia Cevallos por su paciencia y dedicación con nosotros en todo momento, han sido claves para culminar esta etapa.

Por último, quiero agradecer a todos mis amigos que han estado conmigo siendo el apoyo en cada triunfo y el consuelo en cada derrota que he afrontado, gracias a ser las personas que han motivado a terminar esta etapa.

Yeimy Alexandra Pino Urquizo

ÍNDICE GENERAL;

DERECHOS DE AUTORÍA

DICTAMEN FAVORABLE DEL PROFESOR TUTOR

CERTIFICADO DEL TUTOR Y MIEMBROS DEL TRIBUNAL

CERTIFICADO DEL SISTEMA ANTIPLAGIO

DEDICATORIA

AGRADECIMIENTO

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

ABSTRACT

CAPÍTULO I. INTRODUCCIÓN.....	21
1.1 Planteamiento del Problema	23
1.2 Objetivos.....	23
1.2.1 General.....	23
1.2.2 Específicos.....	23
2.2 Comunicaciones Moleculares.....	26
2.3 Movimiento Browniano.....	27
2.3.1 Proceso de Wiener	27
2.3.2 Proceso de Wiener con arrastre (drift).....	29
2.4 Leyes de Fick.....	30
2.5 Definición de Nanomáquina.....	32
2.6 Ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN) como nanomáquinas biológicas.....	33
2.6.1 Roles y Funciones del ADN y ARN.....	33

2.7 Modelo de comunicación molecular basado en la expresión genética.....	37
2.8 La expresión genética y transporte de proteínas analizada como un sistema de comunicación digital.....	39
2.8.1 El Núcleo Celular como Equipo Terminal de Datos Biológicos (DTE)	39
2.8.2 Los ribosomas y el retículo endoplásmico (RE) como equipo de comunicación de datos biológicos (DCE)	41
2.8.3 El aparato de Golgi (AG) como enrutador de frontera de una Interred.....	45
2.8.4. Utilización de la información en el receptor	46
2.9 Tipos de Modulación	50
2.9.1 Modulación por desplazamiento de concentración (CSK).....	51
2.9.2 Modulación por desplazamiento de molécula (MoSK).....	51
2.9.3 Modulación basada en el radio molecular	52
2.9.4 Modulación por tiempo transcurrido (TEC).....	52
CAPÍTULO III. METODOLOGÍA.....	53
3.1 Tipo de estudio	53
3.2 Métodos de investigación	53
3.3 Fuentes de Información	53
3.4 Población y muestra	54
3.4.1 Población	54
3.4.2 Muestra	54
3.5 Procesamiento y análisis.....	54
3.6 Operacionalización de las variables	54
3.7 Procesamiento y Análisis.....	55
3.8 Modelo de implementación en NS-3	55
3.8.1 Conexión de nodos de comunicación y dispositivos	58
3.8.2 Clase MessageCarrier	59
3.8.3 Clase Perturbación (Perturbation)	60

3.8.4 Clase Motion	60
3.8.5 Clase Campo (Field).....	60
3.9 Representación binaria de la expresión genética.....	61
3.10 Diseño matemático	63
3.10.1 Modelo matemático implementado en el simulador.....	63
3.10.2 Modelo matemático para el contraste de resultados.....	64
CAPÍTULO IV. RESULTADOS Y DISCUSIÓN	66
4.1. Concentración molecular.....	66
4.2 Tiempo de delay	71
4.3 Amplitud máxima de pulso	72
4.4 Ancho del pulso molecular.....	73
4.5 Análisis estadístico	75
4.5.1 Recolección de datos	75
4.5.2 Prueba de normalidad.....	75
4.5.3 Análisis de las moléculas ligadas al receptor	76
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES	82
5.1 Conclusiones.....	82
5.2 Recomendaciones	84
ANEXOS	89
ANEXO I.....	90
ANEXO II	91
ANEXO III.....	94
ANEXO IV.....	95
ANEXO V	97

ÍNDICE DE FIGURAS

Figura 1. Comunicación Molecular.....	26
Figura 2. Proceso de Wiener con y sin arrastre. Velocidad de arrastre $v = 1 \text{ um/s}$ y $v = 0 \text{ um/s}$, respectivamente. En ambos casos, $D= 1 \text{ um}^2/\text{s}$, y $B(0)=0$	30
Figura 3. La transcripción produce ARNm a partir de una secuencia de una hebra de ADN.	35
Figura 4. La traducción produce una proteína a partir de una secuencia de ARNm.....	36
Figura 5. Marco biológico en la transcripción del ADN.....	41
Figura 6. Una interfaz física en un sistema digital puede ser análogo al citosol en un sistema biológico.....	42
Figura 7. Modelo de red en capas DTE-DCE de la transcripción y traducción del ADN..	45
Figura 8. Las funciones del AG como un router de interred.....	46
Figura 9. La equivalencia interna de redes de la expresión genética.....	49
Figura 10. La expresión genética como un modelo de red en capas.....	50
Figura 11. Funcionamiento de una modulación multinivel de CSK.....	51
Figura 12. Procesamiento y análisis de la investigación.....	54
Figura 13. Directorio del estándar p1906.1 y p19061.1.....	56
Figura 14. Diagrama UML del marco de la comunicación molecular.....	58
Figura 15. La expresión genética representada en bits.....	62
Figura 16. Transcripción del ADN, cuadro verde: codón del aminoácido (Prolina), cuadro rojo: decodificación binaria.....	63
Figura 17. Resultado de la consola de la recepción del mensaje.....	67
Figura 18. Concentración molecular en función de la distancia y el tiempo.....	68
Figura 19. Concentración Molecular.....	70
Figura 20. Tiempo de delay con variación de los coeficientes de difusión (D).....	72
Figura 21. Amplitud máxima de pulso en función de la distancia y las moléculas liberadas.	73
Figura 22. Ancho del pulso molecular en función de la distancia y los diferentes coeficientes de fusión.....	74
Figura 23. Mediana de la cantidad de moléculas que se ligan al receptor con respecto a los coeficientes de difusión.....	80
Figura 24. Descarga del simulador NS-3 (3.28.1.tar.bz2).....	90
Figura 25. Proceso para descomprimir un archivo tar.....	90

Figura 26. Descomprimir el archivo p1906.....	91
Figura 27. Inicializar el software NS-3.	92
Figura 28. Instalación de los archivos p1906.	92
Figura 29. Instalación del archivo p1906.	93
Figura 30 Instrucciones para Copiar el Archivo mol-example.cc al Directorio Scratch....	93
Figura 31. Ejecución del escenario.....	94
Figura 32. Creación del archivo .sh para la obtención de resultados.	95
Figura 33. Código del archivo .sh.	96
Figura 34. Extracción de datos.	97
Figura 35. Análisis ANOVA SPS.	97
Figura 36. Cuadro de dialogo del ANOVA de un factor.....	98
Figura 37. Cuadro de dialogo comparaciones post hoc.....	98
Figura 38. Cuadro de dialogo opciones ANOVA de un factor.	99

ÍNDICE DE TABLAS

Tabla 1. Operacionalización de las variables	54
Tabla 2. Codificación de los nucleótidos.....	61
Tabla 3. Parámetros generales de la simulación.....	66
Tabla 4. Valores de concentración obtenidos de los diferentes coeficientes de difusión...	71
Tabla 5. Valores del tiempo de delay obtenidos a los diferentes coeficientes de difusión.	72
Tabla 6. Valores del ancho del pulso molecular en función de los coeficientes de difusión.	74
Tabla 7. Prueba de normalidad Shapiro-Wilk.	76
Tabla 8. Análisis ANOVA.....	77
Tabla 9. Análisis Tukey.....	77

RESUMEN

En los últimos años las comunicaciones moleculares (CM) han ido tomando más relevancia e importancia en el campo de la medicina ya que algunas de las aplicaciones más importantes que han surgido es la administración de fármacos de forma eficiente para evitar efectos secundarios en el tratamiento de enfermedades. Para implementar CM, IEEE creo los estándares IEEE 1906.1 Y 1906.1.1 y se pueden ejecutar en el software NS-3. En este trabajo de investigación se realizó un sistema CM que responde a la expresión genética, haciendo uso de las herramientas que ofrece NS-3 se pudo simular la transcripción y traducción del ácido desoxirribonucleico (ADN) con una modulación MoSK (Molecular On Shift Keying), en el canal de transmisión se aplicó un movimiento browniano con arrastre que sigue las condiciones del flujo sanguíneo, para la recepción de la información que ingresa al canal de transmisión se implementó un receptor de tipo ligando-receptor con ello pudimos realizar el análisis del rendimiento del sistema de comunicación.

Palabras claves: Expresión genética, movimiento browniano con arrastre, modulación molecular MoSK, ligando-receptor, estándares IEEE 1906.1 y 1906.1.1.

ABSTRACT

In recent years, molecular communications (MC) have become increasingly relevant in the medical field, particularly for applications meant at the efficient delivery of drugs to mitigate side effects in disease treatment. The IEEE has played a crucial role in this arena by introducing standards such as IEEE 1906.1 and 1906.1.1, which can be implemented using the NS-3 software platform. This study presents the development of an MC system personalized to respond to gene expression, harnessing NS-3's capabilities to simulate the transcription and translation processes of deoxyribonucleic acid (DNA) with MoSK modulation (Molecular On Shift Keying). The transmission channel was modeled after Brownian motion with drift to mimic conditions akin to blood flow. Furthermore, a ligand-receptor receptor system was incorporated for information reception, facilitating the analysis of the system's performance.

Keywords: Gene expression, Brownian motion with drift, MoSK molecular modulation, ligand-receptor, IEEE 1906.1 and 1906.1.1 standards.

Reviewed and improved by: Professor. Armijos Jacqueline



CAPÍTULO I. INTRODUCCIÓN

Los estudios más recientes sobre la comunicación biológica sugieren que existe una gran similitud con los procesos que ocurren en la comunicación digital, ya que en las comunicaciones moleculares (CM) se utilizan componentes de transmisión similar a los presentados en el teorema de Shannon [1] que son: transmisor, canal de comunicación y receptor, en la expresión genética dichos componentes y sus analogías biológicas intercelulares son: Transmisor comprendido por la célula, el canal de comunicación representado con el torrente sanguíneo y el receptor es representado con el órgano objetivo. En este trabajo de tesis se realizará el estudio y simulación empleando NS-3 del proceso biológico que se conoce como expresión genética mediante el cual la información transportada por el ácido desoxirribonucleico (ADN), se transforma en las proteínas apropiadas para cumplir una función fisiológica en un órgano específico [2]. Este proceso se asemeja al transporte de información en un modelo de capas de red cuando existe un flujo de información de un computador transmisor hacia un computador receptor a través de conmutadores o routers.

Los procesos mediante los cuales la expresión genética se llevan a cabo son: transcripción y traducción del ADN. La transcripción es el proceso mediante el cual se transfiere la información contenida en la secuencia del ADN, divididos en bloques de nucleótidos llamados genes, mediante moléculas de Ácido Ribonucleico (ARN) como intermediario [3]. La traducción es el proceso mediante el cual la información genética se convierte en cadena de aminoácidos con funcionalidad biológica que permite la obtención de proteínas a través de tales aminoácidos [4].

La expresión genética biológicamente se establece en células de organismos eucariotas para transformar la información que contiene el núcleo celular que contiene el ADN y se transmite de varias formas a una célula, tejido y órgano destino fuera de la célula para poder cumplir una función vital para el organismo [5]. Cuando este tipo de comunicación falla existen graves enfermedades de tipo genético que incluso pueden causar la muerte [6].

En el canal de comunicaciones cuando la expresión genética debe enviar información mediante el sistema endocrino para poder alcanzar un órgano receptor de la información del ADN, la información molecular se propaga mediante el torrente sanguíneo el cual básicamente es un fluido que posee cierta velocidad de arrastre y transporta la información biológica [7].

El extremo comunicacional destino existen varias formas de procesar la información recibida, una de estas formas son los receptores biológicos denominados complejo ligando-receptor [8].

1.1 Planteamiento del Problema

Los estándares de nanocomunicaciones, particularmente de comunicaciones moleculares de IEEE correspondientes a 1906.1 y 1906.1.1 establece como software de simulación de redes de comunicación biológica a NS-3. En este trabajo de tesis se propone realizar por tanto la simulación en NS-3 y el análisis de sistemas de nanocomunicaciones que obedecen a la expresión genética, en los que la información es producida en el interior de una célula y debe llegar a un órgano destino por medio de un canal de comunicaciones molecular que en este caso es el torrente sanguíneo, se planea realizar la simulación y análisis en un lapso aproximado de 6 meses.

Teniendo en cuenta que la expresión genética determina las condiciones fisiológicas (variable independiente) para que se produzca el procesamiento, transmisión y recepción (estos tres procesos definen el comportamiento de las señales de información molecular y, por tanto, ya que dependen de las características de la expresión genética se tiene que considerar como variables dependientes) de moléculas de ADN desde el núcleo celular hasta un órgano destino.

1.2 Objetivos

1.2.1 General

- Simular y analizar los resultados de la expresión genética mediante el software de código abierto NS-3.

1.2.2 Específicos

- Simular (mediante el uso de librerías, punteros y demás herramientas de información establecidas en el simulador NS-3) y analizar la transferencia de información desde la fuente biológica mediante la transcripción y traducción de moléculas de ADN mediante modulación del tipo Molecular On Shift Keying (MoSK).
- Simular (mediante el uso de librerías, punteros y demás herramientas de información establecidas en el simulador NS-3) y analizar la transmisión de información en un canal de comunicaciones molecular con arrastre, cuyas condiciones de propagación son las del flujo sanguíneo.
- Simular (mediante el uso de librerías, punteros y demás herramientas de información establecidas en el simulador NS-3) y analizar la recepción de información en el extremo destino comunicacional, a través de receptores biológicos complejo ligando-receptores.

CAPÍTULO II. MARCO TEÓRICO

2.1 Estado del arte

Transmitir información a distancia siempre ha sido un problema presente en la sociedad humana. Hoy en día, se han resuelto algunos de estos problemas con los sistemas de comunicación modernos usando señales eléctricas o electromagnéticas de radio a bandas ópticas. Sin embargo, existen muchas aplicaciones especialmente en el campo de la medicina en las que estas tecnologías no son muy útiles o apropiadas. Un escenario que se plantea es como se podría administrar fármacos dentro del cuerpo humano y llevarlos exactamente al lugar que lo necesitan (por ejemplo, administrar el fármaco directamente a un tumor maligno), para ello se propone usar miles de robots del tamaño de un glóbulo sanguíneo (o nanobots), los mismos que se deben comunicar y cooperar entre sí para llegar a su objetivo que es identificar el tumor y liberar sus fármacos para acabar con él. En este proceso los dispositivos no tienen la suficiente reserva de energía por lo que tienen que tomarla de su entorno, como todo este proceso se realiza dentro del cuerpo humano los dispositivos deben funcionar sin causar ningún daño a los tejidos sanos y ser destruidos por nuestro sistema inmunológico, todo esto representa un reto a la hora de ejecutarlo, lo que lleva a los organismos a intercambiar señales compuestas de moléculas, o comunicaciones moleculares (CM).

Las CM son una tecnología de comunicación bioinspirada para permitir la comunicación en escala nano-micro, que ha estimulado un gran interés tanto en el ámbito académico como en el médico [9], siendo la manera de comunicación predominante en la naturaleza entre los organismos vivos. Una nanored (nanodispositivos) en base a CM puede utilizar a través de microorganismos existentes, como bacterias y células, como parte de la red y es mayormente compatible con el cuerpo humano, una característica necesaria para las aplicaciones biomédicas [10], además es un sistema usual y general para la comunicación de nanomáquinas biológicas, el tipo de CM basado en difusión [11] es el que más se usa en el mundo académico, el estudio teórico y modelado de CM se ha hecho bajo la perspectiva de teorías de la información lo que derivó en varias investigaciones sobre el modelamiento de canales de CM, también en el ámbito del flujo estacionario se toma en cuenta condiciones reales en un ambiente líquido con una determinada velocidad de flujo.

Las nanocélulas y los nanodispositivos solo pueden realizar simples operaciones debido a su pequeña escala física y limitación de recursos. En base a las teorías de información de Shannon, Toby Berger señala el hecho de que los sistemas vivos emplean estructuras simples

para codificar y decodificar información, y tienen "poca o ninguna necesidad de los elegantes teoremas y técnicas de codificación convolucional y de bloques de teorías de la información", Berger argumenta (principalmente en el contexto de las redes neuronales) que esto se debe a que el medio de comunicación se ha adaptado a las fuentes de datos en el proceso evolutivo [10]. Por lo tanto, la optimización de codificadores y decodificadores con estructuras simples se debe al hecho de que el canal de transmisión y la información están emparejados. La comunicación no codificada, en particular, es una estrategia atractiva para las aplicaciones biológicas y se ha demostrado que es óptima en algunos entornos [12].

Las CM se han estudiado ampliamente desde varios aspectos, incluidos los modelos de canal, los esquemas de modulación y los protocolos de comunicación, la mayoría de los cuales se basan en la detección de una señal ideal. El aumento de la distancia de comunicación conducirá a un retraso de propagación inaceptable y a la variación de los tiempos entre llegada de las moléculas, lo que plantea desafíos para la retransmisión de información y la sincronización del sistema [11].

El objetivo de un sistema de comunicación es transferir información a través del espacio y el tiempo. Para lograr esto, un transmisor genera una señal de acuerdo con la información que se pretende transferir a un receptor. Esta señal luego se propaga a un receptor, donde la información deseada se decodifica a partir de la señal recibida. Por lo tanto, cualquier sistema de comunicación se puede dividir en tres componentes principales: el transmisor, el receptor y el canal [10], tanto el transmisor como el receptor adoptan diferentes tipos de procesamiento de información. En CM un transmisor molecular es una célula biológica o producida artificialmente cuyas acciones influyen en la densidad de moléculas en su entorno. Un canal molecular se refiere a un medio físico en él se propagan las moléculas. Un receptor molecular (o sensor) es una célula biológica o de ingeniería que está influenciada por la densidad de las moléculas en el entorno en su vecindad [10].

Como se puede observar en toda la teoría descrita antes, las CM actualmente son un tema muy interesante y desafiante, debido a todas las aplicaciones en las que se pueden emplear siendo el campo de la medicina el más atractivo por que supone una manera más sencilla y eficiente a la hora de tratar enfermedades, todas estas investigaciones nos aportan la información necesaria para el desarrollo de la presente investigación.

2.2 Comunicaciones Moleculares

Las CM corresponden a la forma en que las entidades biológicas se transmiten información mediante señales fisicoquímicas. Dado que tales entidades se componen de materiales biológicos y no son compatibles con los medios de comunicación tradicionales (p. ej., ondas electromagnéticas), las comunicaciones moleculares proporcionan un mecanismo para que las entidades biológicas se comuniquen mediante la transmisión, propagación y recepción de moléculas que representan información [13].

Este proceso biológico ha ocurrido desde el origen de la humanidad, como por ejemplo para comunicar órganos, este mismo concepto se puede reinterpretar en los sistemas de comunicación actuales, ya que poseen elementos similares en cada etapa de la comunicación, como son transmisor, receptor y canal de transmisión.

En la Figura 1 se aprecia el proceso de CM en donde se añaden máquinas biológicas (en Transmisión y Recepción respectivamente) ubicados en alguna parte del cuerpo humano. Estas nanomáquinas podrían ser gobernadas para cumplir la función básica de comunicación de información e intercambiarla. Este intercambio se lo puede llevar a cabo de distintas maneras como son: Difusión, haciendo uso de moléculas; proteínas motoras, usando vesículas; y movilidad bacteriana, mediante bacterias [9]. La información recolectada se puede usar en aplicaciones médicas.

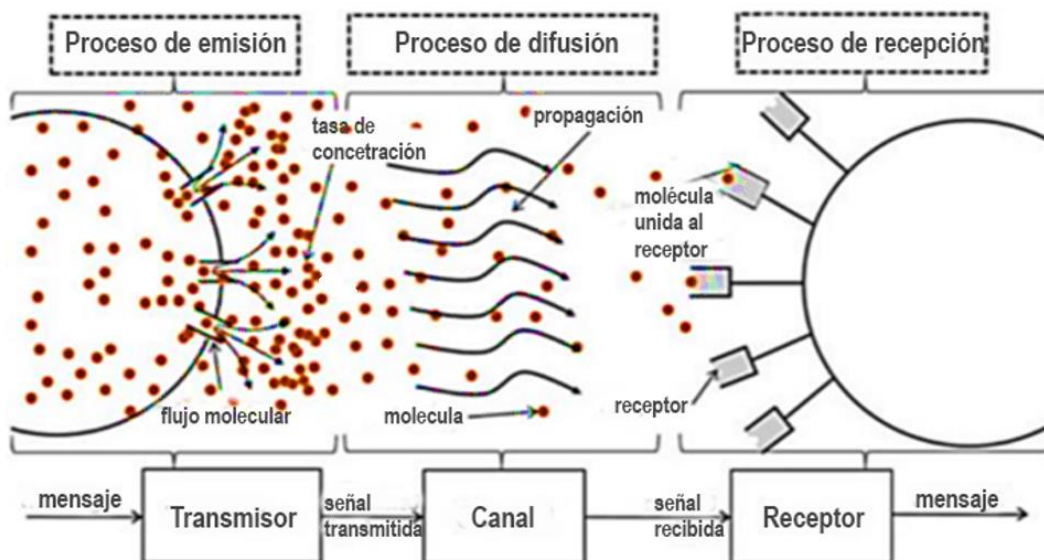


Figura 1. Comunicación Molecular.

Fuente: [14]

En esta investigación se proporcionará una visión holística de un sistema biológico con un sistema de comunicación de extremo a extremo con todos sus respectivos elementos, concretamente, se analiza la expresión genética desde la perspectiva de un sistema de comunicación digital. Usamos teorías de networking, como direccionamiento, control de flujo, control de errores, control de tráfico y el teorema de Shannon para diseñar modelos de capas de red que representan la transcripción y traducción del ADN y la transmisión de proteínas de extremo a extremo a un órgano objetivo. Estos modelos aprovechan las características clave de los sistemas de comunicación digital para transmitir información biológica en sistemas de expresión genética. Además, al analizar ambas clases de modelos de comunicación, establecemos la dualidad entre los sistemas de comunicación digital y biológicos y proporcionamos soluciones para superar los desafíos que enfrentan ambos tipos de sistemas y se puede entonces encontrar alternativas para el tratamiento de enfermedades, (como el cáncer). Por ejemplo, este análisis puede evitar efectos secundarios al transmitir información biológica a un destino específico (es decir, solo a un órgano objetivo) para lograr tratamientos óptimos [5].

2.3 Movimiento Browniano

Llamado así en honor al botánico Robert Brown. Fue descubierto mientras Robert Brown estudiaba la vida microscópica en un fluido, y notó que pequeñas partículas de polen se movían en el fluido estudiado [15].

Se define el movimiento browniano como el movimiento aleatorio de una partícula inducido por las colisiones con las moléculas presentes en un fluido. El movimiento browniano no necesita energía externa para ser efectuado. Además, es un fenómeno bien estudiado y existe suficiente literatura para implementar simulaciones de alguno de sus modelos matemáticos [6].

2.3.1 Proceso de Wiener

El proceso de Wiener es uno de los procesos estocásticos más importantes. Se lo considera como una representación matemática simple del movimiento browniano. Este proceso describe un movimiento aleatorio y continuo de una partícula, sujeto a la influencia de un largo número de colisiones entre moléculas moviéndose en un líquido [16].

El proceso de Wiener se define en términos de una distribución Gaussiana, con media m y varianza σ^2 , tal como se muestra en la Ecuación (2.1) [17].

$$x \sim N(m, \sigma^2) \quad (2.1)$$

Donde su Función de Densidad de Probabilidad (PDF) está dada por la Ecuación (2.2).

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (2.2)$$

Mediante la Ecuación (2.2), se plantea $B(t)$ como la posición en el MB, donde $B(0)$ es la posición inicial para $t \geq 0$. Note que $B(t)$ es un proceso de Wiener en una dimensión, si se cumplen las siguientes condiciones:

1. Para cualquier tiempo t_1 y t_2 (dónde $t_2 > t_1 \geq 0$), y alguna constante σ^2 , (ver Ecuación (2.3)).

$$B(t_2) - B(t_1) \sim N(0, \sigma^2(t_2 - t_1)) \quad (2.3)$$

2. Para dos intervalos $[t_1 - t_2]$ y $t_3 - t_4$, los incrementos $B(t_4) - B(t_3)$ y $B(t_2) - B(t_1)$ son estadísticamente independientes si los intervalos no se solapan.

Ahora, se considera que el proceso inicia en $B(0) = 0$ y que se revisa el proceso una vez por segundo, por ejemplo: $B(1), B(2), B(3)$, y así sucesivamente. Entonces $B(k)$, posición en el tiempo k , es una variable aleatoria Gaussiana con una distribución normal representada por la Ecuación (2.4).

$$B(k) \sim N(0, k\sigma^2) \quad (2.4)$$

Donde cada incremento es $B(k) - B(k-1)$ y tienen una distribución normal dada por la Ecuación (2.5).

$$B(k) - B(k-1) \sim N(0, k\sigma^2) \quad (2.5)$$

Además, los incrementos son independientes para cada valor de k . Como resultado se obtiene una varianza que incrementa con el tiempo, pero los incrementos son independientes e idénticamente distribuidos. También, se cumple para intervalos de tiempo $\Delta_t < 1$. En este caso, la distribución se expresa mediante la Ecuación (2.6), y los incrementos están dados $B(k\Delta_t) - B((k-1)\Delta_t) \sim N(0, \Delta_t\sigma^2)$.

$$B(k\Delta_t) \sim N(0, \Delta_t\sigma^2) \quad (2.6)$$

Por otro lado, en el MB la varianza σ^2 está dada por $\sigma^2 = \alpha D$. Donde D es el coeficiente de libre difusión de la propagación de las moléculas en el medio. El valor de $\alpha = 2, 4, 6$, dependiendo si el sistema está en 1, 2 o 3 dimensiones, respectivamente. El valor del coeficiente de difusión está establecido por la Ecuación (2.7).

$$D = \frac{k_B T}{6\pi\eta R_H} \quad (2.7)$$

Donde k_B es la constante de Boltzman [J/K], T es la temperatura en [K], η es la viscosidad dinámica del fluido, y R_H es el radio hidráulico de la molécula [17].

2.3.2 Proceso de Wiener con arrastre (drift)

El proceso de Wiener con arrastre está relacionado al MB con arrastre. Este proceso toma en cuenta la dirección del fluido por el cual se propagan las moléculas. Como ejemplo, se tiene un nanotransmisor ubicado en una vena, arteria o capilar que libera una molécula, la dirección de la molécula va a ser influenciada por el flujo sanguíneo [17].

En un proceso de Wiener con deriva, la distribución del incremento mencionada en la Ecuación (2.3) se reemplaza con la Ecuación (2.8).

$$B(t_2) - B(t_1) \sim N(v(t_2 - t_1), \sigma^2(t_2 - t_1)) \quad (2.8)$$

Donde v es la velocidad de arrastre. Para simplificar el análisis, se considera a la velocidad de arrastre constante en el tiempo.

Las propiedades estadísticas del proceso de Wiener con arrastre, incluida la propiedad de Markov, son idénticas al proceso de Wiener sin arrastre. Esto se debe a que el componente de arrastre es determinista y se puede restar del movimiento [17]; por ejemplo, si $B(t)$ es un proceso de Wiener con velocidad de arrastre v , entonces se forma un proceso $B'(t) = B(t) - vt$; la cual es un proceso de Wiener sin arrastre. La Figura 2, muestra un ejemplo de un proceso Wiener con y sin arrastre.

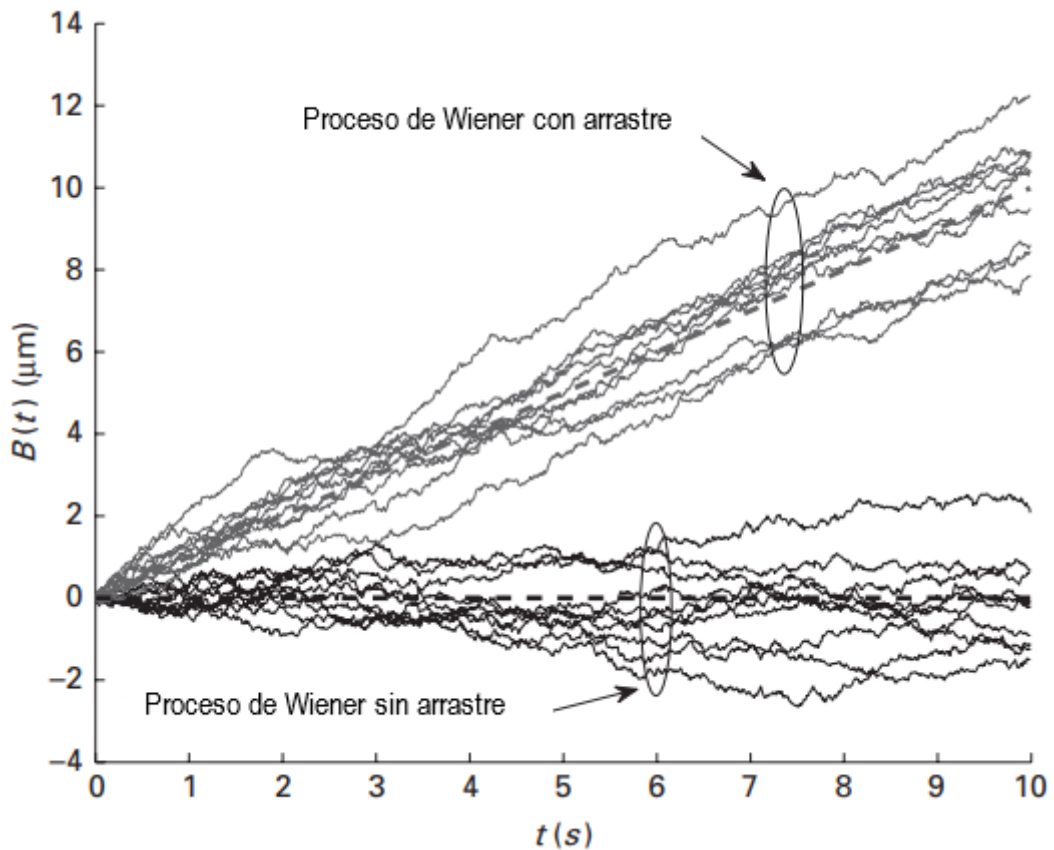


Figura 2. Proceso de Wiener con y sin arrastre. Velocidad de arrastre $v = 1 \mu\text{m/s}$ y $v = 0 \mu\text{m/s}$, respectivamente. En ambos casos, $D = 1 \mu\text{m}^2/\text{s}$, y $B(0) = 0$.

Fuente: [17]

2.4 Leyes de Fick

Las leyes de Fick se establecen a partir de la relación entre el proceso de difusión y el proceso de conducción de calor. Por un lado, la difusión es el proceso por el cual la materia es transportada de una parte del sistema a otro como resultado del movimiento aleatorio de las moléculas [18]. El movimiento aleatorio de las moléculas da como resultado la conducción de calor, esto se descubrió en 1855 por Adolf Fick, el mismo que adoptó la ecuación matemática de conducción de calor para el proceso de difusión.

Por ello, según [18] el proceso de difusión para un medio isotrópico se basa en la hipótesis de que la velocidad de transferencia de la sustancia en difusión a través del área unitaria de una sección es proporcional al gradiente de concentración normal a la sección. Esto está dado por la Ecuación (2.9).

$$F = -D \frac{dC}{ds} \quad (2.9)$$

Donde F es la tasa de transferencia por unidad de área de sección, C es la concentración de la sustancia difusiva, s es la coordenada espacial normal a la sección y D es llamado el coeficiente de difusión. De la Ecuación (2.9) se puede obtener la ecuación diferencial fundamental de difusión en un medio isotrópico, expresada mediante la Ecuación (2.10) [18]. El coeficiente de difusión en esta ecuación se le considera constante.

$$\frac{dC}{dt} = D \left(\frac{d^2C}{dx^2} + \frac{d^2C}{dy^2} + \frac{d^2C}{dz^2} \right) \quad (2.10)$$

La Ecuación (2.10) se la puede reducir a una dimensión si el gradiente de concentración está solo en dirección del eje x (ver Ecuación (2.11)).

$$\frac{dC}{dt} = D \left(\frac{d^2C}{dx^2} \right) \quad (2.11)$$

A las Ecuaciones (2.9) y (2.11), usualmente se las conoce como la primera y segunda Ley de Fick de difusión. Para determinar la solución a la Ecuación (2.11) en dos dimensiones, se considera dos condiciones de frontera: la fuente de emisión es un punto y el espacio es un plano infinito. En [11] se plantea la solución mediante la Ecuación (2.12).

$$C = \frac{M}{4\pi Dt} e^{-\frac{r^2}{4Dt}} \quad (2.12)$$

Por otro lado, para sistema de tres dimensiones con las condiciones de frontera: la fuente es un punto y el medio es un volumen infinito. La solución planteada se muestra en la Ecuación (2.13).

$$C = \frac{M}{(4\pi Dt)^{\frac{3}{2}}} e^{-\frac{r^2}{4Dt}} \quad (2.13)$$

Donde:

M es el número de moléculas liberadas por el transmisor.

D es el coeficiente de transmisión.

t el tiempo transcurrido desde la emisión de las moléculas.

r la distancia entre el Tx y Rx.

2.5 Definición de Nanomáquina

Para comprender el contexto en el que de forma básica se desarrollan todas las investigaciones referentes a sistemas de CM se debe definir el término nanomáquina (nanomachine) [19]. La Nanotecnología facilita el desarrollo de dispositivos en una escala en el orden de uno a algunos cientos de nanómetros. A escala nano, la unidad más básica funcional es justamente una nano-máquina, la cual se define como un elemento mecánico artificial cuyos componentes poseen dimensiones a nivel nano. El término máquina molecular (molecular machine) se define como un elemento mecánico provisto de una estructura molecular que ejecuta funciones útiles mediante el uso de componentes a nano-escala; ambos tipos de nanomáquinas, artificiales y naturales se pueden encontrar interactuando con sistemas biológicos [20], [21].

En entornos acuosos, es decir, en los que las comunicaciones a nano-escala dentro de organismos vivientes ocurren (mediante el intercambio de moléculas) es más propicio hablar de nanomáquinas biológicas (bio-nanomachines) [22]. Comúnmente las nanomáquinas biológicas en las investigaciones hacen alusión a elementos: biológicos (como células, tejidos, órganos, etc.), biológicos híbridos, o no biológicos, que son capaces de realizar funciones químicas simples. Algunos ejemplos de nanomáquinas biológicas son moléculas de ácido desoxirribonucleico diseñadas para llevar a cabo operaciones lógicas, motores proteicos construidos para transportar moléculas, proteínas y células modificadas genéticamente (las células en general se consideran como provistas de dimensiones a nivel micro) [5], [22]. En este texto se empleará (implícita y explícitamente cuando se las mencionen) este tipo de nanomáquinas para describir el proceso comunicacional que establece la expresión genética y el transporte de proteínas entre transmisores y receptores moleculares [19].

En términos generales las nanomáquinas se conceptualizan como capaces de ejecutar una función específica a nivel nano, entre las cuales se cuentan [2], [22]:

1. Tareas simples de cómputo.
2. Almacenamiento.
3. Detección de señales (sensor).
4. Actuación.

Las tareas ejecutadas por las nanomáquinas son muy simples y se restringen al entorno que las circunda debido a su baja complejidad y diminuto tamaño. Sin embargo, el nivel de

complejidad de una nanomáquina de acuerdo con el número de sus componentes va desde simples conmutadores moleculares hasta nanorobots [20], [21]. Las nanomáquinas más completas están provistas de una arquitectura formada por [20]:

- 1.Unidad de control.
- 2.Unidad de comunicaciones.
- 3.Unidad de reproducción.
- 4.Unidad de provisión de energía.
- 5.Unidad de detección de señales y actuación.

2.6 Ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN) como nanomáquinas biológicas

Como se sabe el ADN y ARN son moléculas que codifican la información genética de un organismo.

2.6.1 Roles y Funciones del ADN y ARN

En el proceso de la expresión génica, el ADN funciona como un almacén de información genética y el ARN participa en la decodificación de la información para sintetizar proteínas. La expresión genética está descrita por el dogma central de la biología molecular con dos procesos: el ADN se copia en ARN a través de la transcripción y el ARN sintetiza proteínas a través de la traducción [17]. Por lo que nos enfocaremos en dichos procesos, transcripción y traducción, y al final podremos observar como el ADN y el ARN trabajan para producir proteínas.

2.6.1.1 Transcripción del ADN

La transcripción del ADN es el proceso de copiar una parte de una de sus moléculas en una molécula de ARN, llamada ARN mensajero (ARNm). El proceso de transcripción es catalizado por una enzima proteica, ARN polimerasa (ARN pol), la misma que se mueve en dirección 3' a 5' a lo largo de una de las dos cadenas de ADN durante el proceso. Primero, la ARN pol reconoce y se une a una secuencia específica de una hebra de ADN llamada promotor (alberga una secuencia de inicio, que es donde ARN pol comienza a agregar nucleótidos para crear una secuencia de ARN mensajero (ARNm) complementario) para iniciar el proceso de transcripción. Luego, la ARN pol comienza a leer la cadena de ADN una núcleo-base a la vez [17].

Siguiendo las reglas de emparejamiento de bases para la núcleo-base que se encuentra en la hebra de ADN, la ARN pol produce una núcleo-base para el ARNm al consumir un trifosfato nucleósido complementario (NTP) disponible en el medio ambiente; es decir, trifosfato de adenosina (ATP), trifosfato de uridina (UTP), trifosfato de citidina (CTP) o trifosfato de guanosina (GTP). Por ejemplo, si se encuentra A (Adenina) en la cadena de ADN, entonces se consume UTP (Figura 3a)) para producir el compañero de apareamiento de bases de A, U (Uracilo) (Figura 3b)). La ARN pol, de esta forma, produce la cadena núcleo-bases para alargar el ARNm (Figura 3c)). La ARN pol puede terminar el proceso de transcripción cuando encuentra una secuencia específica de ADN llamada terminador (secuencia de finalización apropiada). El resultado del proceso de transcripción es un ARNm con una secuencia complementaria a una parte de la hebra de ADN [17].

Dentro del proceso de transcripción primaria (pre-ARNm) se producen también tres modificaciones que son [17]:

- 1.Splicing. Los segmentos en la molécula de pre-ARNm que no proporcionan información (intrones) se eliminan.
- 2.Capping. Procesamiento postranscripcional del final 5' del ARN se produce a través del proceso de terminación 5'. En la parte final de este proceso, el extremo 5' de la transcripción de ARN contiene un grupo trifosfato libre porque este nucleótido se incorporó primero en la cadena. El proceso de finalización reemplaza el grupo trifosfato con otra estructura llamada "capping", la cual se agrega por la enzima guanil transferasa. Esta enzima cataliza la reacción entre el extremo 5' de la transcripción de ARN y una molécula de trifosfato de guanina.
- 3.Poliadenilación. Procesamiento de ARN postranscripcional en el extremo opuesto (3'extremo) de la transcripción se produce a través de una cadena de aproximadamente 250 adeninas unidas al final de la cadena de ARN sintetizada. Esta cadena de adenina se denomina "cola poli-A". La adición de las adeninas es catalizada por la enzima poli (A) polimerasa.

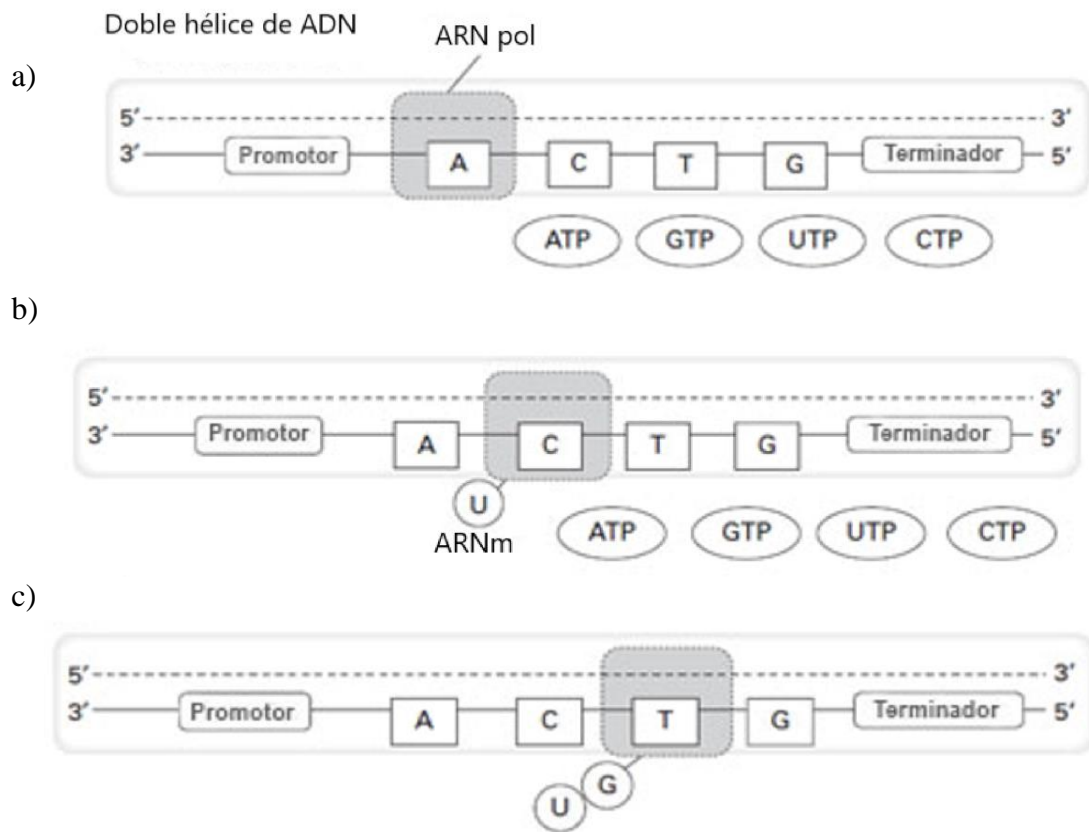


Figura 3. La transcripción produce ARNm a partir de una secuencia de una hebra de ADN.

Fuente: [17]

2.6.1.2 Traducción del ADN

La traducción es el proceso de sintetizar una proteína a partir de una secuencia de ARNm producida a partir de la transcripción. La secuencia de ARNm determina la estructura primaria de una proteína donde una secuencia de tres nucleótidos llamada codón codifica un aminoácido particular. Por ejemplo, la secuencia GAA (Guanina Adenina Adenina) codifica un ácido glutámico. Existirán $4^3 = 64$ codones, cada uno de los cuales codifica uno de los 20 aminoácidos con algunas excepciones, incluido un codón de inicio (en la mayoría de los casos, AUG) donde se inicia la traducción y tres codones de parada donde no se produce ningún aminoácido y se termina la traducción. En la traducción, la codificación de un codón en un aminoácido está mediada por moléculas específicas de ARN de transferencia (ARNt). Cada ARNt contiene una secuencia de tres nucleótidos, llamada anticodón, que es complementaria a un codón particular de acuerdo con las reglas de apareamiento de bases; y también lleva un aminoácido específico del anticodón [17].

El proceso de traducción es catalizado por un gran complejo ARN-proteína llamado ribosoma (lee la información contenida en la secuencia biológica utilizando un sistema de codones (un codón es un triplete de nucleótidos)) que tiene dos sitios catalíticos para promover reacciones. Primero, un tipo específico de ARNt se une al codón de inicio en el ARNm a través del apareamiento de bases codón anticodón (Figura 4a)). De manera similar, otro tipo de ARNt se une al siguiente codón a través del apareamiento de bases codón-anticodón (Figura 4b)). Luego, el ribosoma cataliza una reacción química para concatenar el aminoácido del primer ARNt y el del siguiente ARNt a través de un enlace peptídico (Figura 4c)). El ribosoma también elimina el primer ARNt del ARNm y se mueve en la dirección 3' por un codón (Figura 4d)), para que esté listo para unirse al siguiente ARNt. Este proceso se repite hasta que el ribosoma encuentra un codón de terminación donde termina la traducción y se libera una proteína [17].

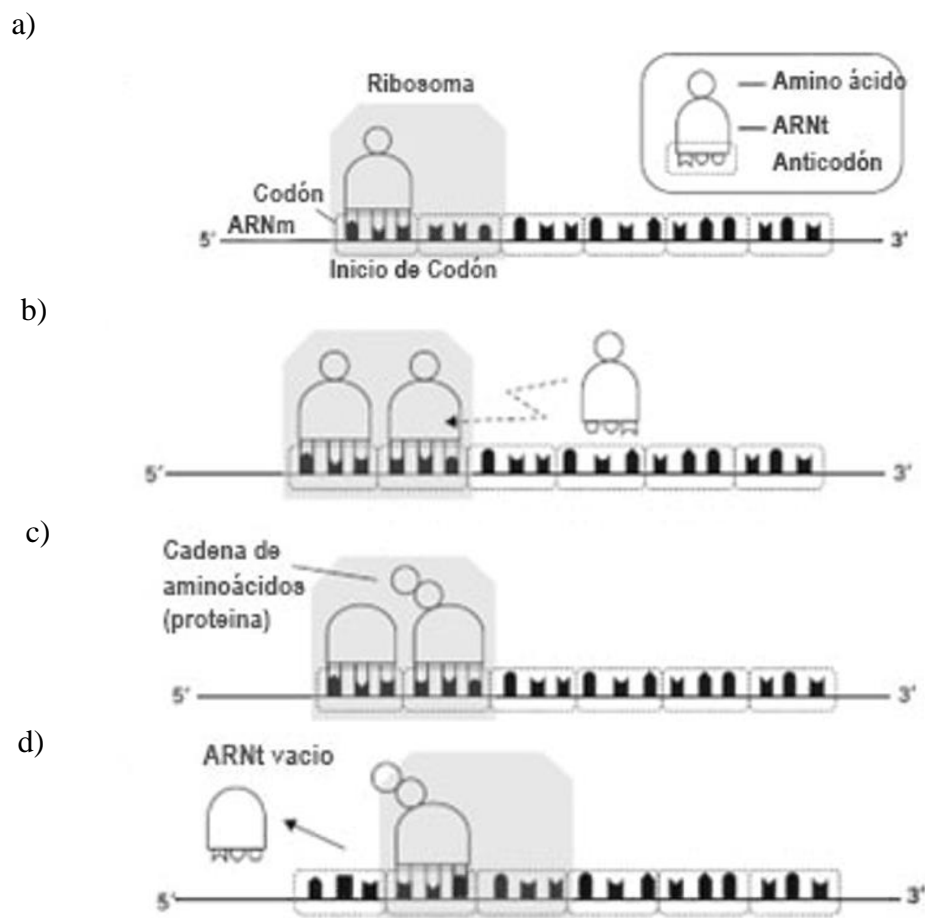


Figura 4. La traducción produce una proteína a partir de una secuencia de ARNm.

Fuente: [17]

2.7 Modelo de comunicación molecular basado en la expresión genética

Un sistema biológico se modela como un sistema de comunicación digital. Así en [23] este sistema de comunicación biológica, las bases de nucleótidos en el ADN sirven como transmisores, el ADN sirve como medio de comunicación y los organismos vivos sirven como receptores. Se supone que los errores ocurren en el canal de transmisión y producen mutaciones en el ADN. El desarrollo de algoritmos robustos que funcionan en una de dos regiones (es decir, ADN no codificante de proteínas o ADN codificante de proteínas) se presenta como un mecanismo para prevenir mutaciones y evitar consecuencias letales, como enfermedades en los receptores biológicos.

La transmisión intracelular de la información genética, a través de un sistema de comunicación biológica se representa como un sistema de comunicación digital en [24]. En este sistema de comunicación biológica, la identificación, reproducción y clasificación matemática de las secuencias de nucleótidos son necesarias para identificar, reproducir y clasificar matemáticamente las secuencias de ADN posteriormente [5]. Los errores en el sistema se corrigen eficientemente utilizando códigos lineales G. Según el estudio en [24], la metodología utilizada para caracterizar este sistema de comunicación biológica también podría usarse para prevenir mutaciones en el ADN, producir nuevos medicamentos y crear mejoras genéticas.

Debido a las similitudes entre la transmisión de información en los sistemas de comunicación y la transmisión de medicamentos en el cuerpo humano en [12], se analizan varias plataformas y modelos de transmisión para administrar medicamentos solo a objetivos específicos para minimizar los efectos secundarios. Estos sistemas de transmisión especializados, que se denominan sistemas de administración de fármacos, podrían utilizarse en el tratamiento médico del cáncer, el VIH, las enfermedades genéticas, las enfermedades gastrointestinales y las enfermedades neurológicas. Las vías de transmisión (es decir, administración) de fármacos en el cuerpo incluyen la ingestión oral, la inyección y la aplicación en la piel [12].

En los sistemas biológicos, el proceso evolutivo utilizado por los organismos vivos durante miles de millones de años [9] y características como la eficiencia energética, la biocompatibilidad, la capacidad de trabajar en un medio acuoso y la capacidad de penetración son las características más importantes [25]. Además, en los sistemas digitales, el rendimiento y la eficiencia son las dos características más importantes [2].

Los dos tipos de sistemas tienen en común desafíos a vencer como; atenuación, el ruido, la interferencia, el retraso y la pérdida de información pueden ocasionar problemas de comunicación fatales. Por ejemplo, la falla en los sistemas biológicos podría resultar en enfermedades letales, como el cáncer, en los seres humanos [26]. Sin embargo, la comunicación biológica (o comunicación molecular) es un campo científico prometedor. Su aplicación al tratamiento médico podría mejorar significativamente la calidad de vida humana; por ejemplo, las terapias convencionales podrían acceder a partes del cuerpo pequeñas y delicadas de forma no invasiva [27]. Los actuales avances en comunicación digital y su aplicación en los sistemas biológicos ha convertido ácido desoxirribonucleico en un medio de almacenamiento de alta densidad y tamaño pequeño (1 g de ADN seco ocupa 1 cm y tiene un potencial de almacenamiento de 455 EB de información) [28]. Esta enorme capacidad de almacenamiento podría utilizarse como base de datos biológica lo que ha sido propuesto por los avances en la investigación informática [29].

En [15] se propone un modelo de red de capas apiladas que explica la expresión génica (es decir, el proceso mediante el cual la información transportada por el ADN se transforma en las proteínas apropiadas) y el papel del aparato de Golgi en la transmisión de hormonas peptídicas a un órgano objetivo, lo cual es similar al proceso de comunicación digital en el que un dispositivo transmisor en una red envía información a un dispositivo de destino en otra red a través de un router. En [15] se plantea el empleo de las teorías de redes (por ejemplo: funciones independientes de una capa en una pila, direccionamiento, control de flujo, control de errores y control de tráfico) y aplicarlas al tratamiento de enfermedades para reducir los efectos secundarios de los medicamentos y mejorar la calidad del tratamiento de los pacientes desde la perspectiva sanitaria y socioeconómica.

En Shirazi [15] se propone una pila de protocolos para nanoredes, y el diseño consta de tres capas (aplicación, red y física) que representan comunicaciones dentro de nanoredes de corto alcance a través de motores moleculares y señalización de calcio y dentro de nanoredes de largo alcance usando feromonas. En el primer caso, el modelo de comunicación usa rieles moleculares que consisten en microtúbulos que poseen la misma función que los cables en las redes tradicionales de comunicación punto a punto. En el segundo caso, el modelo de comunicación transfiere iones de calcio a células adyacentes a través de uniones comunicantes, como una conexión por cable en las redes de comunicación tradicionales. En el tercer caso, la información se transmite a través de feromonas que pueden modelarse como comunicación inalámbrica en las redes tradicionales.

En [4] se proporciona una perspectiva arquitectónica de CM empleando un enfoque de arquitectura en capas, que se utiliza tradicionalmente en las redes de comunicación. De manera similar al modelo de referencia de interconexión de sistemas abiertos (ISO/OSI) y la arquitectura de Internet TCP/IP, una arquitectura en capas (Aplicación, Transporte, Enlace de Datos y Física) y describe problemas en la CM y, por lo tanto, facilita el diseño y desarrollo de la entrega de fármacos y otras aplicaciones de comunicaciones moleculares en transporte activo y pasivo, mediante el empleo de funciones independientes entre dichas capas.

2.8 La expresión genética y transporte de proteínas analizada como un sistema de comunicación digital

2.8.1 El Núcleo Celular como Equipo Terminal de Datos Biológicos (DTE)

En [2] supone que el núcleo celular se comporta como un DTE biológico, y que el núcleo representa el transmisor que contiene la fuente de información compuesta por bloques de nucleótidos llamados genes. Esta información biológica debe ser procesada intracelular o extracelularmente. Biológicamente, un gen se define comúnmente como un conjunto de nucleótidos que almacena la información necesaria para llevar a cabo una función (mediante una proteína o ácido ribonucleico (ARN)) a realizar en un destino [28]. Como se menciona en [2], el contenido de un gen puede entenderse como una dirección en la capa de red, en razón de que si en la expresión genética se emplea proteínas que van a zonas corporales distantes entonces, se trata de un direccionamiento a larga distancia a nano-escala, de forma análoga y direccionamiento que ocurre en la capa red.

Las moléculas de ADN contienen información digital ya que está codificada por cuatro valores discretos (cuatro nucleótidos). Los nucleótidos son monómeros de ácidos nucleicos (ADN y ARN) que comprenden una base nitrogenada, un azúcar de cinco carbonos (desoxirribosa en el ADN y ribosa en el ARN) y al menos un grupo fosfato. Las bases nitrogenadas incluyen adenina (A), timina (T), citosina (C), guanina (G) y uracilo (U) [5]. La doble hélice del ADN comprende nucleótidos que contienen las bases A, T, C y G [30] y conserva su estructura debido a la complementariedad entre las bases nitrogenadas de cada hebra de la hélice [31] (es decir., la afinidad de la adenina por la timina y la de la citosina por la guanina) [32]. Como la información en el ADN se divide en bloques de nucleótidos llamados genes [7], [33] que poseen secuencias de inicio y terminación, entonces podemos interpretar que la información biológica se divide en segmentos de datos [2]. En las redes de

conmutación de paquetes, la información digital se divide en unidades más pequeñas conocidas como paquetes para facilitar el procesamiento. Por lo tanto, un paquete de red digital puede ser análogo a un gen en una red de comunicación biológica [2].

Al comienzo de la transcripción, el motor molecular ARN polimerasa II (la enzima RNAP II) reconoce una región de la secuencia de ADN de la región promotora [34]. El promotor alberga la secuencia de inicio, que es cuando RNAP II comienza a agregar nucleótidos para crear una secuencia de ARN mensajero (ARNm) complementario. Durante el inicio de la transcripción, RNAP II produce una copia complementaria de ARNm monocatenario de una de las dos cadenas de ADN. La única diferencia entre el ARN y el ADN es que RNAP II utiliza uracilo (U) en lugar de timina (T) durante este proceso [35], [36]. No hay necesidad de copiar ambas cadenas de ADN porque las cadenas son complementos exactos. Además, este proceso biológico se asemeja a la compresión de datos digitales porque la misma cantidad de datos está contenida en un espacio más pequeño [2]. De hecho, la compresión máxima del código genético para el uso óptimo de sus nucleótidos viene dada por $H(pv) / \log_2 a$ [40], donde $H(pv)$ es la entropía del vector de probabilidad p_v relativo a un nucleótido específico, y a es el número de letras de un alfabeto; por lo que los cuatro nucleótidos consisten en el alfabeto [5].

El **enhancer** es otro elemento de la cadena de ADN que controla la cantidad de proteína producida en función de la cantidad de ARNm. Dado que el enhancer controla la cantidad de información enviada al receptor, este proceso puede entenderse como un control de flujo en el extremo emisor [2]. La capa de enlace de datos es responsable del control de flujo para garantizar que un emisor rápido no pueda saturar a un receptor lento con información de los que pueda procesar [37]. La transcripción procede unidireccionalmente a lo largo de una de las cadenas de ADN desde el 5'P hasta el 3'OH de la columna vertebral del fosfato de desoxirribosa. Este orden es esencial para garantizar que la información genética se copie adecuadamente; de forma similar, en las comunicaciones digitales, por ejemplo, el orden lsb (bit menos significativo) o msb (bit más significativo) es fundamental; precisamente, en una transmisión en serie, el lsb suele ser el primero en enviarse [2].

La detención de la transcripción se produce cuando una secuencia de finalización apropiada es reconocida por la RNAP II. [4]. En la molécula de transcripción primaria (es decir, pre-ARNm) ocurren las siguientes modificaciones (maduración) [38]:

1. Splicing.
2. Capping.

3. Poliadenilación.

La información añadida durante el capping y la poliadenilación puede ser equivalente a las banderas delimitadoras de datos utilizadas en los sistemas de comunicación digital, como las cabeceras y los trailers que encapsulan la información en la capa de enlace de datos en las jerarquías de protocolo en el software de red como se observa en la Figura 5 [2], [17]. Estas banderas se utilizan para el procesamiento y el control de errores [4]; en el mismo periodo de tiempo, con usar banderas digitales, las maduraciones anteriormente mencionadas proporcionan estabilidad (control y procesamiento posterior) a las moléculas de ARNm y evitan la degradación del ARNm por enzimas en el citosol (fluido intracelular), por lo tanto, de esta manera permite que las moléculas avancen a las fases posteriores de procesamiento biológico [2].

El transporte mecánico de moléculas de ARNm a través del citosol puede ser análogo a la transmisión de información en las comunicaciones por cable (tarea de la capa física) [2], [4], [8], [30].

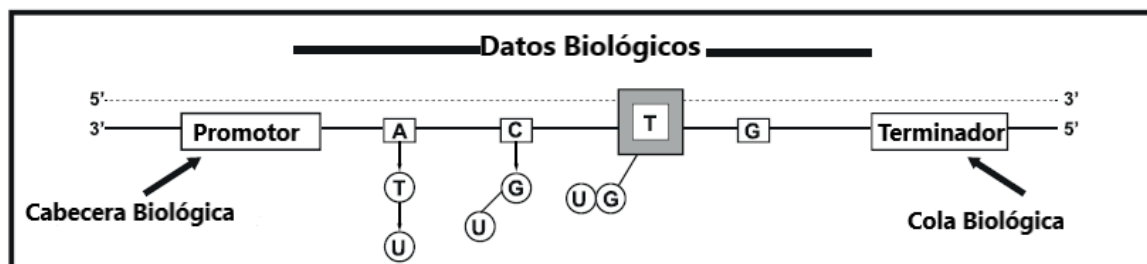


Figura 5. Marco biológico en la transcripción del ADN.

Fuente: [5]

2.8.2 Los ribosomas y el retículo endoplásmico (RE) como equipo de comunicación de datos biológicos (DCE)

Se permite que el proceso de transcripción copie la información biológica del ADN al ARN; esto es necesario porque las moléculas de ADN no pueden salir del núcleo de la célula. Por lo tanto, en este punto, el DTE debe transmitir los datos biológicos al DCE a través de una interfaz física (como en los sistemas de comunicaciones convencionales) [5]. En el caso de [2], el citosol puede representar la interfaz física Figura 6.

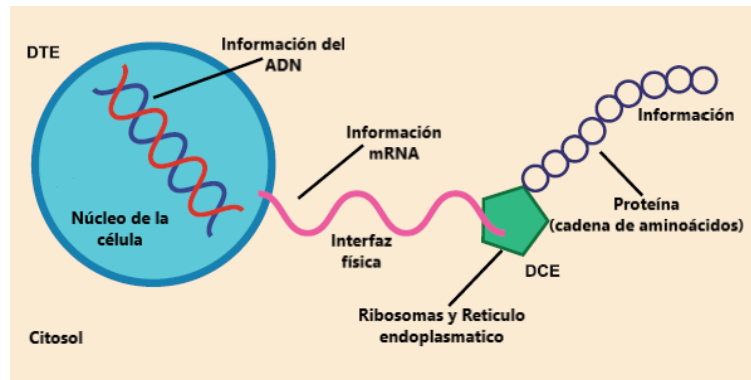


Figura 6. Una interfaz física en un sistema digital puede ser análogo al citosol en un sistema biológico.

Fuente: [2]

Se considera que los ribosomas y el retículo endoplasmático representan el DCE biológico [2] porque, a través de estos orgánulos, la información genética adquiere una estructura funcional (o formato cuando se refiere a datos) que luego se libera en el canal de comunicación biológica y finalmente llega al receptor biológico. Durante la traducción, el "formato apropiado" de los datos biológicos ocurre cuando la información se convierte en cadenas de aminoácidos para obtener funcionalidad dentro y fuera de la célula [2].

Así, el DCE biológico procesa (codifica) la información a través de traducción y proporciona una secuencia de entrada específica (datos en el ARNm) que está asociada con una salida específica (secuencia de aminoácidos); desde la perspectiva de los sistemas digitales, este tipo de proceso de codificación corresponde a la codificación convencional [2]. Como en todo estado de los procesos biológicos, a lo largo de la traducción, la madre naturaleza ha establecido el direccionamiento oportuno; por lo tanto, el ARNm que sale del núcleo tiene una dirección adyacente implícito que es comparable al direccionamiento de capa de enlace de datos para facilitar la comunicación en un rango directo de comunicación [2], [4]; por lo tanto, el ARNm se une a los ribosomas en el citosol o aquellos asociados con el RE rugoso (RER). La transmisión (movimiento) de información biológica desde el núcleo a los ribosomas o ER a través del citosol representa la llegada y el movimiento de información a un canal de comunicación biológico, lo que se considera una tarea en la capa física [2], [4]. Los ribosomas, que son estructuras que sirven como motores moleculares, leen la información contenida en la secuencia biológica mediante un sistema de codones (un codón es un triplete de nucleótidos) [2], [34]. En el ribosoma, los codones del ARNm son reconocidos por los ARN de transferencia (ARNt) que poseen un anticodón (una secuencia

complementaria a un codón particular) asociado con un aminoácido único que se une específicamente a la estructura molecular de ese ARNt [5].

La síntesis de proteínas ocurre en el ribosoma a través de la señalización de los ARNt que indican a los ribosomas el inicio (codón de inicio) y el final (codón de terminación) del proceso para asegurar la correcta lectura de la información biológica [32]. En un paradigma de comunicación digital, el inicio y final de los codones pueden corresponder a señales de sincronismo. El sincronismo entre el origen y el destino se lleva a cabo en transmisiones síncronas a través de una bandera de inicio [5]. En este tipo de comunicación, el transmisor envía los datos y el receptor debe recopilar y procesar los mismos, el final del codón en la señalización biológica puede ser equivalente a una bandera de finalización en comunicaciones síncronas que se utiliza en el destino para indicar el final de la comunicación [38]. Las señales para iniciar y detener la traducción del ADN permiten que los relojes biológicos presentes en las células proporcionen retroalimentación durante los procesos celulares [2], [4]. Porque el procesamiento de aminoácidos específicos genera proteínas (es decir, cadenas polipeptídicas), el orden de los aminoácidos depende del aminoácido anterior; por lo tanto, esta característica biológica puede corresponder a la entrada en la fuente de Markov [5]. Las proteínas constan de 20 aminoácidos diferentes; por lo tanto, la entropía de orden k es [2]:

$$I_k(x) = - \sum_{i=1}^{20} \sum_s^{20^{k-1}} p_i p(i|s) \cdot \log_2 p(i|s) \quad (2.14)$$

donde p_i es la probabilidad de encontrar el aminoácido i th, y $p(i|s)$ es la probabilidad condicional de que el aminoácido i th ocurra después de la cadena de aminoácidos s [2], [8]. Si todos los codones (aminoácidos) ocurren con igual probabilidad, el número de secuencias posibles (nps) en cadenas polipeptídicas de longitud N puede calcularse como:

$$nps = 20^N \quad (2.15)$$

Dado que muchos aminoácidos no tienen la misma probabilidad, la Ecuación (2.15) debe modificarse. Así, si consideramos una secuencia larga de N símbolos seleccionados de un alfabeto de codones o aminoácidos, podemos determinar la probabilidad de esa secuencia de la siguiente secuencia como:

$$P = \prod_i p(i)^{N \cdot p(i)} \quad (2.16)$$

donde $N \cdot p(i)$ es la probabilidad del símbolo i th [15]. Considerando el logaritmo a ambos lados en la Ecuación (2.16),

$$\log_2 P = N \cdot \sum_i p(i) \cdot \log_2 p(i) = -N \cdot H \quad (2.17)$$

donde $H = -\sum_i p(i) \cdot \log_2 p(i)$ es la información de Shannon del espacio de probabilidad que contiene los sucesos i . En consecuencia, la probabilidad de obtener una secuencia larga de N símbolos o sucesos independientes a partir de un alfabeto finito es $P = 2^{-NH}$ y el número de secuencias de longitud N es casi 2^{NH} [5].

La tarea principal de SRP es permitir que la proteína nascente llegue a una proteína de canal en RE que supervisa la translocación de la proteína dentro del RE. Luego, el SRP se separa de la proteína y se recicla en el citosol [2], [32]. Por lo tanto, después de que la información de procesamiento y la información de control se utilizan en los sistemas de comunicación digital, se descartan [5]. En este punto dentro del RE, las proteínas se pliegan y adquieren la estructura tridimensional funcional necesaria para que cumplan sus funciones biológicas específicas [32] (equivalente a la información digital después de ser procesada por el DCE, es decir, teniendo el formato apropiado) [2].

En los sistemas biológicos, pueden producirse errores de información durante la transcripción y traducción del ADN; del mismo modo, en los sistemas de comunicación convencionales, pueden aparecer errores en la transmisión. Los errores en el procesamiento celular y su transporte son responsables de muchos trastornos médicos como el cáncer, la autoinmunidad y la diabetes [8]. En [2], la Figura 7, se presenta como un modelo de red en capas de la transcripción y traducción del ADN desde la perspectiva de una red de comunicación típica. La estructura de un modelo por capas descompone un sistema a gran escala en un conjunto de unidades más pequeñas (es decir, capas) que son funcionalmente independientes unas de otras y especifica las interacciones entre las capas [4], [39]. Por tanto, una ventaja de emplear una pila de capas es el uso fundamental de la capa de enlace de datos para transformar un canal imperfecto en una línea de transmisión libre de errores o informar de problemas no resueltos a la capa superior [2], [40]. Por lo tanto, la aplicación de un modelo de este tipo a los sistemas biológicos (por ejemplo, en la administración de fármacos) podría proporcionar una alta confiabilidad [2].

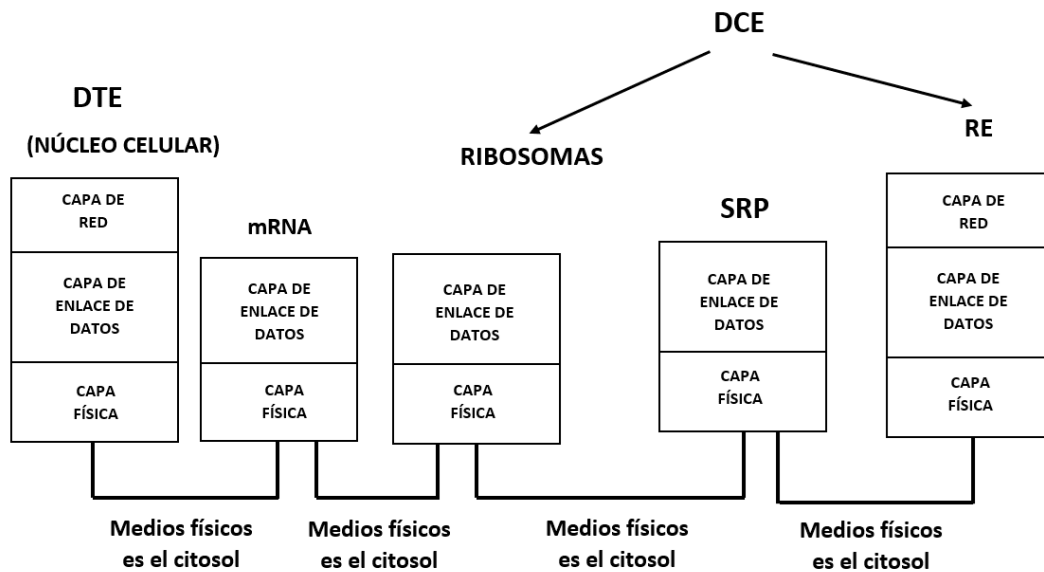


Figura 7. Modelo de red en capas DTE-DCE de la transcripción y traducción del ADN.
Fuente: [5]

2.8.3 El aparato de Golgi (AG) como enrutador de frontera de una Interred

Cuando las proteínas son funcionales, el RER transfiere las proteínas a través de motores moleculares para el AG. Porque cada proteína contiene una dirección adyacente implícita a través del etiquetado molecular, es comparable a una dirección de capa de enlace de datos que facilita la comunicación dentro de un rango directo de comunicación [2], [4], [17], las proteínas se enrutan a los destinos intercelulares apropiados; sin embargo, el AG determina si las proteínas permanecen o no dentro de la célula [32].

Durante este proceso, las proteínas y su contenido de información se mueven del RER al AG, donde la información se deposita en vesículas que se unen a la cara cis AG. Luego, se generan nuevas vesículas que contienen la información de la proteína y se agregan otros componentes celulares necesarios para procesar las proteínas. Las vesículas nuevas depositan su contenido en la cara medial de GA y, se forman vesículas nuevas que contienen la proteína y los elementos pertinentes para el procesamiento posterior. Finalmente, las vesículas alcanzan la cara trans AG, donde ocurre un proceso idéntico al definido anteriormente; por lo tanto, las proteínas se insertan en vesículas nuevas, pero se dirigen a la membrana endoplásmica para ser secretadas fuera de la célula [5].

Las funciones mencionadas son comparables a las que, son realizadas por un router de frontera en una interred, como se ilustra en la topología presentada en la Figura 8. Un ruter

decide si la información reside dentro de la red o sale de ella; por lo tanto, los procesos de enrutamiento y rutas son esenciales para enviar información al destino adecuado. Además, las acciones de depositar proteínas, formar vesículas y adjuntar información para especificar el destino de una proteína son análogas a las requeridas en el procesamiento de unidades de datos de protocolo (PDU) en las capas del router [2].

La utilidad de las teorías digitales que se han aplicado al análisis biológico actual podría ser significativa para el tratamiento de enfermedades porque las proteínas transmitidas desde el extremo emisor, que corresponden a la expresión genética, definen el desarrollo de organismos pluricelulares como los humanos, y errores en este proceso pueden causar el crecimiento de anomalías patológicas, como el paladar hendido o el cáncer [41].

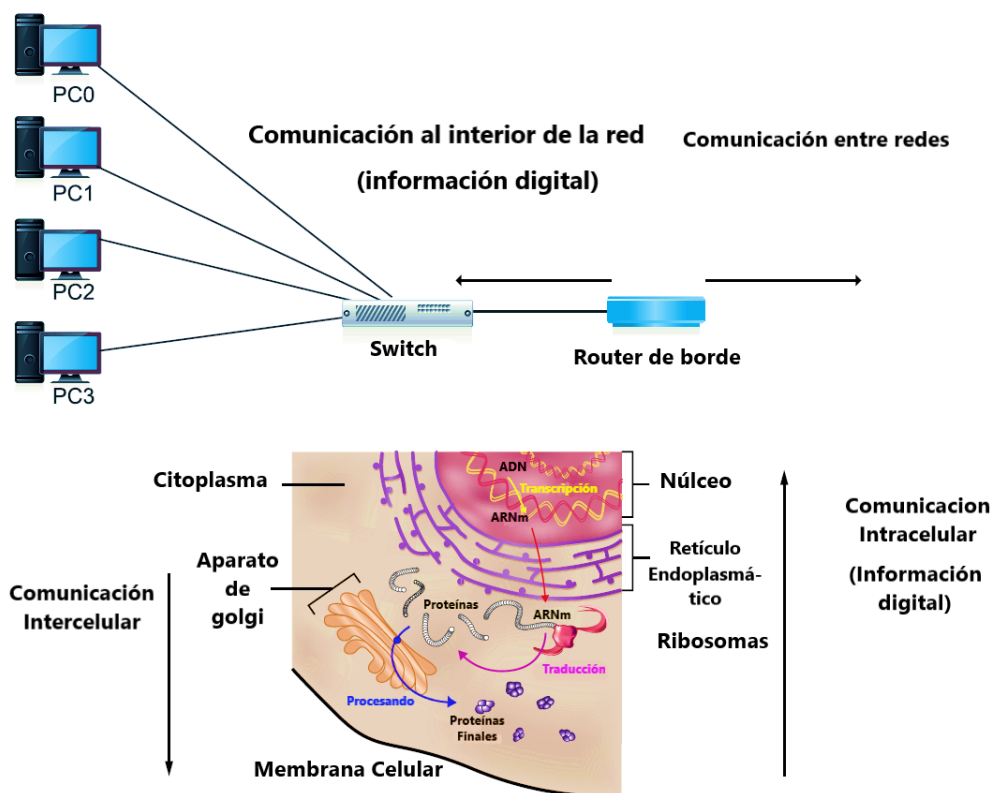


Figura 8. Las funciones del AG como un router de interred.

Fuente: [2]

2.8.4. Utilización de la información en el receptor

En el cuerpo humano, la comunicación de información biológica de un emisor a un receptor se realiza a través del torrente sanguíneo y una célula objetivo. El transmisor envía la información usando los datos almacenados en las moléculas de ADN para que en el destino

se pueda reconocer la célula, tejido u órgano objetivo. Por otro lado, en función del tipo de hormona proteínica implicada, el receptor procesa la información recibida. Aquí describimos brevemente un caso en el que este procesamiento se realiza a través de ligandos y sus receptores. Los ligandos pueden considerarse señales unidas a los receptores en la superficie de un receptor biológico [5]. Las señales de otros receptores son amplificadas e integradas por diversos receptores, y las señales resultantes se transmiten a la célula de destino. Los siguientes cuatro rasgos caracterizan la transducción de señales procesada por la arquitectura de comunicación del receptor: especificidad, amplificación, desensibilización e integración [8].

- La especificidad determina la afinidad entre ligandos y sus receptores en el extremo biológico destino. Los ligandos específicos se unen a receptores específicos (o complementarios). La especificidad está relacionada con la molécula señal que coincide con su receptor molecular complementario. Como concepto de red, la especificidad puede emplearse como mecanismo de direccionamiento cuando, por ejemplo, en redes Ethernet y a nivel de la capa de enlace de datos, se envía una difusión para reconocer una dirección de Protocolo de Internet (IP). En este caso, todos los dispositivos de la red podrían recibir el mensaje de difusión, pero será procesado por el dispositivo cuya MAC coincida con su dirección IP [2], [40].
- El aumento de la intensidad de la señal se denomina amplificación. En esta fase fisiológica del procesamiento biológico, la respuesta a la primera señal activa una segunda señal, la respuesta a la segunda señal provoca la activación de una tercera señal, y así sucesivamente. Desde el punto de vista switches, esta función celular es similar a la de los dispositivos de red (es decir, switches, routers, etc.) que amplifican la señal recibida antes de procesarla [42].
- La capacidad de eliminar la señal una vez recibida es una acción de desensibilización. La sensibilidad de los sistemas receptores es la función por la cual la señal se atenúa después de un cierto periodo umbral (es decir, el receptor se vuelve insensible a la señal). Tras este periodo umbral, el receptor vuelve a ser sensible. Este proceso puede ser comparable al caso de red mencionado anteriormente, ya que cuando un ordenador recibe una difusión con una IP que no coincide con la suya, ignora la información, pero si se envía un nuevo broadcast, el ordenador la procesa [2], [40].
- La capacidad de interpretar adecuadamente e incorporar la señal recibida con otras señales se denomina Integración. La integración se define como la capacidad de un sistema para

recibir múltiples señales y ensamblar una respuesta unificada adecuada a las necesidades de la célula u organismo. Por lo que, el procesamiento simultáneo de las señales recibidas en el destino en los sistemas de comunicación digital y los sistemas biológicos pueden ser análogo [9].

Al igual que en los sistemas de comunicación digital, la recepción de información biológica a través de receptores puede ser descrita por estados discretos, es decir, unido (B) o desunido (U). En el estado U, el receptor espera la llegada de información molecular; una vez que la información arriba entra en el estado B. En el estado B, el receptor no puede recibir otra información molecular (lo que lo vuelve insensible a otra señal), y se requiere cierta cantidad de tiempo de procesamiento antes de que el receptor regrese al estado U [43].

Este proceso de unión puede ser representado por una cadena de Markov en tiempo discreto utilizando pasos de longitud Δt y una matriz de probabilidad de transición de estados (en el i -ésimo paso) [43]:

$$P_i = \begin{bmatrix} 1 - \Delta t \cdot r_{UB} \cdot C(i\Delta t) & \Delta t \cdot r_{UB} \cdot C(i\Delta t) \\ \Delta t \cdot r_{BU} & 1 - \Delta t \cdot r_{BU} \end{bmatrix} \quad (2.18)$$

Donde $r_{UB} \cdot C(i\Delta t)$ es la tasa de transición de U a B, la cual es proporcional a la concentración del ligando $C(i\Delta t)$, y la tasa de transición de B a U es r_{BU} , la cual es independiente de la concentración del ligando [43].

En la ecuación (2.18), si $C_i = C(i\Delta t)$, $\alpha_{c_i} = \Delta t r_{UB} C(i\Delta t)$ y $\beta = \Delta t r_{BU}$, se convierte en:

$$P_i = \begin{bmatrix} 1 - \alpha_{c_i} & \alpha_{c_i} \\ \beta & 1 - \beta \end{bmatrix} \quad (2.19)$$

Donde α_{c_i} es la probabilidad de transición de U a B, que es sensible a la concentración de ligando de entrada c_i , y β es la probabilidad de transición de B a U, que es insensible a la transición. Si consideramos la concentración extrema de ligando desde la concentración mínima permitida $C_i = L$ hasta la concentración máxima permitida $C_i = H$, el proceso de transducción de señal es una cadena de Markov no homogénea en el tiempo según sea (2.20) o (2.21) [43].

$$P_H = \begin{bmatrix} 1 - \alpha_H & \alpha_H \\ \beta & 1 - \beta \end{bmatrix} \quad (2.20)$$

$$P_L = \begin{bmatrix} 1 - \alpha_L & \alpha_L \\ \beta & 1 - \beta \end{bmatrix} \quad (2.21)$$

La elección entre (2.20) o (2.21) está determinada por la concentración de ligando [43].

Como se ha mencionado anteriormente, la expresión genética y el transporte de proteínas podría considerarse una interred en la que la comunicación intracelular en el extremo emisor emula una red LAN en la que la distancia del RE al AG (d_{LAN} en la Figura 9) es de $2\mu\text{m}$. No consideramos en d_{LAN} la distancia entre los ribosomas y el RE porque los ribosomas que participan en este proceso se encuentran en el citosol y en el RE; por lo tanto, esta distancia no es fija [5]. Sin embargo, si esta distancia pudiera considerarse d_{LAN} , seguiría representando una distancia de cobertura de LAN biológica, porque representa una distancia intracelular. De forma similar a las redes típicas, la WAN biológica cubre una distancia más significativa que las LAN biológicas [44]. Además, de forma similar a las redes convencionales, las LAN biológicas están interconectadas por routers [45].

Por lo tanto, en el extremo transmisor, el AG actúa como un router biológico. Consideramos que el extremo receptor, es decir, el órgano objetivo, también se comporta como un router porque entre un conjunto de receptores que son probables destinos conectados por un canal de transmisión común (es decir, el torrente sanguíneo), los receptores a los que va dirigida la información (es decir, con el direccionamiento adecuado) reaccionan a la información recibida. Una vez más, el caso descrito puede ser equivalente al de las redes Ethernet en las que se envía una difusión a través del canal de transmisión (con una topología de bus lógico) a cada ordenador de la red, pero sólo el ordenador con la dirección IP adecuada procesa la comunicación recibida [5]. Por lo tanto, el direccionamiento de extremo a extremo podría ser comparable al direccionamiento de red en la capa de red [2] como se puede ver en la Figura 9.

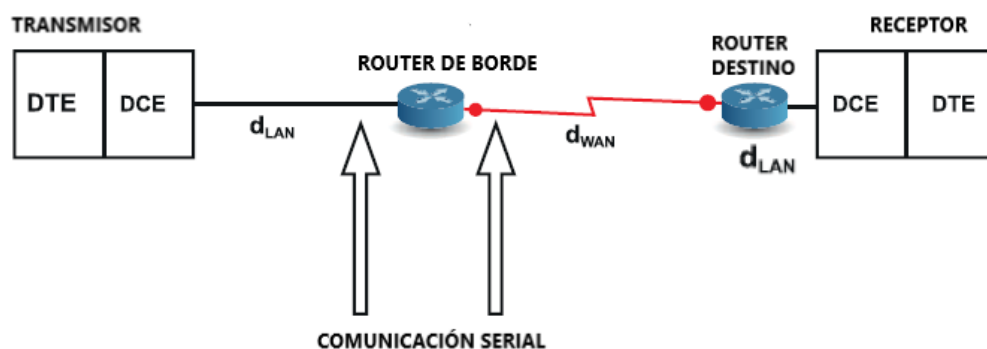


Figura 9. La equivalencia interna de redes de la expresión genética.

Fuente: [8]

Una vez que la célula de destino recibe los datos biológicos, la información se comunica a otros orgánulos empleando una dirección adyacente implícita que es comparable a la dirección de la capa de enlace de datos utilizada para facilitar la comunicación dentro de un

rango directo de comunicación [1], [32]. El mensaje biológico recibido se transmite físicamente a la célula de destino. La Figura 10 presenta un modelo de red en capas de extremo a extremo de la expresión genética y el transporte de proteínas desde la perspectiva de una red de comunicación típica para la expresión genética como una interred [1]. La estructura de un modelo por capas descompone un sistema a gran escala en un conjunto de unidades más pequeñas (es decir, capas) que son funcionalmente independientes entre sí y especifica las interacciones entre las capas. Una ventaja de una pila de capas es el uso fundamental de la capa de enlace de datos para transformar un canal imperfecto en una línea libre de errores de transmisión o informar de problemas no resueltos a la capa superior [40], por lo tanto, la aplicación de un modelo de este tipo modelo a los sistemas biológicos (por ejemplo, la administración de fármacos) es altamente confiable [2].

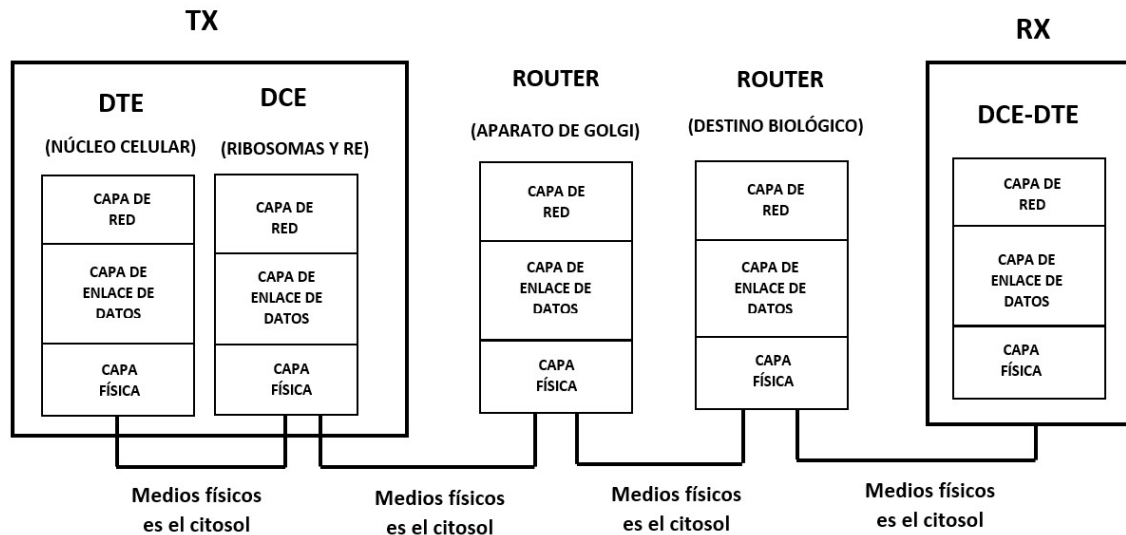


Figura 10. La expresión genética como un modelo de red en capas.

Fuente: [8]

2.9 Tipos de Modulación

En CM la propagación de moléculas de información se basa en la difusión principalmente, los símbolos que son enviados sean decodificados en el receptor. Existen diferentes técnicas que se pueden usar para llegar a dicho objetivo, el símbolo puede ser modulado tomando en cuenta las propiedades de la molécula mensajera (o de información) que llega al receptor (ejemplo: concentración, frecuencia, tipo de molécula) para formar o reconstruir la señal de información [46].

2.9.1 Modulación por desplazamiento de concentración (CSK)

Consiste en la emisión de un número de moléculas de un mismo tipo que conllevan la información. En CSK la concentración de la molécula recibida representa la amplitud de la señal. CSK es el análogo a ASK (Modulación por desplazamiento de amplitud) en los sistemas de comunicación convencional. CSK puede ser implementada en la práctica como BCSK (caso binario) o multinivel MCSK dependiendo de la velocidad por símbolo escogida y del número de moléculas liberadas. Esta técnica también contiene, la forma más simple OOK (Modulación de encendido y apagado) que es el desplazamiento de encendido y apagado que representa el envío o no de partículas o moléculas de información [47]. El receptor decodifica los símbolos utilizando diferentes niveles de concentración de acuerdo a la cantidad de bits por símbolo en un sistema multinivel MCSK, como se muestra en la Figura 11. CSK es afectada por interferencia entre símbolos (ISI por sus siglas en inglés), debido a los excedentes de moléculas anteriores, retardos en la llegada de las partículas de información y/o reacciones químicas que cambian los niveles de concentración. En el intento por mejorar el efecto de la ISI se propone la modificación Zebra-CSK [48], la cual usa 2 tipos de moléculas: las moléculas mensajeras para codificar la información en símbolos y las moléculas para inhibir los residuos de símbolos anteriores. En esa búsqueda por reducir la ISI se propone la variación de CSK utilizando 2 tipos de moléculas conocida como desplazamiento por transición de molécula (MTSK) [49].

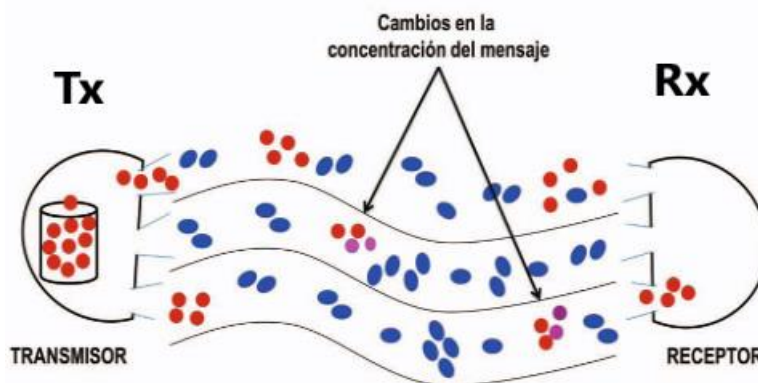


Figura 11. Funcionamiento de una modulación multinivel de CSK.

Fuente: [46]

2.9.2 Modulación por desplazamiento de molécula (MoSK)

La modulación por Desplazamiento de Molécula (MoSK) es una técnica en la que se emiten diferentes tipos de moléculas las cuales representan la información, $2n$ tipos de moléculas o

partículas de información son necesarias para enviar n bits por símbolo y obtener a multinivel MoSK (MMoSK). La emisión o liberación de partículas depende del símbolo a enviar y la detección en el receptor de la concentración en un intervalo específico. Al igual que en CSK, los excedentes de cualquier partícula de símbolos anteriores causan ISI, aunque MMoSK presenta este problema en menor grado que CSK [47]. Su mejor rendimiento frente a la ISI complica los diseños del receptor y transmisor debido a que el mecanismo molecular es más complejo y requiere de técnicas de codificación y detección bioquímicas certeras para reducir o eliminar errores. Según el símbolo deseado, el transmisor libera una de estas moléculas. Según el tipo y la concentración de moléculas recibidas dentro de un intervalo de tiempo, el receptor decodifica el símbolo deseado. Si la concentración de un solo tipo de molécula en el receptor supera el umbral, el símbolo se decodifica a partir de la secuencia de bits relacionada con esa molécula. Si el error de concentración, es decir, si la concentración de cualquier tipo de molécula no supera el umbral, o si la concentración de más de un tipo de molécula supera el umbral el símbolo no se decodifica [49].

2.9.3 Modulación basada en el radio molecular

Adicional a las técnicas propuestas, existe una nueva técnica de modulación basada en el radio molecular (IRSK por sus siglas en inglés) utilizando isómeros como partículas de información, específicamente aldohexosas [50]. La información es codificada utilizando el radio de las moléculas mensajeras. El esquema más simple desarrollado requiere el uso de 2 tipos de moléculas, siendo el mismo más robusto frente a ISI y presenta un diseño del receptor menos complejo en comparación con CSK bajo el mismo tipo de moléculas [46].

2.9.4 Modulación por tiempo transcurrido (TEC)

Es un esquema para sistemas de propagación muy lento como comunicación de bacterias dentro de un chip. En esta técnica la información es codificada mediante el intervalo de tiempo entre dos pulsos consecutivos, específicamente en se demuestra mediante fluorescencia en bacterias [51].

CAPÍTULO III. METODOLOGÍA

3.1 Tipo de estudio

Para la ejecución de la presente tesis se simulará un escenario real de un sistema de comunicación molecular que responde a la expresión genética, visto desde la perspectiva de un sistema de comunicación digital. En particular, las teorías de red, como el direccionamiento, el control de errores, el control de flujo, el control de tráfico y el teorema de Shannon, todos estos elementos se usan para diseñar un sistema de comunicación digital de extremo a extremo que se va a representar con la expresión genética.

Se propone un modelo de red en capas que en términos de la expresión genética que representa la transcripción y traducción ADN y por ende la transmisión de extremo a extremo de proteínas a un órgano objetivo, dicho modelo aprovecha las características más importantes de los sistemas de comunicación digital, como la eficiencia y el rendimiento para transmitir la información biológica, por lo que corresponde a un tipo de investigación experimental ya que, se centra en la observación de los resultados obtenidos a partir de la manipulación de las variables independientes del estudio, esto dentro de un ambiente controlado, simulador de eventos NS-3, de este modo se espera dar un análisis a las señales de rendimiento de la comunicación resultantes.

Con los datos obtenidos de la simulación podremos realizar un análisis de cuáles son las mejores condiciones que se deben usar para que la transmisión de información biológica al órgano objetivo llegue sin errores y con la información completa.

3.2 Métodos de investigación

- **Cuantitativo**

Se deberá realizar una medición de las variables involucradas en la investigación permitiéndonos entender mediante métodos estadísticos que nos permitan realizar un análisis de los resultados para determinar si existen condiciones ideales para que la transmisión de información biológica sea la más eficiente posible.

3.3 Fuentes de Información

Se ejecuta una revisión sistemática de varios documentos, revistas, libros, papers, artículos científicos entre otros, con el fin de recaudar información importante relacionada con el tema de estudio. Las fuentes de información tienen que ser de entidades científicas conocidas y confiables.

3.4 Población y muestra

3.4.1 Población

La población son los datos que se obtendrán luego de ir modificando los parámetros que forman la comunicación molecular, al momento de ejecutar las simulaciones forman la población de estudio.

3.4.2 Muestra

Es un grupo representativo de la población, que con las modificaciones que se realizan se analizan las diferentes métricas haciendo uso de los datos recopilados.

3.5 Procesamiento y análisis

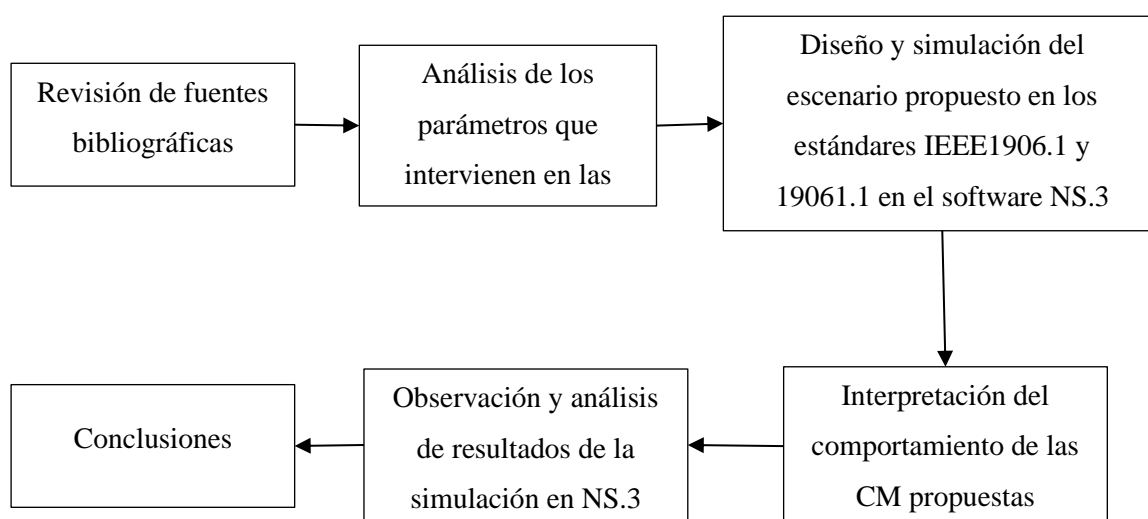


Figura 12. Procesamiento y análisis de la investigación.

Fuente: Autores

3.6 Operacionalización de las variables

Tabla 1. Operacionalización de las variables.

VARIABLES	CONCEPTO	INDICADORES
Clase molecular	Variable independiente, es la que se pone a prueba a nivel experimental, la misma que se manipula en la simulación.	<ul style="list-style-type: none">• Distancia entre nodos• Coeficiente de difusión• Tipo de moléculas.• Intervalo del pulso molecular• Velocidad de transmisión

Rendimiento de transmisión y recepción	Variables dependientes, son aquellas que van a ser estudiadas en función de la manipulación de las variables independientes.	<ul style="list-style-type: none"> • Pulso molecular • Concentración Molecular • Pulso de retardo • Capacidad del canal • Ancho del pulso molecular
--	--	--

3.7 Procesamiento y Análisis

Para analizar el comportamiento de las CM, se exploran los factores que inciden en cada una de las partes del proceso. Esto posibilita comprender de qué manera ciertos elementos impactan en la eficacia de la comunicación a nivel molecular. Además, se lleva a cabo la evaluación de parámetros de la simulación utilizando datos reales recopilados de [5], facilitando la comparación de diversos escenarios simulados. Basándonos en lo mencionado anteriormente, se implementan clases que permiten simular las distintas partes de los estándares, lo que habilita la medición del rendimiento de la comunicación molecular al modificar variables como:

- Coeficiente de difusión.
- La distancia entre nodos.

Con estas modificaciones, podremos evaluar las características de rendimiento en la comunicación a través de elementos como:

- Concentración molecular.
- Tiempo de delay.
- Amplitud máxima del pulso.
- Ancho del pulso molecular.

3.8 Modelo de implementación en NS-3

Los archivos .cc y .h forman la base del módulo IEEE P1906, siendo elementos cruciales para la implementación en C++. Mientras que los archivos .h incluyen declaraciones de clases, estructuras y funciones para la comunicación entre componentes, los archivos .cc contienen la implementación del código, definiendo funciones y métodos. La estructura de directorios se organiza como una biblioteca dinámicamente utilizada por el archivo principal mol-example.cc, en la Figura 13 se puede apreciar la estructura de estos archivos. Este último

ejecuta la simulación en NS-3, define la topología y configuración de escenarios, y realiza ajustes en parámetros fundamentales y funciones comunes.

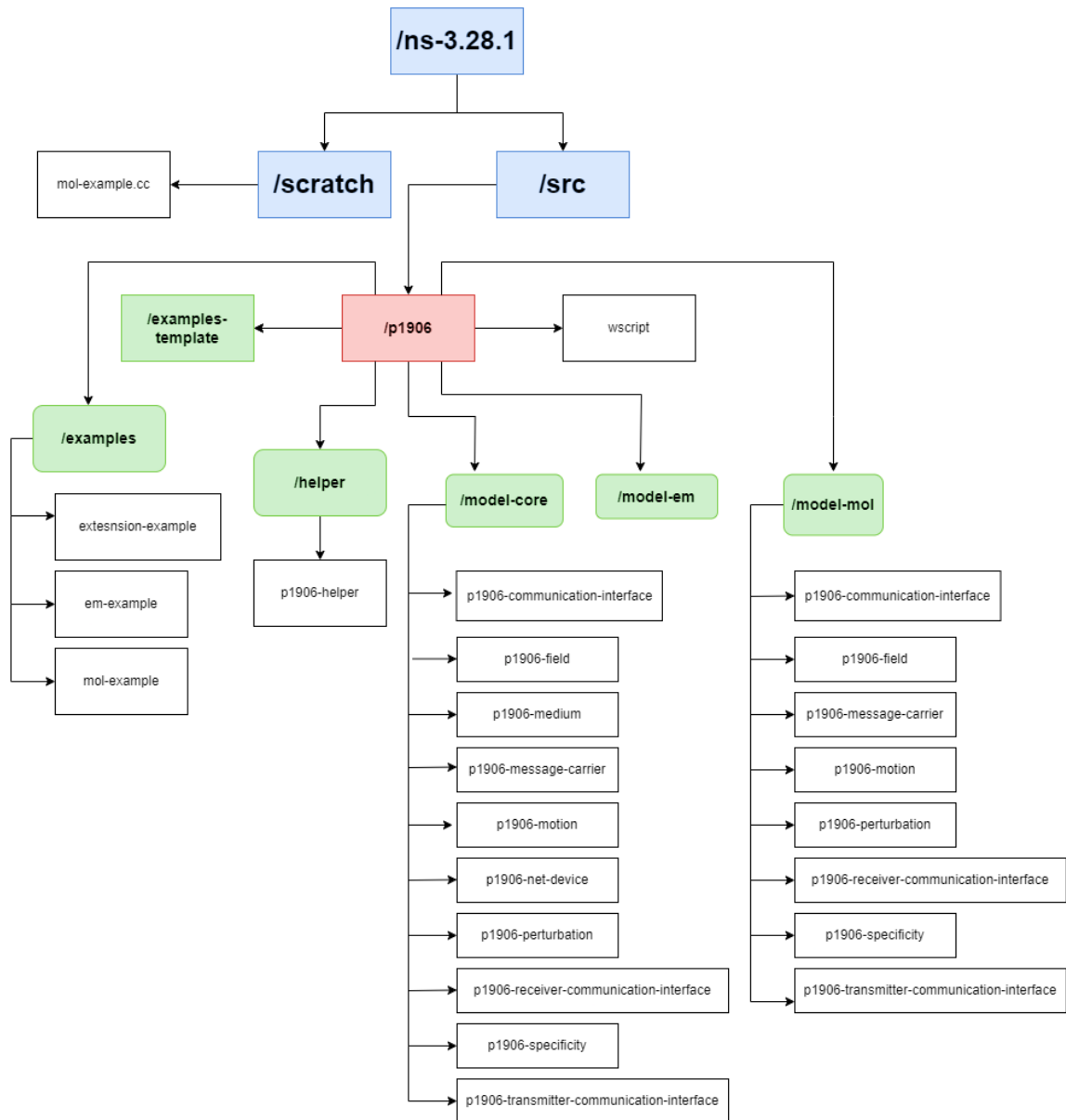


Figura 13. Directorio del estándar p1906.1 y p1906.1.1.

Fuente: Autores

Las clases que implementan los componentes del marco IEEE 1906.1 y 1906.1.1 [52] son:

- *P1906MOLMessageCarrier.*
- *P1906MOLMotion.*
- *P1906Field.*
- *P1906Perturbation.*
- *P1906MOLSpecificity.*

Además de las clases que se describen anteriormente existen otras clases que implementan otras entidades que, a pesar de estar involucradas en el proceso de la comunicación no se clasifican como “componentes”:

- *P1906CommunicationInterface*.
- *P1906TransmitterCommunicationInterface*.
- *P1906ReceiverCommunicationInterface*.
- *P1906Medium*.
- *P1906NetDevice*.

El grupo de investigación IEEE P1906 utiliza el Lenguaje Unificado de Modelado (UML) para representar el proceso de Comunicaciones Moleculares (CM). En la Figura 14, se destacan los pasos principales del módulo [52].

1. *NetDevice* (emisor-fuente) obtiene un mensaje desde las capas superiores, el mismo que se transfiere a la interfaz de comunicación.
2. El componente Perturbación (*Perturbation*) crea el componente del portador de mensajes proporcionado por (*Message Carrier*).
3. La propagación en el medio se activa con la interfaz de comunicación del transmisor y pasa por los demás componentes a la portadora de mensajes, perturbación y campo (*Field*).
4. Las propiedades de la portadora de mensajes se modifica con el componente Movimiento (*Motion*), como, por ejemplo, el retardo y la perdida de propagación.
5. El componente Especificidad verifica la compatibilidad de la portadora que se entrega en el receptor.
6. El mensaje se entrega únicamente cuando la compatibilidad ocurre y se entrega a las capas superiores.
7. Finalmente las capas superiores obtienen el mensaje final.

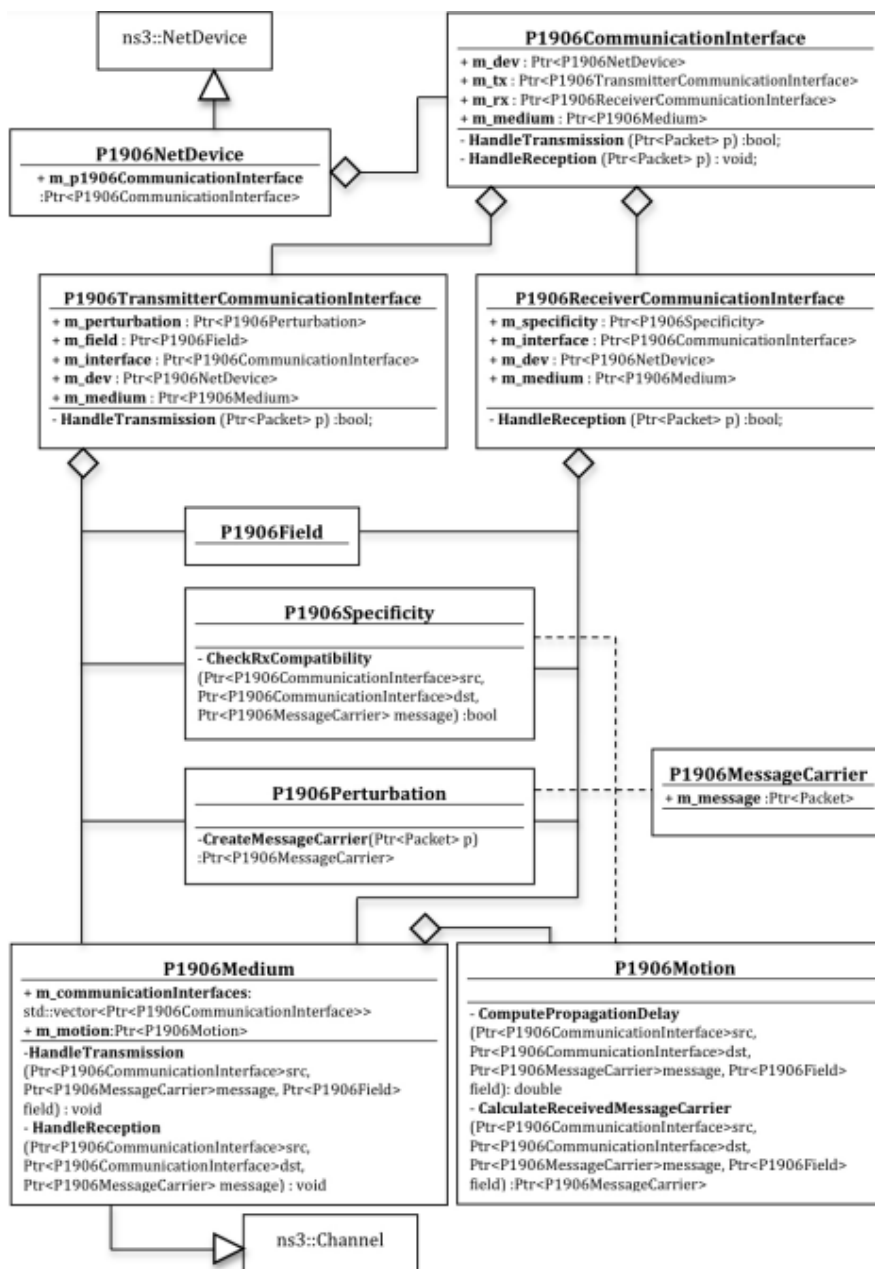


Figura 14. Diagrama UML del marco de la comunicación molecular.
Fuente: [52]

3.8.1 Conexión de nodos de comunicación y dispositivos

NetdeviceContainer o NodeContainer, son las dos clases con las que se pueden crear los nodos de comunicación, para simplificar este proceso la clase helper cuenta con todos los métodos necesarios para establecer métodos específicos del módulo P1906NetDevice, de igual manera, cuenta con funciones para conectar todo el sistema y crear instancias.

En la clase NodeContainer usando el puntero *Ptr<Node>*, se puede acceder a los métodos y atributos del objeto Node, con ellos podemos modificar las propiedades como la distancia,

modelo de movilidad y posición. Dentro la simulación con el puntero es mucho más sencilla acceder y configurar los métodos y atributos de cada nodo según el escenario que se nos presente [52].

La clase P1906NetDevice modela el dispositivo de red que participa en el proceso de comunicación. Ha sido concebida como un contenedor para el conjunto completo de protocolos y la interfaz física. Al extender la clase general ns3::NetDevice disponible en la versión oficial de NS-3, ya cuenta con los métodos básicos que gestionan: 1) la identificación única del dispositivo en la red, 2) el manejo de la movilidad, 3) proporcionar la interfaz a la capa de aplicación, y 4) ofrecer acceso directo a cualquier parte del conjunto de protocolos y la capa física del dispositivo modelado. Todas estas características son muy útiles para construir aplicaciones específicas sobre el marco de comunicación general desarrollado por el grupo de trabajo IEEE 1906.1 [52].

3.8.2 Clase MessageCarrier

La clase P1906MessageCarrier implementa el mensaje intercambiado entre dispositivos. Por defecto, tiene solo un miembro de datos: un puntero al objeto ns3::Packet, que representa la cadena de bits almacenada con el mensaje. Se requieren parámetros adicionales para personalizar la definición del componente Message Carrier según un modelo de comunicación específico y un caso de uso determinado [52].

En la clase extendida ns3::P1906MOLMessageCarrier se encuentran los siguientes parámetros:

- Número de moléculas
- Intervalo de pulso
- Tiempo de inicio
- Duración
- Mensaje por transmitir (paquete)

Todos estos parámetros son establecidos por el componente de perturbación antes de la transmisión física.

A través de los métodos SetMessage y GetMessage, funciones heredadas del /model-core, capturan el mensaje y lo deposita en el portador [52].

3.8.3 Clase Perturbación (Perturbation)

La clase perturbación corresponde a la componente *ns3::P1906MOLPerturbation*, esta es la encargada de realizar la modulación del paquete. Se puede hacer la comparación con la modulación de señales digitales, donde la variación principal puede ser el movimiento o concentración de las moléculas, además se puede obtener y retener cierta información como el número de moléculas por cada pulso y el intervalo de tiempo entre pulsos sucesivos que se transmiten. El método *CreateMessageCarrier* es usado para codificar la información que se transmitirá y asociarla con la portadora correspondiente. Se creó el objeto de mensaje portadora *P1906MOLMessageCarrier* a partir del paquete *Packet*, el método comienza creando la función y asignado una variable de tipo *Ptr<P1906MOLMessageCarrier>* que es empleado como almacenamiento y transporte de información de la modulación que se esté empleando, iterando sobre el tamaño del paquete realizando la operación de verificación de bit y agregando la cadena correspondiente [52].

La modulación MoSK es la que se implementara en esta clase, se utilizan cuatro moléculas diferentes (Adenina, Guanina, Citosina y Uracilo) que se pueden combinar para formar diferentes codones que, a su vez forman los 20 aminoácidos de la expresión genética, una vez formado el mensaje se envía al medio de transmisión.

3.8.4 Clase Motion

En este componente se implementa el modelo de propagación que define cómo se transmitirá el mensaje, usando específicamente un enfoque basado en difusión. Con el objetivo de lograr un escenario lo más realista posible, se emplea el movimiento browniano con arrastre para simular el desplazamiento de las moléculas a través del torrente sanguíneo. En la simulación, se considera que no hay obstáculos entre el transmisor y el receptor, lo que implica que el movimiento es ideal y no se ve afectado por otras entidades [52].

3.8.5 Clase Campo (Field)

La clase campo corresponde a la componente *ns3::P1906MOLField*, su función es actuar como una guía de onda para proporcionar la capacidad de direccionar una portadora de mensaje (message carrier) hacia su receptor. Según la necesidad de estudio se puede implementar de manera interna o externa en relación con el medio molecular de transmisión, en este trabajo de investigación se utilizará una implantación externa ya que, la función del movimiento browniano con arrastre usara los parámetros que son enviados al medio por esta

clase, para luego ser la encargada de enviar estos mismos parámetros al componente movimiento [52].

3.7.6 Clase Especificidad (Specificity)

En la recepción la función `HandleReception` activa el procedimiento de las clases `ns3::MOLReceiverCommunicationInterface` para comenzar con verificando que la interfaz `m_interface` pueda recibir los mensajes transmitidos. Por lo que, se llama al método `CheckRxCompatibility` de la función `P1906Specificity` para determinar si el mensaje llega o no correctamente, esto se implementa en la componente `ns3::MOLSpecificity` para la recepción de un portador de mensaje y lo hace comprobando si la capacidad del canal es mayor que la tasa de transmisión, en cuyo caso se realiza la recepción y demodulación de lo contrario no se realizan esta acción.

La familia de los estándares IEEE P1906 retorna un valor booleano (Verdadero o Falso) para indicar si se cumple o no el límite de Fick (umbral que establece la familia de los estándares) [52].

3.9 Representación binaria de la expresión genética

Se utiliza un alfabeto de cuatro letras para codificar la información genética; cada letra corresponde a una base de nucleótidos. El alfabeto de aminoácidos, por otro lado, se compone de 20 letras distintas, una para cada aminoácido. Como resultado, de acuerdo con la teoría de la información, la cantidad total de información transportada por la base de un solo nucleótido es de 2 bits $\left[-\log_2\left(\frac{1}{4}\right)\right]$ y la cantidad de información necesaria para elegir claramente un aminoácido de una colección de 20 sin confusiones es de 4,3 bits $\left[-\log_2\left(\frac{1}{20}\right)\right]$ [5], [53].

Las codificaciones digitales en [5] de los nucleótidos se muestran en la Tabla 2.

Tabla 2. Codificación de los nucleótidos

BASE DE NUCLEÓTIDOS	CODIFICACIÓN
U	00
C	01
A	10
G	11

El uso de 6 bits para designar un aminoácido representa un exceso de información; sin embargo, esta capacidad de información excedente puede explicar la redundancia del código genético. La redundancia es un método utilizado en los sistemas de comunicación digital para controlar la información y detectar fallas. De manera similar, el código genético contiene información redundante para reducir los errores en la expresión genética [5], [40]. El código genético utiliza tres nucleótidos (codones) para representar un aminoácido específico [32], lo que permite el uso de un decodificador con 64 salidas para describir un aminoácido. Debido a que hay 20 aminoácidos, se requiere agrupar los cuatro nucleótidos en grupos de al menos tres para codificar los 20 aminoácidos en 64 combinaciones viables (es decir, $4^3 = 64$ porque $4^2 = 16$ es insuficiente para codificar 20 aminoácidos) [5]. Para simular el proceso de traducción del ADN, comenzamos por procesar los tripletes de nucleótidos específicos empleando un selector de codones Figura 15 para generar una representación de seis bits para cada codón individual. Después, para finalizar el proceso de traducción, es esencial relacionar estos seis bits con la representación digital correspondiente de un aminoácido, que se encuentra en la Tabla 2. Para este propósito, empleamos un decodificador digital que convierte los seis bits de entrada en una salida de 64 bits [5].

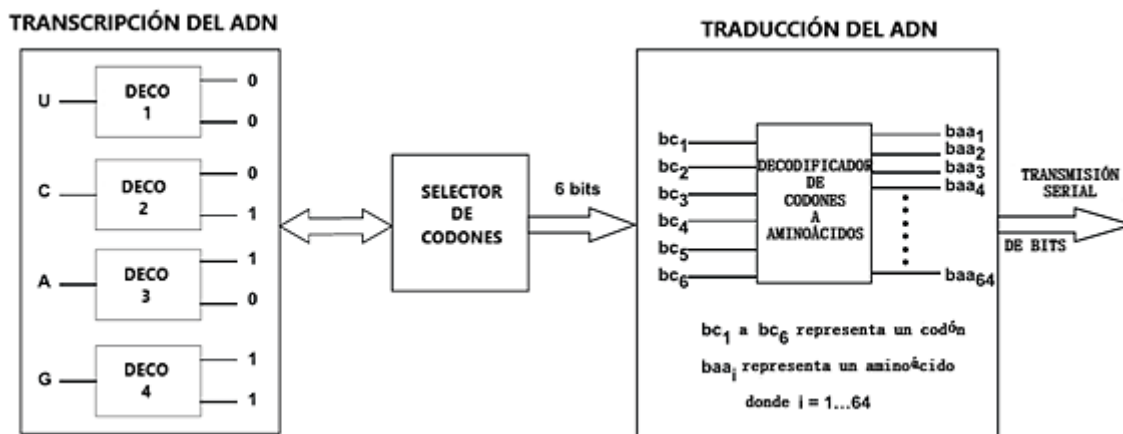


Figura 15. La expresión genética representada en bits.
Fuente: [5]

En la Figura 16 se puede observar el proceso que se describió anteriormente donde conformamos el triplete de nucleótidos (codon), y luego utilizamos el decodificador digital para formar su correspondencia en binario y poder enviar la información en el medio.

```

yeimy@yeimy-VirtualBox: ~/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1
P1906CommunicationInterface:SetP1906Medium(0x8067c0)
P1906TransmitterCommunicationInterface:SetP1906Medium(0x803f80)
P1906ReceiverCommunicationInterface:SetP1906Medium(0x80d0e0)
P1906CommunicationInterface:GetP1906TransmitterCommunicationInterface(0x8067c0)
P1906TransmitterCommunicationInterface:SetP1906Perturbation(0x803f80)
P1906CommunicationInterface:GetP1906TransmitterCommunicationInterface(0x8067c0)
P1906TransmitterCommunicationInterface:SetP1906Field(0x803f80)
P1906CommunicationInterface:GetP1906ReceiverCommunicationInterface(0x8067c0)
P1906ReceiverCommunicationInterface:SetP1906Specificity(0x80d0e0)
P1906Specificity:SetP1906CommunicationInterface(0x8a6d40)
P1906Medium:AddP1906CommunicationInterface(0x86a8e0)
- Prolina: CCU
Packet: 010100
P1906CommunicationInterface:HandleTransmission(0x869110, "Transmitting a packet
[id,size]", 0, 1)
P1906TransmitterCommunicationInterface:HandleTransmission(0x7f7b00)
P1906MOLPerturbation>CreateMessageCarrier(0x831c80)
P1906MessageCarrier:P1906MessageCarrier(0x81fbd0)
P1906MOLMessageCarrier:P1906MOLMessageCarrier(0x81fbd0)
P1906MessageCarrier:SetMessage(0x81fbd0)
P1906MOLPerturbation>CreateMessageCarrier(0x831c80, "[t,bits,pulseI,duration]",
0, 8, +7000000000.0ns, 56)

```

Figura 16. Transcripción del ADN, cuadro verde: codón del aminoácido (Prolina), cuadro rojo: decodificación binaria.

Fuente: Autores

3.10 Diseño matemático

Esta sección la dividiéremos en dos partes: la primera será para el modelo matemático implementado en el simulador, y el segundo para el contraste de los resultados, se determinarán las ecuaciones para el movimiento browniano con drift que se usaron en el simulador y las ecuaciones que validaran los resultados experimentales. Todas estas ecuaciones se obtuvieron de la segunda ley de Fick.

3.10.1 Modelo matemático implementado en el simulador

Para simular un movimiento browniano con arrastre se parte de la ecuación 2.8, al resolver esta ecuación para dos dimensiones en coordenadas rectangulares, se obtiene:

$$x_i = x_{i-1} + \sqrt{4D\Delta t} \cdot \cos(\theta_i) + v_x \Delta t \quad (3.1)$$

$$y_i = y_{i-1} + \sqrt{4D\Delta t} \cdot \sin(\theta_i) + v_y \Delta t \quad (3.2)$$

Las ecuaciones (3.1) y (3.2) se considera la velocidad del medio y se suma sus componentes (v_x, v_y) respectivamente, la velocidad que se empleó en las simulaciones se tomó de [17].

Con estas ecuaciones se simula el movimiento de las moléculas por el medio hasta llegar a su receptor, en cada paso dado el receptor verifica si las moléculas llegaron o se perdieron durante la transmisión. Esto se contrasta con la concentración de moléculas obtenida en el receptor con el modelo establecido por las leyes de Fick.

3.10.2 Modelo matemático para el contraste de resultados

Se parte de la segunda ley de Fick ecuación (2.10) en [18] se puede obtener la solución de la ecuación diferencial para un plano infinito, el resultado para dos dimensiones se obtiene en la ecuación (2.13). Se modela teóricamente la concentración molecular en el receptor para un movimiento browniano con drift. Los parámetros que se usaran para el análisis de los resultados son: tiempo de delay, amplitud máxima de pulso y ancho del pulso mínimo.

3.10.2.1 Tiempo de delay

El tiempo de delay es el que representa el tiempo medido en el receptor desde que se contabiliza la primera molécula hasta donde la concentración de moléculas es máxima. Para determinar este parámetro se deriva la ecuación (2.12) e igualando a 0 [53] y se obtiene:

$$\frac{dc(r, t)}{dt} = \frac{d}{dt} \frac{Q e^{-\frac{r^2}{4Dt}}}{(4\pi Dt)^{\frac{3}{2}}} = 0 \quad (3.3)$$

Despejando la variable t obtenemos el tiempo de delay.

$$t_d = \frac{r^2}{4D} \quad (3.4)$$

Como podemos ver en la ecuación (3.4) el delay se expresa en función de la distancia al cuadrado entre el receptor y transmisor, y el coeficiente de difusión. Por lo que, podemos decir que el retardo del pulso aumenta cuadráticamente cuando la distancia también lo hace.

3.10.2.2 Amplitud máxima de pulso

La variación de la amplitud del pulso en el espacio, que puede interpretarse como la atenuación del canal, es un aspecto crucial para analizar. Para determinar su valor, lo calcularemos evaluando la concentración molecular en el momento en que alcanza su máximo [53]. A continuación, se presenta su representación en forma de ecuación:

$$C_{max} = \left(\frac{3}{2\pi e} \right)^{\frac{3}{2}} \frac{Q}{r^3} \quad (3.5)$$

En este caso, no observaremos que el coeficiente de difusión afecte la amplitud del pulso, es decir, que el coeficiente de difusión no causará atenuación en los pulsos moleculares a lo largo del espacio.

3.10.2.3 Ancho de pulso molecular

Una medida que restringe el ancho de banda alcanzable es el ancho del pulso. Siguiendo la práctica común en comunicaciones electromagnéticas, calculamos el ancho del pulso al 50% de su capacidad, es decir, en el intervalo de tiempo en el que el pulso tiene una amplitud mayor a la mitad de su valor máximo [53]:

$$\frac{C_{max}}{2} = \frac{1}{2} \left(\frac{3}{2\pi e} \right)^{\frac{3}{2}} \frac{Q}{r^3} \quad (3.6)$$

Obtenemos la siguiente expresión al despejar la variable del tiempo:

$$t = - \frac{r^2}{6DW \left(-\frac{1}{2^{2/3}e} \right)} \quad (3.7)$$

La ecuación presenta dos soluciones, que indican los dos puntos en el tiempo en los que la amplitud del pulso alcanza la mitad de su máximo valor. Estos puntos son calculados a través de la función Lambert W, dichas soluciones son:

$$t_1 = -\frac{0.0728}{D} r^2, t_2 = -\frac{0.5229}{D} r^2 \quad (3.8)$$

Por último, podemos encontrar la expresión para el ancho del pulso t_w al restar estos dos momentos en el tiempo.

$$t_w = t_2 - t_1 = \frac{0.4501}{D} r^2 \quad (3.9)$$

En esta situación, al igual que sucedió con el retardo del pulso, la amplitud del pulso disminuye cuando el coeficiente de difusión del medio aumenta. En resumen, un coeficiente de difusión mayor resulta en pulsos recibidos más estrechos [53].

CAPÍTULO IV. RESULTADOS Y DISCUSIÓN

Para el análisis de los resultados se realizaron 30 simulaciones con los parámetros establecidos en la Tabla 3.

Para validar los resultados que se obtuvieron en las diferentes simulaciones, se utilizaron los parámetros del canal molecular descritas en el capítulo III.

Tabla 3. Parámetros generales de la simulación

Número de moléculas enviadas	Distancia de transmisión	Coefficiente de difusión	Pulso del intervalo	Velocidad de arrastre
5000	2.1 μm	1 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$
5000	2.1 μm	3 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$
5000	2.1 μm	5 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$
5000	2.1 μm	7 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$
5000	2.1 μm	10 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$
5000	2.1 μm	100 $\mu\text{m}^2/\text{s}$	7000	1 $\mu\text{m}/\text{s}$

4.1. Concentración molecular

Los estándares IEEE 1906 en NS-3 tiene como resultado principal la concentración molecular que se calcula en función del tiempo y la distancia. Se establece un transmisor que libera moléculas en el medio, las mismas que siguen un movimiento browniano con arrastre, en el extremo receptor la nanomáquina receptora puede contar el número de moléculas en un volumen cercano, con lo que se puede calcular la concentración molecular. Los estándares IEEE 1906.1 y 1906.1.1 se rigen a las leyes de Fick que se describieron en la sección II, estas establecen la distancia máxima a la que se puede establecer la comunicación entre el transmisor y receptor esto para prevenir errores durante la transmisión. La evaluación de la comunicación depende de la concentración molecular, entre más concentración exista mejor es la comunicación, sin embargo, si la distancia entre el nodo transmisor y receptor es muy grande la concentración no alcanzara el umbral impuesto en el simulador por las leyes de Fick lo que ocasionaría la pérdida de información que se envió.

En la Figura 17 a) el (cuadro verde) nos indica que la ley de Fick se ha respetado lo que sugiere que cumplió con el umbral establecido, por ende, el mensaje se recibió correctamente (cuadro azul). La Figura 17 b) permite visualizar que con una distancia diferente (0.001 m) a la que se estableció en la tabla 3 la ley de Fick no se respeta (cuadro rojo), lo que manifiesta que no cumplió con el umbral establecido, con ello en mensaje no se recibió correctamente (cuadro gris).

a)

```

yeimy@yeimy-VirtualBox: ~/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1
P1906MOLMotion:CalculateReceivedMessageCarrier(0x2683d90, "Do nothing for the Fick's law")
P1906Medium:HandleReception(0x26868e0)
P1906CommunicationInterface:GetP1906ReceiverCommunicationInterface(0x26227c0)
P1906MOLReceiverCommunicationInterface:HandleReception(0x26290e0)
P1906ReceiverCommunicationInterface:GetP1906Specificity(0x26290e0)
P1906MessageCarrier:GetMessage(0x263bbd0)
P1906MOLSpecificity:CheckRxCompatibility(0x26c2d40)
P1906MessageCarrier:GetPulseInterval(0x263bbd0)
P1906CommunicationInterface:GetP1906NetDevice(0x2685110)
P1906NetDevice:GetNode(0x2628ff0)
P1906CommunicationInterface:GetP1906NetDevice(0x26227c0)
P1906NetDevice:GetNode(0x26bb2b0)
P1906MOLSpecificity:CheckRxCompatibility(0x26c2d40, "testcapacity: [distance, txRate, channelCapacity]",
2.1e-06, 0.142857, 0.00198494)
P1906MOLSpecificity:CheckRxCompatibility(0x26c2d40, "Fick's bound has been respected")
P1906MessageCarrier:GetAminoacido(0x263bbd0)
P1906MOLReceiverCommunicationInterface:HandleReception(0x26290e0, "message received correctly")
P1906MessageCarrier:GetMessage(0x263bbd0)
P1906MessageCarrier:GetModulation(0x263bbd0)

```

b)

```

yeimy@yeimy-VirtualBox: ~/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1
P1906ReceiverCommunicationInterface:GetP1906Specificity(0x12cd0e0)
P1906MessageCarrier:GetMessage(0x12dfbd0)
P1906MOLSpecificity:CheckRxCompatibility(0x1366d40)
P1906MessageCarrier:GetPulseInterval(0x12dfbd0)
P1906CommunicationInterface:GetP1906NetDevice(0x1329110)
P1906NetDevice:GetNode(0x12ccff0)
P1906CommunicationInterface:GetP1906NetDevice(0x12c67c0)
P1906NetDevice:GetNode(0x135f2b0)
P1906MOLSpecificity:CheckRxCompatibility(0x1366d40, "testcapacity: [distance, txRate, channelCapacity]",
0.001, 0.142857, 450.1)
P1906MOLSpecificity:CheckRxCompatibility(0x1366d40, "Fick's bound has NOT been respected --> transmission
failed")
P1906MessageCarrier:GetAminoacido(0x12dfbd0)
P1906MOLReceiverCommunicationInterface:HandleReception(0x12cd0e0, "message NOT received correctly")
P1906MessageCarrier:GetMessage(0x12dfbd0)
P1906MOLMessageCarrier::~P1906MOLMessageCarrier(0x12dfbd0)
P1906MessageCarrier::SetMessage(0x12dfbd0)
P1906MessageCarrier::~P1906MessageCarrier(0x12dfbd0)
P1906NetDevice:DoDispose(0x12ccff0)
P1906NetDevice:DoDispose(0x135f2b0)
P1906Medium:DoDispose(0x132a8e0)
P1906MOLMotion::~P1906MOLMotion(0x1327d90)
P1906Motion::~P1906Motion(0x1327d90)

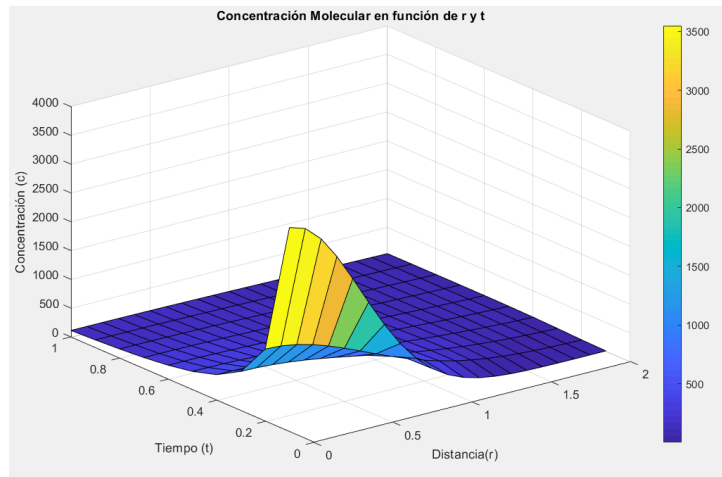
```

Figura 17. Resultado de la consola de la recepción del mensaje.

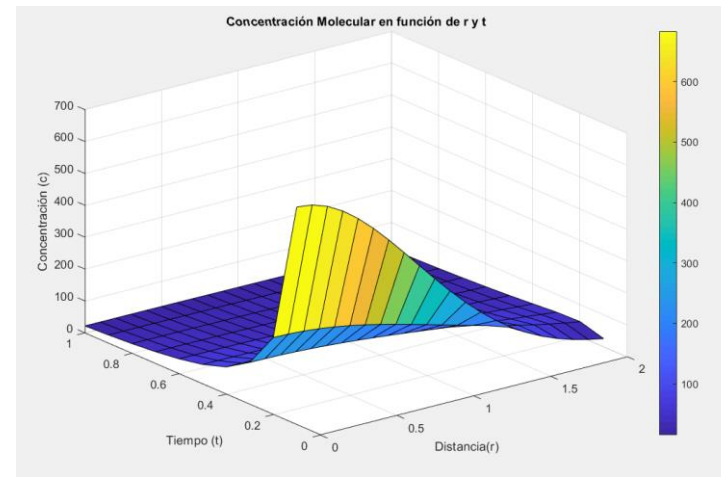
Fuente: Autores

La Figura 18 a) denota que la concentración molecular obtenida con el coeficiente de difusión de $1\mu\text{m}^2/\text{s}$ como esta se basa en una difusión ideal en el medio de propagación cada una de las moléculas que si liberan tienen la misma probabilidad dar el siguiente paso a la derecha o izquierda desde la posición inicial que tenía. Se puede visualizar como la concentración es más intensa en el primer instante de tiempo y a mitad que el tiempo avanza y con ello la distancia va disminuyendo debido al efecto de difusión, el pulso de retardo se considera este instante donde se registra el pico de concentración. Para los coeficientes $3\mu\text{m}^2/\text{s}$ y $5\mu\text{m}^2/\text{s}$ los resultados de la concentración molecular se muestran en la Figura 18 b) y 18 c) respectivamente.

a)



b)



c)

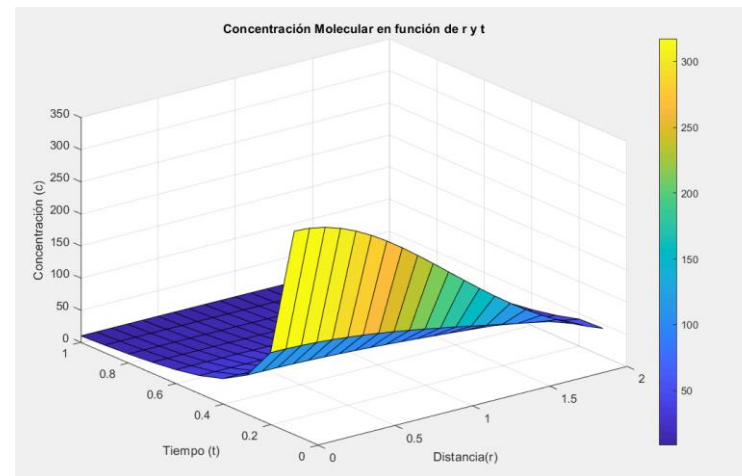
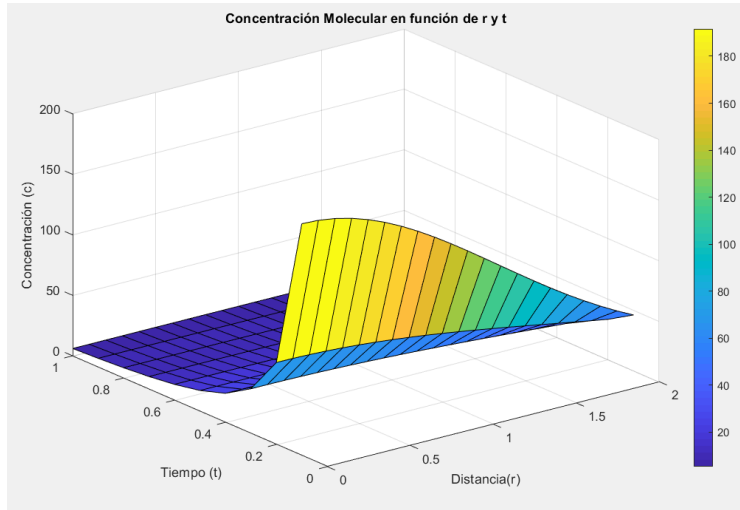


Figura 18. Concentración molecular en función de la distancia y el tiempo.

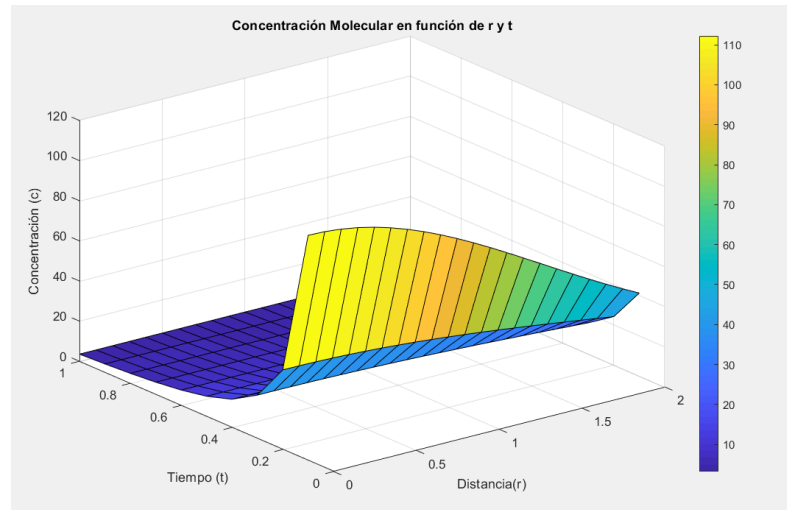
Fuente: Autores

En la Figura 19 se ilustran las concentraciones moleculares de los coeficientes más grandes que se usaron en la simulación (7,10,100 $\mu\text{m}^2/\text{s}$) que corresponden a la Figura 19 a), 19 b) y 19 c) respectivamente, se puede notar como las concentraciones son menores con relación a las que se obtuvieron en la Figura anterior. Como se describió en la Figura 18 el pulso de retardo no solo nos muestra el instante en donde la concentración alcanza su máximo valor, sino que también nos proporciona información sobre la velocidad de propagación de la señal molecular. Por lo que podemos decir que, entre más corto sea el pulso de retardo más rápida será la difusión y la propagación de la señal molecular será más eficiente.

a)



b)



c)

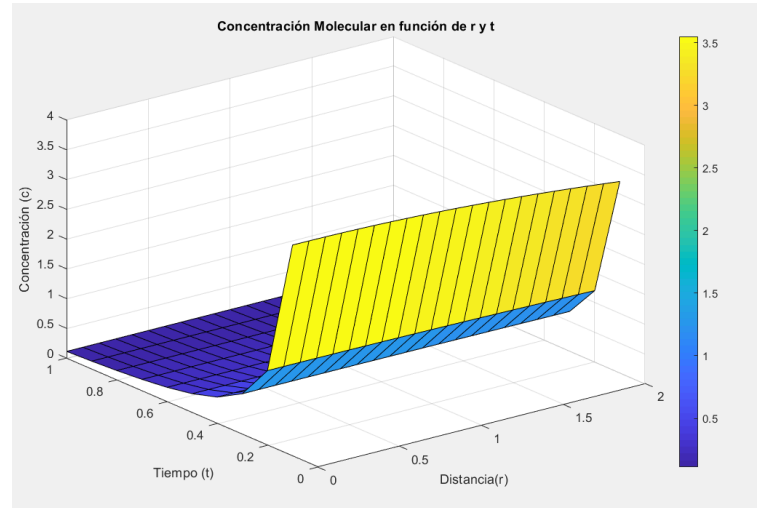


Figura 19. Concentración Molecular.
Fuente: Autore

En la Tabla 4 se muestran los valores obtenidos de la concertación molecular para los diferentes coeficientes de difusión, podemos observar como la concentración disminuye a medida que el coeficiente de difusión es más grande.

Los resultados nos muestran la distancia máxima a la que se puede ubicar el nodo de transmisor y receptor depende del coeficiente de difusión del medio, para transmitir a mayores distancias siempre se deberá establecer una relación entre la distancia y el coeficiente de fusión para el que receptor sea capaz de detectar un determinado umbral de concentración molecular.

Tabla 4. Valores de concentración obtenidos de los diferentes coeficientes de difusión.

Coefficiente de difusión ($\mu\text{m}^2/\text{s}$)	Concentración Molecular ($\text{mol}/\mu\text{m}^3$)
1	3549,4022
3	683,0828
5	317,4682
7	191,6497
10	112,242
100	3,5494

4.2 Tiempo de delay

Para las simulaciones del tiempo de delay se usó una misma distancia entre el transmisor y el receptor de $2.1 \mu\text{m}$ y 6 coeficientes diferentes los mismos que se especifican en la Tabla 3.

La Figura 20 esquematiza como el tiempo de delay es mayor con los coeficientes más pequeños y aumenta a mediad que se incrementa la distancia, por otro lado, con los coeficientes más grandes el delay baja a mediad que la distancia aumenta. Por lo que se puede decir que los coeficientes más grandes tienen una mejor respuesta en distancias mayores.

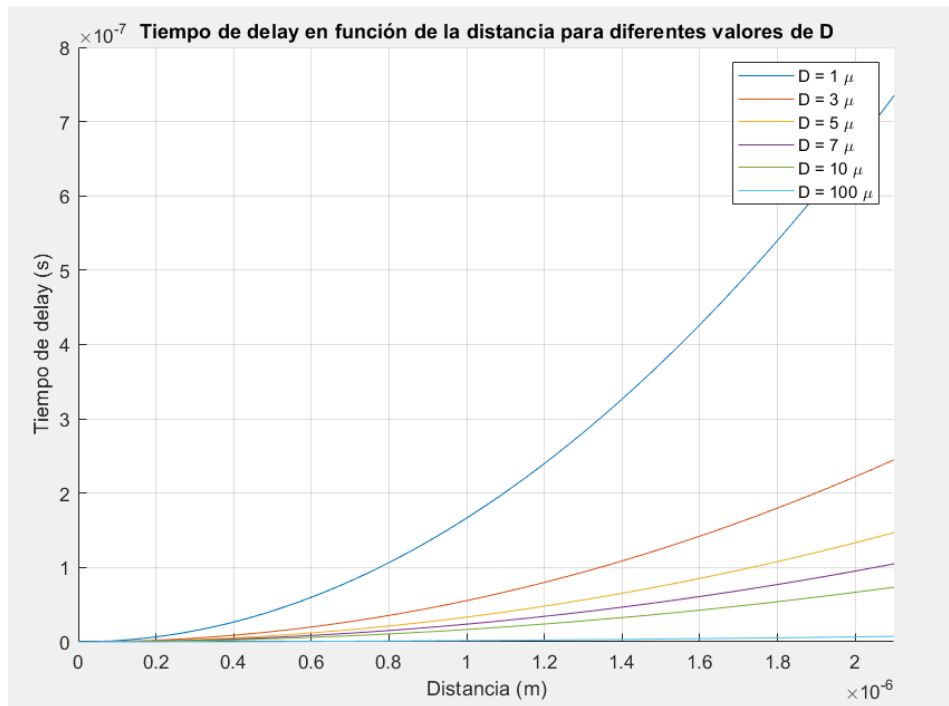


Figura 20. Tiempo de delay con variación de los coeficientes de difusión (D).
Fuente: Autores

En la tabla 5 se detallan los valores del tiempo de delay para los diferentes coeficientes, y en ella se observa como los valores aumentan con los coeficientes de difusión mayores.

Tabla 5. Valores del tiempo de delay obtenidos a los diferentes coeficientes de difusión.

Distancia (μm)	Coefficiente de difusión ($\mu\text{m}^2/\text{s}$)	Tiempo de delay (s)
2.1	1	7.3500e-07
2.1	3	2.4500e-07
2.1	5	1.4700e-07
2.1	7	1.0500e-07
2.1	10	7.3500e-08
2.1	100	7.3500e-09

4.3 Amplitud máxima de pulso

Para obtener la amplitud máxima de pulso se usó la ecuación (3.5) con 5000 moléculas que se liberaron desde el transmisor hacia el receptor.

En la Figura 21 se puede notar como la amplitud máxima de pulso disminuye a medida que aumenta la distancia, en su pico más alto es cuando podemos ver la mayor cantidad de moléculas y de igual manera van disminuyendo a mitad que crece la distancia.

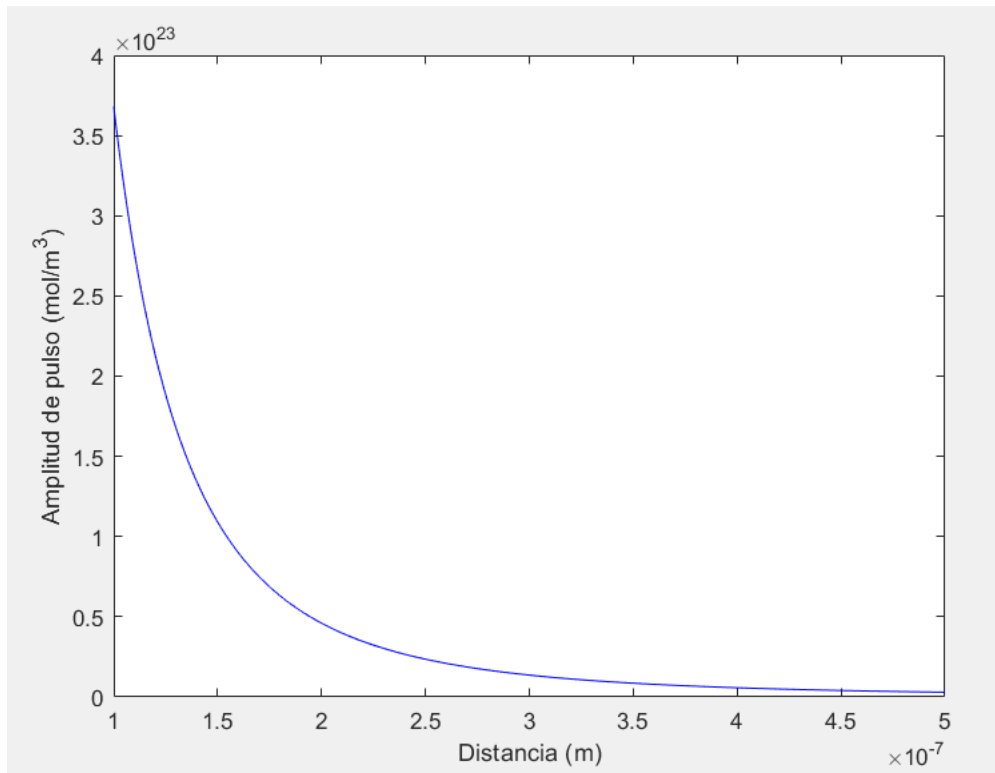


Figura 21. Amplitud máxima de pulso en función de la distancia y las moléculas liberadas.

Fuente: Autores

4.4 Ancho del pulso molecular

Para calcular los datos del ancho del pulso molecular se usó la ecuación (3.9) con los diferentes coeficientes de difusión y una distancia de $2.1\mu\text{m}$.

La Figura 22 esquematiza como el ancho de pulso molecular aumenta a mitad que crece la distancia debido a la duración del pulso de la señal de la molécula, y en comparación con los diferentes coeficientes de difusión este disminuye con los coeficientes más grandes.

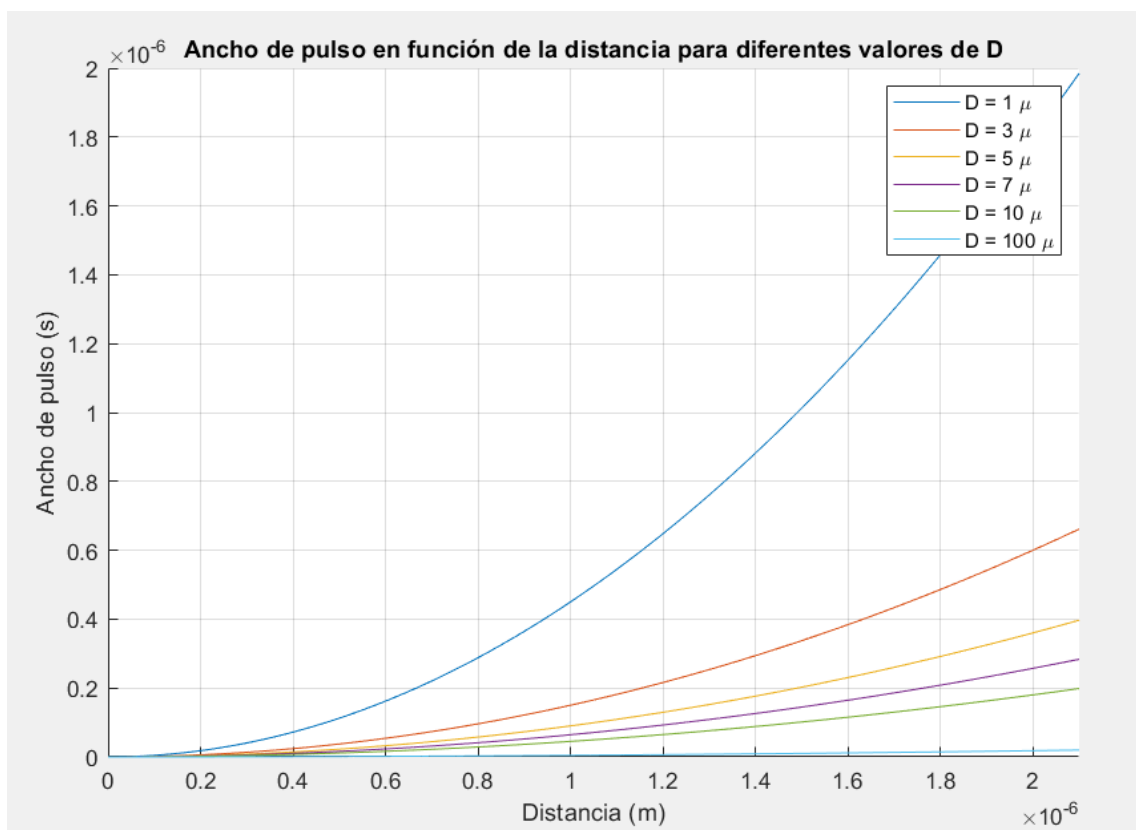


Figura 22. Ancho del pulso molecular en función de la distancia y los diferentes coeficientes de difusión.

Fuente: Autores

En la Tabla 6 se describen los valores obtenidos del ancho del pulso molecular para los diferentes coeficientes de difusión, en ella se observa cómo los datos entre el coeficiente menor disminuyen en relación con el mayor.

Tabla 6. Valores del ancho del pulso molecular en función de los coeficientes de difusión.

Distancia (um)	Coefficiente de difusión (um²/s)	Ancho del pulso molecular (s)
2.1	1	1.9849e-06
2.1	3	6.6165e-07
2.1	5	3.9699e-07
2.1	7	2.8356e-07
2.1	10	1.9849e-07
2.1	100	1.9849e-08

4.5 Análisis estadístico

Para evaluar la cantidad de moléculas que se ligan al receptor biológico en relación con el coeficiente de difusión que se empleó en las distintas simulaciones, se recolectan los datos obtenidos de las mismas que se realizaron por cada coeficiente de difusión en nuestro caso 6 (1, 3, 5, 7, 10, 100 $\mu\text{m}^2/\text{s}$).

4.5.1 Recolección de datos

Para la recolección de datos se realizaron 45 simulaciones en NS-3 con los parámetros descritos en la Tabla 3, se crea un archivo .sh en NS-3 para guardar los datos obtenidos de cada simulación estos archivos se describirán en el Anexo IV y luego elaborar una base de datos en el software estadístico IBM SPSS Statistics para su posterior análisis.

4.5.2 Prueba de normalidad

Es una prueba estadística empleada para determinar si un conjunto de datos sigue una distribución normal o gaussiana. La distribución normal es una distribución continua en la que la mayoría de los datos se concentran alrededor de la media, y la forma de la distribución es simétrica alrededor de la media [54].

- Planteamiento de hipótesis:

H_0 = Los datos siguen una distribución normal.

H_1 = Los datos no siguen una distribución normal.

- Estadísticos de prueba:

Shapiro-Wilk $n \leq 50$

Kolmogorov - $n \geq 50$

- Zona de rechazo

Si $p - \text{valor} > 0.05$ se acepta H_0 .

Si $p - \text{valor} < 0.05$ se acepta H_1 .

Debido a que tenemos 45 muestras realizamos la prueba de normalidad de Shapiro-Wilk los resultados se muestran en la Tabla 7.

Tabla 7. Prueba de normalidad Shapiro-Wilk.

	Coeficiente de difusión	Shapiro-Wilk		
		Estadístico	gl	Sig.
Moléculas Ligadas	1 $\mu\text{m}^2/\text{s}$,981	45	,675
	3 $\mu\text{m}^2/\text{s}$,980	45	,624
	5 $\mu\text{m}^2/\text{s}$,970	45	,280
	7 $\mu\text{m}^2/\text{s}$,958	45	,103
	10 $\mu\text{m}^2/\text{s}$,981	45	,655
	100 $\mu\text{m}^2/\text{s}$,984	45	,781

En la Tabla 7 se puede observar los diferentes datos obtenidos para cada coeficiente de difusión. El p-valor(sig) de cada caso es mayor a 0.05, por lo que, se acepta la hipótesis nula esto quiere decir que durante el proceso estadístico se utilizará una estadística paramétrica ya que, los resultados corresponden a una distribución normal.

4.5.3 Análisis de las moléculas ligadas al receptor

En base a los resultados de la prueba de normalidad se usará la prueba paramétrica ANOVA para comparar 3 o más muestras y determinar si existen diferencias significativas.

- Planteamiento de hipótesis

El promedio de cantidad de llegada de algún coeficiente de difusión es significativamente mayor al de otro coeficiente de difusión.

H_0 = Las medias de los 6 coeficientes son iguales

H_1 = Las medias son diferentes en al menos 2 coeficientes

Si $p - \text{valor} > 0.05$ se acepta H_0 .

Si $p - \text{valor} < 0.05$ se acepta H_1 .

Tabla 8. Análisis ANOVA.

ANOVA					
Moléculas Ligadas					
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	219289975,859	5	43857995,172	78643,511	<,001
Dentro de grupos	147227,796	264	557,681		
Total	219437203,655	269			

En la Tabla 8 podemos notar que el p-valor es menor que 0,05 por lo cual se rechaza la hipótesis nula, y se acepta la alterna. Con ello podemos demostrar que al menos 2 de las medias de los coeficientes de difusión tienen un comportamiento diferente entre ellos.

Tabla 9. Análisis Tukey.

Comparaciones múltiples						
Variable dependiente: Moléculas Ligadas						
HSD Tukey						
(I) Coeficiente de difusión	(J) Coeficiente de difusión	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
1 $\mu\text{m}^2/\text{s}$	3 $\mu\text{m}^2/\text{s}$	690,12222 [*]	4,97854	<,001	675,8300	704,4145
	5 $\mu\text{m}^2/\text{s}$	980,70000 [*]	4,97854	<,001	966,4078	994,9922
	7 $\mu\text{m}^2/\text{s}$	1301,06667 [*]	4,97854	<,001	1286,7744	1315,3589

	10 $\mu\text{m}^2/\text{s}$	1600,557 78*	4,97854	<,001	1586,2655	1614,8500
	100 $\mu\text{m}^2/\text{s}$	2923,060 00*	4,97854	<,001	2908,7678	2937,3522
3 $\mu\text{m}^2/\text{s}$	1 $\mu\text{m}^2/\text{s}$	- 690,1222 2*	4,97854	<,001	-704,4145	-675,8300
	5 $\mu\text{m}^2/\text{s}$	290,5777 8*	4,97854	<,001	276,2855	304,8700
	7 $\mu\text{m}^2/\text{s}$	610,9444 4*	4,97854	<,001	596,6522	625,2367
	10 $\mu\text{m}^2/\text{s}$	910,4355 6*	4,97854	<,001	896,1433	924,7278
	100 $\mu\text{m}^2/\text{s}$	2232,937 78*	4,97854	<,001	2218,6455	2247,2300
5 $\mu\text{m}^2/\text{s}$	1 $\mu\text{m}^2/\text{s}$	- 980,7000 0*	4,97854	<,001	-994,9922	-966,4078
	3 $\mu\text{m}^2/\text{s}$	- 290,5777 8*	4,97854	<,001	-304,8700	-276,2855
	7 $\mu\text{m}^2/\text{s}$	320,3666 7*	4,97854	<,001	306,0744	334,6589
	10 $\mu\text{m}^2/\text{s}$	619,8577 8*	4,97854	<,001	605,5655	634,1500
	100 $\mu\text{m}^2/\text{s}$	1942,360 00*	4,97854	<,001	1928,0678	1956,6522
7 $\mu\text{m}^2/\text{s}$	1 $\mu\text{m}^2/\text{s}$	- 1301,066 67*	4,97854	<,001	-1315,3589	-1286,7744

	3 $\mu\text{m}^2/\text{s}$	- 610,9444 4*	4,97854	<,001	-625,2367	-596,6522
	5 $\mu\text{m}^2/\text{s}$	- 320,3666 7*	4,97854	<,001	-334,6589	-306,0744
	10 $\mu\text{m}^2/\text{s}$	299,4911 1*	4,97854	<,001	285,1989	313,7834
	100 $\mu\text{m}^2/\text{s}$	1621,993 33*	4,97854	<,001	1607,7011	1636,2856
10 $\mu\text{m}^2/\text{s}$	1 $\mu\text{m}^2/\text{s}$	- 1600,557 78*	4,97854	<,001	-1614,8500	-1586,2655
	3 $\mu\text{m}^2/\text{s}$	- 910,4355 6*	4,97854	<,001	-924,7278	-896,1433
	5 $\mu\text{m}^2/\text{s}$	- 619,8577 8*	4,97854	<,001	-634,1500	-605,5655
	7 $\mu\text{m}^2/\text{s}$	- 299,4911 1*	4,97854	<,001	-313,7834	-285,1989
	100 $\mu\text{m}^2/\text{s}$	1322,502 22*	4,97854	<,001	1308,2100	1336,7945
100 $\mu\text{m}^2/\text{s}$	1 $\mu\text{m}^2/\text{s}$	- 2923,060 00*	4,97854	<,001	-2937,3522	-2908,7678
	3 $\mu\text{m}^2/\text{s}$	- 2232,937 78*	4,97854	<,001	-2247,2300	-2218,6455
	5 $\mu\text{m}^2/\text{s}$	- 1942,360 00*	4,97854	<,001	-1956,6522	-1928,0678

	7 $\mu\text{m}^2/\text{s}$	- 1621,993 33*	4,97854	<,001	-1636,2856	-1607,7011
	10 $\mu\text{m}^2/\text{s}$	- 1322,502 22*	4,97854	<,001	-1336,7945	-1308,2100
*. La diferencia de medias es significativa en el nivel 0.05.						

En Tabla 9 se puede apreciar como la variable tiene siempre un comportamiento diferente independientemente del coeficiente de difusión que se evalué. La cantidad de moléculas que se ligan al receptor son mayores con el coeficiente de difusión de 1 $\mu\text{m}^2/\text{s}$, es decir se tiene un rendimiento mayor de la variable y a medida que aumenta el coeficiente de difusión el rendimiento disminuye.

En la Figura 23 se puede observar que para el coeficiente de difusión 1 $\mu\text{m}^2/\text{s}$ se tiene la mayor cantidad de moléculas ligadas al receptor 3280,24 MOL y 357,18 MOL para la cantidad más baja que pertenecer al coeficiente de difusión 100 $\mu\text{m}^2/\text{s}$.

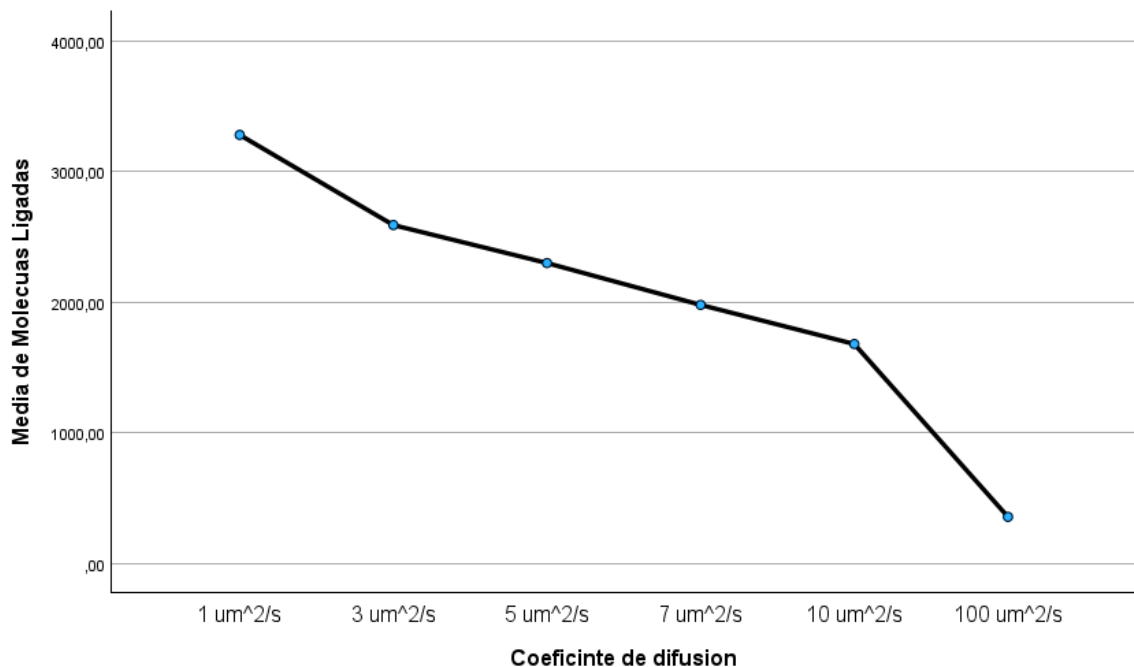


Figura 23. Media de la cantidad de moléculas que se ligan al receptor con respecto a los coeficientes de difusión.

Fuente: Autores

Luego de analizar los datos obtenidos podemos decir que, el promedio de llegada de moléculas para coeficientes de difusión menores tiene una mejor respuesta que los coeficientes de difusión mayores.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

- Se ha llevado a cabo una simulación de los estándares de la familia protocolaria IEEE 1906, que representan los lineamientos para las nanocomunicaciones aplicadas a las comunicaciones moleculares. Este proceso involucró la implementación de los componentes teóricos y de programación del marco de referencia de estos estándares mencionados específicamente en las comunicaciones moleculares.
- Se llevó a cabo la simulación del movimiento browniano con drift utilizando el módulo *Motion*, que es el encargado de modificar las propiedades de la portadora del mensaje. Implementamos la función *BrownianMotionDrift*, que reproduce un movimiento libre con arrastre, emulando el flujo sanguíneo. La velocidad de drift establecida en $1 \mu\text{m}/\text{s}$ nos permitió analizar el movimiento y la pérdida de moléculas. Observamos que con un coeficiente de difusión de $1 \mu\text{m}^2/\text{s}$ se maximiza la cantidad de moléculas dentro del radio de recepción del receptor.
- El módulo *perturbation* despliega un papel crucial al generar la portadora del mensaje. En este módulo, llevamos a cabo tanto la transcripción como la traducción del ADN, lo que nos permite crear los diferentes aminoácidos a partir de los nucleótidos Adenina, Guanina, Citosina y Uracilo. Este proceso culmina en la creación del mensaje de la portadora utilizando MoSK que emplea los 4 nucleótidos, liberándolo posteriormente en el medio. Este enfoque nos proporciona una plataforma eficaz para la creación y transmisión de mensajes en el contexto de la comunicación molecular.
- El módulo *communication-interface* se encarga de gestionar los procesos de recepción del mensaje dentro del marco de IEEE p1906.1 y p1906.1.1. Este módulo nos permitió implementar un receptor biológico del tipo ligando-receptor, utilizando un proceso estocástico basado en una cadena de Markov. Esta implementación nos permitió evaluar el estado del receptor, ya sea libre o unido a una molécula. Durante este proceso, observamos la pérdida de moléculas, ya que el receptor requiere un tiempo para liberarse y estar disponible para recibir otra molécula después de unirse a una.
- Las simulaciones realizadas revelaron que los mejores resultados se obtuvieron con un coeficiente de difusión de $1 \mu\text{m}^2/\text{s}$. Este coeficiente permitió mitigar significativamente las pérdidas de moléculas en comparación con el coeficiente de $100 \mu\text{m}^2/\text{s}$, que mostró una

mayor pérdida en el canal de transmisión. Este último es el punto crítico donde se registra la mayoría de las pérdidas durante la transmisión.

5.2 Recomendaciones

- Se recomienda utilizar los paquetes oficiales tanto del software NS-3 como del framework proporcionado por el grupo de investigación IEEE 1906 para evitar posibles incompatibilidades. Además, es importante revisar detenidamente la documentación proporcionada por los desarrolladores del software NS-3 para comprender los posibles errores que podrían surgir al ejecutar este simulador.
- Debido a los parámetros que se definieron en la simulación es recomendable no excederse en el número de moléculas que se vayan a enviar o en dicho caso que se requiera usar más moléculas, se recomienda usar un súper computador.
- Para trabajos futuros se podría centrar en el estudio de la comunicación molecular basado en motores moleculares como lo menciona en el libro de Tadashi Nakano “Molecular Communication” como medio de análisis de propagación de las moléculas, debido a que este tipo de estructuras afectan en la expresión genética.

BIBLIOGRAFÍA

- [1] C. E. Shannon, “A mathematical theory of communication,” *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.
- [2] Y. Cevallos, L. Molina, A. Santillán, F. De Rango, A. Rushdi, and J. B. Alonso, “A digital communication analysis of gene expression of proteins in biological systems: A layered network model view,” *Cognit Comput*, vol. 9, no. 1, pp. 43–67, 2017.
- [3] W. P. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, “How cells read the genome: from DNA to protein.,” *New York: Garland Science*, 2014.
- [4] T. Nakano, T. Suda, Y. Okaie, M. J. Moore, and A. V Vasilakos, “Molecular communication among biological nanomachines: A layered architecture and research issues,” *IEEE Trans Nanobioscience*, vol. 13, no. 3, pp. 169–197, 2014.
- [5] Y. Cevallos *et al.*, “Modeling Gene Expression and Protein Delivery as an End-to-End Digital Communication System,” *Open Bioinforma J*, vol. 14, no. 1, 2021.
- [6] Y. Chahibi, “Molecular communication for drug delivery systems: A survey,” *Nano Commun Netw*, vol. 11, pp. 90–102, 2017.
- [7] S. R. Goodman, *Medical cell biology*. Academic Press, 2007.
- [8] S. F. Bush, *Nanoscale communication networks*. Artech House, 2010.
- [9] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, “A comprehensive survey of recent advancements in molecular communication,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1887–1919, 2016.
- [10] A. Gohari, M. Mirmohseni, and M. Nasiri-Kenari, “Information theory of molecular communication: Directions and challenges,” *IEEE Trans Mol Biol Multiscale Commun*, vol. 2, no. 2, pp. 120–142, 2016.
- [11] J. Wang, M. Peng, Y. Liu, X. Liu, and M. Daneshmand, “Performance analysis of signal detection for amplify-and-forward relay in diffusion-based molecular communication systems,” *IEEE Internet Things J*, vol. 7, no. 2, pp. 1401–1412, 2019.
- [12] M. Gastpar, B. Rimoldi, and M. Vetterli, “To code, or not to code: Lossy source-channel communication revisited,” *IEEE Trans Inf Theory*, vol. 49, no. 5, pp. 1147–1158, 2003.
- [13] T. Nakano, T. Suda, T. Koujin, T. Haraguchi, and Y. Hiraoka, “Molecular communication through gap junction channels,” in *Transactions on Computational Systems Biology X*, Springer, 2008, pp. 81–99.

- [14] J. Suzuki, T. Nakano, and M. J. Moore, *Modeling, methodologies and tools for molecular and nano-scale communications: modeling, methodologies and tools*, vol. 9. Springer, 2017.
- [15] A. Z. Shirazi, S. M. Mazinani, and S. K. Eghbal, "Protocol stack for nano networks," in *2012 International Symposium on Computer, Consumer and Control*, IEEE, 2012, pp. 849–853.
- [16] T. Szabados, "An elementary introduction to the Wiener process and stochastic integrals," *arXiv preprint arXiv:1008.1510*, 2010.
- [17] T. Nakano, *Molecular communication*. Cambridge University Press, 2013.
- [18] J. CRANK, "The mathematic of diffusion," *2nd. Edition*, pp. 129–135, 1975.
- [19] Y. Cevallos *et al.*, "Modelamiento comunicacional de la expresión genética y el transporte de proteínas mediante un sistema de transmisión digital extremo a extremo," 2022.
- [20] Y. Cevallos, L. Tello-Oquendo, D. Inca, D. Ghose, A. Z. Shirazi, and G. A. Gomez, "Health applications based on molecular communications: a brief review," in *2019 IEEE International Conference on E-health Networking, Application & Services (HealthCom)*, IEEE, 2019, pp. 1–6.
- [21] Y. Cevallos *et al.*, "On the efficient digital code representation in DNA-based data storage," in *Proceedings of the 7th ACM International Conference on Nanoscale Computing and Communication*, 2020, pp. 1–7.
- [22] Y. Cevallos, L. Tello-Oquendo, D. Inca, C. Palacios, and L. Rentería, "Genetic expression in biological systems: A digital communication perspective," *Open Bioinforma J*, vol. 12, no. 1, 2019.
- [23] D. Haughton and F. Balado, "BioCode: Two biologically compatible Algorithms for embedding data in non-coding and coding regions of DNA," *BMC Bioinformatics*, vol. 14, no. 1, pp. 1–16, 2013.
- [24] L. C. B. Faria, A. S. L. Rocha, and R. Palazzo Jr, "Transmission of intra-cellular genetic information: A system proposal," *J Theor Biol*, vol. 358, pp. 208–231, 2014.
- [25] P. Raut and N. Sarwade, "Study of environmental effects on the connectivity of molecular communication based Internet of Nano things," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, IEEE, 2016, pp. 1123–1128.

- [26] Q. A. Acton, *Cancer Gene Therapy: New Insights for the Healthcare Professional: 2013 Edition: ScholarlyBrief*. ScholarlyEditions, 2013.
- [27] Q. H. Abbasi *et al.*, “Nano-communication for biomedical applications: A review on the state-of-the-art from physical layers to novel networking concepts,” *IEEE Access*, vol. 4, pp. 3920–3935, 2016.
- [28] S. Shrivastava and R. Badlani, “Data storage in DNA,” *International Journal of Electrical Energy*, vol. 2, no. 2, pp. 119–124, 2014.
- [29] W.-L. Chang and A. V Vasilakos, “DNA algorithms of implementing biomolecular databases on a biological computer,” *IEEE Trans Nanobioscience*, vol. 14, no. 1, pp. 104–111, 2014.
- [30] B. Alberts, *Molecular biology of the cell*. Garland science, 2017.
- [31] L. P. G. Păun and M. J. P.-J. T. Song, “Bio-Inspired Computing–Theories and Applications”.
- [32] B. Alberts *et al.*, *Essential cell biology*. Garland Science, 2015.
- [33] C. Starr and B. McMillan, *Human biology*. Cengage Learning, 2013.
- [34] K. Sneppen, *Models of life*. Cambridge University Press, 2014.
- [35] L. B. Jorde, J. C. Carey, and M. J. Bamshad, *Medical genetics e-Book*. Elsevier Health Sciences, 2015.
- [36] R. Nussbaum, R. R. McInnes, and H. F. Willard, *Thompson & Thompson genetics in medicine e-book*. Elsevier Health Sciences, 2015.
- [37] T. Nakano, Y. Okaie, and A. V Vasilakos, “Transmission rate control for molecular communication among biological nanomachines,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 835–846, 2013.
- [38] P. Y. De Silva and G. U. Ganegoda, “New trends of digital data storage in DNA,” *Biomed Res Int*, vol. 2016, 2016.
- [39] A. S. Tanenbaum, *Computer networks*. Pearson Education India, 2021.
- [40] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, “Layering as optimization decomposition: A mathematical theory of network architectures,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, 2007.
- [41] H. Lodish, A. Berk, C. A. Kaiser, M. Krieger, M. P. Scott, and P. Matsudaira, *Molecular Cell Biology*. Macmillan, 2008.
- [42] T. Uhlig, F. Sellmaier, and M. Schmidhuber, *Spacecraft operations*. Springer, 2015.

- [43] A. W. Eckford and P. J. Thomas, "Information theory of intercellular signal transduction," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, IEEE, 2015, pp. 119–122.
- [44] C.-H. J. Wu and J. D. Irwin, *Introduction to computer networks and cybersecurity*. CRC Press, 2016.
- [45] E. Coll, *Telecom 101 Telecommunications Reference Book: 2016 Fourth Edition. High-Quality Reference Book and Study Guide Covering All Major Telecommunications Topics... in Plain English*. 2016. [Online]. Available: <https://books.google.com.ec/books?id=Lj72CwAAQBAJ>
- [46] M. Zambrano, "Comunicación Molecular: Retos y oportunidades," *Prisma Tecnológico*, vol. 6, no. 1, pp. 13–18, 2015.
- [47] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," in *2011 IEEE international conference on communications (ICC)*, IEEE, 2011, pp. 1–5.
- [48] S. Pudasaini, S. Shin, and K. S. Kwak, "Robust modulation technique for diffusion-based molecular communication in nanonetworks," *arXiv preprint arXiv:1401.3938*, 2014.
- [49] B. Tepekule, A. E. Pusane, H. B. Yilmaz, and T. Tugcu, "A novel modulation technique in diffusion based molecular communication and its performance analysis," in *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2014, pp. 1110–1113.
- [50] N.-R. Kim and C.-B. Chae, "Novel modulation techniques using isomers as messenger molecules for nano communication networks via diffusion," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 847–856, 2013.
- [51] D. Kilinc and O. B. Akan, "Receiver design for molecular communication," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, pp. 705–714, 2013.
- [52] "IEEE Recommended Practice for Nanoscale and Molecular Communication Framework," *IEEE Std 1906.1-2015*, pp. 1–64, 2016.
- [53] I. Llatser, E. Alarcón, and M. Pierobony, "Diffusion-based channel characterization in molecular nanonetworks," in *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2011, pp. 467–472.
- [54] C. A. Sánchez, "Las pruebas de normalidad," *Instituto Nacional de Metrología (INM)*, Junio, pp. 1–8, 2023.

ANEXOS

ANEXO I. Descarga e instalación de NS-3

ANEXO II. Descarga e instalación del módulo P1906

ANEXO III. Ejecutar el simulador

ANEXO IV. Creación de un archivo .sh para la obtención de resultados

ANEXO V. Análisis estadístico de los resultados

ANEXO I

Descarga e instalación de NS-3

Para la instalación de NS-3 se requiere el sistema operativo de Ubuntu, en nuestro caso usamos la versión 16.04.7, al ser de código abierto tiene una compatibilidad con Linux para poder ejecutar todas las herramientas de desarrollo que nos ofrece NS-3. El software NS-3 se lo puede obtener directamente de su página oficial nsnam.org en el siguiente enlace (<https://www.nsnam.org/releases/ns-3-29/download/>)

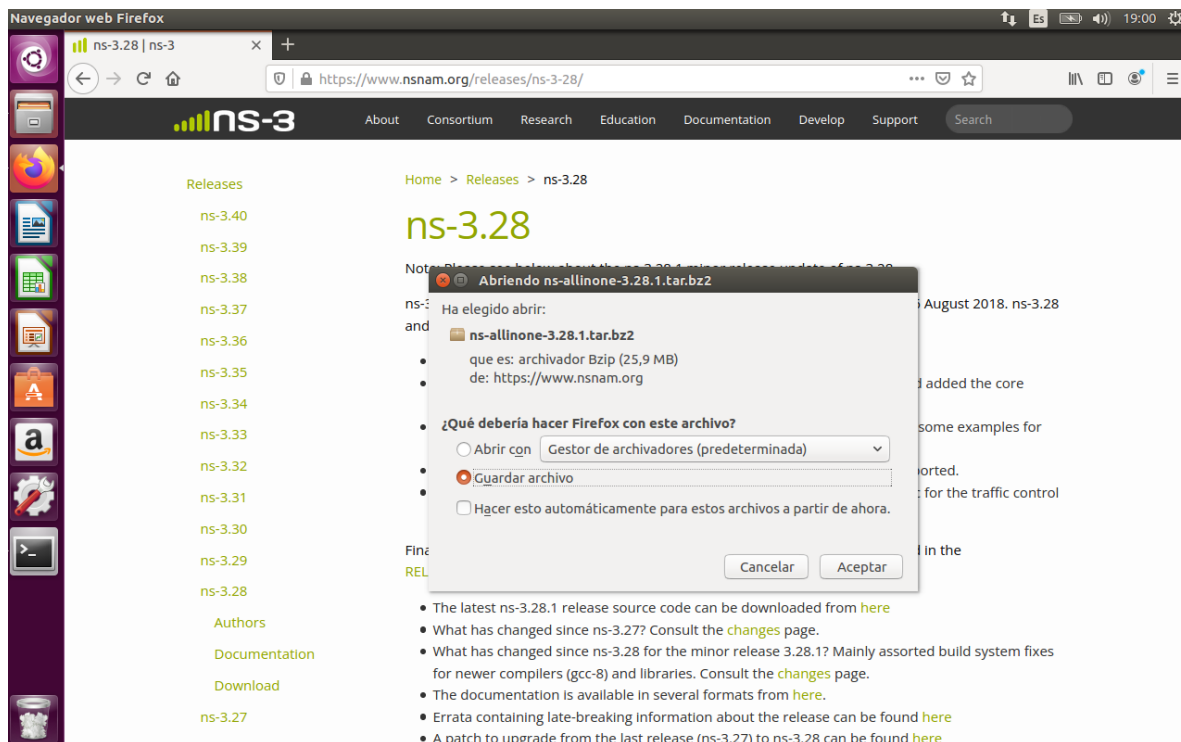


Figura 24. Descarga del simulador NS-3 (3.28.1.tar.bz2).

Fuente: Autores

Una vez descargado el simulador procedemos a descomprimirlo desde el terminal con el siguiente comando `tar -xf ns-allinone-3.28.1.tar.bz2`

```
yeimy@yeimy-VirtualBox:~/Escritorio$ tar -xf ns-allinone-3.28.1.tar.bz2
yeimy@yeimy-VirtualBox:~/Escritorio$
```

Figura 25. Proceso para descomprimir un archivo tar.

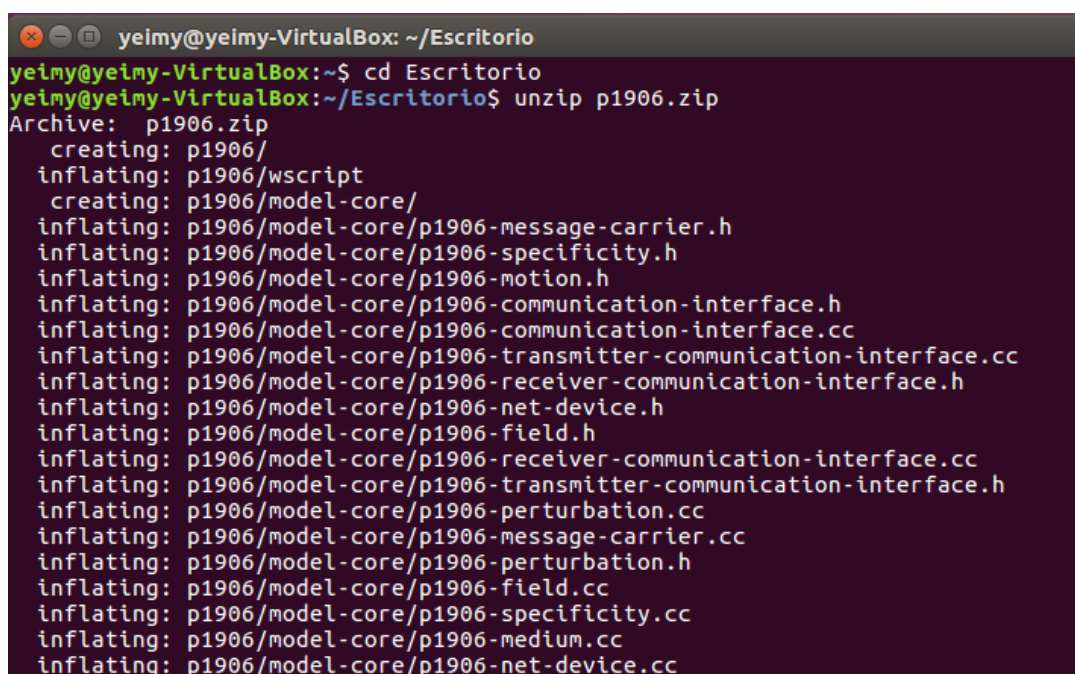
Fuente: Autores

Una vez tengamos descomprimido el archivo está listo para usarlo en lo que se requiera.

ANEXO II

Descarga e instalación del módulo P1906

El módulo P1906 se puede obtener directamente del siguiente enlace (<https://github.com/ieee-p1906-1-reference-code/p1906-dev>) contiene varios archivos entre ellos el README.md donde nos indica como instalar y ejecutar el simulador, el archivo p1906 es el archivo que debemos descargar y descomprimir para su posterior instalación. Este archivo consta de las 9 clases y 21 archivos configurados para la simulación en NS-3. Una vez descargado el archivo necesario procedemos a descomprimir desde el terminal con el comando `unzip p1906.zip`



```
yeimy@yeimy-VirtualBox: ~/Escritorio
yeimy@yeimy-VirtualBox:~$ cd Escritorio
yeimy@yeimy-VirtualBox:~/Escritorio$ unzip p1906.zip
Archive:  p1906.zip
  creating: p1906/
  inflating: p1906/wscript
   creating: p1906/model-core/
  inflating: p1906/model-core/p1906-message-carrier.h
  inflating: p1906/model-core/p1906-specificity.h
  inflating: p1906/model-core/p1906-motion.h
  inflating: p1906/model-core/p1906-communication-interface.h
  inflating: p1906/model-core/p1906-communication-interface.cc
  inflating: p1906/model-core/p1906-transmitter-communication-interface.cc
  inflating: p1906/model-core/p1906-receiver-communication-interface.h
  inflating: p1906/model-core/p1906-net-device.h
  inflating: p1906/model-core/p1906-field.h
  inflating: p1906/model-core/p1906-receiver-communication-interface.cc
  inflating: p1906/model-core/p1906-transmitter-communication-interface.h
  inflating: p1906/model-core/p1906-perturbation.cc
  inflating: p1906/model-core/p1906-message-carrier.cc
  inflating: p1906/model-core/p1906-perturbation.h
  inflating: p1906/model-core/p1906-field.cc
  inflating: p1906/model-core/p1906-specificity.cc
  inflating: p1906/model-core/p1906-medium.cc
  inflating: p1906/model-core/p1906-net-device.cc
```

Figura 26. Descomprimir el archivo p1906.

Fuente: Autores

Previamente ya instalamos el simulador NS-3 nos dirigimos a la carpeta ns-3.28.1 desde el terminal con el siguiente comando `cd ns-allinone-3.28.1/ns3.28.1/` una vez dentro de la carpeta copiamos el archivo p1906 a la carpeta src, luego procedemos a iniciar el simulador con el comando `./waf configure`.

```
yeimy@yeimy-VirtualBox:~/Escritorio/ns-allinone-3.28.1/ns-3.28.1$ ./waf configure
Setting top to : /home/yeimy/Escritorio/ns-allinone-3.28.1/ns-3.28.1
Setting out to : /home/yeimy/Escritorio/ns-allinone-3.28.1/ns-3.28.1/build
Checking for 'gcc' (C compiler) : /usr/bin/gcc
Checking for cc version : 5.4.0
Checking for 'g++' (C++ compiler) : /usr/bin/g++
Checking for compilation flag -Wl,--soname=foo support : ok
```

*Figura 27. Inicializar el software NS-3.
Fuente: Autores*

Dentro del terminal para instalar el archivo p1906 ejecutamos el comando `./waf` al realizar esto estamos listo para iniciar con las simulaciones que se vayan a ejecutar

```
yeimy@yeimy-VirtualBox:~/Escritorio/ns-allinone-3.28.1/ns-3.28.1$ ./waf
Waf: Entering directory `/home/yeimy/Escritorio/ns-allinone-3.28.1/ns-3.28.1/build'
[ 1/1903] Compiling install-ns3-header: ns3/antenna-model.h
[ 2/1903] Compiling install-ns3-header: ns3/isotropic-antenna-model.h
[ 3/1903] Compiling install-ns3-header: ns3/angles.h
[ 4/1903] Compiling install-ns3-header: ns3/parabolic-antenna-model.h
[ 5/1903] Compiling install-ns3-header: ns3/cosine-antenna-model.h
[ 6/1903] Compiling install-ns3-header: ns3/aodv-packet.h
[ 7/1903] Compiling install-ns3-header: ns3/aodv-neighbor.h
[ 8/1903] Compiling install-ns3-header: ns3/aodv-rqueue.h
[ 9/1903] Compiling install-ns3-header: ns3/aodv-routing-protocol.h
[10/1903] Compiling install-ns3-header: ns3/aodv-rtable.h
[11/1903] Compiling install-ns3-header: ns3/aodv-id-cache.h
[12/1903] Compiling install-ns3-header: ns3/aodv-dpd.h
[13/1903] Compiling install-ns3-header: ns3/aodv-helper.h
[14/1903] Compiling install-ns3-header: ns3/udp-client.h
[15/1903] Compiling install-ns3-header: ns3/bulk-send-application.h
[16/1903] Compiling install-ns3-header: ns3/on-off-helper.h
[17/1903] Compiling install-ns3-header: ns3/udp-client-server-helper.h
[18/1903] Compiling install-ns3-header: ns3/udp-trace-client.h
[19/1903] Compiling install-ns3-header: ns3/udp-echo-server.h
[20/1903] Compiling install-ns3-header: ns3/packet-sink.h
```

*Figura 28. Instalación de los archivos p1906.
Fuente: Autores*

Para ver si se instaló de manera correcta el archivo ingresamos a la carpeta `src` y enlistamos los archivos y ahí deberíamos encontrar el archivo `p1906`.

```

yeimy@yeimy-VirtualBox: ~/Escritorio/ns-allinone-3.28.1/ns-3.28.1
Build commands will be stored in build/compile_commands.json
'build' finished successfully (32m23.160s)

Modules built:
antenna          aodv             applications
bridge          buildings       config-store
core            csma            csma-layout
dsdv           dsr             energy
fd-net-device   flow-monitor    internet
internet-apps  lr-wpan         lte
mesh           mobility        mpi
netanim (no Python) network         nix-vector-routing
olsr           p1906          point-to-point
point-to-point-layout propagation     sixlowpan
spectrum       stats          tap-bridge
test (no Python) topology-read traffic-control
uan           virtual-net-device wave
wifi          wimax

Modules not built (see ns-3 tutorial for explanation):
brite          click           openflow
visualizer

yeimy@yeimy-VirtualBox:~/Escritorio/ns-allinone-3.28.1/ns-3.28.1$

```

Figura 29. Instalación del archivo p1906.

Fuente: Autores

Para ejecutar el ejemplo del simulador copiamos el archivo “mol-example.cc” que se encuentra en la carpeta examples del p1906 y lo copiamos a la carpeta scratch desde el terminal con el siguiente comando una vez realizamos esa acción estamos listos para ejecutar las simulaciones.

```

yeimy@yeimy-VirtualBox:~/Escritorio/ns-allinone-3.28.1/ns-3.28.1$ cp examples/mol-example.cc /Escritorio/ns-allinone-3.28.1/ns-3.28.1/scratch/

```

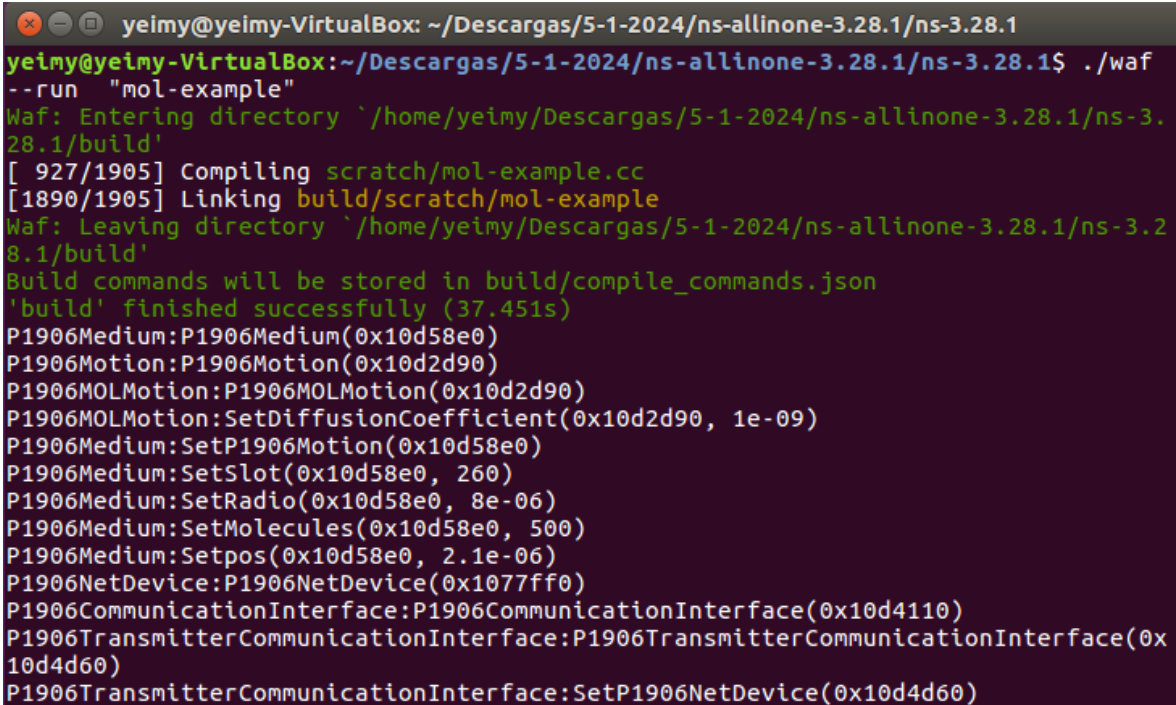
Figura 30 Instrucciones para Copiar el Archivo mol-example.cc al Directorio Scratch.

Fuente: Autores

ANEXO III

Ejecutar el simulador

Cuando ejecutamos el simulador con el comando `./waf -run "mol-example.cc"`, el programa buscará dentro del directorio `/scratch` el archivo de configuración del escenario a simular, en este caso, `"mol-example"`. Al iniciar el simulador, se llevará a cabo la inicialización de todos los componentes del protocolo IEEE 1906, como se detalla en la Figura 27



```
yeimy@yeimy-VirtualBox: ~/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1
yeimy@yeimy-VirtualBox:~/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1$ ./waf
--run "mol-example"
Waf: Entering directory `/home/yeimy/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1/build'
[ 927/1905] Compiling scratch/mol-example.cc
[1890/1905] Linking build/scratch/mol-example
Waf: Leaving directory `/home/yeimy/Descargas/5-1-2024/ns-allinone-3.28.1/ns-3.28.1/build'
Build commands will be stored in build/compile_commands.json
'build' finished successfully (37.451s)
P1906Medium:P1906Medium(0x10d58e0)
P1906Motion:P1906Motion(0x10d2d90)
P1906MOLMotion:P1906MOLMotion(0x10d2d90)
P1906MOLMotion:SetDiffusionCoefficient(0x10d2d90, 1e-09)
P1906Medium:SetP1906Motion(0x10d58e0)
P1906Medium:SetSlot(0x10d58e0, 260)
P1906Medium:SetRadio(0x10d58e0, 8e-06)
P1906Medium:SetMolecules(0x10d58e0, 500)
P1906Medium:Setpos(0x10d58e0, 2.1e-06)
P1906NetDevice:P1906NetDevice(0x1077ff0)
P1906CommunicationInterface:P1906CommunicationInterface(0x10d4110)
P1906TransmitterCommunicationInterface:P1906TransmitterCommunicationInterface(0x10d4d60)
P1906TransmitterCommunicationInterface:SetP1906NetDevice(0x10d4d60)
```

Figura 31. Ejecución del escenario.

Fuente: Autores

ANEXO IV

Creación de un archivo .sh para la obtención de resultados

A través del archivo .sh (Figura 30) que implementamos podemos obtener los resultados de la simulación en el software NS-3 para posteriormente realizar los diferentes análisis estadísticos.

En la Figura 29 podemos observar un ciclo for que nos permitirá realizar un número finito de repeticiones para de esta manera obtener una mayor cantidad de resultados. Para definir este número de simulaciones se deberá modificar la secuencia del for, en este ejemplo se han simulado 15 simulaciones.

Los resultados de cada simulación, en este caso el número de moléculas que han sido ligadas en el receptor, se albergan en un archivo RES_LIG el cual luego se puede pasar al software IBM SPSS Statistics para su análisis.

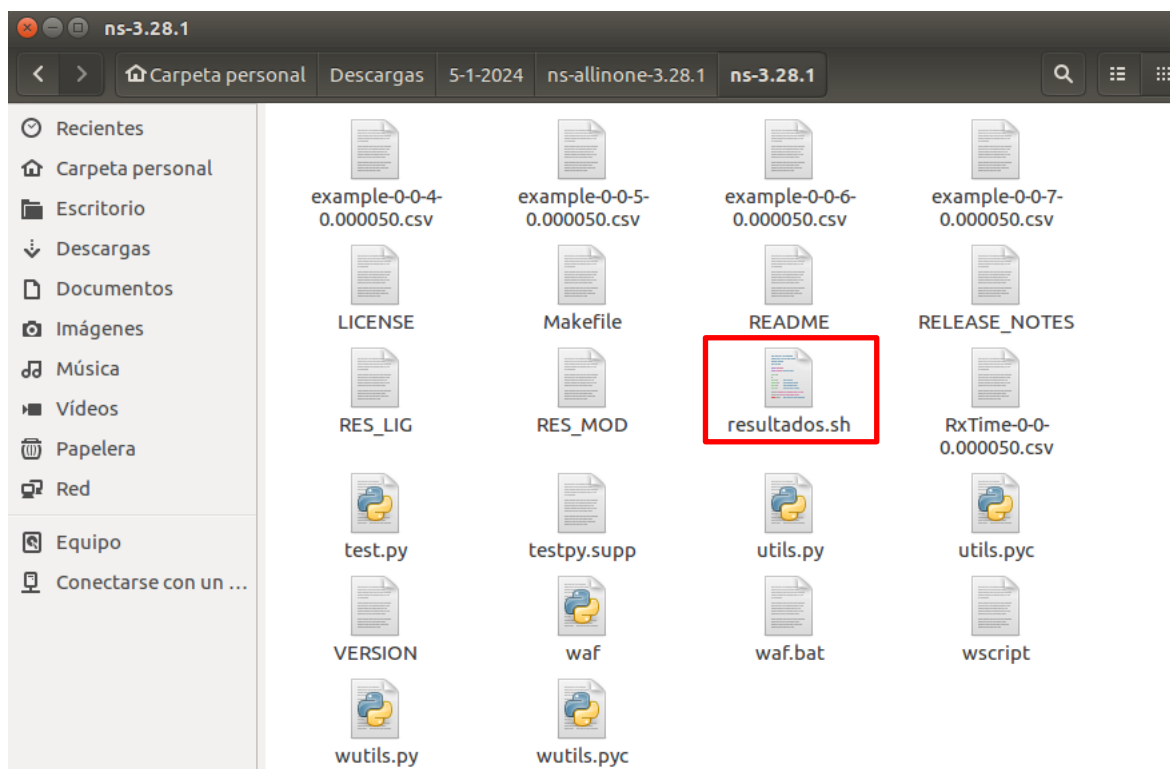


Figura 32. Creación del archivo .sh para la obtención de resultados.

Fuente: Autores

```
#!/bin/bash
touch RES_LIG RES_MOD
rm RES_*

for i in $(seq 1 15)
do
    ./waf --run "mol-example" &> tmp_out
    grep ligadas tmp_out | awk '{print $4}' > tmp_lig
    grep moduladas tmp_out > tmp_mod
    sed 's/,/ /g' tmp_lig > tmp_lig2
    mv tmp_lig2 tmp_lig
    sed 's/)/ /g' tmp_lig > tmp_lig2
    mv tmp_lig2 tmp_lig
    sed 's/\././g' tmp_lig > tmp_lig2
    mv tmp_lig2 tmp_lig
    cat tmp_lig >> RES_LIG

    sed 's/,/ /g' tmp_mod | awk '{print $5}' > tmp_mod2
    mv tmp_mod2 tmp_mod
    sed 's/)/ /g' tmp_mod > tmp_mod2
    mv tmp_mod2 tmp_mod
    sed 's/\././g' tmp_mod > tmp_mod2
    mv tmp_mod2 tmp_mod
    cat tmp_mod >> RES_MOD

    rm tmp_*
done
```

*Figura 33. Código del archivo .sh.
Fuente: Autores*

ANEXO V

Análisis estadístico de los resultados

Una vez extraídos los datos del archivo RES_LIG, detallado en el anexo 4, procedemos a transferirlos A SPSS, tal como se ilustra en la Figura 32.

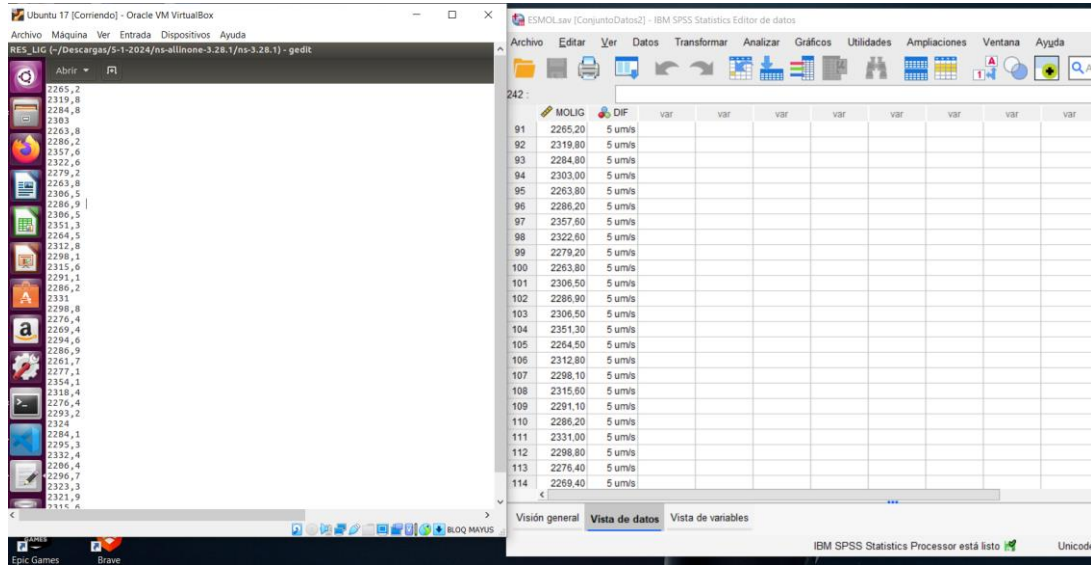


Figura 34. Extracción de datos.

Fuente: Autores

Para examinar las diferencias entre los diversos conjuntos de resultados obtenidos, recurrimos a un análisis ANOVA. Accedemos a esta función siguiendo las indicaciones proporcionadas en la Figura 33.

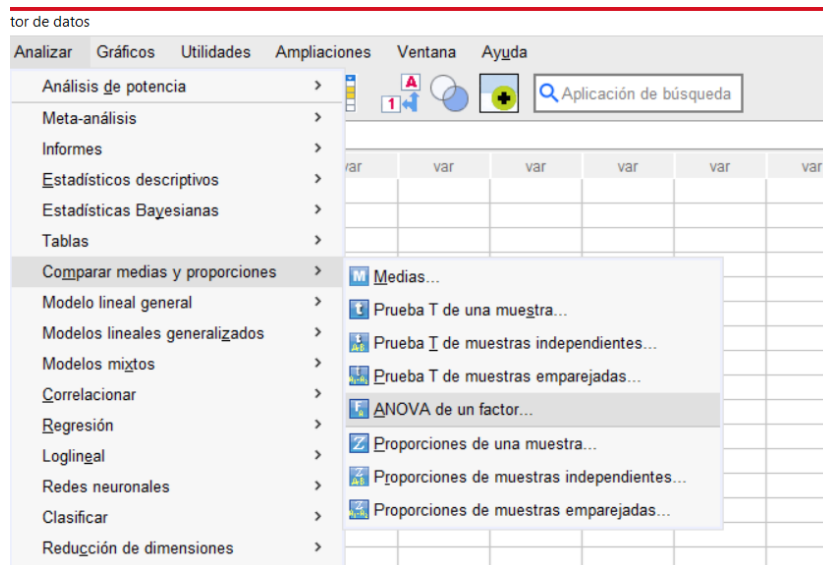


Figura 35. Análisis ANOVA SPS.

Fuente: Autores

Para configurar la ANOVA de un factor lo podemos hacer a través del cuadro de dialogo que se muestra en la Figura 34

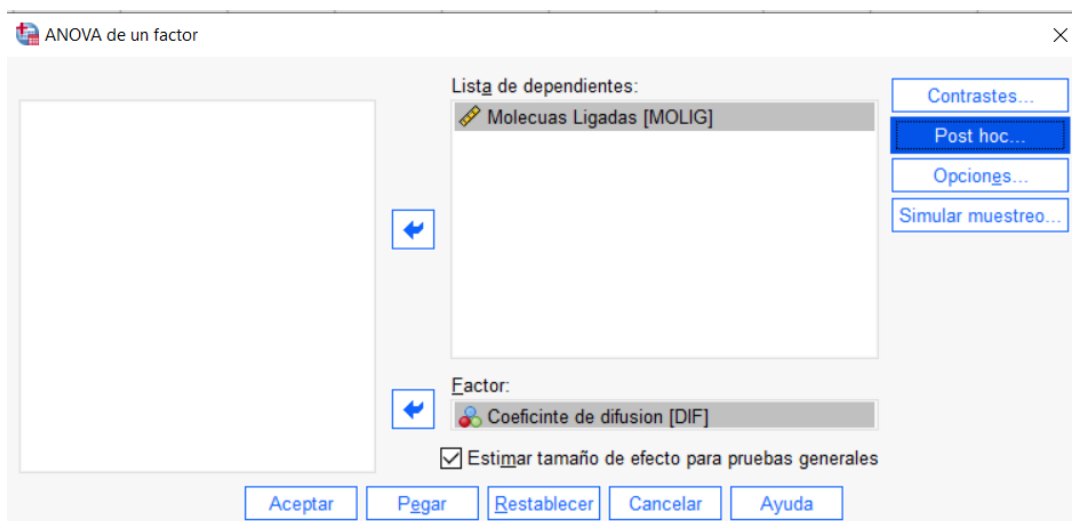


Figura 36. Cuadro de dialogo del ANOVA de un factor.
Fuente: Autores

Para obtener un análisis más riguroso, aplicamos un proceso de comparaciones múltiples tipo Tukey. Esta configuración se ajusta en la pestaña Post hoc y se selecciona Tukey, como se muestra en la Figura 35.

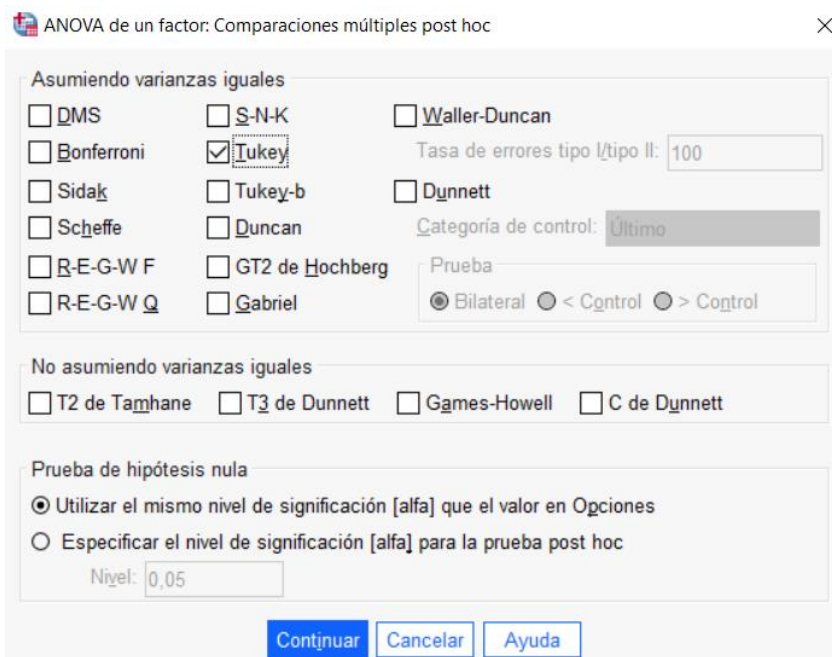


Figura 37. Cuadro de dialogo comparaciones post hoc.
Fuente: Autores

Además, para visualizar la tendencia de las medias y analizar su comportamiento, empleamos un gráfico. Activar esta opción se realiza en opciones del cuadro de dialogo de

la Figura 34 y se nos mostrara el cuadro de dialogo de la Figura 36 y activamos la opción grafico de las medias.

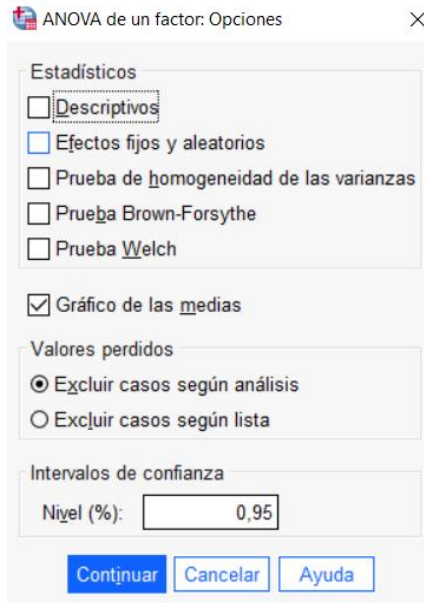


Figura 38. Cuadro de dialogo opciones ANOVA de un factor.
Fuente: Autores