PUCE Virtual

CLASE

3

Teoría de la Información y la Comunicación

Fundamentos de la Teoría de la Información y Modelos de Comunicación

QS WORLD UNIVERSITY RANKINGS

Educación de calidad

INTRODUCCIÓN DE LA CLASE

GRADO / POSGRADO / TECNOLOGÍAS

Estudia a tu tiempo
son interrupciones

EDUCACIÓN EN LÍNEA DE CALIDAD

La gestión segura de la información en los entornos organizacionales modernos requiere no solo comprender cómo se transmiten los datos, sino también medir cuánta información contienen y cuánta incertidumbre generan. En este contexto, la Teoría de la Información proporciona herramientas matemáticas que permiten analizar los flujos de datos desde una perspectiva cuantitativa, superando el enfoque meramente descriptivo de los sistemas de comunicación.

Durante esta semana se introducen los conceptos de **medida de la información** y **entropía**, los cuales permiten caracterizar el comportamiento estadístico de las fuentes de datos que circulan por redes corporativas. Estos conceptos resultan esenciales para evaluar la previsibilidad de los eventos digitales, como accesos al sistema, tráfico de red o generación de registros de auditoría. La entropía, propuesta inicialmente por Shannon, se convierte en un indicador clave para analizar la variabilidad y el desorden presentes en los sistemas de comunicación. En el ámbito de la ciberseguridad, esta métrica permite detectar patrones anómalos, identificar comportamientos sospechosos y fortalecer los mecanismos de monitoreo y respuesta ante incidentes.

3. Entropía y Medida de la Información

La Teoría de la Información surge como una disciplina formal que busca cuantificar y modelar el comportamiento de los sistemas de comunicación desde una perspectiva matemática y probabilística. A diferencia de enfoques tradicionales que entienden la información como un contenido semántico o significativo, esta teoría se centra en medir la información como una magnitud abstracta relacionada con la incertidumbre y la probabilidad de los eventos. En el contexto de los sistemas digitales modernos, especialmente aquellos vinculados a la ciberseguridad organizacional, esta aproximación resulta fundamental para analizar, monitorear y proteger los flujos de datos que circulan en la infraestructura tecnológica.

La entropía y la medida de la información permiten responder preguntas clave como: ¿qué tan predecible es el comportamiento de un sistema?, ¿qué eventos aportan más información?, ¿cómo identificar patrones normales y anómalos en grandes volúmenes de datos?

3.1 Medida de la Información

En la Teoría de la Información, medir la información implica cuantificar el grado de sorpresa asociado a la ocurrencia de un evento. Este enfoque parte del supuesto de que no todos los eventos aportan la misma cantidad de información: algunos son esperados y frecuentes, mientras que otros son raros e inesperados. Cuanto menos probable es un evento, más información aporta cuando ocurre, ya que reduce en mayor medida la incertidumbre del sistema receptor (Alencar, 2015).

Esta concepción rompe con la noción intuitiva y cotidiana de información entendida como “cantidad de datos”. Desde esta perspectiva clásica, un archivo grande parecería contener más información que uno pequeño. Sin embargo, la Teoría de la Información redefine el concepto y lo vincula directamente con la probabilidad: un mensaje corto pero inesperado puede ser más informativo que un mensaje largo y repetitivo.

Desde una perspectiva matemática, la información asociada a un símbolo depende exclusivamente de su probabilidad de ocurrencia. Si un evento ocurre con alta frecuencia, su aparición no genera sorpresa ni reduce significativamente la incertidumbre del sistema. Por el contrario, un evento raro introduce nueva información y obliga al sistema receptor a actualizar su estado de conocimiento.

La información no se mide por el volumen de datos, sino por la improbabilidad del evento observado.

En ciberseguridad organizacional, la medida de la información se aplica al análisis de eventos digitales con el objetivo de identificar aquellos que rompen el patrón esperado. Por ejemplo, un acceso exitoso reiterado al sistema por parte de un usuario autorizado es un evento altamente

probable y, por tanto, poco informativo. En cambio, un acceso fallido desde una ubicación inusual o en un horario atípico constituye un evento improbable que incrementa la información disponible para el sistema de monitoreo.

Desde el punto de vista formal, la medida de la información de un evento puede expresarse mediante una función logarítmica de su probabilidad. El uso del logaritmo responde a la necesidad de cumplir ciertas propiedades deseables, como la aditividad y la proporcionalidad. De este modo, eventos independientes pueden combinarse de forma coherente, y la información total puede interpretarse como la suma de informaciones individuales.

$$I(x_i) = -\log_2 p(x_i)$$

3.2 Entropía de Shannon

La entropía constituye una generalización del concepto de medida de la información aplicada no a un evento aislado, sino a una fuente completa de símbolos. Mientras que la medida de la información evalúa la sorpresa asociada a un único evento, la entropía calcula el promedio ponderado de dicha sorpresa considerando todas las posibles salidas de una fuente de información (Shannon, 1948; Csiszár & Körner, 2011).

Shannon definió la entropía como el valor esperado de la información producida por una fuente aleatoria discreta, estableciendo así una métrica formal para cuantificar la incertidumbre inherente a un sistema de comunicación (Shannon, 1948).

Formalmente, la entropía se expresa como en la Figura 1; esta fórmula permite cuantificar el nivel promedio de incertidumbre presente en una fuente de información. Si todos los símbolos tienen probabilidades similares, la entropía alcanza valores elevados, indicando que el sistema es difícil de predecir. Por el contrario, si uno de los símbolos domina claramente sobre los demás, la entropía disminuye, reflejando un comportamiento más predecible.

La entropía mide cuán impredecible es una fuente de información.

Sea una variable aleatoria discreta Xcon alfabeto finito $\{x_1, x_2, \dots, x_n\}$y función de probabilidad $p(x_i)$. La entropía de Shannon se define como en la figura 1.

Figura 1

Entropía: Formula y conceptos Clave

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$$

Símbolo	Descripción
$H(X)$	Entropía de la variable aletoria X
$p(x_i)$	Probabilidad del símbolo x_i
log_2	Logaritmo en base 2 (unidad de medida en bits)

Nota: Definición de la entropía ce una variable aleatoria discreta como la cantidad promedio de incertidumbre, asociada a un conjunto de posibles símbolos X_i

Desde una perspectiva conceptual, la entropía puede interpretarse como una medida del desorden o variabilidad de un sistema. Un sistema altamente ordenado, donde casi siempre ocurre el mismo evento, presenta baja entropía. En cambio, un sistema con múltiples eventos igualmente probables exhibe alta entropía.

En ciberseguridad, esta métrica resulta particularmente útil para identificar desviaciones respecto a un comportamiento normal previamente aprendido.

Por ejemplo, supongamos un servidor de base de datos en una organización donde, durante condiciones normales, el 90 % de las consultas corresponden a operaciones de lectura estándar (SELECT), el 8 % a actualizaciones (UPDATE) y solo el 2 % a eliminaciones (DELETE).

En este escenario:

- SELECT → p = 0.90
- UPDATE → p = 0.08
- DELETE → p = 0.02

Si aplicamos la medida de información:

- $I(\text{SELECT}) = -\log_2(0.90) \approx 0.15 \text{bits}$
- $I(\text{UPDATE}) = -\log_2(0.08) \approx 3.64 \text{bits}$
- $I(\text{DELETE}) = -\log_2(0.02) \approx 5.64 \text{bits}$

Observamos que una operación DELETE aporta mucha más información que una operación SELECT, porque es considerablemente menos frecuente.

Ahora imaginemos que, de forma repentina, el porcentaje de operaciones DELETE aumenta al 25 %. La distribución de probabilidades cambia y, por tanto, también cambia la entropía total del sistema. Este incremento en la diversidad de eventos genera un aumento de la incertidumbre promedio.

En términos de seguridad, este comportamiento podría indicar:

- Eliminación masiva no autorizada de registros.

- Intento de sabotaje interno.
- Actividad de patrones sospechosos de posibles secuestros digitales (ransomware) que borra información antes de cifrarla.

En este caso, la entropía no aumenta por intentos fallidos de acceso (como en el ejemplo anterior), sino por un cambio abrupto en la distribución de tipos de operaciones dentro del sistema.

Además, la entropía no solo se aplica a eventos discretos como accesos o alertas, sino también a características de los datos, como la distribución de puertos utilizados, la longitud de los paquetes o la diversidad de direcciones de origen. En todos estos casos, un cambio significativo en la entropía puede servir como indicador temprano de actividad maliciosa.

3.3 Propiedades de la Entropía

La entropía presenta un conjunto de propiedades fundamentales que facilitan su aplicación práctica en sistemas digitales y permiten interpretar sus valores de manera consistente.

No negatividad.: La entropía siempre es mayor o igual a cero. Este principio indica que no existe incertidumbre negativa: como mínimo, un sistema puede ser completamente predecible, pero nunca más que eso. En términos prácticos, una entropía cercana a cero señala un comportamiento altamente determinista.

Enunciado formal

$$H(X) \geq 0$$

Dado que:

- $0 \leq p(x_i) \leq 1$
- $\log_2 p(x_i) \leq 0$(porque el logaritmo de un número entre 0 y 1 es negativo)

Entonces:

$$-p(x_i)\log_2 p(x_i) \geq 0$$

Como la suma de términos no negativos es no negativa:

$$H(X) \geq 0$$

Por tanto, la entropía nunca puede ser negativa. No existe “incertidumbre negativa”. El mínimo absoluto es cero, lo que representa un sistema completamente predecible.

Supongamos un sistema de monitoreo donde siempre ocurre el mismo evento:

- Evento: “Heartbeat del servidor”
- Probabilidad: $p = 1$

Entonces:

$$H(X) = -1 \cdot \log_2(1) = 0$$

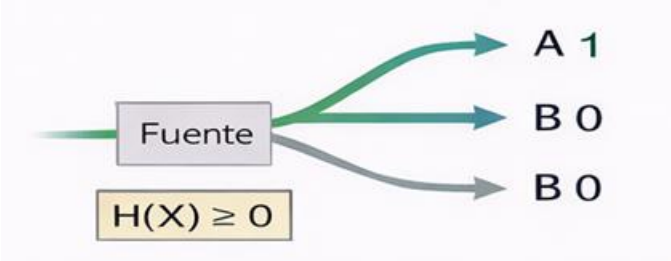
Esto significa que el sistema es totalmente determinista. Cada vez que se consulta el log, aparece el mismo evento. En términos de seguridad, esto podría representar:

- Un proceso automático repetitivo
- Un sistema sin variabilidad

Sin embargo, si un sistema crítico muestra entropía cero en eventos donde normalmente debería existir variabilidad (por ejemplo, tráfico de red), podría indicar manipulación o filtrado artificial de registros.

Figura 2

Ejemplo de No Negatividad en Entropía



Nota: La figura muestra que la entropía es siempre mayor o igual a cero, sin importar como se distribuyan las probabilidades. Elaboración propia adaptado de Gorzalka & Deloumeaux (2012). Miguel Ortiz Navarrete

Máximo de entropía.: La entropía alcanza su valor máximo cuando todos los eventos posibles son equiprobables. En este escenario, el sistema no presenta preferencias ni patrones dominantes, lo que dificulta la predicción de eventos futuros.

Enunciado formal

Para una variable discreta con n símbolos:

$$H(X) \leq \log_2 n$$

La igualdad se cumple cuando hay n eventos posibles y todos tienen la misma probabilidad:

$$p(x_i) = \frac{1}{n} \forall i$$

Si todos los eventos son equiprobables, remplazamos en la definición:

$$H(X) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n}$$

Aquí ocurre algo clave:

- Todos los términos de la suma son iguales.
- Tenemos exactamente n términos idénticos.

Entonces la suma se convierte en:

$$H(X) = -n\left(\frac{1}{n}\log_2\frac{1}{n}\right)$$

Observa que:

$$n \cdot \frac{1}{n} = 1$$

Entonces:

$$H(X) = -\log_2\frac{1}{n}$$

Sabemos que:

$$\log_2\frac{1}{n} = -\log_2 n$$

Por lo tanto:

$$H(X) = -(-\log_2 n)$$

$$H(X) = \log_2 n$$

Esta expresión final nos dice: La entropía máxima depende únicamente del número de eventos posibles, no de su contenido.

En otras palabras:

- *Si existen más resultados posibles, hay mayor incertidumbre.*
- *Si el número de opciones se duplica, la entropía aumenta en 1 bit.*

Tabla 1

Incremento de los bits con el incremento de posibilidades

n	*H(X)*
2	*1 bit*
4	*2 bits*
8	*3 bits*
16	*4 bits*

Nota: La tabla muestra que: cada vez que duplicamos el número de posibilidades, aumentamos 1 bit de incertidumbre.

Supongamos un análisis de tráfico DNS donde se observan 4 dominios consultados con igual frecuencia:

Probabilidad:

$$p = 0.25$$

Entonces:

$$H(X) = -4(0.25\log_2 0.25)$$
$$H(X) = 2 \text{ bits}$$

En seguridad, una distribución uniforme puede indicar:

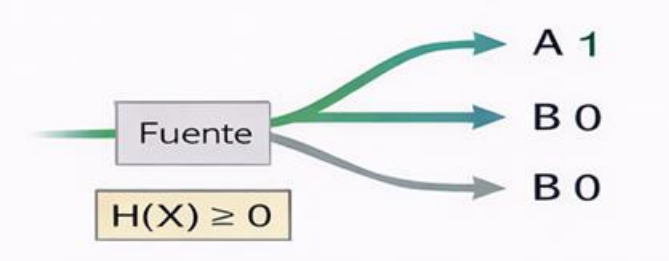
- Tráfico legítimo diversificado
- O técnicas de ocultamiento (por ejemplo, malware que distribuye consultas equitativamente para evitar patrones detectables)

Muchos sistemas de detección de exfiltración utilizan análisis de entropía en tráfico DNS, ya que datos cifrados tienden a generar distribuciones más uniformes (mayor entropía).

La entropía es máxima cuando el sistema es completamente impredecible. La entropía máxima ocurre cuando no existe patrón dominante en la distribución de probabilidades.

Figura 3

Ejemplo de Máximo de entropía



Nota: La figura muestra una distribución equiprobable de eventos, indicando el estado de máxima incertidumbre. Elaboración propia adaptado de Gorzalka & Deloumeaux (2012). Miguel Ortiz Navarrete

Entropía mínima.: La entropía es mínima cuando un evento ocurre con probabilidad uno. En este caso, no existe incertidumbre alguna, ya que el resultado es completamente predecible. Aunque esto puede parecer deseable, una entropía excesivamente baja también puede ser sospechosa, ya que puede indicar automatización o comportamiento artificial.

Enunciado formal

$$H(X) = 0 \text{ si y solo si } p(x_k) = 1 \text{ para algún } k$$

Si un solo evento tiene probabilidad 1:

$$H(X) = -1\log_2(1) = 0$$

No existe incertidumbre. La entropía mínima indica predictibilidad absoluta.

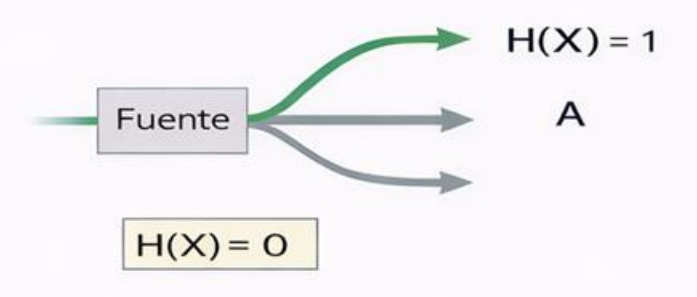
Supongamos un firewall donde solo se registra el evento: “Conexión permitida” y nunca aparece ningún otro tipo de evento. Esto genera entropía cero.

En un sistema real, esto puede significar:

- Configuración mal definida
- Logs incompletos
- Manipulación de registros
- Automatización excesiva

Figura 4

Ejemplo de entropía mínima



Nota: La figura indica una fuente de información que siempre emite el mismo símbolo, reflejando un estado de mínima entropía. Elaboración propia adaptado de Csiszár & Körner (2011). Miguel Ortiz Navarrete

Aditividad.: La entropía conjunta de dos fuentes discretas depende de la independencia de estas. Si las fuentes son independientes, la entropía total es la suma de las entropías individuales. Si existe dependencia, la entropía conjunta se reduce, reflejando la correlación entre los eventos.

Sea:

- X= variable aleatoria discreta 1
- Y= variable aleatoria discreta 2

La entropía conjunta se define como:

$$H(X,Y) = -\sum_{x,y} p(x,y) \log_2 p(x,y)$$

Esta cantidad mide la incertidumbre total del sistema combinado (X,Y).

Existe una identidad fundamental (Descomposición) en teoría de la información:

$$H(X,Y) = H(X) + H(Y \mid X)$$

Esto siempre es cierto, sin importar si hay independencia o no.

¿Qué significa esta exprecióin?

- $H(X)$= incertidumbre de X.
- $H(Y \mid X)$= incertidumbre de Y sabiendo el valor de X.

Es decir: La incertidumbre total es igual a la incertidumbre de la primera variable más la incertidumbre de la segunda después de conocer la primera. Esta es una especie de "regla de suma" de incertidumbre.

Caso especial: Independencia

Dos variables son independientes si:

$$p(x,y) = p(x)p(y)$$

Cuando eso ocurre, conocer Xno cambia la probabilidad de Y.

Entonces:

$$H(Y \mid X) = H(Y)$$

Sustituyendo en la fórmula general:

$$H(X,Y) = H(X) + H(Y)$$

Relación con Información Mutua ($I(X;Y)$)

El término: $I(X;Y)$ se llama Información Mutua entre las variables aleatorias Xy Y.

La diferencia entre ambos casos se expresa mediante la información mutua:

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$I(X;Y)$ mide: Cuánta incertidumbre de una variable se reduce al conocer la otra.

Es decir: Si conocer Yno cambia en nada lo que sabemos sobre X, entonces:

$$I(X;Y) = 0$$

Si conocer Yelimina totalmente la incertidumbre sobre X, entonces la información mutua es máxima:

$$I(X;Y) > 0$$

Por tanto:

$$H(X,Y) = H(X) + H(Y) - I(X;Y)$$

Esto significa que:

- Si no hay relación → simplemente se suman las incertidumbres.
- Si hay relación → debemos restar la parte compartida.

Esa "parte compartida" es precisamente $I(X;Y)$.

La información mutua también puede escribirse como:

$$I(X;Y) = \sum_{x,y} p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

Esta expresión revela algo clave:

- Si $p(x,y) = p(x)p(y)$→ el logaritmo es 0 → información mutua 0.
- Si difieren → hay dependencia → información mutua positiva.

Esto mide cuánto se aparta la distribución conjunta de la independencia.

Supongamos:

Variable X = Tipo de evento (LOGIN_OK / LOGIN_FAIL)

Variable Y = Dirección IP

Caso normal:

- Los eventos están distribuidos aleatoriamente entre IPs.
- No hay patrón.
- $I(X;Y) = 0$

Caso ataque:

- Todos los LOGIN_FAIL provienen de una IP específica.
- Conocer la IP reduce mucho la incertidumbre sobre el tipo de evento.
- Entonces:

$$I(X;Y) > 0$$

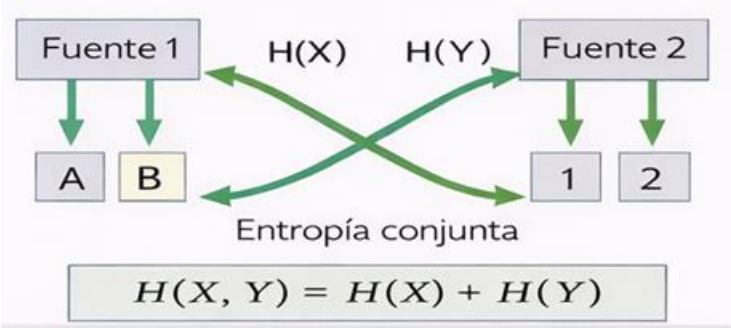
Esto indica correlación estructural.

Los sistemas de detección pueden usar esta métrica para identificar:

- Ataques dirigidos
- Escaneos coordinados
- Exfiltración focalizada

Figura 5

Ejemplo de Aditividad en entropía



Nota: La figura representa dos fuentes de información independientes, donde la entropía conjunta es la suma de las entropías individuales. Elaboración propia adaptado de Csiszár & Körner (2011).

Las propiedades de la entropía permiten distinguir entre comportamiento normal y comportamiento anómalo (Csiszár & Körner, 2011).

Figura 6

Propiedades fundamentales de la entropía

Nota: La figura muestra las principales propiedades de la entropía en la teoría de la información: no negatividad, entropía mínima, aditividad e información mutua, junto con ejemplos aplicados a ciberseguridad. Elaboración propia adaptado de Claude E. Shannon (1948), Imre Csiszár y János Körner (2011), y Marcelo Sampaio de Alencar (2015). Miguel Ortiz Navarrete

Taba 2

Matriz comparativa de las propiedades de la entropía

Propiedad	Definición	Expresión Matemática	Interpretación	Ejemplo en Ciberseguridad
No negatividad	La entropía nunca puede ser negativa	H(X) ≥ 0	La incertidumbre mínima es cero	Logs donde siempre aparece el mismo evento
Entropía mínima	Ocurre cuando un evento tiene probabilidad 1	H(X)=0	Sistema completamente predecible	Firewall que solo registra “Conexión permitida”

Entropía máxima	Se produce cuando todos los eventos son equiprobables	$H(X)=\log_2(n)$	Máxima incertidumbre del sistema	Tráfico DNS distribuido uniformemente
Aditividad	La entropía conjunta depende de la independencia de variables	$H(X,Y)=H(X)+H(Y)$ si son independientes	Permite combinar fuentes de información	Eventos de autenticación independientes del origen
Relación con información mutua	La dependencia entre variables reduce la entropía conjunta	$H(X,Y)=H(X)+H(Y)-I(X;Y)$	La información compartida reduce la incertidumbre total	IP específica asociada a múltiples intentos de login fallidos

Nota: Matriz elaborada con base en los fundamentos de la teoría de la información descritos por Shannon (1948), Csiszár y Körner (2011) y Alencar (2015).

3.4 Aplicaciones de la Entropía en Ciberseguridad

En el ámbito de la ciberseguridad, la entropía se emplea como una métrica transversal para analizar patrones de comportamiento en sistemas digitales complejos. Su principal ventaja radica en que permite detectar anomalías sin necesidad de conocer previamente la naturaleza exacta del ataque, basándose únicamente en cambios estadísticos en la distribución de los eventos.

Un incremento inesperado de entropía puede indicar actividad anómala, como exploraciones de red, tráfico cifrado malicioso o generación de datos aleatorios para evadir mecanismos de detección. Por el contrario, una reducción excesiva de entropía puede revelar automatización, repetición sistemática o intentos de explotación mediante scripts.

Entre las principales aplicaciones de la entropía en ciberseguridad se encuentran:

- **Detección de ataques de fuerza bruta.** La repetición de intentos fallidos modifica la distribución de eventos de autenticación, generando patrones detectables mediante entropía.
- **Análisis de tráfico de red.** Cambios en la entropía de direcciones IP, puertos o tamaños de paquetes pueden señalar escaneos, exfiltración de datos o comunicación con servidores de comando y control.
- **Evaluación de entropía en contraseñas.** La entropía permite medir la fortaleza de una contraseña en función de la diversidad y distribución de sus caracteres.
- **Identificación de malware polimórfico.** Este tipo de malware modifica su estructura para evadir firmas.

- Alencar, M. S. (2015). Information theory. Momentum Press.
- Csiszár, I., & Körner, J. (2011). Information theory: Coding theorems for discrete memoryless systems (2nd ed.). Cambridge University Press.
- Gorzalka, J. D., & Deloumeaux, P. (2012). Information theory: New research. Nova Science Publishers.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

PUCE Virtual

síguenos en:

www.pucevirtual.puce.edu.ec