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

APLICACIÓN DE UN MODELO DE APRENDIZAJE BASADO EN LA EXPERIENCIA A JUEGOS DE CLASIFICACIÓN DE ADVERSARIOS

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

JUAN ANDRÉS MUÑOZ OLIVEROS

PROFESOR GUÍA:
RICHARD WEBER HAAS

MIEMBROS DE LA COMISIÓN:
NICOLÁS FIGUEROA GONZÁLEZ
JUAN ESCOBAR CASTRO

SANTIAGO DE CHILE
2018

RESUMEN DE LA MEMORIA PARA OPTAR
AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: JUAN ANDRÉS MUÑOZ OLIVEROS
FECHA: 2018
PROF. GUÍA: RICHARD WEBER HAAS

APLICACIÓN DE UN MODELO DE APRENDIZAJE BASADO EN LA EXPERIENCIA A JUEGOS DE CLASIFICACIÓN DE ADVERSARIOS

Un juego de clasificación de adversarios típico considera un clasificador y un adversario, que puede ser de tipo regular o malicioso. El clasificador debe intentar clasificar bien al adversario, sin conocer su tipo; mientras que el adversario conoce las preferencias del clasificador, y puede adaptar sus jugadas (tipo de mensaje enviado) para burlar la clasificación.

La literatura se ha centrado en modelar este juego desde distintos enfoques, siempre buscando encontrar la estrategia óptima del adversario. Luego, con ella, se deduce la estrategia óptima que debe seguir el clasificador. Las pruebas con datos reales han arrojado resultados muy superiores a los algoritmos típicos de clasificación, que no incorporan técnicas de la teoría de juegos.

En esta investigación se plantea un modelo basado en la estructura de los juegos de señalización, que deja completamente de lado los supuestos de información pública sobre el clasificador, y la capacidad de los jugadores de observar las acciones del otro.

Para ello, se introduce un algoritmo de aprendizaje mediante la regla de elección aleatoria Logit, que los induce a adaptarse desde el ensayo y error. De esta manera los jugadores son capaces de adaptar sus estrategias turno a turno, observando únicamente sus propias estrategias y las utilidades obtenidas en el pasado.

Utilizando este modelo, los jugadores son capaces de converger rápidamente al equilibrio bayesiano perfecto del juego, de manera mixta: los adversarios de tipo regular juegan estrategias puras sobre su mensaje preferido, mientras que los de tipo malicioso juegan estrategias mixtas entre los distintos mensajes disponibles. Por su parte, en el equilibrio las estrategias del clasificador se han ajustado a la proporción de adversarios maliciosos que envía cada mensaje en el equilibrio.

En el equilibrio de este juego, los adversarios maliciosos se mueven dinámicamente entre los mensajes que escogen enviar, buscando burlar la clasificación. El error de clasificación asociado a ellos oscila constantemente, incluso en el equilibrio; lo que demuestra un comportamiento de *gato y ratón* constante entre el clasificador y los adversarios maliciosos.

La mayor contribución del modelo, es que logra capturar la evolución hacia el equilibrio, las estrategias, el dinamismo del juego y la persecución constante entre los jugadores; sin que estos se puedan observar directa o indirectamente en todo el juego.

A mis padres.

Agradecimientos

A mis padres, Juan Muñoz y Violeta Oliveros, quienes sin haber tenido la oportunidad, lo dieron todo para que yo pudiera estudiar una carrera universitaria sin que nada me hiciera falta.

A Richard Weber y Nicolás Figueroa, por todo el apoyo, la confianza y la libertad que me dieron durante el desarrollo de esta investigación.

A Claudio Letelier e Ignacio Calisto. Profesores que con simples conversaciones me han enseñado mucho más de lo que se puede aprender en una sala de clases.

A Mary Farías, por todo el cariño, la compañía y la ayuda incondicionales en este proceso.

A mis compañeros y amigos, que hicieron tan único mi paso por la universidad.

Tabla de Contenido

Índice de Ilustraciones

Introducción

Un problema de clasificación corresponde a un juego en el que un jugador, llamado *clasificador*, intenta designar la clase correcta a los distintos jugadores (no clasificados) a los que se enfrenta. Para ello utiliza la información pública y privada disponible en las condiciones iniciales del juego, y la que va recopilando durante el desarrollo del mismo.

Sin embargo, muchas de las aplicaciones más interesantes en esta área, como la detección de spam en correo electrónico o la detección de fraudes, tienen en común a jugadores con intenciones indeseadas, llamados *adversarios maliciosos*. Estos jugadores aprenden a ocultar su comportamiento entre el de aquellos cuyas intenciones no son indeseadas, llamados *adversarios regulares*. De esta manera logran eludir la clasificación y aumentar su beneficio.

El caso de detección de spam en correo electrónico es el más ampliamente desarrollado por la literatura, debido a que sus características son representativas de las otras aplicaciones de este tipo de modelos. En este caso, el clasificador se enfrenta a adversarios regulares, que envían correos electrónicos; y a maliciosos, que envían spam. El clasificador busca filtrar los mensajes enviados por los adversarios maliciosos, mientras que estos aprenden a burlar los filtros cambiando el mensaje enviado para así, imitar el comportamiento de los adversarios regulares.

Desde la teoría de juegos se ha abordado este problema, con el fin de entender la interacción entre el clasificador y los adversarios maliciosos. Para ello, en la literatura se han estudiado modelos con información completa o incompleta sobre las preferencias, siempre escogiendo estrategias óptimas en función de las decisiones del otro.

En este trabajo se propone un modelo donde, tanto el clasificador como los adversarios maliciosos, optimizan su estrategia en función de su propia información privada, sin observar ni definir creencias sobre las preferencias del otro.

En el capítulo 1 se sientan las bases teóricas más relevantes para este trabajo. En el capítulo 2 se presenta el modelo propuesto, incorporando aprendizaje de los jugadores a partir de su información privada, y se complementa con un ejemplo práctico que grafica la dinámica del juego en este contexto. Finalmente, en el capítulo 3 se presentan los principales resultados del juego al simular la interacción entre los jugadores con datos sintéticos, y los efectos de los parámetros más relevantes sobre sus resultados.

Capítulo 1

Marco Teórico

La detección de ataques es modelada normalmente como un problema de clasificación, en aplicaciones como la detección de spam, fraude, intrusión a sistemas informáticos, vigilancia contra el terrorismo, etc.

Sin embargo, en este tipo de aplicaciones, los algoritmos de clasificación estándar entregan resultados deficientes. Esto se debe a que se enfrentan a atacantes que modifican activamente su comportamiento con el objetivo de burlar la clasificación.

La dificultad se presenta en que los datos con los que se intenta aplicar los algoritmos de clasificación, no cumplen el típico supuesto de ser idénticos y uniformemente distribuidos.

Por el contrario,los datos se encuentran sucios producto de un atacante, llamado *adversario malicioso*, que adapta constantemente su estrategia para burlar al clasificador.

Para agregar dificultad, este jugador convive en el mismo ambiente que los agentes no atacantes, llamados *adversarios regulares*, quienes se comportan de acuerdo a reglas completamente diferentes. Estos agentes no intentan influir de ninguna manera en el rendimiento del clasificador.

Por su parte, el clasificador debe no clasificar erróneamente como atacantes a los adversarios regulares, mientras intenta clasificar correctamente a los adversarios maliciosos. Todo esto sin conocer realmente quiénes son sus adversarios, ni a cuántos se enfrenta.

Desde la teoría de juegos se ha desarrollado una línea de trabajo que intenta modelar este tipo de escenarios y estudiar sus propiedades. A este tipo de modelos se le llama *juegos de clasificación de adversarios*.

1.1. Avances en la literatura

El primer acercamiento a los juegos de clasificación de adversarios es el trabajo de Dalvi et al. (2004). Ellos proponen representar el juego como la interacción entre dos agentes: un adversario malicioso y un clasificador.

El adversario malicioso intenta burlar al clasificador imitando las acciones de los adversarios regulares, para así aumentar su utilidad. El clasificador, por su parte, intenta identificar tantas acciones del adversario malicioso como pueda. Sin embargo, las acciones del adversario malicioso le inducen aumentar la tasa de error de tipo I; clasificar falsamente como maliciosa la acción de un adversario regular. El clasificador debe identificar correctamente las acciones maliciosas cuidándose del error, maximizando así su utilidad esperada.

El modelo de Dalvi et al. (2004) considera una fase de entrenamiento en la que el clasificador desarrolla una regla *bayesiana inocente*. Con esta regla clasifica a los adversarios en el desarrollo posterior del juego. El modelo se soporta en cuatro supuestos:

1. Información completa. Tanto el adversario como el clasificador conocen todos los parámetros relevantes del juego. Estos parámetros son: conjuntos de acciones, pagos, utilidades y la regla de decisión desarrollada por el clasificador en la fase de entrenamiento.
2. El adversario asume que el clasificador no sabe que se encuentra en presencia del adversario. En otras palabras, el adversario asume que el clasificador usará siempre la regla de decisión desarrollada en la fase de entrenamiento.
3. El clasificador asume que el adversario usa su estrategia óptima para modificar sus elecciones.
4. Los datos usados en la fase de entrenamiento provienen de la distribución real entre adversarios maliciosos y regulares. Es decir, no es manipulada por el adversario.

Bajo estos supuestos, Dalvi et al. (2004) define los conjuntos de decisión y las utilidades del adversario y del clasificador, con el objetivo de deducir la estrategia óptima del adversario. Luego, con esta estrategia, recupera la estrategia de mejor respuesta del clasificador.

Los autores testean con datos reales las estrategias óptimas del clasificador, y concluyen que presenta resultados mucho mejores que los que se obtienen con algoritmos típicos de clasificación que no consideran adversarios. Sin embargo, aún con los buenos resultados empíricos, los supuestos sobre los que se construye el modelo son muy alejados de la realidad.

Lowd y Meek (2005) continúan con la línea de investigación de Dalvi et al. (2004). Exploran los resultados que se obtienen al relajar el supuesto de información completa de los agentes. Para ello introducen una experimentación activa de los adversarios para mejorar sus estrategias, a pesar de la limitación de información.

Los autores definen un sistema de aprendizaje del adversario, llamado *ingeniería reversa del clasificador* (ACRE[1]). Este sistema permite que el adversario obtenga información

[1]Por sus siglas en inglés: adversarial classifier reverse engineering.

suficiente sobre el comportamiento del clasificador como para construir sus estrategias de ataque. Aquí el objetivo del adversario no es conocer a la perfección al clasificador, sino que, encontrar suficientes características de los mensajes que le permitan no ser clasificados como maliciosos; para así mejorar su estrategia.

De esta manera, el trabajo de Lowd y Meek (2005) permite encontrar una estrategia óptima del adversario más realista. Debido a que no necesita del supuesto de información completa, ni supuestos sobre la distribución de los datos con que el clasificador decide sus reglas de clasificación. Con este resultado se estima la estrategia óptima del clasificador y se prueba con datos reales de filtros de spam; y se obtienen buenos resultados.

Sin embargo, el trabajo de Lowd y Meek (2005) no considera efectos como un posible ajuste en las estrategias del clasificador como respuesta óptima a las acciones del adversario. Otros autores toman otros enfoques del mismo problema, relajando estos supuestos.

Liu y Chawla (2009) modelan la interacción entre el clasificador y el adversario como un juego no cooperativo a la Stackelberg, donde el adversario actúa como líder; y el clasificador como su seguidor. Este juego se trata como un problema de optimización, derivando así las estrategias de cada jugador.

De esta forma los autores logran mejorar el modelo de Lowd y Meek (2005), al relajar el supuesto de que el adversario conoce la función de pagos del clasificador.

Como resultado de sus simulaciones, Liu y Chawla (2009) logran concluir que los jugadores pueden alcanzar el equilibrio de Stackelberg al jugar sus estrategias óptimas de manera simultánea.

Otros autores utilizan la estructura de los juegos de señales desarrollada por Fudenberg y Tirole (1991), y Gibbons (1992). En este modelo se presenta un juego con dos jugadores: el primero con información privada y el segundo con información pública. En un primer turno, el primer agente elige una acción, la que es observada por el segundo, quien elige su acción basado en esta información. En el contexto de clasificación de adversarios, el primer jugador corresponde a un adversario malicioso, y el segundo, al clasificador.

El trabajo de Figueroa et al. (2017) utiliza la estructura de los juegos de señales para modelar el juego, y aplica el modelo a datos reales de correo electrónico, en el marco de la aplicación en detección de spam.

Para ello, utiliza algoritmos de la minería de datos para formar clusters de correos electrónicos, que pasan a formar el conjunto de acciones que puede tomar el adversario. El clasificador, en tanto, debe decidir con qué probabilidad frenar (filtrar) cada uno de los correos electrónicos pertenecientes a cada cluster. En su modelo, se debe aplicar la misma estrategia (probabilidad de filtrado) a todos los correos pertenecientes al mismo cluster.

Figueroa et al. (2017) deriva las estrategias óptimas del adversario y del clasificador como un equilibrio de Nash bayesiano, y obtiene resultados competitivos con la literatura al aplicarlos en la detección de spam con datos reales.

1.2. Aprendizaje basado en la experiencia

En el contexto de juegos de clasificación de adversarios, la literatura se ha centrado en modelar el juego desde distintos enfoques. Siempre mantiene el objetivo de modelar la estrategia óptima de los adversarios maliciosos, y con esta información, derivar la estrategia óptima del clasificador.

En este trabajo se amplía la línea de investigación de Figueroa et al. (2017), utilizando la estructura de los juegos dinámicos bayesianos. Sin embargo, se dejan completamente de lado los supuestos de información pública sobre el clasificador, y la capacidad de los jugadores de observar las acciones de cualquier otro.

Para determinar las decisiones de cada jugador, en cada turno, se utiliza un sistema de *aprendizaje desacoplado*[2] similar a la *experimentación dinámica simple* definida por Marden et al. (2007).

En su modelo, Marden et al. (2007) considera juegos en el que los jugadores deciden de manera simultánea y repetida sus acciones, entre un conjunto finito de estrategias. Define distintas reglas de ajuste de estrategias, que se basen solo en la experiencia privada de cada jugador: sus propias acciones y pagos. En particular, define la dinámica de experimentación simple como un sistema de ajuste de estrategias que sigue las siguientes fases:

1. Inicialización: En $t = 0$, cada jugador elige de manera aleatoria la acción que jugará en el primer turno. Esta acción es considerada la *acción base* del jugador, y su utilidad asociada se considera la *utilidad base*.
2. Selección de acciones: Para los turnos siguientes, cada jugador elige su acción base con probabilidad $(1-\varepsilon)$, o experimenta con una acción nueva con probabilidad ε. La acción con la que se experimenta se elige con probabilidad uniforme entre todas las acciones disponibles.
3. Actualización de la información: Si el jugador experimentó y obtuvo una mayor utilidad que la base, entonces define como acción base la acción jugada, y su utilidad base se actualiza al valor de la utilidad obtenida. Si experimentó pero la utilidad obtenida fue menor, entonces mantiene su acción base y utilidad base. Por último, si el jugador no experimentó, entonces mantiene su acción base, y actualiza su utilidad base al valor obtenido en este turno.
4. Volver al paso 2 y repetir.

Este algoritmo permite que los jugadores actualicen sus decisiones basados en su propio historial de acciones y utilidades obtenidas, por lo que es posible prescindir de supuestos de información completa del juego, o de creencias que puedan definir los jugadores sobre las acciones o pagos del resto.

[2]Hart y Mas-Collel (2003) definen el *aprendizaje desacoplado* de un jugador, como cualquier sistema de aprendizaje que no utilice la información de los pagos de sus contrincantes, aunque puede depender de sus decisiones.

Sin embargo, el modelo de aprendizaje de Marden et al. (2007) define una probabilidad de experimentación fija, lo que no considera aspectos más sutiles de la dinámica de un juego de clasificador de adversarios. En estos juegos, por ejemplo, un adversario puede ver reducida su utilidad debido a simple mala suerte, y no porque el clasificador haya descubierto sus intenciones.

Por otro lado, Cominetti et al. (2010) utiliza una regla de decisión Logit aplicada a la elección de rutas de tránsito de los jugadores. En su modelo, se asume que cada jugador tiene una percepción o estimación previa del rendimiento de cada ruta, y decide en función de ella la ruta a tomar. Luego observa su utilidad resultante y actualiza sus creencias respecto a esa ruta en particular, modificando las probabilidades de elección de cada ruta según la regla Logit.

Un juego de clasificación de adversarios y un juego de elección de rutas en el tránsito tienen una característica clave en común: cuando un jugador elige su acción a tomar, no tiene la posibilidad de saber cómo habrían resultado sus pagos de haber tomado otra decisión. Además, cada jugador es incapaz de observar las decisiones que tomaron los otros jugadores, y sin embargo estas decisiones afectan directamente su utilidad.

Por esta razón, en este trabajo se adapta el algoritmo de aprendizaje descrito por Marden et al. (2007) a la clasificación de adversarios, mediante la regla de elección aleatoria Logit. De esta forma, se le da libertad a la probabilidad de experimentación para que evolucione en el tiempo según la dinámica particular del juego, y se mantiene la característica de ajuste de las estrategias a partir de la experiencia propia de cada jugador.

Capítulo 2

Modelo

2.1. Modelo en estrategias puras

Este juego es modelado como un juego bayesiano, repetido e infinito, donde dos jugadores, el adversario y el clasificador, juegan en forma simultánea sus estrategias.

En todo el modelo, llamamos t a cada periodo en el tiempo.

1. Jugadores:

$$I = \{A, C\}$$

   Donde A corresponde al adversario y C al clasificador.

2. Estrategias:

$$\begin{array}{lrcl} \text{Adversario:} & S_t^A & \in & K = \{1, \ldots, k\},\ \forall t \\ \text{Clasificador:} & S_t^C & \in & \{0,1\}^k,\ \forall t \end{array}$$

   Se define que el clasificador *detiene* el mensaje $\hat{j}$ escogido por el adversario, si $S_t^C = 0$, y que lo *permite* si $S_t^C = 1$.

3. Tipos de cada jugador:

$$\begin{array}{lrcl} \text{Adversario:} & \theta^A & = & (\tau, \bar{j}) \in \{R, M\} \times \{1, ..., k\} \\ \text{Clasificador:} & \theta^C & = & \{C\} \end{array}$$

   $\tau = R$ representa que el adversario es de tipo regular, y $\tau = M$, que es de tipo malicioso. La asignación $\bar{j} \in \{1, ..., k\}$ corresponde a la acción que reporta el máximo de utilidades para el adversario i, en caso de no ser frenada por el clasificador.

   Por su parte, el clasificador es de un tipo único, C, y será omitido en el futuro.

4. Utilidades:

   Las utilidades de cada jugador, en cada momento t del juego, están dadas por:

   (a) Adversario:

$$u_t^A = \begin{cases} b_{\hat{j}\bar{j}} & \text{si} \quad S_{t\hat{j}}^C = 1 \\ -c & \text{si} \quad S_{t\hat{j}}^C = 0 \end{cases}$$

   Donde $\hat{j}$ es la estrategia escogida por el adversario en el turno t.

   El adversario obtiene cierto beneficio $b_{\hat{j}}$ cuando pasa el filtro del clasificador, y un costo c cuando no. El beneficio depende del mensaje escogido (en particular, si $\hat{j} = \bar{j}$, el beneficio es máximo), pero es constante en el tiempo, mientras que el costo es constante tanto en j como en t.

   (b) Clasificador:

$$u_t^C = \begin{cases} -L_M & \text{si} \quad S_{t\hat{j}}^C = 1 \wedge \theta^A = \left(M, \hat{j}\right),\ \forall \hat{j} \in \{1, \ldots, k\} \\ -L_R & \text{si} \quad S_{t\bar{j}}^C = 0 \wedge \theta^A = \left(R, \bar{j}\right),\ \forall \bar{j} \in \{1, \ldots, k\} \\ 0 & \sim \end{cases}$$

   El clasificador obtiene utilidad nula si clasifica bien al adversario. Sin embargo, si lo clasifica mal y el adversario era de tipo malicioso, pierde L_M, y si era de tipo regular, pierde L_R.

   Típicamente se entiende que el error de Tipo I es más grave que el error de Tipo II, por lo que se define $-L_R < -L_M < 0$.

5. Distribución de los tipos:

   El tipo del adversario se escoge acorde a la distribución de probabilidad $p(\theta^A)$, donde $p(M, \bar{j})\ ,\ p(R, \bar{j}) > 0$ con $\sum_{\theta^A} p(\theta^A) = 1,\ \forall \bar{j} \in \{1, ..., k\}$.

   Se usa la notación $p(\tau, \bar{j}) = p(\theta^A = (\tau, \bar{j}))$.

2.1.1. Extensión mixta

1. Estrategias mixtas:

$$\begin{array}{lrcl} \text{Adversario:} & \sigma_t^A & = & \Delta_t\left(K\right),\ \forall t \\ \text{Clasificador:} & \sigma_t^C & = & [0,1]^k,\ \forall t \end{array}$$

   El adversario escoge en cada turno el espacio de probabilidades con las que jugar cada acción $j \in K$, mientras que el clasificador escoge de manera simultánea la probabilidad con la que *permite* cada mensaje.

   Cabe notar que $\sigma_{t\hat{j}}^A = 1$ es equivalente a $S_t^A = \hat{j}$, y que $\sigma_t^C = 1$ es equivalente a $S_t^C = 1$ en la versión de estrategias puras del juego.

2. Utilidad esperada:

   (a) Adversario:

$$\overline{u^A} = \sum_{t=1}^{\infty} \sum_{\hat{j}=1}^{k} \sigma_{t\hat{j}}^A \cdot p\left(\theta^A\right) \cdot \left(\sigma_{t\hat{j}}^C \cdot b_{\hat{j}\bar{j}} - \left(1 - \sigma_{t\hat{j}}^C\right) \cdot c\right)$$

   (b) Clasificador:

$$\overline{u^C} = -\sum_{t=1}^{\infty} \sum_{\hat{j}=1}^{k} \sigma_{t\hat{j}}^A \cdot \left(1 - \sigma_{t\hat{j}}^C\right) \cdot \left(p\left(M, \bar{j}\right) \cdot L_M + p\left(R, \bar{j}\right) \cdot L_R\right)$$

2.2. Desarrollo del juego

2.2.1. Definiciones previas

1. Desde este punto en adelante, se interpretará que existen N adversarios a los que la naturaleza asigna en $t = 0$ un único tipo θ_{i}^A, $\forall \mathrm{i} \in \{1, \ldots, N\}$, según la distribución de probabilidad $p\left(\theta^A\right)$ definida en la sección 2.1.

   El resultado son n_M adversarios de tipo malicioso y n_R adversarios de tipo regular, con $n_M + n_R = N$. El clasificador no puede observar n_M ni n_R, pero sí N.

   Esto es equivalente a tener un solo adversario al que la naturaleza le asigna su tipo según la distribución de probabilidad anterior, sin que el clasificador pueda observarlo.

   Por lo tanto, desde este punto en adelante se interpretan los parámetros del juego relacionados con el adversario, definidos en las secciones 2.1. y 2.1.1., como vectores cuya componente $\mathrm{i} \in \{1, \ldots, N\}$ corresponde a la definición del parámetro para el adversario i.

2. Se llamará *acción* a la jugada pura ($\hat{j} \in K$) que percibe el contrincante respectivo de cada jugador, al aplicarse las estrategias mixtas escogidas en cada turno.

   En otras palabras, cada adversario y el clasificador juegan su estrategia mixta, pero solo pueden observar los resultados de las acciones puras en las que se traducen las estrategias de su contrincante. Esta información es la que luego pueden usar para decidir su estrategia del turno siguiente en función de los resultados que han obtenido en el pasado.

   Se definen Φ_t^A y Φ_t^C como los vectores que contienen las acciones de cada turno, de los adversarios y del clasificador respectivamente, donde cada elemento i de Φ_t^A corresponde a la acción de cada adversario, y cada elemento i de Φ_t^C corresponde a la acción que percibe cada adversario como resultado del filtro (estrategia mixta) que aplica el clasificador sobre cada mensaje $j \in K$.

3. Por último, para incorporar el aprendizaje de los jugadores, se definen las siguientes matrices, en las que cada jugador guarda la historia de resultados que ha obtenido con cada acción $\hat{j} \in K$ en el desarrollo del juego:[1]

[1] Solo es relevante definir las matrices de guardado de información del clasificador y de los adversarios maliciosos, ya que los adversarios regulaes tienen una estrategia fija, que no depende de sus resultados acumulados del pasado. Esto es equivalente a decir que cada jugador i de tipo regular también guarda su historia de la misma manera que uno de tipo malicioso, pero no la usa para decidir su estrategia del turno siguiente, por lo que es innecesario mostrarla.

(a) Adversario malicioso: sea V_t^{i} el vector que guarda los resultados acumulados hasta el turno t por el adversario i, de tipo malicioso, al jugar cada acción posible:

$$V_t^{\mathrm{i}} = \begin{pmatrix} v_{1t}^{\mathrm{i}} \\ v_{2t}^{\mathrm{i}} \\ \vdots \\ v_{kt}^{\mathrm{i}} \end{pmatrix}, \text{ donde } v_{\hat{j}t}^{\mathrm{i}} = \begin{cases} v_{\hat{j},t-1}^{\mathrm{i}} & \text{si} \quad \Phi_{\mathrm{i},t-1}^{A} \neq \hat{j} \\ \left(1 - \alpha_{\hat{j}t}^{\mathrm{i}}\right) \cdot v_{\hat{j},t-1}^{\mathrm{i}} + \alpha_{\hat{j}t}^{\mathrm{i}} \cdot u_{\hat{j}t}^{\mathrm{i}} & \text{si} \quad \Phi_{\mathrm{i},t-1}^{A} = \hat{j} \end{cases}$$

Donde $u_{\hat{j}t}^{\mathrm{i}}$ es la utilidad obtenida por el jugador i en el turno t del juego, al jugar la acción $\hat{j}$.

Se define el parámetro $\alpha_{\hat{j}t}^{\mathrm{i}} = \dfrac{1}{1 + \eta_{\hat{j},t-1}^{\mathrm{i}}}$, donde $\eta_{\hat{j},t-1}^{\mathrm{i}}$ es el número de veces que i ha jugado la acción $\hat{j}$ hasta el turno $t-1$.[2]

(b) Clasificador: sea W_t el vector que guarda los resultados acumulados hasta el turno t por el clasificador, al jugar cada acción posible, $\phi \in \{0, 1\}$:

$$W_t = \begin{pmatrix} w_{1t}^{0} & w_{1t}^{1} \\ w_{2t}^{0} & w_{2t}^{1} \\ \vdots & \vdots \\ w_{kt}^{0} & w_{kt}^{1} \end{pmatrix}, \text{ donde } w_{\hat{j}t}^{\phi} = \left(1 - \frac{1}{t}\right) \cdot w_{\hat{j},t-1}^{\phi} + \frac{1}{t} \cdot \sum_{\mathrm{i} | \Phi_{\mathrm{i},t-1}^{A} = \hat{j}} \mu_{\mathrm{i}\hat{j}t}^{\phi}$$

Si el clasificador aplicó la acción ϕ en el mensaje j con al menos un jugador i en el turno t. En caso contrario, $w_{\hat{j}t}^{\phi} = w_{\hat{j},t-1}^{\phi}$.

$\mu_{\mathrm{i}\hat{j}t}^{\phi}$ corresponde a la utilidad que obtiene en el turno t el clasificador, al aplicar la acción ϕ sobre el jugador i, cuya acción fue el mensaje $\hat{j}$.

En otras palabras, el clasificador guarda en su matriz W_t los resultados acumulados de jugar su acción $\phi \in \{0, 1\}$ (columnas) en cada mensaje $j \in K$ (filas) a lo largo del juego. Para esto, en cada nuevo turno t revisa si aplicó la acción ϕ con algún adversario en el mensaje j. De ser así, aplica la regla de actualización anterior. En caso contrario, mantiene el valor que la matriz ya tenía guardado.

[2] Con esta definición de α_{jt}^{i} se captura mejor el historial que construye el jugador sobre cada acción tomada, independiente de cuándo la ha jugado. Por un lado, por ejemplo, si ha jugado muchas veces la misma acción con buenos resultados, y a la siguiente oportunidad que la juega, es detenido por el clasificador, puede ignorar (hasta cierto punto) este mal resultado en función del historial que ha tenido, y considerar que fue una anomalía. Por otro lado, si es la primera vez que juega cierta acción $\hat{j}$, entonces $\alpha_{\hat{j}t}^{\mathrm{i}} = 1$, independiente del momento t en que la juegue, ya que aún no ha construido un historial de los resultados con esa jugada.

2.2.2. Dinámica del Juego

1. En $t = 0$, la naturaleza asigna el tipo θ_{i}^A de cada adversario $\mathrm{i} \in \{1, \ldots, N\}$, según la distribución de probabilidad $p\left(\theta_{\mathrm{i}}^A\right)$, donde $p\left(\theta_{\mathrm{i}}^A = (M, \overline{j})\right)$, $p\left(\theta_{\mathrm{i}}^A = (R, \overline{j})\right) > 0$, con $\sum_{\theta_{\mathrm{i}}^A} p(\theta_{\mathrm{i}}^A) = 1, \ \forall \overline{j} \in \{1, ..., k\}$.

   Los tipos son información privada de cada adversario y se mantienen constantes a lo largo de todo el juego.

2. Los adversarios observan su tipo y juegan, en cada turno t, su estrategia:

   (a) Adversarios regulares: para todo adversario de tipo regular, la estrategia será jugar $\sigma_t^A = 1$ en su componente $\overline{j}$ y $\sigma_t^A = 0$ en el resto, $\forall t$.

   Es decir, este tipo de adversario juega siempre la acción de utilidad máxima que le fue asignada por la naturaleza en $t = 0$.

   (b) Adversarios maliciosos: cada adversario i de tipo malicioso juega la estrategia $\sigma_{\mathrm{i}t}^C$ escogida al final del turno anterior.

   Para $t = 1$, la estrategia $\sigma_{\mathrm{i}1}^A$ corresponde a una distribución uniforme sobre todos los posibles mensajes, $j \in K$.

   (c) Clasificador: De manera simultánea a los adversarios, el clasificador juega la estrategia σ_t^C escogida al final del turno anterior.

   Para $t = 1$, la estrategia σ_1^C corresponde a una distribución uniforme por cada mensaje $j \in K$, sobre sus dos posibles acciones, $\phi \in \{0, 1\}$.

3. Cada adversario observa la utilidad recibida en este turno, $u_{\hat{j}t}^{\mathrm{i}}$.

   En otras palabras, cada adversario puede observar la utilidad que recibe por haber jugado la acción $\hat{j}$ en el turno t, pero no puede observar la estrategia del clasificador, ni puede tampoco saber cómo le habría ido de haber jugado otra estrategia (u otra acción), ya que no puede observar las utilidades recibidas por los otros adversarios ni sus estrategias (o acciones).

   Por su parte, el clasificador no puede observar las estrategias jugadas por los adversarios ni sus acciones; solo puede observar la utilidad que recibió por cada acción tomada por él ($\phi \in \{0, 1\}$) en cada mensaje $j \in K$. Es decir, el clasificador solo observa:

$$\mu_{\hat{j}t}^{\phi} = \sum_{\mathrm{i} | \Phi_{\mathrm{i},t-1}^A = \hat{j}} \mu_{\mathrm{i}\hat{j}t}^{\phi}$$

4. Los adversarios maliciosos y el clasificador actualizan respectivamente sus matrices V_t^{i} y W_t utilizando la información observada, bajo las reglas descritas en la sección 2.2.1., punto 3.

5. Los adversarios maliciosos y el clasificador escogen su estrategia del turno siguiente, $t+1$:[3]

   (a) Adversarios maliciosos: cada adversario de tipo malicioso consulta su matriz V_t^{i} de guardado de la información, y elige su estrategia mixta del turno siguiente, $\sigma^A_{\mathrm{i},t+1}$, según la regla:

   $$\sigma^A_{\mathrm{i},t+1,\hat{j}} = \mathbb{P}\left(S^A_{\mathrm{i},t+1} = \hat{j}\right) = \frac{\exp(v^{\mathrm{i}}_{\hat{j}t})}{\sum_{j=1}^{k}\exp(v^{\mathrm{i}}_{jt})}.$$

   (b) Clasificador: de manera simultánea a los adversarios, el clasificador consulta su matriz W_t de guardado de la información, y elige su estrategia mixta del turno siguiente, σ^C_{t+1}, según la regla:

   $$\sigma^C_{t+1,\hat{j}} = \mathbb{P}\left(S^C_{t+1,\hat{j}} = 1\right) = \frac{\exp(w^1_{\hat{j}t})}{\exp(w^0_{\hat{j}t})) + \exp(w^1_{\hat{j}t}))}$$

   Las matrices V_t^{i} y W_t se inicializan en $t=0$ como matrices nulas, por lo que en $t=1$, las estrategias mixtas de los adversarios maliciosos y del clasificador se inicializan como una distribución uniforme.

   Estas definiciones de σ_t^A y σ_t^C a partir de una función logística, permiten que los adversarios maliciosos y el clasificador, respectivamente, jueguen con mayor probabilidad aquella acción que les ha reportado mayores utilidades acumuladas en el pasado, a la vez que dejan espacio (una probabilidad no nula) para probar acciones nuevas, y así burlar las creencias de su contrincante.

   De esta forma, se captura el hecho de que ni los adversarios ni el clasificador pueden observar realmente las jugadas del otro, y sin embargo pueden aprender mutuamente de su contrincante a lo largo del juego, mediante las utilidades que reciben en cada turno, hasta llegar al EBP del juego.

[3]En esta parte se introduce dinamismo al modelo de aprendizaje, al incorporar la actualización de las probabilidades a partir de un modelo Logit.

2.2.3. Ejemplo: juego con aprendizaje basado en la experiencia

Para ejemplificar y asentar las características del juego recientemente descrito, se desarrollará en esta sección una versión simple del juego, en solo los primeros $t = 4$ turnos, explorando los distintos resultados que pueden obtener los jugadores, y sus implicancias tanto en su aprendizaje como en las interacciones entre ellos.

Se toma un juego con $N = 4$ jugadores y $k = 2$ mensajes. Los tipos de cada adversario son los siguientes:

$$\begin{aligned} \theta_1^A &= (M, 1) \\ \theta_2^A &= (R, 1) \\ \theta_3^A &= (R, 2) \\ \theta_4^A &= (M, 2) \end{aligned}$$

Es decir, en este juego hay $n_M = 2$ adversarios de tipo malicioso, y $n_R = 2$ de tipo regular. Durante el juego solo se analizarán las utilidades y estrategias del adversario 1 (malicioso) y del clasificador.

Los pagos de los adversarios maliciosos y del clasificador son:

$$\begin{aligned} b_1 &= 10 \qquad & L_R &= 10 \\ b_2 &= 5 \qquad & L_M &= 8 \\ c &= 8 \end{aligned}$$

Como una consideración previa, y con el objetivo de hacer más ilustrativo este ejemplo, se define una matriz de pagos del clasificador, con las mismas dimensiones de la matriz de información W_t:

$$\Psi_t = \begin{pmatrix} \mu_{1t}^0 & \mu_{1t}^1 \\ \mu_{2t}^0 & w_{2t}^1 \\ \vdots & \vdots \\ \mu_{kt}^0 & \mu_{kt}^1 \end{pmatrix}$$

Donde μ_{kt}^{ϕ} corresponde a la suma de utilidad recibida por el clasificador por cada adversario i al que le aplicó la acción ϕ sobre el mensaje j, en cada turno t.

Si en algún turno t no hubo ningún adversario que recibió la acción ϕ por parte del clasificador, en el mensaje j, entonces $\mu_{jt}^{\phi} = w_{j,t-1}^{\phi}$. De esta forma, siempre la matriz de información del clasificador se actualiza de la siguiente manera:

$$w_{jt}^{\phi} = \left(1 - \frac{1}{t}\right) \cdot w_{j,t-1}^{\phi} + \frac{1}{t} \cdot \mu_{jt}^{\phi} \ , \ \forall j \in K, \ \forall \phi \in \{0, 1\}, \ \forall t.$$

Primer turno, $t = 1$

1. Jugadas iniciales:[4]

$$\sigma_1^A = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix} \qquad \sigma_1^C = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix}$$

2. Acciones tomadas:[5]

$$\Phi_1^A = (1,\ 1,\ 2,\ 2) \qquad \Phi_1^C = (1,\ 0,\ 1,\ 0)$$

   Es decir, luego de aplicar las probabilidades definidas por sus estrategias mixtas, todos los adversarios juegan su mensaje preferido.

   Por su parte, el clasificador, luego de aplicar las probabilidades definidas por su estrategia mixta, permite los mensajes enviados por los Adversarios 1 y 3, y detiene los mensajes enviados por los adversarios 2 y 4.

3. Utilidades: la utilidad recibida por el adversario malicioso es $u_{11}^1 = 10$. La utilidad que recibe el clasificador por cada acción, corresponde a:

$$\Psi_1 = \begin{pmatrix} -10 & -8 \\ 0 & 0 \end{pmatrix}$$

4. Actualización de la información:

   (a) Adversario 1:

$$V_1^1 = \begin{pmatrix} v_{11}^1 \\ v_{21}^1 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{1}{1+0}\right) \cdot 0 + \frac{1}{1+0} \cdot 10 \\ 0 \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

   (b) Clasificador:

$$W_1 = \begin{pmatrix} w_{11}^0 & w_{11}^1 \\ w_{21}^0 & w_{21}^1 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{1}{1}\right) \cdot 0 + \frac{1}{1} \cdot -10 & \left(1 - \frac{1}{1}\right) \cdot 0 + \frac{1}{1} \cdot -8 \\ \left(1 - \frac{1}{1}\right) \cdot 0 + \frac{1}{1} \cdot 0 & \left(1 - \frac{1}{1}\right) \cdot 0 + \frac{1}{1} \cdot 0 \end{pmatrix} = \begin{pmatrix} -10 & -8 \\ 0 & 0 \end{pmatrix}$$

[4] Cabe recordar que en $t = 1$, las estrategias se inicializan como una distribución uniforme.

[5] A pesar de que los jugadores no pueden observar las estrategias ni acciones de sus contrincantes, se explicitarán en este ejemplo las acciones en las que se concretan las estrategias tomadas por cada jugador, para entender el origen de las utilidades percibidas, los cambios aplicados por cada jugador en sus estrategias, y el análisis a partir de estos resultados.

5. Estrategias: En este punto, los jugadores eligen la estrategia del siguiente periodo, basados en los resultados acumulados del pasado (resumido en sus matrices de información):

   (a) Adversario 1:

$$\sigma_2^A = \begin{pmatrix} \frac{\exp\left(v_{11}^1\right)}{\exp\left(v_{11}^1\right)+\exp\left(v_{21}^1\right)} \\ \frac{\exp\left(v_{21}^1\right)}{\exp\left(v_{11}^1\right)+\exp\left(v_{21}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(10)}{\exp(10)+\exp(0)} \\ \frac{\exp(0)}{\exp(10)+\exp(0)} \end{pmatrix} = \begin{pmatrix} 0{,}9999 \\ 0{,}0001 \end{pmatrix}$$

   (b) Clasificador:

$$\sigma_2^C = \begin{pmatrix} \frac{\exp\left(w_{11}^1\right)}{\exp\left(w_{11}^0\right)+\exp\left(w_{11}^1\right)} \\ \frac{\exp\left(w_{21}^1\right)}{\exp\left(w_{21}^0\right)+\exp\left(w_{21}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(-8)}{\exp(-10)+\exp(-8)} \\ \frac{\exp(0)}{\exp(0)+\exp(0)} \end{pmatrix} = \begin{pmatrix} 0{,}8808 \\ 0{,}5 \end{pmatrix}$$

En primer lugar, es fácil observar que la primera vez que juega cada mensaje $\hat{j}$ un Adversario de tipo Malicioso, siempre tendrá mucho valor en su estrategia siguiente, ya que como no tiene una historia construida, toda la información que guarda corresponde a la utilidad recibida en ese turno, para bien o para mal.

El clasificador, por su parte, clasifica mal al adversario 1, al permitir que realice su acción. Sin embargo, actualiza su estrategia aumentando la probabilidad con la que permite el mensaje 1. Esto se explica porque el Clasificador además comete un error al detener al adversario 2, que es de tipo regular, lo cual se castiga con mayor pérdida que el error relacionado al jugador 1. Por esta razón mueve su estrategia en la dirección de evitar volver a cometer el error de detener al jugador 2, sin haber observado realmente a los jugadores.

Para el mensaje 2, el clasificador detiene el mensaje del adversario 4 (malicioso), y permite el del adversario 3 (regular). Como obtiene un resultado óptimo en este turno, mantiene su estrategia para el siguiente.

Segundo turno, $t = 2$

1. Jugadas Iniciales:

$$\sigma_2^A = \begin{pmatrix} 0{,}9999 \\ 0{,}0001 \end{pmatrix} \qquad \sigma_2^C = \begin{pmatrix} 0{,}8808 \\ 0{,}5 \end{pmatrix}$$

2. Acciones tomadas:

$$\Phi_2^A = (1,\ 1,\ 2,\ 1) \qquad \Phi_2^C = (0,\ 1,\ 0,\ 0)$$

   El adversario 1 (malicioso) juega el mensaje 1. Por su parte, el clasificador detiene los mensajes enviados por los adversarios 1, 3 y 4, y permite el mensaje enviado por el adversario 2.

3. Utilidades: la utilidad del adversario 1 es $u_{12}^1 = -8$. La utilidad del clasificador por cada acción tomada, corresponde a:

$$\Psi_2 = \begin{pmatrix} 0 & 0 \\ -10 & 0 \end{pmatrix}$$

4. Actualización de la información:

   (a) Adversario 1:

$$V_2^1 = \begin{pmatrix} v_{12}^1 \\ v_{22}^1 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{1}{1+1}\right) \cdot 10 + \frac{1}{1+1} \cdot -8 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

   (b) Clasificador:

$$W_2 = \begin{pmatrix} w_{12}^0 & w_{12}^1 \\ w_{22}^0 & w_{22}^1 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{1}{2}\right) \cdot -10 + \frac{1}{2} \cdot 0 & \left(1 - \frac{1}{2}\right) \cdot -8 + \frac{1}{2} \cdot 0 \\ \left(1 - \frac{1}{2}\right) \cdot 0 + \frac{1}{2} \cdot -10 & 0 \end{pmatrix} = \begin{pmatrix} -5 & -4 \\ -5 & 0 \end{pmatrix}$$

5. Estrategias: en este punto, los jugadores eligen la estrategia del siguiente periodo, basados en sus matrices de información:

   (a) Adversario malicioso:

$$\sigma_3^A = \begin{pmatrix} \frac{\exp\left(v_{12}^1\right)}{\exp\left(v_{12}^1\right) + \exp\left(v_{22}^1\right)} \\ \frac{\exp\left(v_{22}^1\right)}{\exp\left(v_{12}^1\right) + \exp\left(v_{22}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(1)}{\exp(1) + \exp(0)} \\ \frac{\exp(0)}{\exp(1) + \exp(0)} \end{pmatrix} = \begin{pmatrix} 0{,}7311 \\ 0{,}2689 \end{pmatrix}$$

   (b) Clasificador:

$$\sigma_3^C = \begin{pmatrix} \frac{\exp\left(w_{12}^1\right)}{\exp\left(w_{12}^0\right) + \exp\left(w_{12}^1\right)} \\ \frac{\exp\left(w_{22}^1\right)}{\exp\left(w_{22}^0\right) + \exp\left(w_{22}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(-4)}{\exp(-5) + \exp(-4)} \\ \frac{\exp(0)}{\exp(-5) + \exp(0)} \end{pmatrix} = \begin{pmatrix} 0{,}7311 \\ 0{,}9933 \end{pmatrix}$$

En este turno, el adversario 1 ve detenida su acción por el clasificador, por lo que adapta su estrategia disminuyendo la probabilidad con la que jugar el mensaje 1 nuevamente. Se puede notar que el cambio no es tan extremo como en $t = 1$, lo que se debe a que el adversario ya tiene una historia de éxito guardada para $j = 1$, por lo que el efecto de haber perdido utilidad repitiendo la acción, se ve compensado positivamente por lo que ocurrió en $t = 1$.

El clasificador, por su parte, falla en detener la acción en $j = 2$, por lo que adapta su estrategia aumentando la probabilidad de permitir ese mensaje. Por otro lado, detiene (sin saberlo) a los Adversarios de tipo Malicioso en $j = 1$, y observa los resultados como un aumento en su utilidad (relativo a lo que ocurrió en $t = 1$). Por lo tanto, reacciona disminuyendo su probabilidad de *permitir* los mensajes enviados en $j = 1$.

Tercer turno, $t = 3$

1. Jugadas iniciales:

$$\sigma_3^A = \begin{pmatrix} 0{,}7311 \\ 0{,}2689 \end{pmatrix} \qquad \sigma_3^C = \begin{pmatrix} 0{,}7311 \\ 0{,}9933 \end{pmatrix}$$

2. Acciones tomadas:

$$\Phi_3^A = (2,\ 1,\ 2,\ 1) \qquad \Phi_3^C = (1,\ 1,\ 1,\ 1)$$

   El adversario 1 juega el mensaje 2, mientras que el clasificador permite los mensajes enviados por todos los adversarios.

3. Utilidades: la utilidad del adversario 1 es $u_{23}^1 = 5$. La utilidad del clasificador por cada acción tomada, corresponde a:[6]

$$\Psi_3 = \begin{pmatrix} -5 & -8 \\ -5 & -8 \end{pmatrix}$$

4. Actualización de la información:

   (a) Adversario 1:

$$V_3^1 = \begin{pmatrix} v_{13}^1 \\ v_{23}^1 \end{pmatrix} = \begin{pmatrix} 1 \\ \left(1 - \frac{1}{1+0}\right) \cdot 0 + \frac{1}{1+0} \cdot 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

   (b) Clasificador:

$$W_3 = \begin{pmatrix} w_{13}^0 & w_{13}^1 \\ w_{23}^0 & w_{23}^1 \end{pmatrix} = \begin{pmatrix} -5 & \left(1 - \frac{1}{3}\right) \cdot -4 + \frac{1}{3} \cdot -8 \\ \left(1 - \frac{1}{3}\right) \cdot -5 + \frac{1}{3} \cdot 0 & \left(1 - \frac{1}{3}\right) \cdot 0 + \frac{1}{3} \cdot -8 \end{pmatrix} = \begin{pmatrix} -5 & \frac{-16}{3} \\ \frac{-10}{3} & \frac{-8}{3} \end{pmatrix}$$

5. Estrategias: en este punto, los jugadores eligen la estrategia del siguiente periodo, basados en sus matrices de información:

   (a) Adversario 1:

$$\sigma_4^A = \begin{pmatrix} \frac{\exp\left(v_{13}^1\right)}{\exp\left(v_{13}^1\right) + \exp\left(v_{23}^1\right)} \\ \frac{\exp\left(v_{23}^1\right)}{\exp\left(v_{13}^1\right) + \exp\left(v_{23}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(1)}{\exp(1) + \exp(5)} \\ \frac{\exp(5)}{\exp(1) + \exp(5)} \end{pmatrix} = \begin{pmatrix} 0{,}0179 \\ 0{,}9820 \end{pmatrix}$$

[6] Cabe recordar que si el clasificador no juega cierta acción ϕ para el mensaje j, entonces esa coordenada de la matriz Ψ_t se mantiene con el valor que tenía su correspondiente en W_{t-1}, del turno pasado. En este caso el clasificador solo juega permitir ($\phi = 1$), por lo que la primera columna (correspondiente a $\phi = 0$) mantiene los valores que tenía en W_2.

(b) Clasificador:

$$\sigma_4^C = \begin{pmatrix} \frac{\exp\left(w_{13}^1\right)}{\exp\left(w_{13}^0\right)+\exp\left(w_{13}^1\right)} \\ \frac{\exp\left(w_{23}^1\right)}{\exp\left(w_{23}^0\right)+\exp\left(w_{23}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp\left(\frac{-16}{3}\right)}{\exp\left(\frac{-16}{3}\right)+\exp(-5)} \\ \frac{\exp\left(\frac{-8}{3}\right)}{\exp\left(\frac{-8}{3}\right)+\exp\left(\frac{-10}{3}\right)} \end{pmatrix} = \begin{pmatrix} 0{,}4174 \\ 0{,}6607 \end{pmatrix}$$

En este turno, el adversario 1 ve permitida su acción por el Clasificador, por lo que ajusta su estrategia aumentando la probabilidad de volver a jugar el mensaje $j = 2$ en el turno siguiente.

El clasificador, por su parte, permite la acción de todos los Adversarios. De esta forma, percibe que su utilidad asociada a los dos mensajes disminuye, producto de que los adversarios maliciosos habían jugado en ambos (aunque el clasificador no lo sepa), y reacciona bajando la probabilidad de permitirlos en el siguiente turno.

Cuarto turno, $t = 4$

1. Jugadas iniciales:

$$\sigma_4^A = \begin{pmatrix} 0{,}0179 \\ 0{,}9820 \end{pmatrix} \qquad \sigma_4^C = \begin{pmatrix} 0{,}4174 \\ 0{,}6607 \end{pmatrix}$$

2. Acciones tomadas:

$$\Phi_4^A = (2,\ 1,\ 2,\ 1) \qquad \Phi_3^C = (0,\ 1,\ 1,\ 0)$$

El adversario 1 juega el mensaje 2, mientras que el clasificador detiene su mensaje y el del adversario 4, y permite los mensajes enviados por los adversarios 2 y 3.

3. Utilidades: la utilidad del adversario 1 es $u_{24}^1 = -8$. La utilidad del clasificador por cada acción tomada, corresponde a:

$$\Psi_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

4. Actualización de la información:

(a) Adversario 1:

$$V_4^1 = \begin{pmatrix} v_{14}^1 \\ v_{24}^1 \end{pmatrix} = \begin{pmatrix} 1 \\ \left(1 - \frac{1}{1+1}\right) \cdot 5 + \frac{1}{1+1} \cdot -8 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{-3}{2} \end{pmatrix}$$

(b) Clasificador:

$$W_4 = \begin{pmatrix} w_{14}^0 & w_{14}^1 \\ w_{24}^0 & w_{24}^1 \end{pmatrix} = \begin{pmatrix} \left(1 - \frac{1}{4}\right) \cdot -5 + \frac{1}{4} \cdot 0 & \left(1 - \frac{1}{4}\right) \cdot \frac{-16}{3} + \frac{1}{4} \cdot 0 \\ \left(1 - \frac{1}{4}\right) \cdot \frac{-10}{3} + \frac{1}{4} \cdot 0 & \left(1 - \frac{1}{4}\right) \cdot \frac{-8}{3} + \frac{1}{4} \cdot 0 \end{pmatrix} = \begin{pmatrix} \frac{-15}{4} & -4 \\ \frac{-10}{4} & -2 \end{pmatrix}$$

5. Estrategias: en este punto, los jugadores eligen la estrategia del siguiente periodo, basados en sus matrices de información:

   (a) Adversario 1:

$$\sigma_5^A = \begin{pmatrix} \frac{\exp\left(v_{14}^1\right)}{\exp\left(v_{14}^1\right)+\exp\left(v_{24}^1\right)} \\ \frac{\exp\left(v_{24}^1\right)}{\exp\left(v_{14}^1\right)+\exp\left(v_{24}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(1)}{\exp(1)+\exp\left(\frac{-3}{2}\right)} \\ \frac{\exp\left(\frac{-3}{2}\right)}{\exp(1)+\exp\left(\frac{-3}{2}\right)} \end{pmatrix} = \begin{pmatrix} 0{,}9241 \\ 0{,}0759 \end{pmatrix}$$

   (b) Clasificador:

$$\sigma_5^C = \begin{pmatrix} \frac{\exp\left(w_{14}^1\right)}{\exp\left(w_{14}^0\right)+\exp\left(w_{14}^1\right)} \\ \frac{\exp\left(w_{24}^1\right)}{\exp\left(w_{24}^0\right)+\exp\left(w_{24}^1\right)} \end{pmatrix} = \begin{pmatrix} \frac{\exp(-4)}{\exp(-4)+\exp\left(\frac{-15}{4}\right)} \\ \frac{\exp(-2)}{\exp(-2)+\exp\left(\frac{-10}{4}\right)} \end{pmatrix} = \begin{pmatrix} 0{,}4378 \\ 0{,}6224 \end{pmatrix}$$

   En este turno, el adversario 1 ve detenida su acción por el clasificador, por lo que ajusta su estrategia para aumentar la probabilidad de jugar $j = 1$ en el turno siguiente. El cambio radical de estrategia se explica por el hecho de que $c > b_2$. Es decir, los costos por no pasar el filtro del clasificador al no jugar su opción preferida, son mayores a las ganancias acumuladas hasta este punto del juego, donde ha ganado y perdido solo una vez respectivamente.

   Por su parte, el clasificador tuvo éxito en todas sus acciones, por lo que ajusta sus estrategias correspondientemente.

En este juego simple, de dos adversarios de tipo malicioso, dos opciones de mensaje y una ventana de tiempo de solo $t = 4$ turnos, se puede apreciar cómo los jugadores pueden ir adaptando sus estrategias a partir de su experiencia a lo largo del tiempo. De esta manera logran avanzar hacia el óptimo a través de ensayo y error.

Lo más llamativo de este modelo, es que ambos jugadores adaptan su estrategia completamente *ciegos* a la estrategia o las acciones de su contrincante. Sin embargo, son capaces de actualizar sus estrategias en la dirección correcta (de acuerdo a su racionalidad), solo dotados de información sobre sus propias utilidades, turno a turno.

Por esta razón es importante que el juego se repita por un largo periodo de tiempo (potencialmente infinito). De esta manera los jugadores pueden probar distintas estrategias, que les permitan aprender de los resultados buenos y malos que enfrentan en el transcurso del juego, y así adaptarse a su contrincante cada vez mejor.

En el siguiente capítulo se analiza lo que ocurre con un conjunto mayor de adversarios (de tipo malicioso y regular), que pueden tener distintos mensajes preferidos; en una ventana de tiempo lo suficientemente larga como para aprender y adaptarse a las jugadas de su contrincante.

Capítulo 3

Simulación

En este capítulo se simula el juego planteado en la parte teórica, con jugadores que interactúan y aprenden en tiempo real y de manera simultánea el uno del otro.[1]

En primer lugar se presentan los resultados de simular un caso con los parámetros definidos en valores estándar. Luego se presentan las diferencias que surgen al mover los valores de los parámetros clave para el desarrollo del juego: la proporción de adversarios maliciosos, el costo de los adversarios maliciosos por ver detenida su acción, y el costo del clasificador por cometer error de tipo II. Este último en proporción al costo por el error de tipo I.

Los valores de los parámetros que cambian en cada simulación se definen en la sección correspondiente a cada una.

En todas las simulaciones se tienen $N = 100$ adversarios y $k = 5$ opciones de mensajes. De esta manera es directo interpretar el número de adversarios maliciosos como una proporción del total de adversarios. Además el total de mensajes disponibles es suficiente para que exista una cantidad significativa de adversarios que prefieren cada uno de los mensajes.

Las preferencias de los adversarios sobre los mensajes son repartidas entre ellos a partir de distribuciones normales de varianza 1, con $\mu = 2$ para los adversarios de tipo malicioso, y $\mu = 4$ para los adversarios de tipo regular.

De esta manera, y sin pérdida de generalidad, las preferencias se concentran alrededor de $\overline{j} = 2$ para los adversarios maliciosos y de $\overline{j} = 4$ para los regulares. Permitiendo así, que existan adversarios de distinto tipo que coinciden en su mensaje preferido.

Los pagos para un adversario[2] se definen a partir de la *distancia* entre el mensaje jugado y el preferido. Esta distancia se obtiene ordenando de manera circular los mensajes disponibles.

[1] El código de la simulación en Matlab se encuentra en el Anexo 4.2.

[2] Los pagos y costos se definen para todos los adversarios. Sin embargo, como los adversarios de tipo regular no cambian el mensaje jugado turno a turno, en la práctica los pagos son solo relevantes para los adversarios de tipo malicioso.

Por ejemplo, si el mensaje preferido del adversario es $\bar{j} = 2$, entonces $j = 1$ y $j = 3$ están a una distancia 1; $j = 5$ y $j = 4$ están a una distancia 2. Los pagos en las simulaciones se definen de esta manera:

1. Pago por el mensaje preferido: $b_1 = 1$.
2. Pago por los mensajes a distancia 1: $b_2 = \frac{1}{2}$.
3. Pago por los mensajes a distancia 2: $b_3 = \frac{1}{3}$.

El costo para el clasificador por cometer error de tipo I se define como $L_R = 1$.

Se utiliza una ventana de tiempo $T = 1000$ para las iteraciones del juego.

Los resultados de las simulaciones se muestran de tres maneras:

1. Evolución temporal, a lo largo de toda la ventana de tiempo del juego.
2. Estado estacionario: se toma como periodo de estado estacionario, una ventana temporal que corresponde al 20 % final de los periodos totales del juego. En este periodo ya se observa un comportamiento en equilibrio del juego.
3. Medias móviles: para cada instante del periodo de estado estacionario, se muestra el promedio de las 100 iteraciones inmediatamente anteriores del juego. Esto se hace para suavizar los gráficos de utilidades y errores de tipo I y II, e interpretar de mejor manera sus resultados.

3.1. Simulación de caso estándar

En esta sección se presentan los principales resultados de simular un juego estándar. Para ello, se definen los parámetros clave del juego de la siguiente manera:

1. Cantidad de adversarios de tipo malicioso: $n_M = 50$. Es decir, corresponden al 50 % del total de adversarios.
2. Costo del adversario malicioso por ver detenida su acción: $c = 0{,}5$. Es decir, el costo se fija igual al ingreso por jugar una acción a un paso de la preferida: $c = b_2$.
3. Costo del clasificador por cometer error de tipo II: $L_M = 0{,}8$.

Estado inicial del juego

Las preferencias de los adversarios se definen a partir de la discretización de una distribución normal para cada tipo de adversario: para los adversarios regulares, la distribución es $\mathcal{N}(1, 4)$, y para los maliciosos, $\mathcal{N}(1, 2)$. El resultado se puede observar en la figura 3.1:

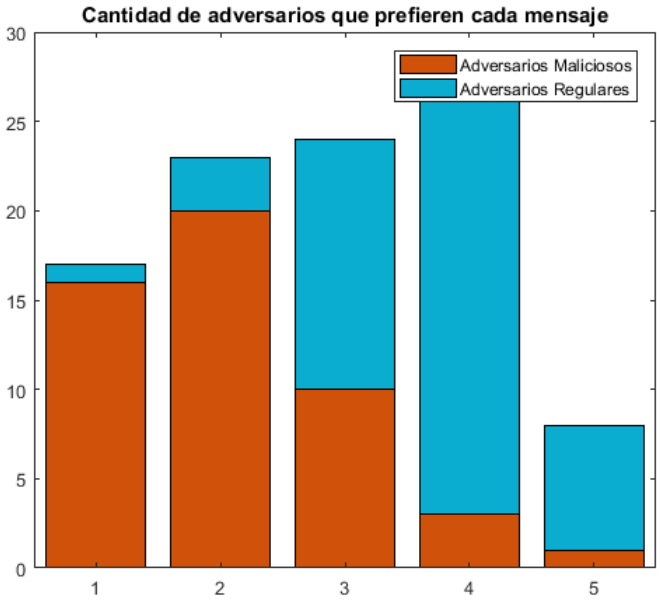


Figura 3.1: Preferencias de los adversarios.

Producto de la distribución utilizada, la preferencia de los adversarios maliciosos se concentra alrededor de $\overline{j} = 2$ y la de los regulares alrededor de $\overline{j} = 4$. Sin embargo, ambos tipos de adversarios coexisten en todos los mensajes.

3.1.1. Equilibrio

Adversarios maliciosos

En la figura 3.2 se presentan las estrategias seguidas por los adversarios maliciosos para cada uno de los mensajes en la parte final de la simulación, una vez alcanzado el equilibrio del juego.

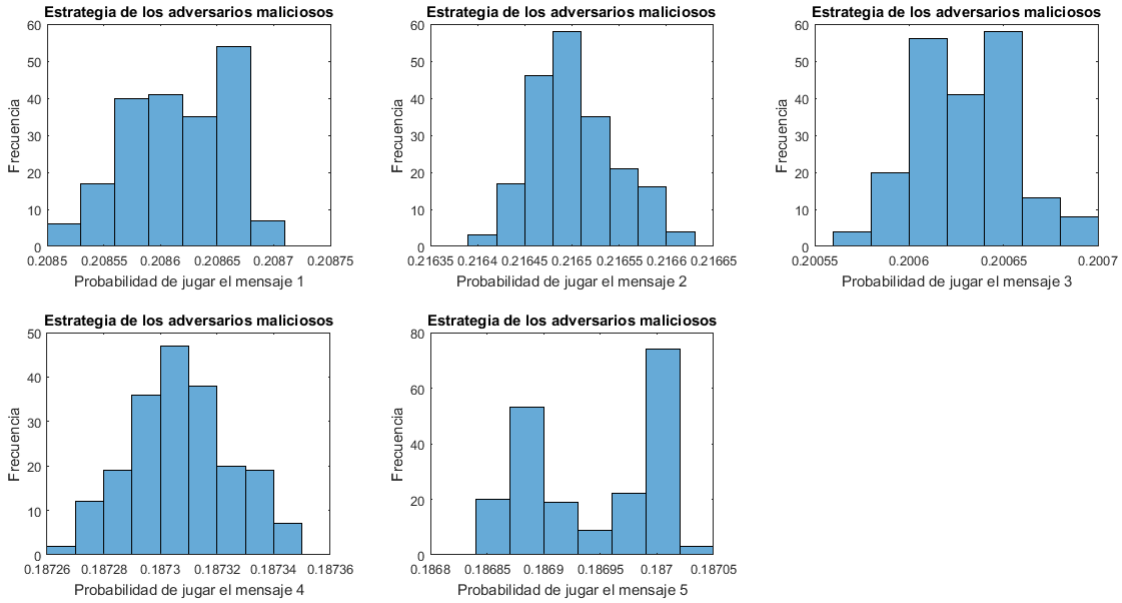


Figura 3.2: Distribución de las estrategias finales de los adversarios maliciosos.

En primer lugar, destaca la baja variabilidad en las estrategias de cada mensaje. A estas alturas del juego, si bien los adversarios siguen actualizando su información y sus estrategias, el rango en que estas se mueven no supera los $0,03$ *pp* en cada caso.

Tal como se observa en la figura 3.3, las estrategias de los adversarios maliciosos, como resultado de su proceso de aprendizaje, se acercan a una distribución uniforme.

Se puede inferir entonces, que en su afán por burlar al clasificador, los adversarios maliciosos buscan ser completamente impredecibles en sus estrategias, lo que los lleva a acercarse a una distribución uniforme.

Sin embargo, no pueden ser completamente uniformes, dadas las diferencias en los beneficios que reciben por cada mensaje, lo que explica el pequeño sesgo hacia sus preferencias iniciales.

Por último, en la figura 3.4 se observan las estrategias que siguen los adversarios maliciosos para cada uno de los mensajes, en el equilibrio.

Recordando que los adversarios regulares siguen una estrategia separadora, se puede inferir que el equilibrio de este juego es mixto:

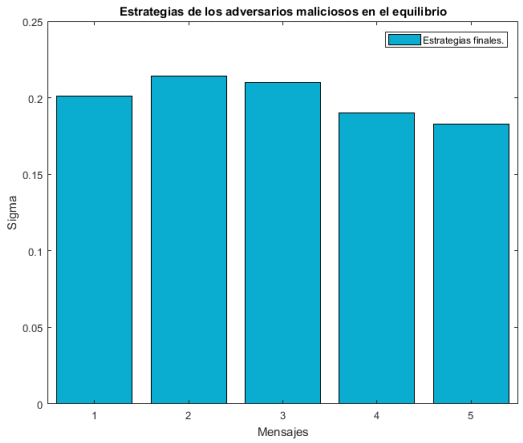


Figura 3.3: Estrategias finales de los adversarios maliciosos.

Para cada mensaje $\hat{j}$, las estrategias de equilibrio (promedio) de los adversarios, son las siguientes:

1. Adversarios de tipo $\left(R, \hat{j}\right)$: $\sigma^A(\hat{j}) = 1$.
2. Adversarios de tipo $\left(M, \hat{j}\right)$: $\sigma^A(\hat{j}) = 0{,}25$.
3. Adversarios de tipo $(R, \bar{j})$, con $\bar{j} \neq \hat{j}$: $\sigma^A(\hat{j}) = 0{,}187$.

Es decir, los adversarios de tipo regular juegan con 100 % de probabilidad su mensaje preferido, mientras que los de tipo malicioso juegan con un 25 % de probabilidad su mensaje preferido, y juegan de manera aleatoria uniforme el resto de los mensajes.

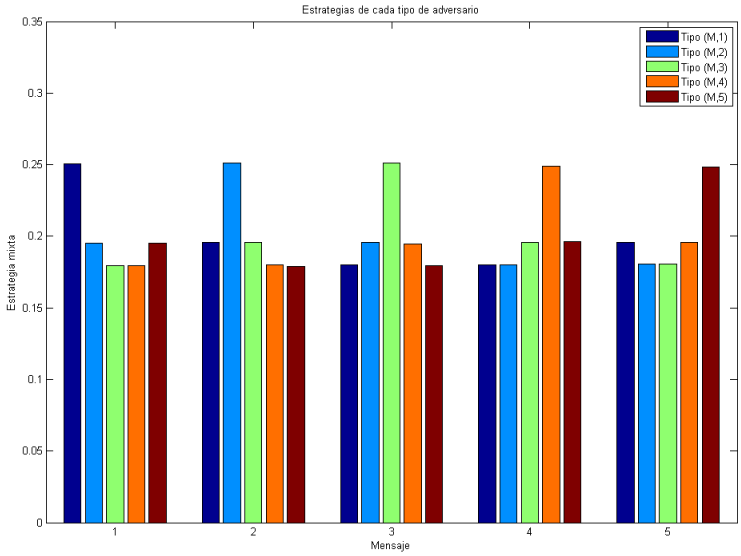


Figura 3.4: Estrategias de los adversarios maliciosos en el equilibrio.

Clasificador

En la figura 3.5 se presenta la evolución temporal de las estrategias seguidas por el clasificador para cada uno de los mensajes, a lo largo de toda la ventana de tiempo.

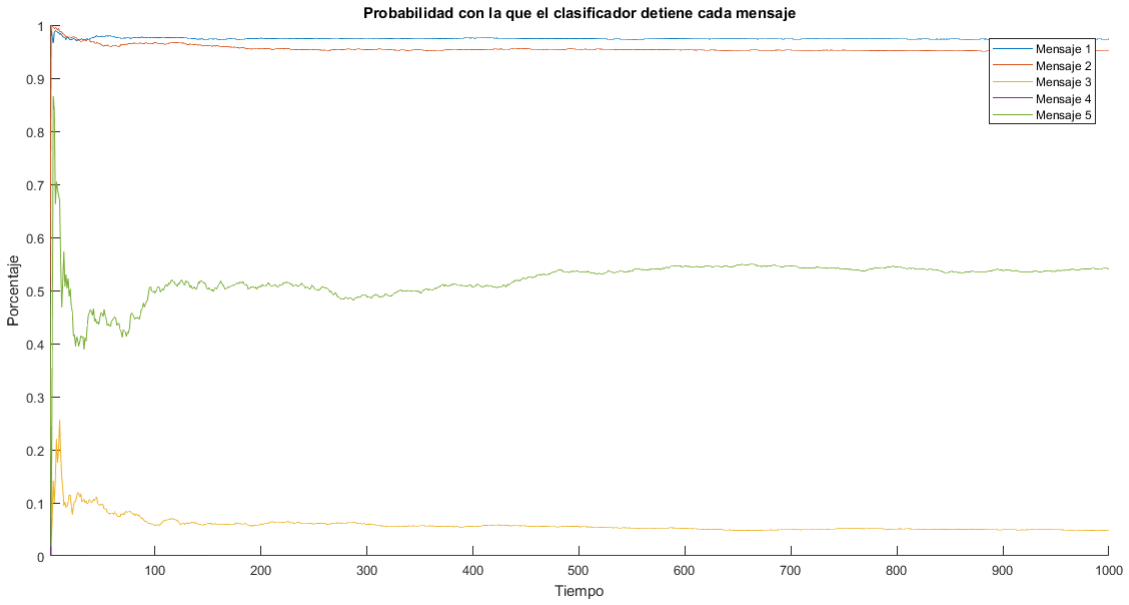


Figura 3.5: Estrategias del clasificador para cada mensaje.

En primer lugar, se puede observar que si bien existen pequeñas fluctuaciones a lo largo de todo el proceso, aproximadamente en $t = 40$ ya las estrategias se ubican en el orden de magnitud del equilibrio final, y desde ese momento solo se refinan, hasta que en $t = 500$ ya se estabilizan del todo.

Cabe recordar que la estrategia del clasificador se inicializa en $t = 1$ de manera uniforme, deteniendo cada mensaje con una probabilidad del 50 %. Es por esta razón que destaca la rapidez con la que el clasificador lleva sus estrategias a niveles diferentes para cada mensaje: en forma prácticamente inmediata, las estrategias quedan en un orden establecido, que se mantiene a lo largo de todo el juego.

Esto se debe a que el mecanismo de aprendizaje le permite distinguir rápidamente qué mensajes debe detener con más frecuencia que los otros, y adaptar su estrategia cen forma correspondiente.

Interacción

En cada opción de mensaje, el clasificador se encuentra con un *trade-off* importante: aumentar la probabilidad con la que detiene el mensaje, y así disminuir el error tipo II,

o disminuir la probabilidad, permitiendo que más adversarios regulares puedan enviar el mensaje, y disminuir así el error de tipo I.

El clasificador aprende a lo largo del juego a equilibrar estos costos, hasta converger a las estrategias finales para cada mensaje.

Esta situación es especialmente notoria en este caso, ya que provoca una separación evidente y prácticamente inmediata entre las estrategias seguidas para cada mensaje:
$\sigma_1^C > \sigma_2^C > \sigma_5^C > \sigma_3^C > \sigma_4^C$.

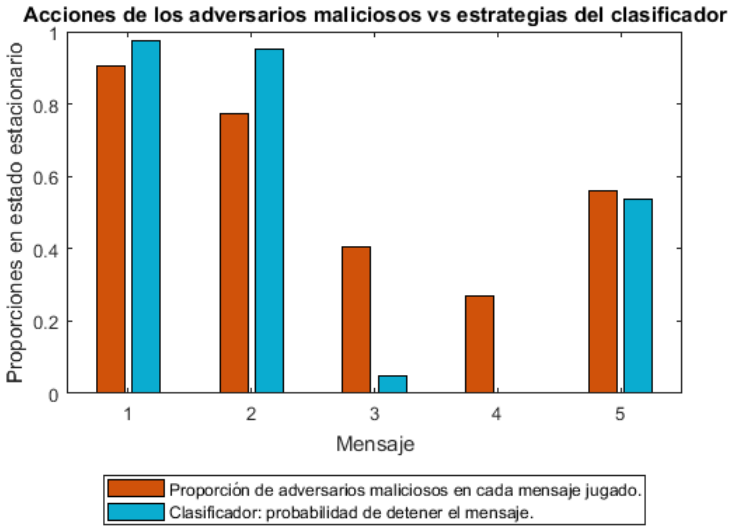


Figura 3.6: Proporción de adversarios en cada mensaje vs estrategias del clasificador.

En la figura 3.6 se muestra, del total de adversarios que juega cada mensaje en cada turno, la proporción promedio que corresponde a adversarios maliciosos, comparado con el promedio de la probabilidad con la que el clasificador detiene cada uno de los mensajes; en el estado estacionario del juego.

Destaca la manera en la que el clasificador, sin observar ni definir creencias respecto a las preferencias o a las jugadas de sus adversarios, es capaz de ajustar a ellas sus estrategias en cada mensaje:

1. Tal como se observa en la figura 3.1, la cantidad de adversarios regulares que prefieren el mensaje 1 es mínima. El clasificador, entonces, ajusta su estrategia hasta detener este mensaje con una probabilidad cercana al 100 %.
2. El mensaje 2 es el preferido por la mayor cantidad de adversarios maliciosos. En este caso, al igual que para el mensaje 1, el clasificador asigna una probabilidad alta de detención del mensaje, con el objetivo de detener a la mayor cantidad de maliciosos posible.
3. El mensaje 3 es preferido por una mayor cantidad de adversarios regulares, que lo juegan turno a turno, por lo que la estrategia del clasificador cae drásticamente, en comparación a sus estrategias para los primeros mensajes.

4. En el mensaje 4 se presenta la situación contraria a la del mensaje 1: aquí el clasificador no detiene los mensajes (la probabilidad que asigna es nula). Esto se debe a que, si bien existen adversarios maliciosos que juegan esta opción, este es el mensaje más jugado por los adversarios regulares. Por lo tanto es costoso para el clasificador detener los mensajes con una probabilidad mayor a cero, debido al alto costo del error tipo I.
5. Como se observa en la figura 3.6, la proporción de adversarios maliciosos que juega el mensaje 5 es similar a la del mensaje 3. Sin embargo la probabilidad con la que detiene este mensaje es mayor. Tal como se observa en la figura 3.1, la cantidad de adversarios regulares que juegan el mensaje 5 es menor que la que juega el mensaje 3. Es por este motivo que el clasificador se ajusta aumentando la probabilidad de detención.

En resumen, la estrategia final del clasificador en cada uno de los mensajes se ajusta a la proporción de adversarios maliciosos que juega cada uno. Además tiene en cuenta la cantidad de adversarios regulares en ellos: cuando la cantidad es menor, aumenta la probabilidad de detención del mensaje, y viceversa.

Este ajuste en las estrategias del clasificador se hace en total ausencia de creencias o información sobre las preferencias, o las jugadas de los adversarios; y corresponde a un proceso de aprendizaje que toma en cuenta solo la información privada de los costos y utilidades propias del clasificador.

3.1.2. Utilidades

En la figura 3.7 se observa la media móvil de las utilidades del clasificador y de la media de los adversarios maliciosos, para los últimos 200 periodos del juego (donde ya se ha alcanzado el equilibrio), y tomando como ventana móvil 100 periodos hacia el pasado.

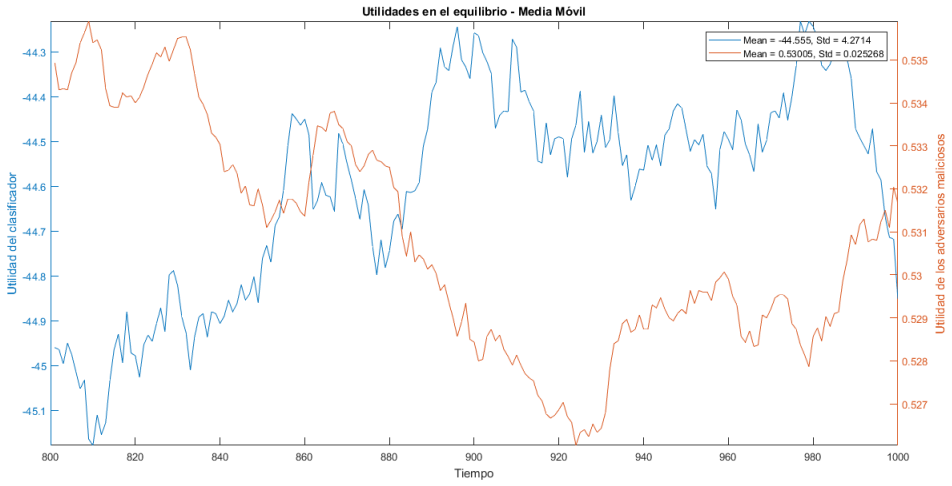


Figura 3.7: Utilidad del clasificador vs utilidad de los adversarios maliciosos (media móvil).

La utilidad media del clasificador en el equilibrio es de $-44{,}555$ y la de los adversarios maliciosos, 0,53. Estas utilidades se mantienen constantes en el equilibrio, lo que se demuestra en sus bajas desviaciones estándar, de valores 4,27 y 0,025 respectivamente.

La correlación entre la utilidad del clasificador y la de los adversarios maliciosos, es de $-0{,}73$. Es decir, existe cuando la utilidad del clasificador disminuye, la de los adversarios maliciosos aumenta, y viceversa.

Esto era de esperar, y se puede explicar con el comportamiento de *gato y ratón* que el clasificador y los adversarios maliciosos demuestran en este tipo de juegos. El clasificador persigue en forma constante a los adversarios maliciosos, y estos a su vez, adaptan su estrategia para ocultarse del clasificador e inducir el error. Esto provoca que aumente su utilidad y en consecuencia, disminuya la del clasificador.

3.1.3. Errores de clasificación

En la figura 3.8 se observa la media móvil de los errores de tipo I y de tipo II[3], para los últimos 200 periodos del juego (donde ya se ha alcanzado el equilibrio), y tomando como ventana móvil 100 periodos hacia el pasado.

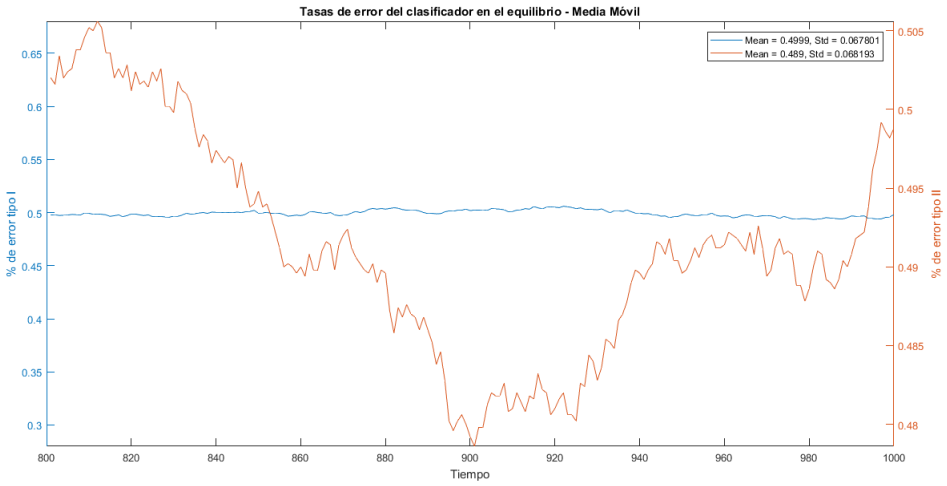


Figura 3.8: Errores tipo I y II en el equilibrio (media móvil).

A estas alturas del juego, una vez alcanzado el equilibrio, los niveles de error se mantienen con poca variabilidad alrededor del 50 %, con errores estándar de 0,067 y 0,068 para los errores de tipo I y II respectivamente.

Destaca en forma especial el nivel prácticamente constante del error de tipo I, de media 49,9 %, mientras que el error de tipo II fluctúa alrededor de 48,9 %.

Esto se explica en las diferentes estrategias que siguen los dos tipos de adversario. Mientras un adversario regular juega siempre el mismo mensaje, el adversario malicioso va cambiando su mensaje jugado, siguiendo las estrategias mixtas que actualiza cada turno, para camuflarse entre los adversarios regulares e inducir el error del clasificador, lo que provoca las fluctuaciones.

[3]El error de tipo I corresponde a que el clasificador se equivoque en la clasificación de los adversarios regulares, y el error de tipo II, a que se equivoque en clasificar a los adversarios maliciosos.

Es decir, el clasificador es capaz de distinguir entre las estrategias puras de los adversarios regulares, y las mixtas de los maliciosos, en el equilibrio; manteniendo constante el error asociado a los primeros, y oscilando con el segundo.

Por último, se puede apreciar que el alza y las fluctuaciones del error de tipo II en la figura 3.8 explican el comportamiento a la baja y fluctuante de la utilidad del clasificador en la figura 3.7, lo que refuerza el hecho de que el clasificador y los adversarios maliciosos tienen un comportamiento de gato y ratón en sus estrategias.

3.2. Variación de parámetros relevantes

En esta sección se presentan los cambios más importantes que surgen en los resultados del juego, al hacer variaciones en los parámetros clave:

1. En el caso estándar se analiza el caso en que la proporción de adversarios regulares y maliciosos es la misma. En esta sección se estudian los casos en que los adversarios de tipo malicioso son mayoría y minoría, respecto al total de adversarios.
2. En el caso estándar se define el costo del adversario malicioso como igual a b_2. En esta sección se analiza el caso en que el costo es más bajo (igual a b_3), y el caso extremo en que el costo es mayor a su máximo ingreso ($c > b_1$).
3. En el caso estándar se define el costo del clasificador por error de tipo II, como cercano al error de tipo I ($L_M = 0{,}8 \cdot L_R$). En esta sección se analiza el caso en que este valor es menor y, si bien se define en el modelo que $-L_M < -L_R < 0$, se estudia también el caso en que la desigualdad se revierte: $-L_R < -L_M < 0$.

3.2.1. Variaciones en la proporción de adversarios maliciosos

En esta sección se presentan los resultados más relevantes de alterar la proporción de adversarios maliciosos con los que se enfrenta el clasificador, con respecto a la simulación presentada en la sección 3.1.

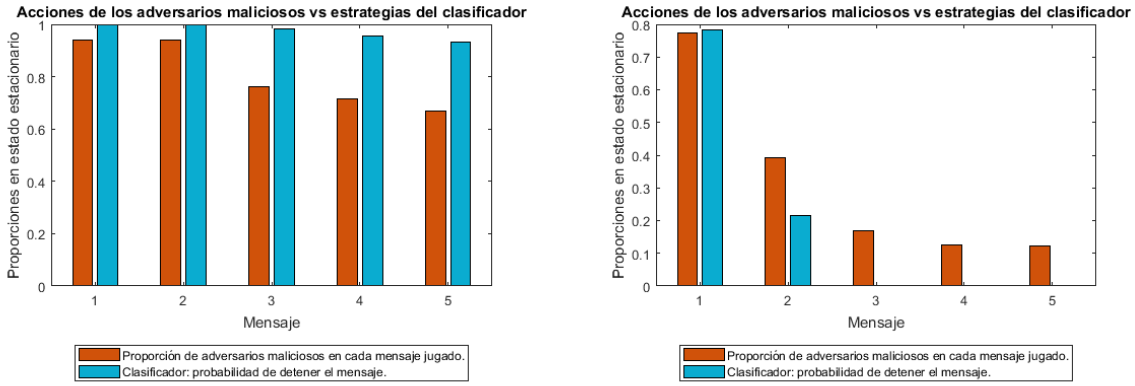
En primer lugar se aumenta la proporción a un 80 %, presentando un escenario donde el clasificador se enfrenta a adversarios que en su mayoría son maliciosos. En segundo lugar se disminuye la proporción a un 20 %, cambiando el escenario a uno donde los adversarios maliciosos son minoría.

En la figura 3.9 se presenta la proporción de adversarios que son de tipo malicioso entre todos los que juegan cada uno de los mensajes, en la fase de equilibrio del juego; comparada con la estrategia del clasificador (probabilidad de detener cada mensaje) para la misma ventana de tiempo. Se muestran estos resultados para los dos escenarios simulados.

Lo que más destaca de los resultados es la capacidad del clasificador de adaptarse a la proporción de adversarios maliciosos a los que se enfrenta, llegando a estrategias completamente diferentes en ambos escenarios.

Por un lado, cuando la mayoría de los adversarios son maliciosos, el clasificador reacciona aumentando la probabilidad de detención en todos los mensajes. Mientras que en el escenario opuesto, cuando los adversarios maliciosos son minoría, el clasificador disminuye drásticamente la probabilidad de detener cada mensaje.

En caso especial son los mensajes 1 y 2, que siempre son jugados en su mayoría por adversarios maliciosos. En estos casos el clasificador es capaz de aprender que la mejor estrategia es mantener una probabilidad de detención alta en estos mensajes.



(a) Mayoría de adversarios maliciosos.

(b) Minoría de adversarios maliciosos.

Figura 3.9: Estrategias del clasificador al variar la proporción de adversarios maliciosos.

En la figura 3.10 se presenta la evolución de las utilidades del clasificador y de los adversarios maliciosos (como media móvil) durante la fase de equilibrio del juego.

La utilidad promedio de los adversarios maliciosos se mantiene en los mismos niveles que en el caso estándar (0,53), solo aumentando ligeramente su desviación estándar (de 0,02 a 0,04) en el caso en que existe una minoría de adversarios maliciosos.

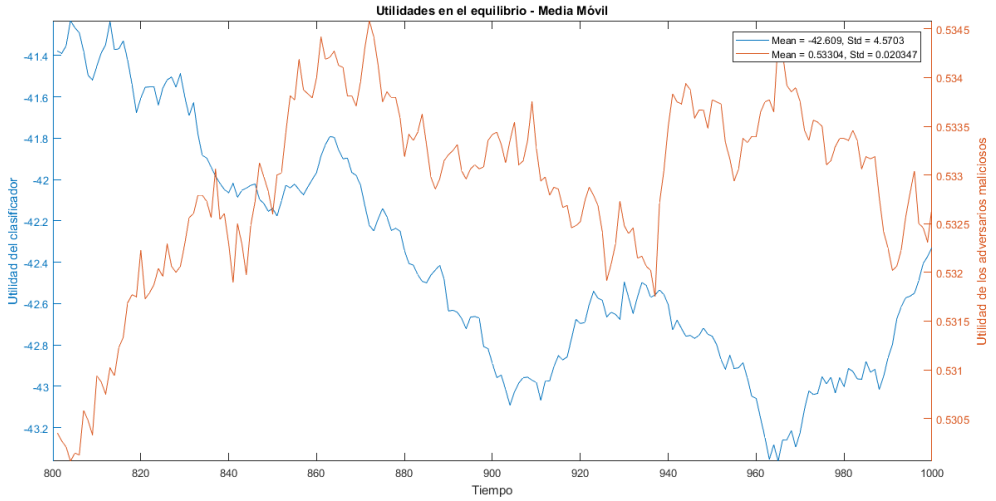
Por otro lado, la utilidad del clasificador aumenta (con respecto al caso estándar) cuando aumenta la proporción de adversarios maliciosos en el juego, y disminuye en el caso contrario.

Este fenómeno se debe a que cuando disminuye la cantidad de adversarios maliciosos, el clasificador se ve en la obligación de disminuir la probabilidad con que detiene cada mensaje, aumentando el error de tipo II. Esto ocurre porque el costo del error de tipo I es mayor, por lo que el clasificador prefiere disminuir su rendimiento. Para el caso opuesto el razonamiento es equivalente.

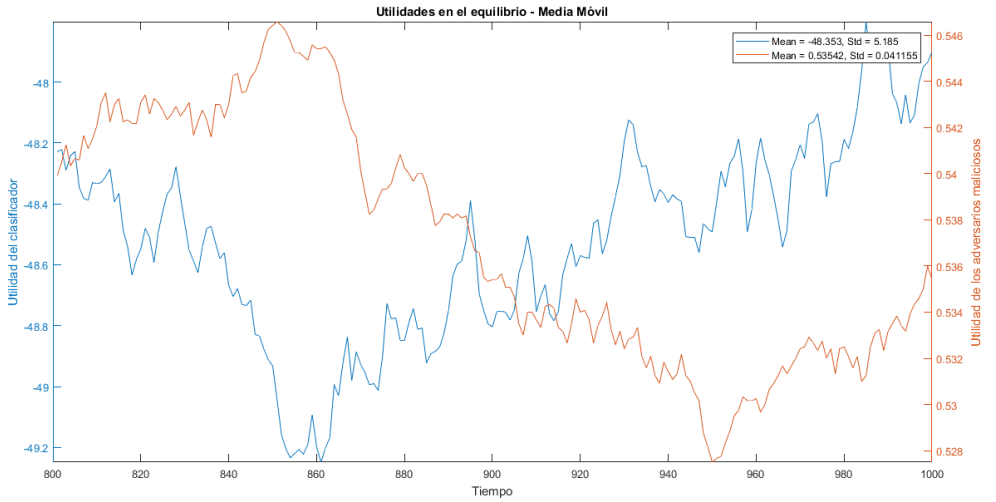
Lo que más destaca de estos resultados, es que para el caso en que los adversarios maliciosos son minoría, su utilidad presenta un comportamiento estrechamente ligado al del clasificador.

Esto se debe a que en este caso el clasificador debe disminuir su capacidad de filtrar a los adversarios maliciosos, para así favorecer la buena clasificación de los adversarios regulares.

Por lo tanto, un aumento (disminución) en la utilidad del clasificador se debe a una mejor (peor) clasificación de los adversarios regulares, y consecuentemente, a una peor (mejor) capacidad de clasificación de los adversarios maliciosos, quienes también ven aumentada (disminuida) su utilidad.



(a) Mayoría de adversarios maliciosos.



(b) Minoría de adversarios maliciosos.

Figura 3.10: Utilidades del clasificador y de los adversarios maliciosos (media móvil), al variar la proporción de adversarios maliciosos.

3.2.2. Variaciones en el costo de los adversarios maliciosos

En esta sección se presentan los resultados más relevantes de alterar el nivel de costo de los adversarios maliciosos, con respecto a la simulación presentada en la sección 3.1.

Por un lado, se sube el costo de los adversarios por sobre su máximo ingreso, a 1,2. Por otro, se disminuye hasta igualar el monto de mínimo ingreso del adversario al pasar al clasificador (0,25).

En la figura 3.11 se presentan las utilidades del clasificador y de los adversarios maliciosos (como media móvil), para la fase de equilibrio del juego.

Al comparar con el caso estándar, se observa que la utilidad media del clasificador no se ve afectada por los cambios en el costo de los adversarios.

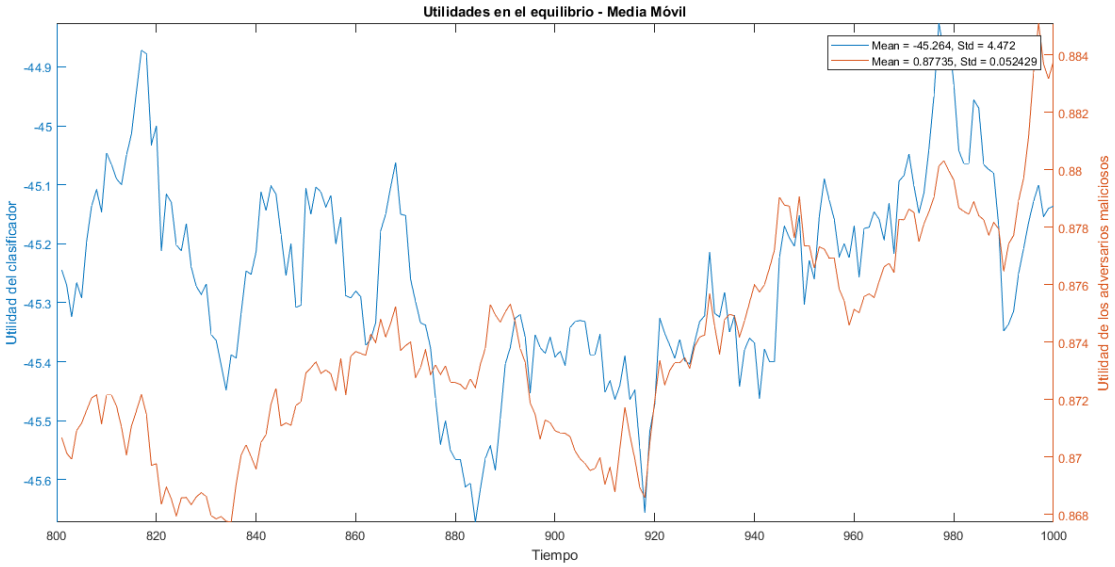
En cambio para los adversarios maliciosos el efecto es directo. Cuando el costo aumenta, su utilidad media también aumenta respecto al caso estándar (de $0,53$ a $0,87$). Por otro lado, cuando el costo es bajo, la utilidad disminuye (de $0,53$ a $0,41$).

Respecto a las tendencias de las utilidades en ambos escenarios, se observa que cuando el costo de los adversarios es mayor, su utilidad evoluciona de la misma manera que la utilidad del clasificador. Por el contrario, si el costo disminuye, la utilidad de los adversarios y del clasificador presentan comportamientos opuestos.

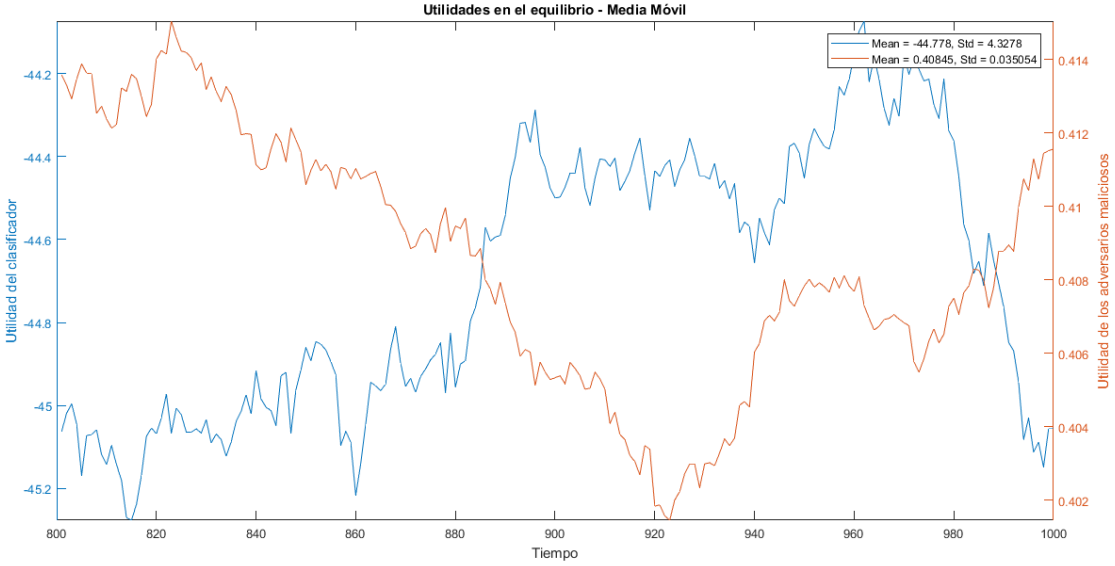
Estos resultados sugieren que a mayor costo del adversario, peor es el rendimiento del clasificador. Sin embargo, la utilidad y las estrategias del clasificador no se ven afectadas por este parámetro (ver anexo 4.1.2. para más detalles).

Por lo tanto, se infiere que los costos de los adversarios, de manera agregada, no presentan mayores efectos en los resultados del juego.

El efecto sobre las utilidades de los adversarios se debe de manera directa al cambio en el valor del costo, y no a una influencia de este sobre las estrategias o la interacción entre los jugadores.



(a) Costo mayor al ingreso máximo.



(b) Costo bajo.

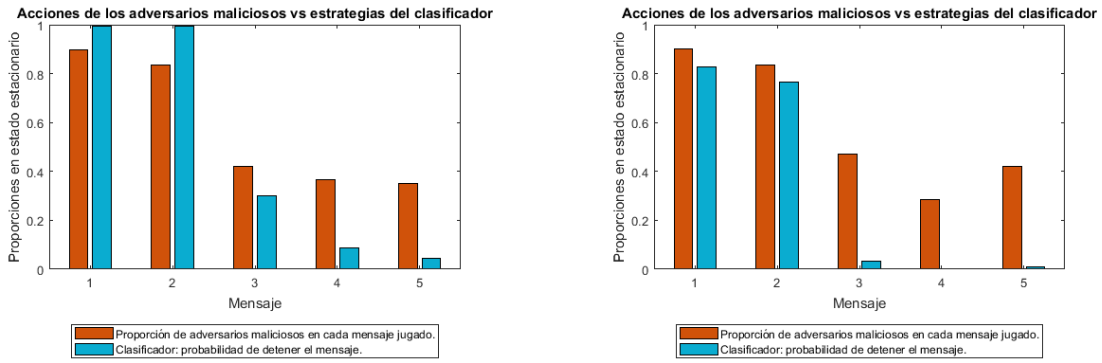
Figura 3.11: Utilidades del clasificador y de los adversarios maliciosos (media móvil), al variar el costo de los adversarios.

3.2.3. Variaciones en el costo del clasificador

En esta sección se modifica la relación entre los costos del clasificador. Se presentan los principales efectos sobre los resultados del juego, comparándolos con los resultados de la simulación del caso estándar en la sección 3.1.

Para cambiar la relación entre los costos del clasificador, se deja fijo (igual a 1) el costo por el error de tipo I y se altera el costo por el error de tipo II. Por un lado se aumenta hasta dejarlo por sobre el costo del error de tipo I (igual a $1,2$). Por otro, se disminuye su valor a $0,4$.

En la figura 3.12 se presenta, para la fase de equilibrio del juego, la estrategia del clasificador para cada uno de los mensajes; comparada con la proporción de adversarios que son maliciosos de entre quienes juegan cada mensaje.



(a) Costo del error tipo II mayor al del error tipo I.

(b) Bajo costo del error tipo II.

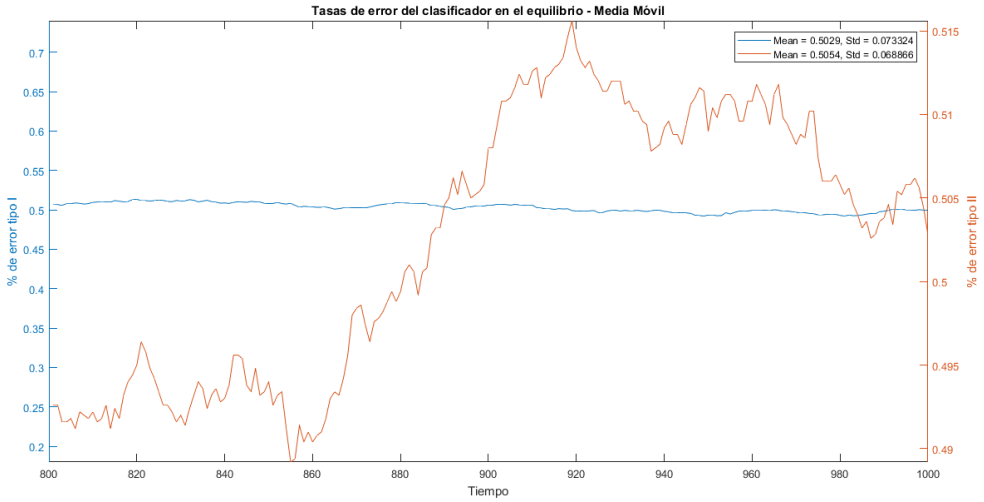
Figura 3.12: Estrategias del clasificador al variar la relación entre sus costos.

Se observa que cuando el costo del error de tipo II es mayor al del error de tipo I, la probabilidad de detener cada mensaje aumenta considerablemente. Por el contrario, cuando el costo por el error de tipo II es bajo, el clasificador disminuye la probabilidad de detener cada mensaje.

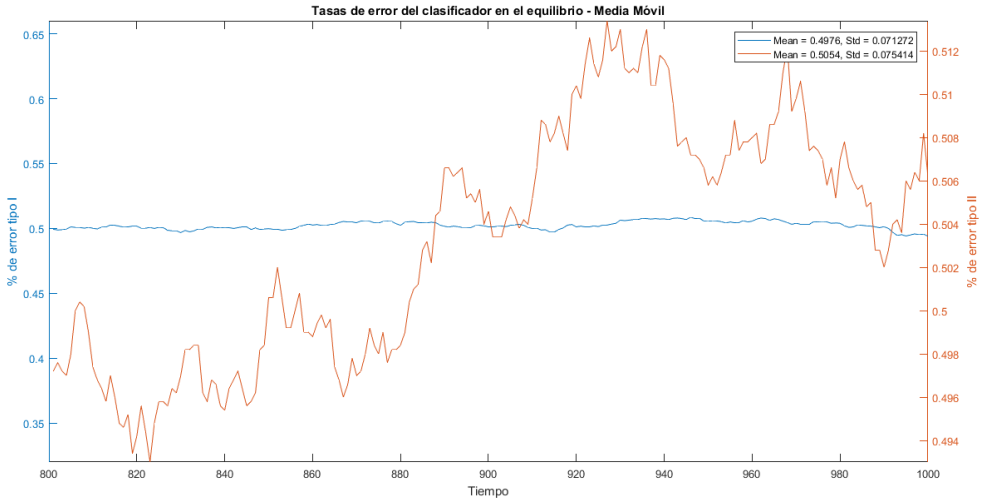
El clasificador sigue siendo capaz de ajustar sus estrategias a la proporción de adversarios maliciosos que juega cada mensaje, sin la necesidad de observarlos o de conocer sus características. Sin embargo, el comportamiento presentado en la figura 3.12, refuerza el hecho de que la estrategia final del clasificador es sensible a la relación entre sus costos, y se ajusta en función de ellos.

En la figura 3.13 se presenta la evolución temporal (como media móvil) de las tasas de error de tipo I y de tipo II del clasificador, para la fase de equilibrio del juego.

No se presentan cambios en las tasas de error del clasificador, al variar la relación entre sus costos.



(a) Costo del error tipo II mayor al del error tipo I.

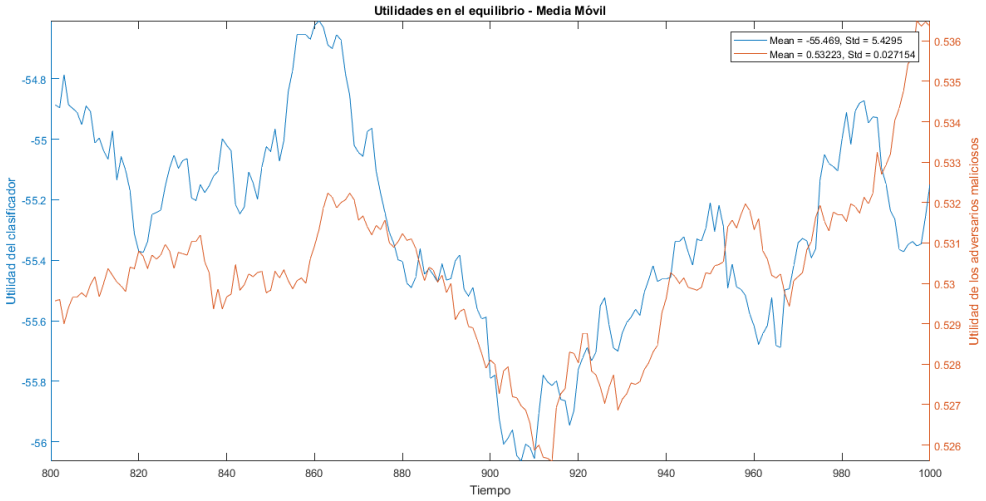


(b) Bajo costo del error tipo II.

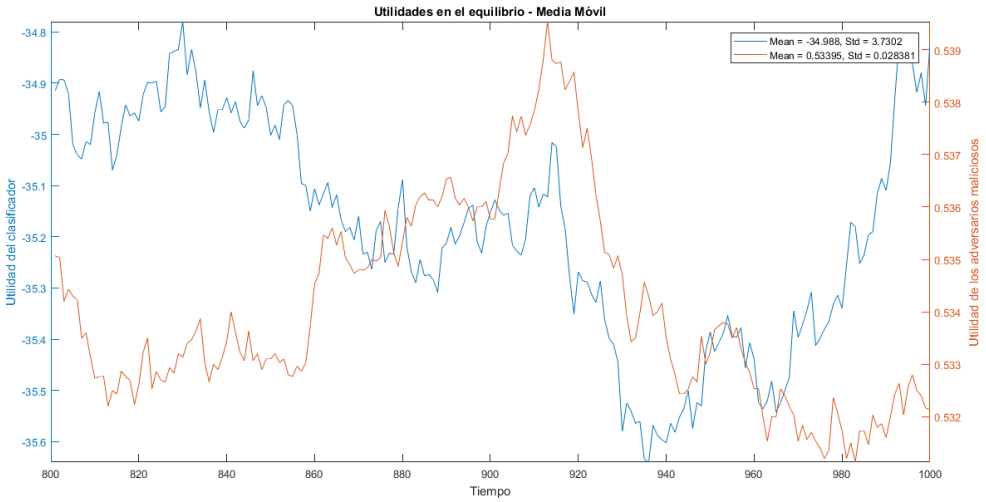
Figura 3.13: Errores tipo I y II (media móvil) al variar la relación entre los costos del clasificador.

En la figura 3.14 se presenta la evolución temporal de las utilidades del clasificador y de los adversarios maliciosos, como una media móvil, en la fase de equilibrio del juego.

Por un lado, se observa que la utilidad media de los adversarios no cambia con respecto al caso estándar. Por otro, la utilidad del clasificador se ve fuertemente afectada. Cuando aumenta el costo del error de tipo II, su utilidad disminuye (de $-44,55$ a $-55,47$). En cambio, cuando el costo es bajo, la utilidad media del clasificador aumenta con respecto al caso estándar (de $-44,55$ a $-34,98$).



(a) Costo del error tipo II mayor al del error tipo I.



(b) Bajo costo del error tipo II.

Figura 3.14: Utilidades del clasificador y de los adversarios maliciosos (media móvil), al variar la relación entre los costos del clasificador.

Estos resultados sugieren que el clasificador empeora su rendimiento a medida que aumenta el costo del error de tipo II, con respecto al error de tipo I.

Sin embargo, como se observa en la figura 3.13, el clasificador no empeora su rendimiento producto del aumento del costo.

Por lo tanto, se deduce que la diferencia en las utilidades del clasificador no depende de la relación entre sus costos. Por el contrario, los cambios en la utilidad promedio se deben a los valores específicos que toman los costos, y no por la influencia de estos en las estrategias.

Conclusiones

En esta investigación se presenta un modelo de clasificación de adversarios desde un enfoque diferente, que no requiere de los típicos supuestos de información completa, o de creencias sobre las preferencias entre los distintos jugadores.

El modelo incorpora aprendizaje desde la propia experiencia. Es decir, los jugadores actualizan sus estrategias observando únicamente sus propias acciones y utilidades del pasado.

El primer gran resultado del estudio es que bajo este modelo, el juego es capaz de llegar rápidamente a las estrategias de equilibrio del juego.

Este equilibrio no responde a un proceso de búsqueda de la estrategia óptima hasta llegar a un equilibrio de Nash (que ha sido el foco de la literatura hasta el momento), sino que se llega a él a través de un proceso de ensayo y error simultáneo entre los jugadores.

Al observar los resultados de las simulaciones en la fase de equilibrio del juego, se extraen los siguientes aprendizajes:

1. Las estrategias agregadas de los adversarios maliciosos siempre se acercan a una distribución uniforme entre los mensajes disponibles. Este resultado no se ve afectado por los valores que tomen los parámetros clave del juego.

2. La estrategia final del clasificador siempre logra seguir a la proporción de adversarios maliciosos que envía efectivamente cada mensaje. Este resultado depende estrechamente de cuál sea la relación entre los costos de error del clasificador.

   Dejando fijo el costo por clasificar mal a un adversario regular: a medida que aumenta el costo por clasificar mal a un adversario malicioso, el clasificador tiende a aumentar la probabilidad con la que detiene los mensajes. El efecto opuesto es equivalente.

3. En general, las utilidades del clasificador y de los adversarios maliciosos se comportan de manera opuesta: cuando una aumenta, la otra disminuye. Esto se debe a que el clasificador aumenta su utilidad mejorando la clasificación de los adversarios maliciosos, quienes ven disminuida su utilidad.

   Sin embargo, este resultado cambia si el costo por error de tipo II es suficientemente bajo comparado con el costo del error de tipo I; o si la cantidad de adversarios mali-

ciosos es muy baja respecto al total de adversarios.

En estos casos el clasificador prioriza aumentar su utilidad mediante clasificar bien a los adversarios regulares (y así evitar caer en los costos del error de tipo I), cayendo en error de tipo II en el proceso; lo que aumenta también la utilidad de los adversarios maliciosos. El caso opuesto sigue una lógica equivalente.

4. El clasificador es capaz de mantener la tasa de error de tipo I a un nivel constante, mientras que la tasa de error de tipo II oscila en torno a su promedio. Esto se debe a que el clasificador y los adversarios maliciosos mantienen una dinámica de gato y ratón, incluso en la fase de equilibrio del juego.

Lo más relevante de los resultados es que los jugadores son capaces de llegar a un equilibrio e interactuar de esta forma a lo largo del juego, sin observar, ni definir creencias sobre las jugadas o las preferencias del otro. Todo se logra observando únicamente la historia de jugadas y utilidades del propio jugador.

En la literatura se ha trabajado con modelos que suponen cierto grado de conocimiento entre los jugadores. Además se han aplicado los resultados a datos históricos, típicamente de correo electrónico y spam, logrando buenos resultados.

Esta investigación logra introducir un modelo que no necesita incorporar supuestos respecto al grado de conocimiento entre los jugadores. Los siguientes pasos son aplicar este modelo a datos reales de correo electrónico y spam. De igual manera, aplicarlo a experimentos con jugadores que interactúen en tiempo real con un clasificador entrenado para aprender, sin conocer a sus adversarios.

Bibliografía

[1] Banks, J. S. y Sobel, J. (1987). Equilibrium Selection in Signaling Games. *Econometrica 55*(3), 647−661.

[2] Brückner, M., Kanzow, C. y Scheffer, T. (2012). Static Prediction Games for Adversarial Learning Problems. *Journal of Machine Learning Research, 13*(1), 2617−2654.

[3] Cho, I.-K. y Sobel, J. (1990). Strategic stability and uniqueness in signaling games. *Journal of Economic Theory, 50*(2), 381−413.

[4] Cominetti, R., Melo, E. y Sorin, S. (2010). A payoff−based learning procedure and its application to traffic games. *Games and Economic Behavior, 70*(1), 7183.

[5] Dalvi, N., Domingos, P., Mausam, Sanghai, S. y Verma, D. (2004). Adversarial classification. *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining − KDD 2004, 99*.

[6] Dritsoula, L., Loiseau, P. y Musacchio, J. (2017). A Game-Theoretic Analysis of Adversarial Classification. *IEEE Transactions on Information Forensics and Security*(12).

[7] Ewerhart, C. y Wichardt, P. (2004). Signaling, Globality, and the Intuitive Criterion. *Institute for Empirical Research in Economics, 189*. University of Zurich.

[8] Figueroa, N., LHuillier, G. y Weber, R. (2017). Adversarial classification using signaling games with an application to phishing detection. *Data Mining and Knowledge Discovery*(31). Springer US.

[9] Fudenberg, D. y He, K. (2017). Learning and Equilibrium Refinements in Signalling Games. Recuperado de http://arxiv.org/abs/1709.01024

[10] Fudenberg, D. y Tirole, J. (1991). *Game theory*. Cambridge: MIT Press.

[11] Gibbons, R. (1992). *Game theory for applied economists*. Princeton: Princeton University Press.

[12] Liu, W. y Chawla, S. (2009). A game theoretical model for adversarial learning. *ICDM Workshops 2009 - IEEE International Conference on Data Mining*, 2530.

[13] Lowd, D. y Meek, C. (2005). Adversarial learning. *Proceeding of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining − KDD 2005, 641.*

[14] Marden, J. R., Young, H. P., Arslan, G. y Shamma, J. S. (2007). Payoff based dynamics for multi-player weakly acyclic games. *Proceedings of the IEEE Conference on Decision and Control*, (January 2014), 34223427.

[15] Mas−Colell, A., Whinston, M. y Green J. (1995). *Microeconomic Theory.* New York: Oxford University Press.

[16] Schipper, B. C. (2017). Strategic Teaching and Learning in Games. Recuperado de http://faculty.econ.ucdavis.edu/faculty/schipper/learning.pdf

[17] Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*(87). The MIT Press.

[18] Vida, P. y Honryo, T. (2015). Strategic Stability of Equilibria in Multi-Sender Signaling Games.

[19] Young, H. P. (2009). Learning by Trial and Error. Recuperado de http://www.econ2.jhu.edu/people/young/Learning5June08.pdf

Apéndices

3.3. Juegos de señalización

Los juegos de señalización comúnmente contienen:

- Jugadores: $\mathrm{i} \in I$. Típicamente $I = \{1, 2\}$.
- Tipos: $\theta_k \in \Theta_\mathrm{i}, \forall \mathrm{i} \in I$.
- Acciones: $a_\mathrm{i} \in A_\mathrm{i}, \forall \mathrm{i} \in I$.
- Creencias: alguna distribución μ sobre Θ.
- Utilidades: $u_{(a_\mathrm{i}, a_{-\mathrm{i}})}, \forall \mathrm{i} \in I$.

Se supone que solo el jugador 1 posee información privada sobre su tipo, y que el tipo del jugador 2 es único y de información pública. Dado esto, el juego en su versión más simple es el siguiente:

$t = 0$: El jugador 1 observa su tipo. El jugador 2 posee creencias iniciales sobre el tipo del jugador 1.

$t = 1$: El jugador 1 realiza una acción a_1.

$t = 2$: El jugador 2 observa a_1 y realiza una acción $a_2(a_1)$. Luego, cada jugador recibe sus utilidades.

En la versión extensiva del juego, se repiten el segundo y tercer paso del algoritmo anterior, indefinidamente. Además, se deben hacer los siguientes refinamientos en las definiciones:

1. El jugador 2 actualiza sus creencias sobre los tipos del jugador 1 mediante una Regla de Bayes, siempre que sea posible.
2. Un **Equilibrio Bayesiano Perfecto (EBP)** es aquel que cumple las siguientes condiciones:
   (a) σ es secuencialmente racional dado μ. Es decir, es la estrategia (mixta) que entrega mayor utilidad esperada al jugador, dadas sus creencias sobre los tipos de los otros jugadores.
   (b) μ se deriva de σ mediante la regla de Bayes, siempre que sea posible.

Existen dos tipos de equilibrio en este tipo de juegos:

1. **Equilibrio separador:** El jugador 1 juegas acciones diferentes según su tipo. En estos

casos, el jugador 2 puede conocer el tipo del jugador 1 gracias a sus acciones, y jugar a_2 de manera correspondiente.

2. **Equilibrio Pooling:** El jugador 1 juega la misma acción (o la misma estrategia mixta) independiente de su tipo. De esta manera, el jugador 2 no puede conocer el tipo del jugador 1, por lo que juega una estrategia tal que maximice su utilidad esperada, acorde a sus creencias sobre los tipos del jugador 1.

Anexos

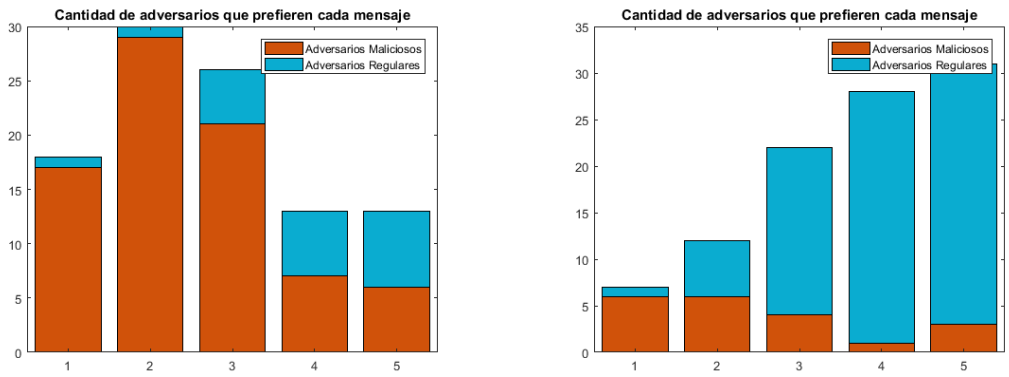
3.4. Variaciones de los parámetros clave del juego

En esta sección se pueden encontrar los gráficos complementarios al análisis de cambios en los valores de los parámetros más relevantes del juego.

Estos gráficos presentan resultados que no agregan un aprendizaje nuevo a lo ya mostrado en la simulación del caso estándar.

3.4.1. Variaciones en la proporción de adversarios maliciosos

En la figura 3.15 se presenta la distribución de preferencias de los adversarios al inicio del juego.



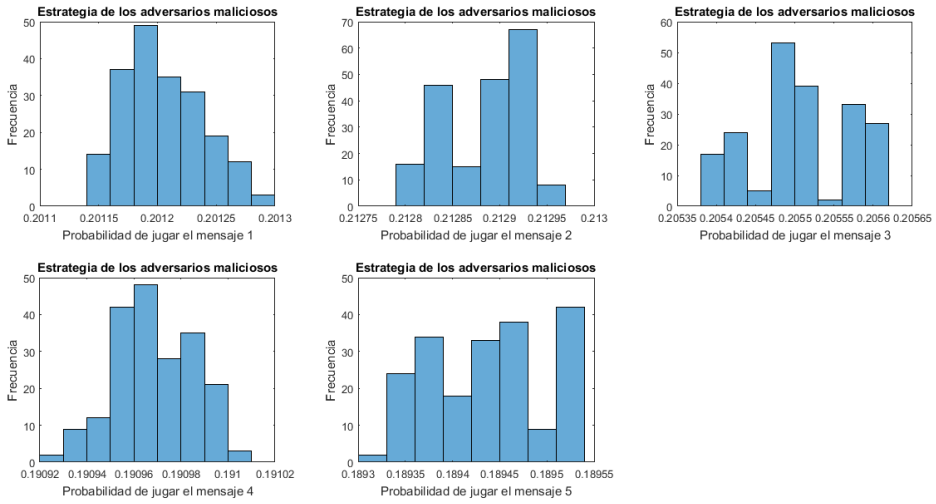
(a) Mayoría de adversarios maliciosos.

(b) Minoría de adversarios maliciosos.

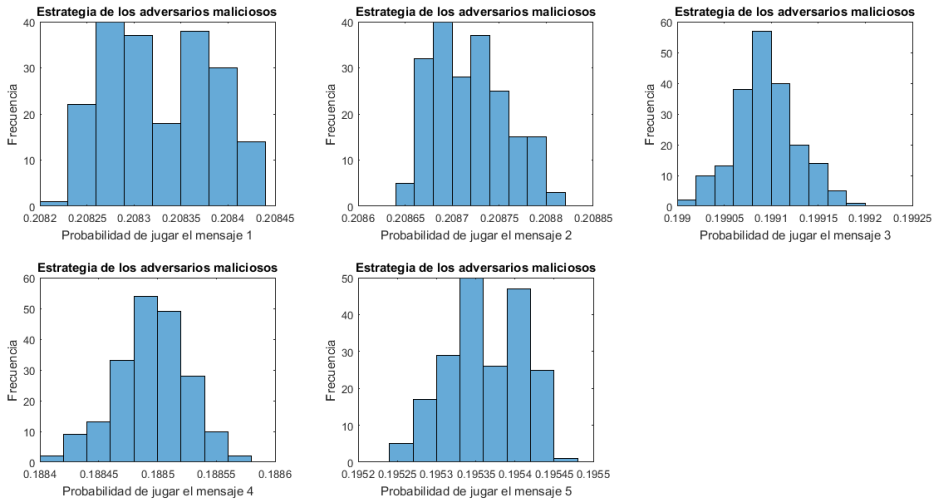
Figura 3.15: Distribución de las preferencias de los adversarios. Caso variación de la proporción de adversarios maliciosos.

En la figura 3.16 se presenta la estrategia de los adversarios maliciosos en la fase de equilibrio del juego. Es decir, la distribución de probabilidad con la que juegan cada mensaje, en promedio, cuando son mayoría y minoría respectivamente, con respecto al total de adversarios en el juego.

Al igual que en el caso estándar, la estrategia promedio de los adversarios maliciosos se acerca a una distribución uniforme en el equilibrio. Este resultado es independiente de la cantidad de adversarios maliciosos presentes en el juego.



(a) Mayoría de adversarios maliciosos.

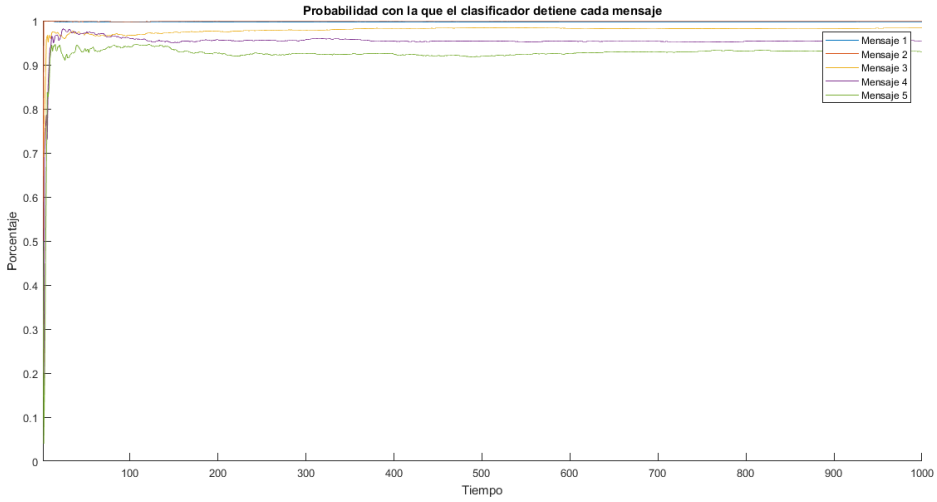


(b) Minoría de adversarios maliciosos.

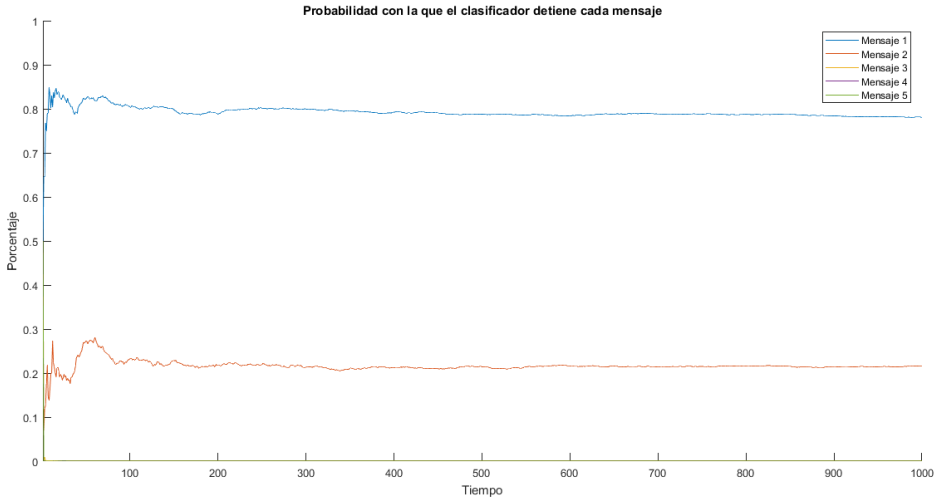
Figura 3.16: Estrategias de los adversarios maliciosos en el equilibrio, al variar su proporción en el juego.

En la figura 3.17 se presenta la evolución de las estrategias del clasificador a lo largo del juego. Es decir, la probabilidad con la que detiene cada mensaje en función del tiempo, para cada uno de los escenarios.

Al igual que en el caso estándar, se puede observar que aproximadamente en $t = 50$ el clasificador ya ha superado la fase de aprendizaje, y sus estrategias se estabilizan.



(a) Mayoría de adversarios maliciosos.



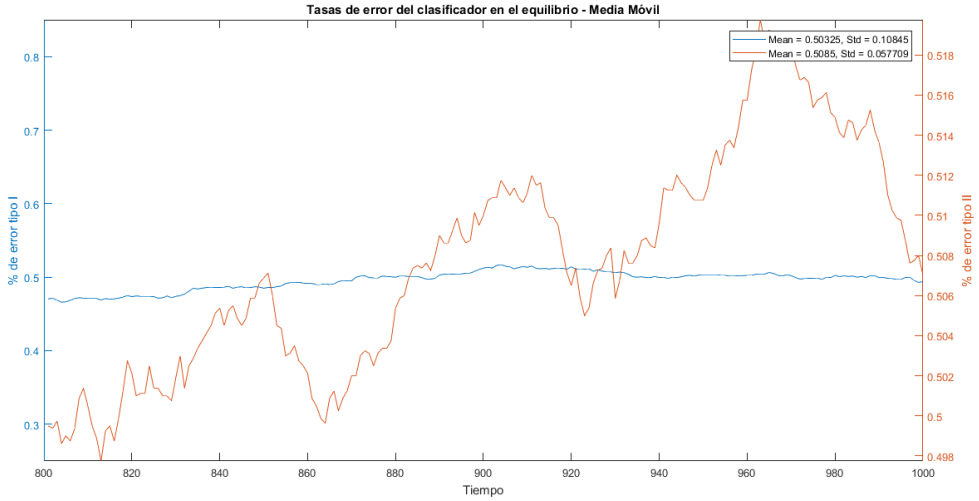
(b) Minoría de adversarios maliciosos.

Figura 3.17: Evolución de las estrategias del clasificador al variar la proporción de adversarios maliciosos.

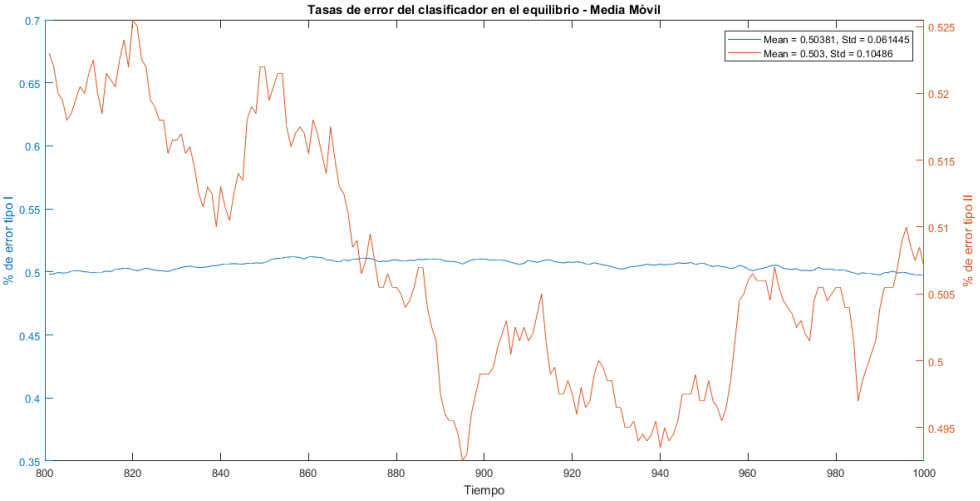
La figura 3.18 presenta la evolución de los errores de tipo I y II (como media móvil) para ambos escenarios, en la fase de equilibrio del juego.

El comportamiento es similar al del caso estándar. El error de tipo I se mantiene en un nivel constante, mientras que el error de tipo II fluctúa de manera anticíclica con respecto a la utilidad del clasificador.

Se puede apreciar que cuando la proporción de adversarios maliciosos aumenta, disminuye la tasa de error en su clasificación. Sin embargo, los cambios son muy pequeños como para ser concluyentes.



(a) Mayoría de adversarios maliciosos.

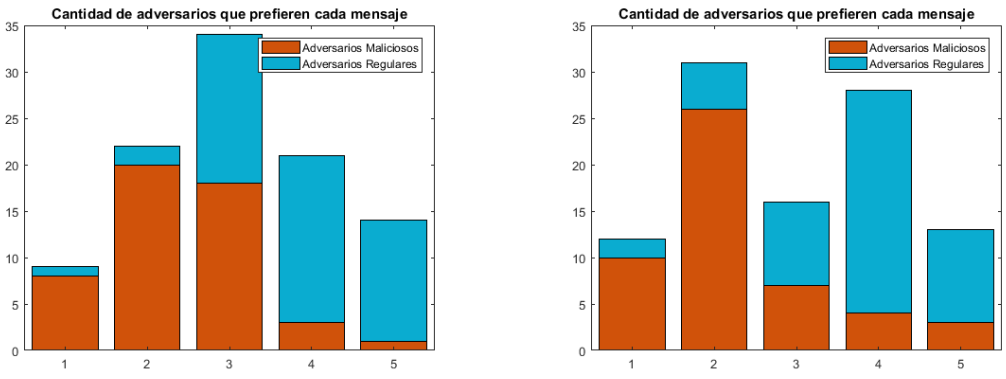


(b) Minoría de adversarios maliciosos.

Figura 3.18: Errores tipo I y II (media móvil) al variar la proporción de adversarios maliciosos.

3.4.2. Variaciones en el costo de los adversarios maliciosos

En la figura 3.19 se presenta la distribución de preferencias de los adversarios al inicio del juego. En ambos escenarios el mensaje 1 es preferido solo por adversarios maliciosos, al igual que en el caso estándar.

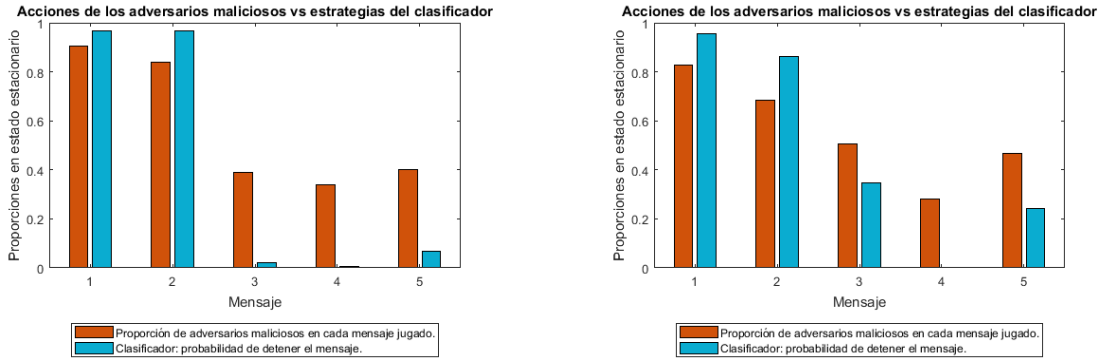


(a) Costo mayor al ingreso máximo. (b) Costo bajo.

Figura 3.19: Distribución de las preferencias de los adversarios. Caso variación de sus costos.

En la figura 3.20 se presenta la comparación en la fase de equilibrio del juego, entre las estrategias del clasificador y la proporción de adversarios que son maliciosos, de entre quienes juegan cada mensaje.

Tal como ocurre en todas las simulaciones, el clasificador es capaz de adaptar sus estrategias a la proporción de adversarios maliciosos que juegan cada mensaje, ajustando por la relación entre los costos de tipo I y II. Es decir, si la cantidad de adversarios regulares que juega un mensaje es alta, entonces baja la probabilidad de detener ese mensaje, y viceversa.

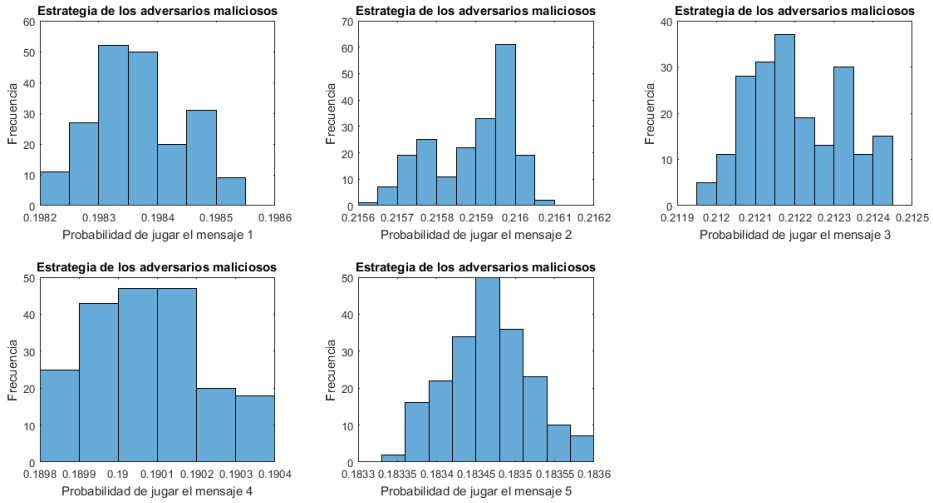


(a) Costo mayor al ingreso máximo. (b) Costo bajo.

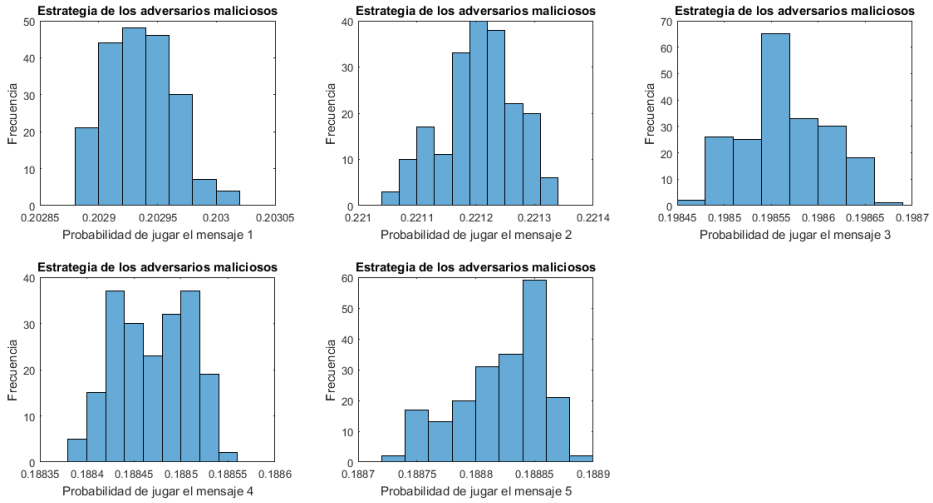
Figura 3.20: Estrategias del clasificador al variar el costo de los adversarios.

En la figura 3.21 se presenta la estrategia de los adversarios maliciosos en la fase de equilibrio del juego. Es decir, la distribución de probabilidad con la que juegan cada mensaje, en promedio, cuando sus costos son muy altos y muy bajos respectivamente.

Al igual que en el caso estándar, la estrategia promedio de los adversarios maliciosos se acerca a una distribución uniforme en el equilibrio. Este resultado es independiente del valor de sus costos.



(a) Costo mayor al ingreso máximo.

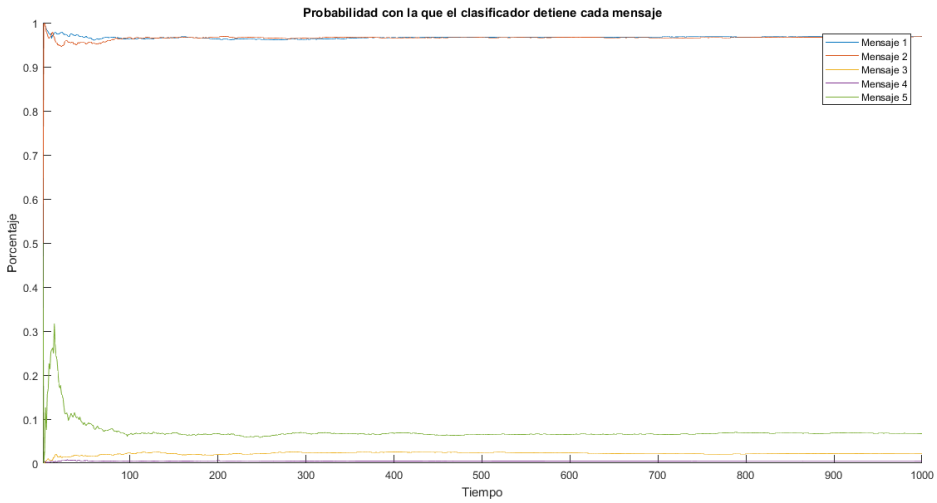


(b) Costo bajo.

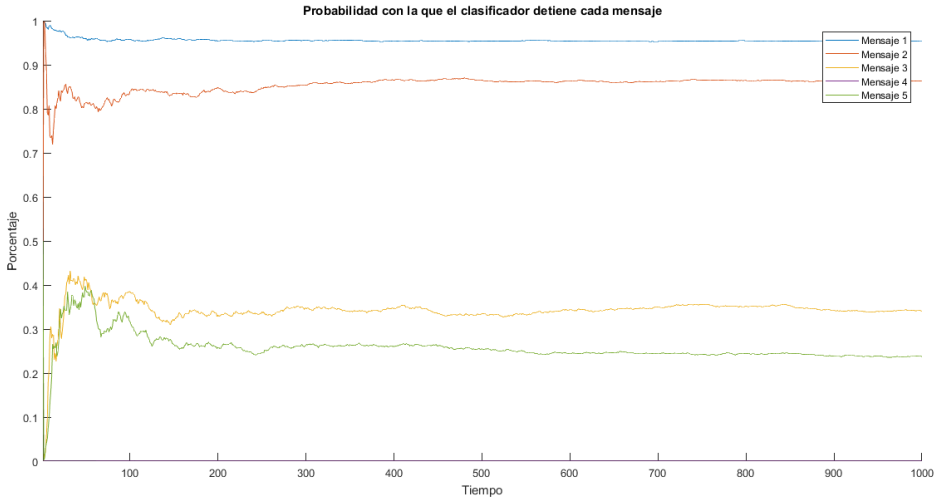
Figura 3.21: Estrategias de los adversarios maliciosos en el equilibrio al variar sus costos.

En la figura 3.22 se presenta la evolución de las estrategias del clasificador a lo largo del juego. Es decir, la probabilidad con la que detiene cada mensaje en función del tiempo, para cada uno de los escenarios.

Al igual que en los casos anteriores, aproximadamente en $t = 50$ sus estrategias se estabilizan.



(a) Costo mayor al ingreso máximo.



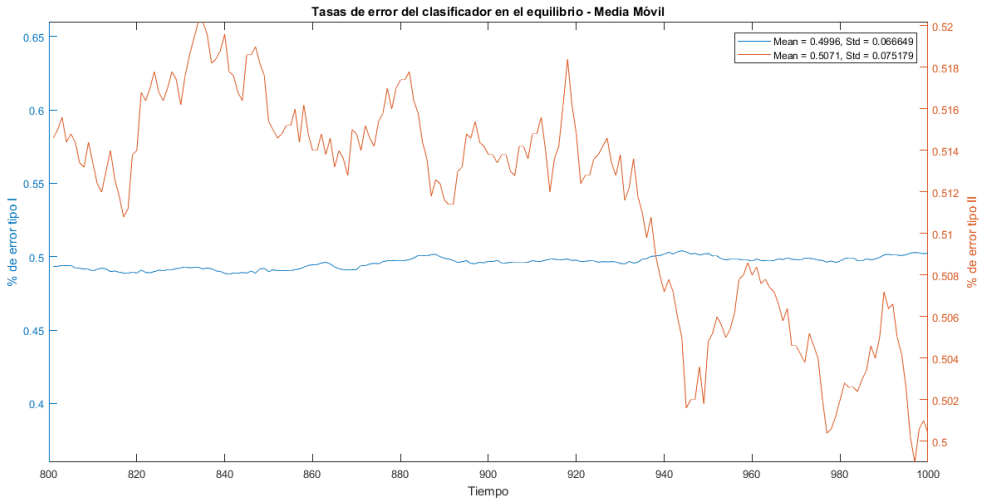
(b) Costo bajo.

Figura 3.22: Evolución de las estrategias del clasificador al variar los costos de los adversarios maliciosos.

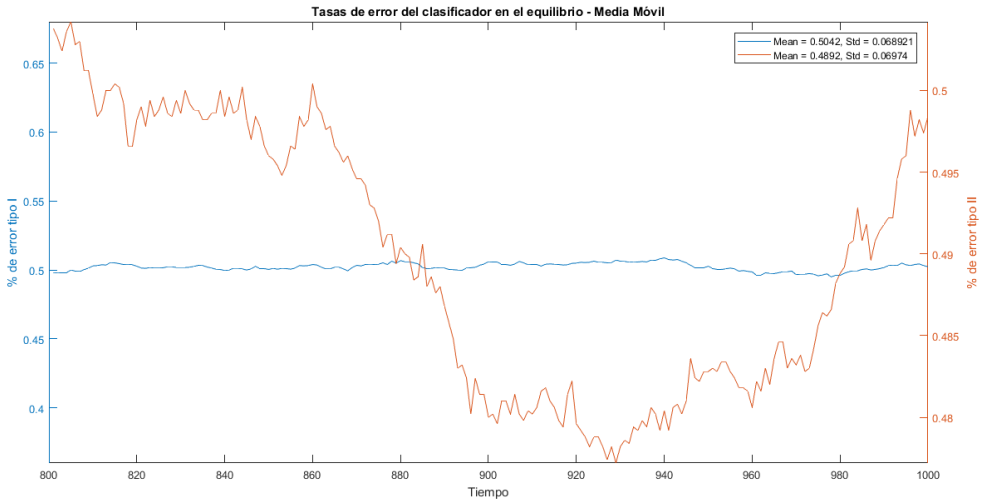
En la figura 3.23 se presenta la evolución temporal (como media móvil) de las tasas de error tipo I y II para el clasificador, en la fase de equilibrio del juego.

Para los dos escenarios el resultado es el mismo: la media del error tipo I es de 0,50, con una desviación estándar de 0,06; y la del error tipo II es 0,49, con una desviación estándar de 0,07. Respecto al caso estándar, estos resultados representan un cambio insignificante.

Esto lleva a concluir que las tasas de error del clasificador no se ven afectadas por el costo de los adversarios.



(a) Costo mayor al ingreso máximo.



(b) Costo bajo.

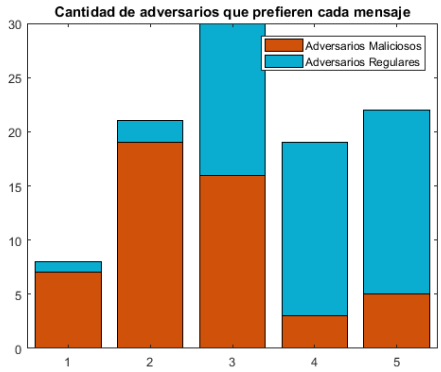
Figura 3.23: Errores tipo I y II (media móvil) al variar el costo de los adversarios.

3.4.3. Variaciones en el costo del clasificador

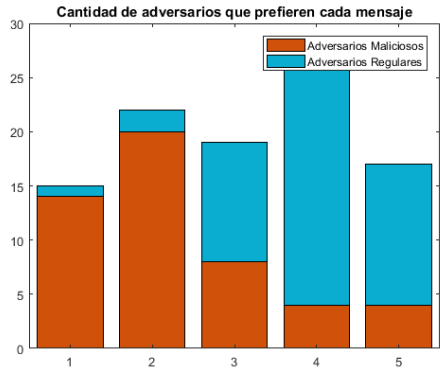
En la figura 3.24 se presenta la distribución de preferencias de los adversarios al inicio del juego.

En el escenario de alto costo para el clasificador, el mensaje 1 es preferido también por adversarios regulares. Esto se debe a la naturaleza aleatoria de la distribución de las preferencias.

Debido a la baja proporción de adversarios regulares en este mensaje, esto no afecta las conclusiones del juego.



(a) Costo del error tipo II mayor al del error tipo I.

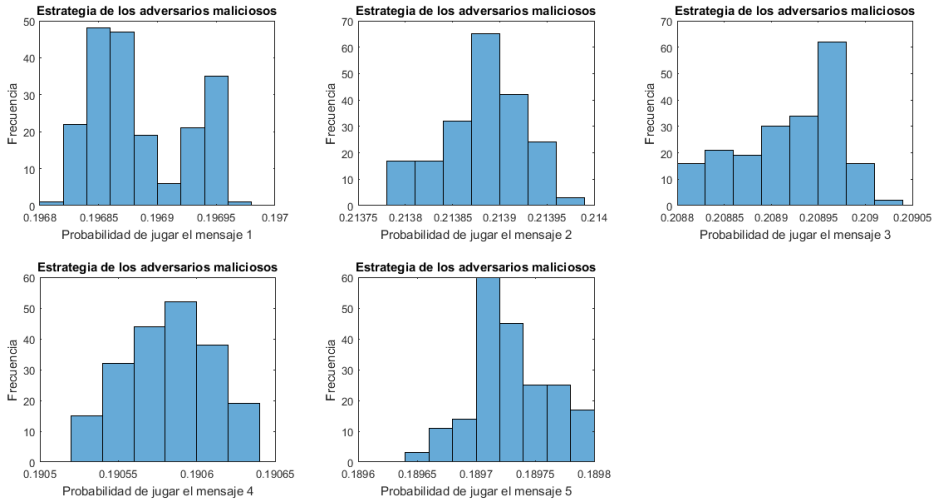


(b) Bajo costo del error tipo II.

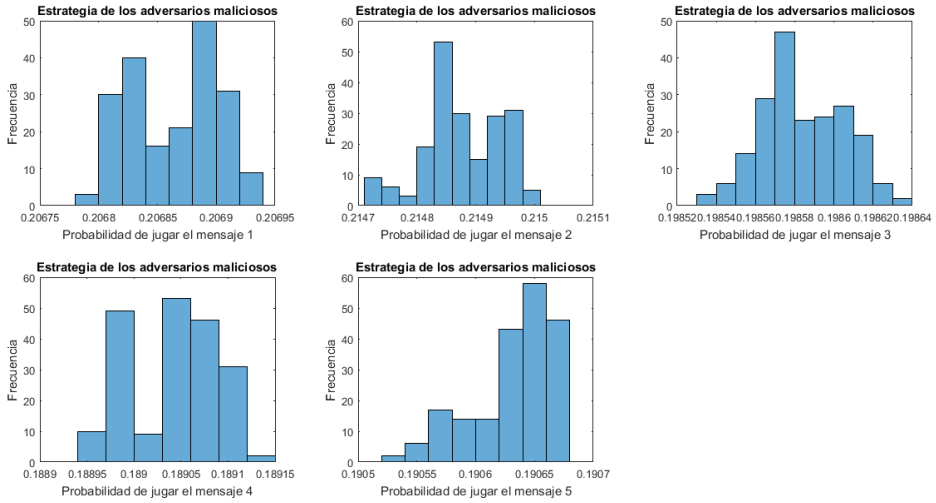
Figura 3.24: Distribución de preferencias de los adversarios al variar los costos del clasificador.

En la figura 3.25 se presenta la estrategia de los adversarios maliciosos en la fase de equilibrio del juego. Es decir, la distribución de probabilidad con la que juegan cada mensaje, en promedio, cuando varía la relación entre los costos del clasificador.

Al igual que en el caso estándar, la estrategia promedio de los adversarios maliciosos se acerca a una distribución uniforme en el equilibrio. Este resultado es independiente de la relación entre los costos del clasificador.



(a) Costo del error tipo II mayor al del error tipo I.

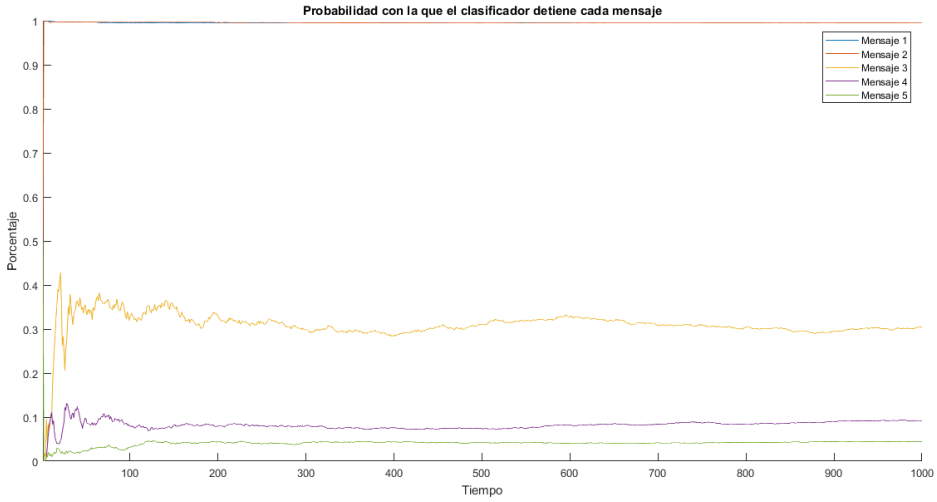


(b) Bajo costo del error tipo II.

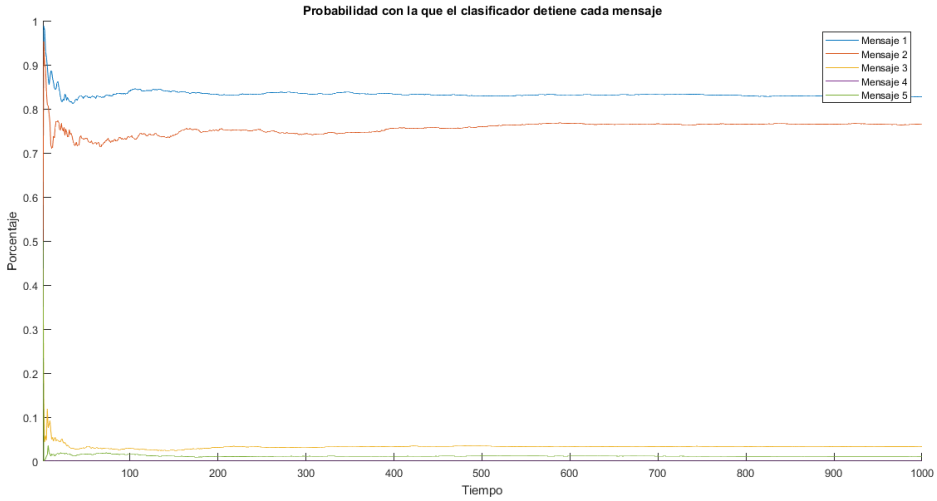
Figura 3.25: Estrategias de los adversarios maliciosos en el equilibrio, al variar los costos del clasificador.

En la figura 3.26 se presenta la evolución de las estrategias del clasificador a lo largo del juego. Es decir, la probabilidad con la que detiene cada mensaje en función del tiempo, para cada uno de los escenarios.

Al igual que en los casos anteriores, aproximadamente en $t = 50$ sus estrategias llegan a los niveles de equilibrio.



(a) Costo del error tipo II mayor al del error tipo I.



(b) Bajo costo del error tipo II.

Figura 3.26: Evolución de las estrategias del clasificador al variar sus costos.

3.5. Código de las simulaciones

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Clasificación de Adversarios %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clc

% Valores de los parámetros según el caso a simular:
%
% 1. Caso Base:
% nm = 50;
% c = 0.5;
% Lm = 0.8;
%
% 2. Variaciones en los parámetros (ceteris paribus):
%
% 2.1. Variaciones en proporción de adversarios maliciosos:
% - Mayoría: nm = 80;
% - Minoría: nm = 20;
%
% 2.2. Variaciones en proporción de costos del clasificador:
% - Muy alto: Lm = 1.2;
% - Bajo: Lm = 0.4;
%
% 2.3. Variaciones en costo de los adversarios maliciosos:
% - Muy alto: c = 1.2;
% - Bajo: c = 0.25;
%
% Parámetros iniciales

% Adversarios
N = 100;
nm = 50;
nr = 100 - nm;

% Mensajes
k = 5;

% Pagos Adversario Malicioso
b1 = 1;
b2 = b1/2;
b3 = b1/3;

% Matriz de pagos: la posición de la columna indica la preferencia del adv
b = [b1 b2 b3 b3 b2;
```

```
45      b2 b1 b2 b3 b3;
46      b3 b2 b1 b2 b3;
47      b3 b3 b2 b1 b2;
48      b2 b3 b3 b2 b1];
49
50  % Costo por ser detenido
51  c = 0.5;
52
53  % Costos Clasificador
54  Lr = 1;
55  Lm = 0.8;
56
57  % Ventana de tiempo
58  T = 1000;
59
60  % Preferencias de los Adversarios
61
62  opc = [1 2 3 4 5];
63  mu_m = 2;
64  mu_r = 4;
65  sigma = 1;
66
67  Ppref_m = normpdf(opc,mu_m,sigma);
68  Ppref_r = normpdf(opc,mu_r,sigma);
69
70  F_m = zeros(1,k);
71  F_r = zeros(1,k);
72
73  F_m(1,1) = Ppref_m(1,1);
74  F_r(1,1) = Ppref_r(1,1);
75
76  for i = 1:(k-1)
77      F_m(1,i+1) = F_m(1,i)+Ppref_m(1,i+1);
78      F_r(1,i+1) = F_r(1,i)+Ppref_r(1,i+1);
79  end
80
81  Pm = zeros(1,nm); % Preferencias de los adversarios maliciosos
82  Pr = zeros(1,nr); % Preferencias de los adversarios regulares
83
84  for i = 1:nm
85      aux = rand();
86      if aux <= F_m(1,1)
87          Pm(1,i) = 1;
88      elseif aux <= F_m(1,2)
89          Pm(1,i) = 2;
90      elseif aux <= F_m(1,3)
91          Pm(1,i) = 3;
```

```
92      elseif aux <= F_m(1,4)
93          Pm(1,i) = 4;
94      else
95          Pm(1,i) = 5;
96      end
97  end
98
99  for i = 1:nr
100     aux = rand();
101     if aux <= F_r(1,1)
102         Pr(1,i) = 1;
103     elseif aux <= F_r(1,2)
104         Pr(1,i) = 2;
105     elseif aux <= F_r(1,3)
106         Pr(1,i) = 3;
107     elseif aux <= F_r(1,4)
108         Pr(1,i) = 4;
109     else
110         Pr(1,i) = 5;
111     end
112 end
113
114 % Gráfico de las preferencias
115 count = [sum(Pm == 1) sum(Pr == 1);...
116     sum(Pm == 2) sum(Pr == 2);...
117     sum(Pm == 3) sum(Pr == 3);...
118     sum(Pm == 4) sum(Pr == 4);...
119     sum(Pm == 5) sum(Pr == 5)];
120
121 figure
122 pbar = bar(count,'stacked');
123 title('Cantidad de adversarios que prefieren cada mensaje')
124 legend('Adversarios Maliciosos','Adversarios Regulares','Location','northe
125 pbar(1).FaceColor = [0.8157 0.3217 0.0392];
126 pbar(2).FaceColor = [0.0392 0.6745 0.8157];
127
128 % Matrices del juego
129
130 % Estrategias
131 sigma_Am = zeros(T,k,nm);
132 sigma_C = zeros(T,k);
133
134 % Acciones
135 acciones_A = zeros(T,N);
136 acciones_C = zeros(T,N);
137
138 % Utilidades
```

```
139 U_Am = zeros(T,nm);
140 U_C = zeros(T,1);
141
142 % Matrices temporales del juego
143
144 % Utilidad en t del clasificador
145 Uc = zeros(k,2);
146
147 % Matrices de guardado de información
148 V = zeros(k,nm);
149 W = zeros(k,2);
150
151 % Nř de veces que cada adversario malicioso ha jugado cada mensaje
152 nu = zeros(k,nm);
153
154 % Mediciones
155
156 % Errores (%)
157 error_T1 = zeros(1,T);
158 error_T2 = zeros(1,T);
159
160 % Inicialización del juego
161
162 % Estrategias
163 for j = 1:k
164     for i = 1:nm
165         sigma_Am(1,j,i) = sigma_Am(1,j,i) + rand();
166     end
167 end
168
169 for j = 1:k
170     sigma_C(1,j) = sigma_C(1,j) + 0.5;
171 end
172
173 % Juego
174 disp('Comienza el juego...');
175 tic
176 for t = 1:T
177     % 1. Acciones
178
179     % 1.1. Acciones de adversarios
180     F_sigma_Am = zeros(1,k,nm);
181
182     for i = 1:nm
183         F_sigma_Am(1,1,i) = sigma_Am(t,1,i);
184     end
185
```

```
186     for i = 1:nm
187         for j = 2:k
188             F_sigma_Am(1,j,i) = F_sigma_Am(1,j-1,i)+sigma_Am(t,j,i);
189         end
190     end
191
192     for i = 1:nm
193         aux = rand();
194         if aux <= F_sigma_Am(1,1,i)
195             acciones_A(t,i) = 1;
196         elseif aux <= F_sigma_Am(1,2,i)
197             acciones_A(t,i) = 2;
198         elseif aux <= F_sigma_Am(1,3,i)
199             acciones_A(t,i) = 3;
200         elseif aux <= F_sigma_Am(1,4,i)
201             acciones_A(t,i) = 4;
202         else
203             acciones_A(t,i) = 5;
204         end
205     end
206
207     for i = 1:nr
208         acciones_A(t,i+nm) = Pr(1,i);
209     end
210
211     % Actualización de parámetros nu
212     for i = 1:nm
213         nu(acciones_A(t,i),i) = nu(acciones_A(t,i),i) + 1;
214     end
215
216     % 1.2. Acciones del clasificador
217     for i = 1:N
218         aux = rand();
219         if aux <= sigma_C(1,acciones_A(t,i))
220             acciones_C(t,i) = 1;
221         else
222             acciones_C(t,i) = 0;
223         end
224     end
225
226     % 2. Utilidades
227
228     % 2.1. Utilidades de los adversarios maliciosos
229     for i = 1:nm
230         U_Am(t,i) = acciones_C(t,i)*b(acciones_A(t,i),Pm(1,i))+(1-acciones
231     end
232
```

```
233     % 2.2. Utilidades del clasificador
234     mal_1 = acciones_C(t,1:nm).*acciones_A(t,1:nm);
235     mal_0 = (1-acciones_C(t,(nm+1):N)).*acciones_A(t,(nm+1):N);
236
237     for j = 1:k
238         Uc(j,1) = -sum(mal_0(1,:) == j)*Lr;
239         Uc(j,2) = -sum(mal_1(1,:) == j)*Lm;
240     end
241
242     % Estadísticas relacionadas
243     U_C(t,1) = sum(sum(Uc));
244
245     error_T1(1,t) = sum(mal_0(1,:) > 0)/nr;
246     error_T2(1,t) = sum(mal_1(1,:) > 0)/nm;
247
248     % 3. Actualización de matrices de información
249
250     % 3.1. Adversarios maliciosos
251     for i = 1:nm
252         V(acciones_A(t,i),i) = (1 - 1/(1 + nu(acciones_A(t,i),i)))*V(acciones_
253     end
254
255     % 3.2. Clasificador
256
257     % Comprobar en qué jugadas aplicó acción el clasificador
258     accion_0 = (1-acciones_C(t,:)).*acciones_A(t,:);
259     accion_1 = acciones_C(t,:).*acciones_A(t,:);
260
261     accion = zeros(k,2);
262
263     for j = 1:k
264         if sum(accion_0(1,:) == j) > 0
265             accion(j,1) = 1;
266         else
267             accion(j,1) = 0;
268         end
269
270         if sum(accion_1(1,:) == j) > 0
271             accion(j,2) = 1;
272         else
273             accion(j,2) = 0;
274         end
275     end
276
277     % Actualizar matriz de información en acciones efectivamente jugadas
278     for j = 1:k
279         W(j,1) = accion(j,1)*((1-1/t)*W(j,1) + (1/t)*Uc(j,1)) + (1-accion(j,1)
```

```
280             W(j,2) = accion(j,2)*((1-1/t)*W(j,2) + (1/t)*Uc(j,2)) + (1-accion(
281         end
282
283         % 4. Elección de las estrategias del siguiente turno
284
285         % 4.1. Estrategias de los adversarios maliciosos
286
287         if t == T
288             disp('Fin del juego.');
289         else
290             for j = 1:k
291                 for i = 1:nm
292                     sigma_Am(t+1,j,i) = exp(V(j,i))/sum(exp(V(:,i)));
293                 end
294             end
295         end
296
297         % 4.2. Estrategias del clasificador
298
299         if t == T
300             disp('Fin del juego.');
301         else
302             for j = 1:k
303                 sigma_C(t+1,j) = exp(W(j,2))/(exp(W(j,1))+exp(W(j,2)));
304             end
305         end
306     end
307
308     % Estadísticas
309
310     % 0. Preparación de las matrices de estadísticas
311
312     % Formato
313     mean_sigma_Am = zeros(k,T);
314     for j = 1:k
315         for t = 1:T
316             mean_sigma_Am(j,t) = mean(sigma_Am(t,j,:));
317         end
318     end
319
320     mean_U_Am = zeros(1,T);
321     for t = 1:T
322         mean_U_Am(1,t) = mean(U_Am(t,:));
323     end
324
325     sigma_C_stop = 1-sigma_C';
326     U_C = U_C';
```

```
327
328 % Periodo estacionario
329 final = 0.2*T;
330
331 % Utilidades final
332 U_Am_final = mean_U_Am(1,(T-final+1):T);
333 U_C_final = U_C(1,(T-final+1):T);
334
335 % Estrategias finales
336 sigma_Am_final = mean_sigma_Am(:,(T-final+1):T);
337 sigma_C_final = sigma_C_stop(:,(T-final+1):T);
338
339 % Errores
340 error_T1_final = error_T1(1,(T-final+1):T);
341 error_T2_final = error_T2(1,(T-final+1):T);
342
343 % Ejes X
344 T_axis = linspace((T-final+1),T,final);
345 k_axis = linspace(1,k,k);
346
347 % Estrategias respecto a los mensajes
348 est_Am_t = zeros(T,k);
349 for i = 1:T
350     for j = 1:k
351         est_Am_t(i,j) = sum(acciones_A(i,1:nm) == j)/sum(acciones_A(i,:) == j)
352     end
353 end
354
355 est_Am_final = est_Am_t((T-final+1):T,:);
356 est_C_t = sigma_C_final';
357
358 est_Am = mean(est_Am_final)';
359 est_C = mean(est_C_t)';
360
361 est = [est_Am est_C];
362
363 % Resumen adversarios maliciosos
364 sigma_Am_k = [mean(sigma_Am_final(1,:)) mean(sigma_Am_final(2,:)) mean(sigma_A
365      mean(sigma_Am_final(4,:)) mean(sigma_Am_final(5,:))]';
366
367 resumen_Am = [count(:,1)/nm sigma_Am_k(:,1)];
368
369 % Medias móviles
370 MM = 100;
371
372 % Utilidades
373 U_Am_MM = zeros(1,T-MM+1);
```

```
374 U_C_MM = zeros(1,T-MM+1);
375
376 for i = 1:(T-MM+1)
377     U_Am_MM(1,i) = mean(mean_U_Am(1,i:(i+MM-1)));
378     U_C_MM(1,i) = mean(U_C(1,i:(i+MM-1)));
379 end
380
381 U_Am_MM_final = U_Am_MM(1,(T-MM+2-final):(T-MM+1));
382 U_C_MM_final = U_C_MM(1,(T-MM+2-final):(T-MM+1));
383
384 % Errores
385 error_T1_MM = zeros(1,T-MM+1);
386 error_T2_MM = zeros(1,T-MM+1);
387
388 for i = 1:(T-MM+1)
389     error_T1_MM(1,i) = mean(error_T1(1,i:(i+MM-1)));
390     error_T2_MM(1,i) = mean(error_T2(1,i:(i+MM-1)));
391 end
392
393 error_T1_MM_final = error_T1_MM(1,(T-MM+2-final):(T-MM+1));
394 error_T2_MM_final = error_T2_MM(1,(T-MM+2-final):(T-MM+1));
395
396 % Estadísticos
397
398 mu_Uc = mean(U_C_final);
399 std_Uc = std(U_C_final);
400
401 mu_UAm = mean(U_Am_final);
402 std_UAm = std(U_Am_final);
403
404 mu_eT1 = mean(error_T1_final);
405 std_eT1 = std(error_T1_final);
406
407 mu_eT2 = mean(error_T2_final);
408 std_eT2 = std(error_T2_final);
409
410 % Gráficos
411
412 % 1. Histogramas
413
414 % Estrategias del clasificador
415
416 figure
417 pbar2 = bar(est,'group');
418 title('Acciones de los adversarios maliciosos vs estrategias del clasifica
419 pbar2(1).FaceColor = [0.8157 0.3217 0.0392];
420 pbar2(2).FaceColor = [0.0392 0.6745 0.8157];
```

```
421 xlabel('Mensaje')
422 ylabel('Proporciones en estado estacionario')
423 legend('Proporción de adversarios maliciosos en cada mensaje jugado.','Clasifi
424
425 % Estrategias de los adversarios maliciosos
426
427 figure
428 pbar1 = bar(k_axis,resumen_Am(:,1));
429 title('Estrategias de los adversarios maliciosos en el equilibrio')
430 xlabel('Mensajes')
431 ylabel('Sigma')
432 pbar1(1).FaceColor = [0.0392 0.6745 0.8157];
433 legend('Estrategias finales.','Location','northeast')
434
435 figure
436 subplot(2,3,1)
437 histogram(sigma_Am_final(1,:));
438 title('Estrategia de los adversarios maliciosos')
439 xlabel('Probabilidad de jugar el mensaje 1')
440 ylabel('Frecuencia')
441
442 subplot(2,3,2)
443 histogram(sigma_Am_final(2,:));
444 title('Estrategia de los adversarios maliciosos')
445 xlabel('Probabilidad de jugar el mensaje 2')
446 ylabel('Frecuencia')
447
448 subplot(2,3,3)
449 histogram(sigma_Am_final(3,:));
450 title('Estrategia de los adversarios maliciosos')
451 xlabel('Probabilidad de jugar el mensaje 3')
452 ylabel('Frecuencia')
453
454 subplot(2,3,4)
455 histogram(sigma_Am_final(4,:));
456 title('Estrategia de los adversarios maliciosos')
457 xlabel('Probabilidad de jugar el mensaje 4')
458 ylabel('Frecuencia')
459
460 subplot(2,3,5)
461 histogram(sigma_Am_final(5,:));
462 title('Estrategia de los adversarios maliciosos')
463 xlabel('Probabilidad de jugar el mensaje 5')
464 ylabel('Frecuencia')
465
466 % 2. Gráficos de avance en el tiempo
467
```

```
% Estrategias de adversarios maliciosos

figure
hold on
plot(est_Am_t(:,1))
plot(est_Am_t(:,2))
plot(est_Am_t(:,3))
plot(est_Am_t(:,4))
plot(est_Am_t(:,5))
xlabel('Tiempo')
ylabel('Porcentaje')
title('% de adversarios que son maliciosos en cada mensaje')
legend('Mensaje 1','Mensaje 2','Mensaje 3','Mensaje 4','Mensaje 5')
axis([1 T 0 1])
hold off

% Estrategias del clasificador

figure
hold on
plot(sigma_C_stop(1,:))
plot(sigma_C_stop(2,:))
plot(sigma_C_stop(3,:))
plot(sigma_C_stop(4,:))
plot(sigma_C_stop(5,:))
xlabel('Tiempo')
ylabel('Porcentaje')
title('Probabilidad con la que el clasificador detiene cada mensaje')
legend('Mensaje 1','Mensaje 2','Mensaje 3','Mensaje 4','Mensaje 5')
axis([1 T 0 1])
hold off

% Utilidades de los adversarios maliciosos
figure
plot(mean_U_Am)
xlabel('Tiempo')
ylabel('Utilidad media')
title('Utilidad de los adversarios maliciosos')

% Utilidades del clasificador
figure
plot(U_C)
xlabel('Tiempo')
ylabel('Utilidad')
title('Utilidad del clasificador')

% Utilidades - media móvil
```

```
figure

yyaxis left
plot(T_axis,U_C_MM_final)
ylabel('Utilidad del clasificador')
ylim([min(U_C_MM_final) max(U_C_MM_final)])

yyaxis right
plot(T_axis,U_Am_MM_final)
ylabel('Utilidad de los adversarios maliciosos')
ylim([min(U_Am_MM_final) max(U_Am_MM_final)])

xlabel('Tiempo')
title('Utilidades en el equilibrio - Media Móvil')

legend(['Mean = ' num2str(mu_Uc) ', Std = ' num2str(std_Uc)],['Mean = ' num2st

% Errores
figure

yyaxis left
plot(T_axis,error_T1_final)
ylabel('% de error tipo I')
ylim([min(error_T1_final) max(error_T1_final)])

yyaxis right
plot(T_axis,error_T2_final)
ylabel('% de error tipo II')
ylim([min(error_T2_final) max(error_T2_final)])

xlabel('Tiempo')
title('Tasas de error del clasificador en el equilibrio')

% Errores - Media Móvil
figure

yyaxis left
plot(T_axis,error_T1_MM_final)
ylabel('% de error tipo I')
ylim([min(error_T1_final) max(error_T1_final)])

yyaxis right
plot(T_axis,error_T2_MM_final)
ylabel('% de error tipo II')
ylim([min(error_T2_MM_final) max(error_T2_MM_final)])

xlabel('Tiempo')
```

```
562 title('Tasas de error del clasificador en el equilibrio - Media Móvil')
563 legend(['Mean = ' num2str(mu_eT1) ', Std = ' num2str(std_eT1)],['Mean = '
564
565 toc
```