Modelamiento Basado en Agentes (ABM) Aplicado a Anti Lavado de Dinero (AML)

Milton Gener Palacin Grijalva
Maestría en Ciencias de Ciencia de la Computación
Universidad Nacional de Ingeniería, Perú
mpalacing@uni.pe

***Resumen*—El propósito de este trabajo es aplicar los temas desarrollados en el curso "MCC639D". El tema seleccionado es "Modelamiento Basado en Agentes como técnica contra el Lavado de Dinero", con este trabajo se pretende demostrar que los resultados, luego de aplicar varias simulaciones de transacciones de dinero, puede ayudar a descubrir patrones que no se puedan detectar utilizando solo data histórica o predicciones corto plazo. El impacto económico de un fraude financiero puede ser un problema crítico cuando los procedimientos de prevención no son robustos. En este trabajo se crea un modelo basado en agentes para detectar transacciones fraudulentas.**

***Palabras Clave*—ABM, AML, Dinero, Agente.**

1. Introducción

El modelamiento y la simulación para predecir el comportamiento humano en la construcción de teorías sociales, económicas y biológicas han evolucionado del modelamiento matemático antes de los 50's (a partir del comportamiento del sistema), desarrollo de la Dinámica de Sistemas en los 50's (Jay W. Forrester), Micro Simulación en los 50's, Autómatas Celulares en los 60's y Modelos Basados en Agentes (ABM - Agent Bases Modelling) en los 90's.

ABM es una técnica potente para el modelamiento y la simulación de complejos sistemas estocásticos, cubriendo brechas de incompletitud de modelos matemáticos y de la inteligencia artificial. ABM considera como unidad básica de medida a los eventos aleatorios contrario a todo lo que se puede determinar bajo un riguroso análisis: " Everything should be made as simple as possible, but no simpler".

Es tipo de fraude que consiste en proceso de ocultamiento de la existencia, la fuente ilegal o la aplicación de ingresos obtenidos provenientes de actividades criminales, y el subsiguiente ocultamiento de la fuente de esos ingresos para hacerla parecer legítimos. También es cualquier tentativa por ocultar o disfrazar la identidad de los fondos obtenidos ilegalmente de manera que aparezcan como originados en fuentes legítimas.

La legitimación de activos no es otra cosa que tomar el producto de una actividad ilícita y darle apariencia de legitimidad. La corrupción y el terrorismo son dos casos que consisten en uso de cierto poder y el terror respectivamente para obtener objetivos económicos, políticos, religioso e ideológicos.

En el presente trabajo se utiliza el Modelamiento Basado en Agente para estudiar las redes Contra el Lavado de Dinero, además se propone una herramienta para realizar simulaciones de transacciones (depósitos, transferencias y retiros). La estructura de presente artículo está organizado de la siguiente manera. La sección 2 se define el problema. En la sección 3 presenta el modelado de sistema Contra el Lavado de Dinero, conocido como 'Anti-Money Laundering" AML (por sus siglas en inglés). En la Sección 4 se describe de forma detallada la experimentación y los resultados. En la Sección 6 presenta las conclusiones y trabajos futuros.

Todo el código y cuadernos de trabajo está disponible de manera directa en:

- https://github.com/miltonpalacin/abm_aml

2. El Problema

2.1. Antecedentes

- Los Informes de Inteligencia Financiera (IIF) de la UIF-Perú desde de enero del 2012 a mayo del 2021 involucran US$ 14,478 millones, procedentes de minería ilegal (57 %), defraudación tributaria (10 %), tráfico ilícito de drogas (10 %), delitos contra la administración pública (10 %) y entre otros (13 %). En el periodo junio 2020 a mayo 2021 fue de UD$ 1,566 millones.
- Perú se encuentra en la posición 104 de 141 países en índice "Anti-lavado de Dinero – AML 2020", que publica el Instituto de Gobernabilidad de Basilea.'
- Las actividades de Lavado de Dinero tienen un costo entre el 2 % y 5 % del PBI mundial (entre US$800 billones a USD$ 2 trillones aproximadamente).
- 95 % de los sistemas de alertas contra el Lavado de Dinero resultan ser falsos positivos.
- Las Naciones Unidas estimó que para 2020-2021, 90 % del dinero lavado permanecerán sin ser detectados.
- Las entidades financieras y reguladores para determinar el fraude utilizan, entre otras técnicas, "Z-Score" y la "Ley de Benford". "Z-Score" es una medida estadística de un valor en relación con la media del grupo de valores, el valor central de la curva Z-Score es 0, el valor es positivo si está a la derecha de la curva y negativo si esta a la izquierda. La Ley de Benford, es una técnica para identificar cuentas o montos sospechosas en base los primeros dígitos de los montos de las transacciones.

2.2. Definición del Problema

En la lucha contra el Lavado de Dinero no se cuenta con herramientas de simulación para la detección de patrones en el comportamiento de los agentes dentro las redes de Lavado de Dinero. Como resultado no se

cuenta con información que permita mejorar políticas y procedimientos para la identificación temprana, acorde a los sectores económicos, de operaciones sospechosas y castigo posterior a los agentes que incurran en delitos. Las herramientas de Business Intelligence así como de IA cubren en su mayoría aspectos determinísticos: datos históricos y predicción a corto plazo, sumando a que todos los datos son confidenciales y protegidas por las leyes, hacen difícil la generación de herramientas que puedan medir y analizar los eventos aleatorios en las redes de Lavado de Dinero.

3. Modelado

3.1. Contexto

3.1.1. Anti-Money Laundering AML. El Lavado de Dinero inicia cuando por medio de un ilícito se produce algún activo del cual es necesario ocultar su procedencia. Se puede reconocer las siguientes etapas principales:

- *Colocación*: Mezclado con dinero lícito es llevado a una entidad financiera para adquirir un producto o servicio financiero.
- *Estratificación*: Dentro del sistema financiero comienza a ocultar su origen ilícito a través de transacciones que vulneran el control de anti-lavado.
- *Integración*: El dinero con apariencias legítima se inserta nuevamente a la economía para adquirir bienes y servicios, mantener vidas lujosas o para re-invertir en actos ilícitos.

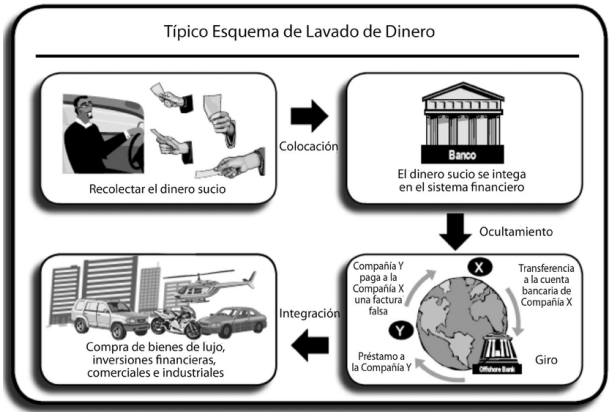


Figura 1. Proceso típico de un esquema de lavado de dinero.
Fuente: Oficinas de las Naciones Unidas

3.1.2. Fraude y Detección. Un fraude puede definirse como "un acto de engaño intencional o deshonesto perpetrado por uno o más individuos, que por lo general con el objetivo de obtener un beneficio económico". Prevención y detección están relacionados con el problema mencionado en la sección anterior, pero estos son dos conceptos diferentes; Prevención engloba políticas, procedimientos, entrenamiento y comunicación dirigido a prevenir que un fraude se concrete. Detección tiene como foco las actividades y técnicas que determinan acciones de forma oportuna si un fraude está sucediendo o por ocurrir. Mientras, la prevención no garantiza que el fraude no ocurra, pero es la primera línea de defensa que minimiza el riesgo de fraude. En la actualidad hay varios métodos de detección de fraude: extracción de conocimiento de los datos por medios estadísticos, inteligencia artificial y modelos de aprendizaje (Machine Learning). Estos métodos intentan descubrir actos fraudulentos basados en anomalías y patrones en las bases de datos.

3.1.3. Modelos de Red de Agentes y Detección. Desde el punto de vista económico; las teorías neo-clásicas ven al individuo (agente) como autor de decisiones independientes. Los agentes negocian en los mercados; los precios son el resultado del comportamiento de otros agentes y las personas afinan sus decisiones basado en la información generada por la dinámica de precios. La "Teoría de Juegos" es usado en la investigación de detección de fraude, debido a que permite el estudio de como las acciones individuales influencias el comportamiento de otros. Es en este punto donde las "Redes" tienen presencia, los cuales han sido tema de estudios empíricos y teóricos para investigar si este tipo de estructura puede influenciar en el comportamiento de sus integrantes. Las redes son útiles cuando los agentes interactúan con grupos pequeños de la población, es decir, cuando interactúan con sus vecinos. Modelado computacional basado en agentes encaja para el estudio de las redes. En el desarrollo del software de simulación, parte de este proyecto, se demuestra, que después de una programación compleja, que se puede visualizar de forma fácil la interacción entre miles de agentes.

3.2. Aplicando ABM a AML

A continuación la descripción de los principales componentes del modelado basado en agentes:

- **Agentes**: existen tres tipos de agentes. Primero; *individuos* comunes. Segundo; *empresas* con fines de lucro, empresas sin fines de lucro, empresas de fondos fiduciario[1] y empresas fantasma. Tercero; *intermediarios* financieros; formales como los bancos o informales como los "Hawala"[2]. Cada tipo de agente se representará como un "Autómata Finito No Determinístico" (AFND).
- **Ambiente**: La topología será una red tipo 'grafo no dirigido" donde cada nodo representa a un agente y cada enlace representa la conexión a sus vecinos más cercanos. Estos enlaces son necesarios para las operaciones de dinero que fluye en la red.
- **Comunicación**: un individuo o empresa solo se comunicará con otros individuos o empresas. Individuos y empresa se comunicarán con los intermediarios entidades financieros simulando poseer una cuenta o billetera.

3.2.1. Agente Individuo y Empresa. Los tipos de agentes individuales y empresas (con su 4 sub-tipos) se pueden modelar con mismo AFND dado que tienen los mismos estados. A continuación se describe el AFND asociado a este tipo de agentes:

- *Estados*:
  1. (q_1) Balance de situación/ estado inicial
  2. (q_2) En espera de alguna operación.
  3. (q_3) Realizando operación.
  4. (q_4) Cuentas congeladas [3].
- *Autonomía*: El agente individuo y empresa tienen una predisposición al fraude ("*predispositionFraud*"),

1. Fondos Fiduciario: cuando un agente cede sus bienes o dinero a una persona jurídica y este recibe un pago por acordado.

2. Hawala: Sistemas de transferencia informal de fondos.

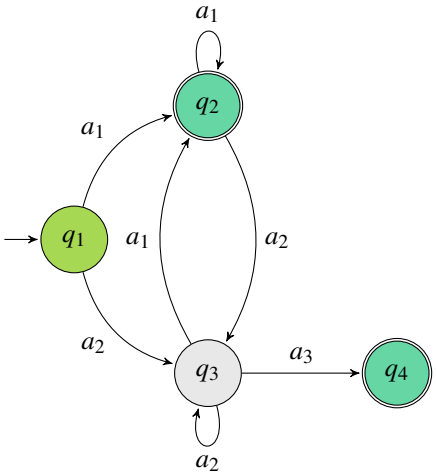
3. El agente no activa el evento de "congelar cuenta" lo hace la red. Sin embargo es un estado básico

además es responsable de su libro contable mayor ("*ledger*"). Con base en estos dos principales atributos toma decisiones para realizar operaciones legítimas o ilegítimas.
- *Atributos*: Sí cuenta esta cancelada ("*isFrozenAccounts*"), libro contable mayor ("*ledger*"), predisposición al fraude ("*predispositionFraud*"), ubigeo de origen ("*place*"), datos que corresponde al AFND.

Datos del AFND.
- Estados: $Q = \{q_1, q_2, q_3, q_4\}$
- Alfabeto: $\sum = \{a_1, a_2, a_3\}$
  - a_1 => Esperar operación.
  - a_2 => Realizar operación.
  - a_3 => Congelar cuenta.
- Palabras aceptadas: $\mathscr{L} = \{(a_1^+), (a_1^+ a_2^+ a_1^+), (a_1^+ a_2^+ a_3), (a_1^+ a_2^+ a_1^+ a_2^+ a_3), ...\}$
- Estado Inicial: q_1
- Estados de aceptación: $\mathscr{F} = \{q_2.q_4\}$
- Transiciones:

$\delta:$	a_1	a_2	a_3
q_1	$\{q_2\}$	$\varnothing$	$\varnothing$
q_2	$\{q_2\}$	$\{q_3\}$	$\varnothing$
q_3	$\{q_2\}$	$\{q_3\}$	$\{q_4\}$
q_4	$\varnothing$	$\varnothing$	$\varnothing$

- Diagrama de estados:

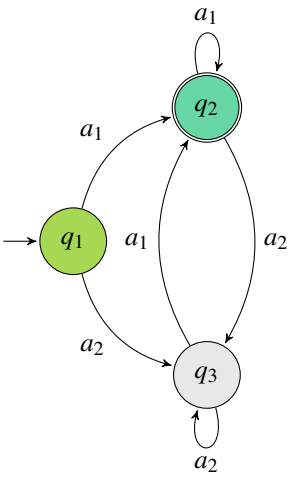


3.2.2. Agente Intermediario. El tipo agente intermediario se enlaza con individuos y empresas para simular una la existencia de una cuenta. A continuación se describe el AFND asociado a este tipo de agentes:
- *Estados*:
  1. (q_1) Actualización/creación de cuentas.
  2. (q_2) En espera de alguna operación.
  3. (q_3) Realizando operación.
- *Autonomía*: El intermediario está asociado a tipo de entidad financiera formal o informal.
- *Atributos*: Entidad financiera asociada ("*finantialEntity*", que puede ser formal o informal), ubigeo de origen ("*place*"), datos AFND.

Datos del AFND.
- Estados: $Q = \{q_1, q_2, q_3\}$
- Alfabeto: $\sum = \{a_1, a_2\}$
  - a_1 => Esperar operación.
  - a_2 => Realizar operación.
- Palabras aceptadas: $\mathscr{L} = \{(a_1^+), (a_1^+ a_2^+ a_1^+), (a_2^+ a_1^+), ...\}$
- Estado Inicial: q_1
- Estados de aceptación: $\mathscr{F} = \{q_2\}$
- Transiciones:

$\delta:$	a_1	a_2
q_1	$\{q_2\}$	$\varnothing$
q_2	$\{q_2\}$	$\{q_3\}$
q_3	$\{q_2\}$	$\{q_3\}$

- Diagrama de estados:



3.2.3. Ambiente.

Creación de la Red.
- Existen dos tipos de enlace: 1) de un intermediario a otros agentes y 2) entre individuos y empresas.
- Existe un parámetro (p.e. 40 %) para que cada intermediario esté enlazado con los otros intermediarios.
- Existe un parámetro (p.e. $[1,4]$) para que cada individuo y empresa se enlace con un cierto número de intermediario.
- Existe un parámetro, de una distribución exponencial de media 2, para que cada individuo y empresa se enlacen entre ellos (vecinos).
- Un individuo o empresa que tiene una predisposición mayor (p.e a 60 %) se le crea un enlace adicional con una empresa fantasma o de fondo fiduciario.
- Un individuo o empresa que tiene una predisposición al fraude mayor (p.e a 90 %) se le crea un enlace adicional con una empresa fantasma o de fondo fiduciario y otra con un agente (individuo o empresa) que tiene una predisposición al fraude mayor (p.e a 60 %).

Transacciones (Movimientos).
- Existe un parámetro (p.e. 5 %) para que individuos o empresa inicie una tracción (operación de dinero).
- Durante el proceso de una simulación el dinero de las transacciones se mantienen contabilizado.
- Existe un parámetro de monto máximo (p.e. S/ 100,000) para las transacciones que realicen los agentes.
- Cada agente origen elige, de sus vecinos cercanos (agentes enlazados), al destinatario. Existe un parámetro (p.e. 10 %) para que el agente elija a otro que no sea su vecino, por ende crea un nuevo enlace.
- Agente origen y destinatario deben estar enlazados al mismo intermediario o a diferentes pero que estén conectados, caso contrario el agente origen creará un enlace hacia un intermediario en común.

- Cada agente origen decide como estructurar la transacción, si ambos tienen baja predisposición o el monto es menor al umbral de "operación sospechosa" se envía todo el dinero en una sola transacción. Caso contrario el agente origen dividirá todo el monto en partes que no superen el umbral y creará tantas transacciones a otros agentes vecinos para hacer llegar todo el dinero al destinatario de manera indirecta. Si el agente origen no tiene el número de vecinos necesario para enviar el monto total, este creará nuevos enlaces con otros agentes (individuos o empresas).
- Se actualizará de manera constante una lista de observación de operaciones sospechosas (**Watchlist**)
- Cuando se creen los agentes en la Red, existe un parámetro (p.e. 10 %) para ponerlos en el Watchlist.
- Existe un parámetro (p.e. $[2-5]$) para congelar todas las cuentas agentes (individuos y empresas) cuando estas pasan mucho tiempo de manera consecutiva en el Watchlist.
- Para el umbral de operación sospechosa se asumirá S/ 9,000.
- Se realizará "N" experimentos, por cada una se ejecutará "D" simulaciones de transacciones, una simulación representa un día.
- Todas las transacciones se llegan a completar en el día, no se deja nada pendiente.

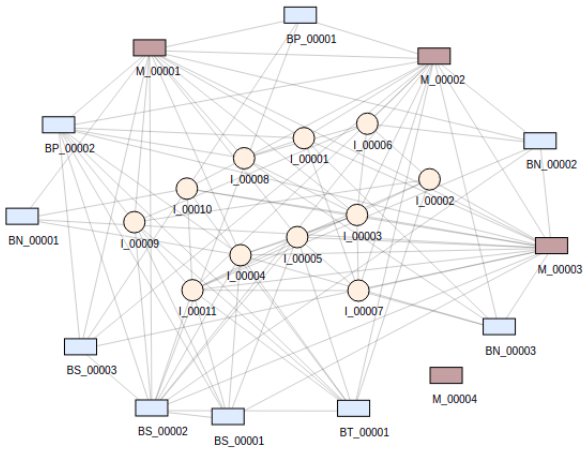


Figura 2. Ejemplo de red: individuos (círculo amarillo), empresas (cuadro celeste), intermediarios (cuadro marrón) y enlaces (línea gris).
Fuente: Creación propia.

4. Experimento y Resultados

4.0.1. Procedimiento. Para realizar el experimento seguiremos el procedimiento que se visualiza en la figura 3. La Fase de Preparación utiliza R para generar las muestras o casos de prueba *sampling*. La Fase de Ejecución utiliza un software preparado para este proyecto ver las figuras 5, 6, 4. La Fase de Análisis utiliza nuevamente R para carga los resultados obtenidos desde el PostgreSQL y generar el Outcome.

- *Generar Muestras LHS*: Utilizaremos la librería *"pse"* (en "R") para generar los casos de pruebas (muestras de LHS). Realizaremos un *sampling* de tamaño 200, con 21 parámetros del modelo. Luego se exporta a un CSV.
- *Iniciar Configuración*: La fase de ejecución empieza con la carga de todos los *sampling* generados en el paso anterior. Se cargan todo en una lista para luego, junto a parámetros fijos, ser leídos por la clase *Setup.ts*, que serán utilizados en los siguientes pasos.
- *Inicializar Agentes*: Con base en los parámetros de los *sampling* se genera las poblaciones para individuos, empresas e intermediarios.
- *Creación de la Red*: Se crean todas las conexiones, de acuerdo a los parámetros de los *sampling*.
- *Simulación de Transacciones*: Con ayuda del parámetro que indica un porcentaje de la población que inicia transacciones cada día.
- *Exportar Resultados*: Exportar el resultado del software de simulación, almacenado en un base de datos PostgreSQL, hacia una hoja de trabajo de R pasar generar el atributo del "Outcome".
- *Proceso de Datos*: Se realiza el análisis de sensibilidad PRCC con la ayuda de la librería *"pse"* .

4.0.2. Principales Parámetros.

4.1. Experimento

- Se realizó 200 (N) experimentos.
- En cada experimento se realiza 120 (D) simulaciones (días) de transacciones.
- Se realizó tres grupos de experimentos; en promedio se tomó 5.5 horas cada una.
- Debido a la cantidad de información generada se reducirá el caso de estudio con el siguiente **fitness function**: $ILD =$ Indicador de lavado de dinero, $DI =$ Dinero Ilegal, $DL =$ Dinero legal.

$$\%ILD = \frac{\sum_{i=1}^{N} DI_i}{\sum_{i=1}^{N} (DI_i + DL_i)} \quad (1)$$

Donde:
$ILD =$ Indicador de lavado de dinero
$DI =$ Dinero Ilegal
$DL =$ Dinero

5. Resultados

- En la figura 7, se puede visualizar la distribución acumulado del Outcome (vendría a ser el valor después de aplicar el *fitness funtion*).
- La tabla 5 muestra los valores de las correlaciones entre las variables y el Outcome. Se sombrea en amarillo los que representan mayor correlación:
  - *rangeMaxTimesWatchList* **[-0.227546195]**: El parámetro tiene correlación negativa con el Outcome. Es decir, el parámetro del tiempo que se debe considerar para que a un agente que está en el Watchlist se le congele sus cuentas, influye de manera negativa en las salidas. Ver la figura 9.
  - *rangeMaxTimesCleanWatchList* **[0.315445893]**: El parámetro tiene correlación positiva con el Outcome. Es decir, el parámetro del tiempo que se debe considerar para que a un agente que está en el Watchlist se le reinicie el conteo, influye de manera positiva en las salidas. Ver la figura 10.
  - *rangeMaxAmountTransaction* **[0.658932723]**: El parámetro tiene correlación positiva con el Outcome. Es decir, el parámetro del monto máximo de operación que elige el agente de manera aleatoria, influye de manera positiva en las salidas. ver la figura 11.

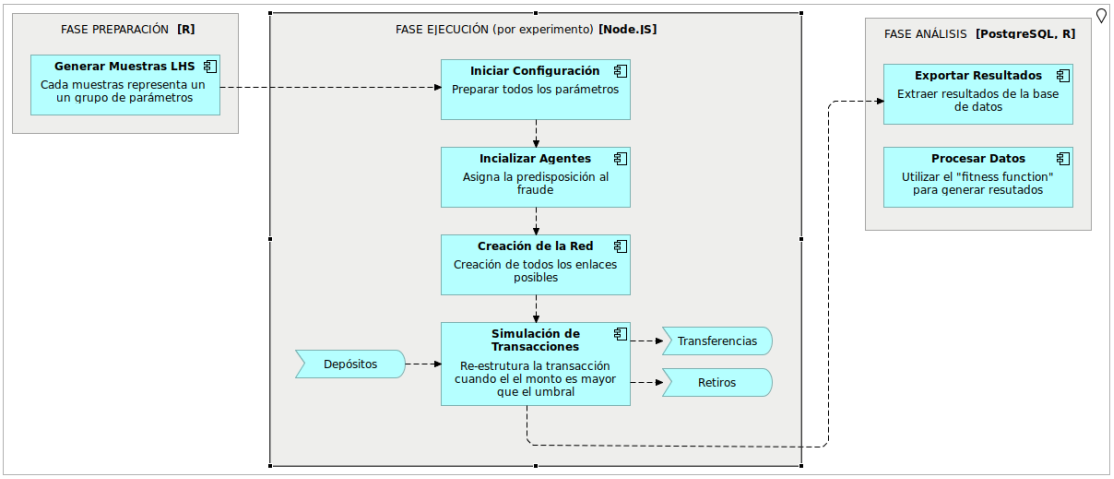


Figura 3. Procedimiento establecido para el experimento.

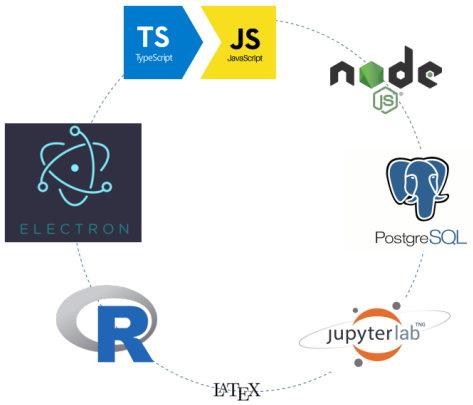


Figura 4. Tecnologías usadas en la construcción del software de simulación.

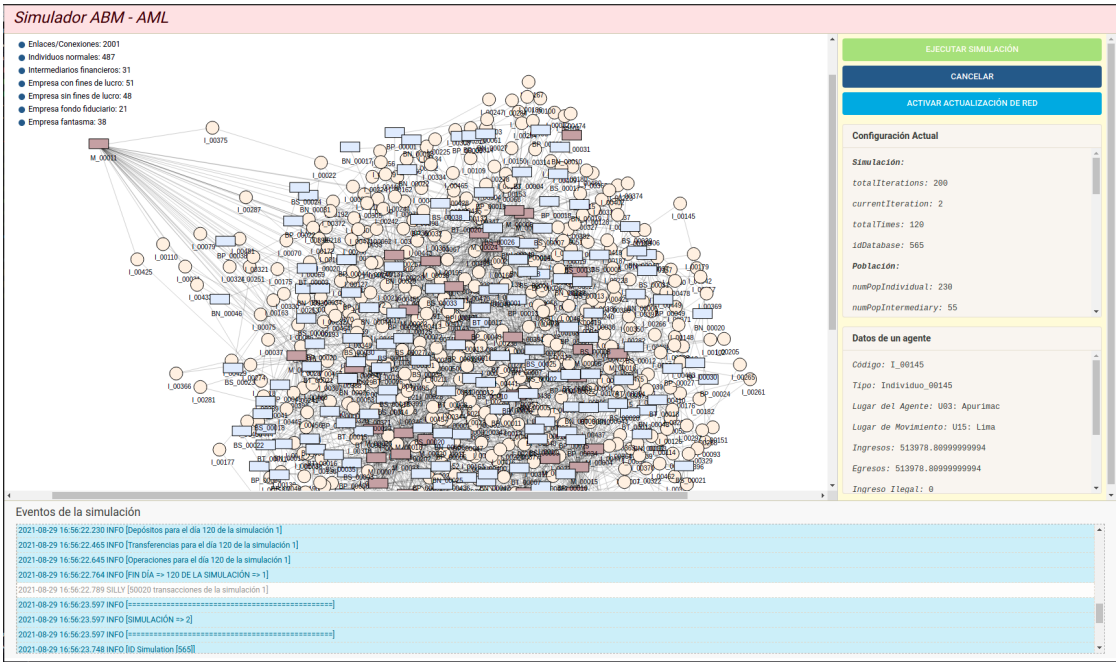


Figura 5. Muestra de una red con más de 600 agentes y 2000 conexiones.

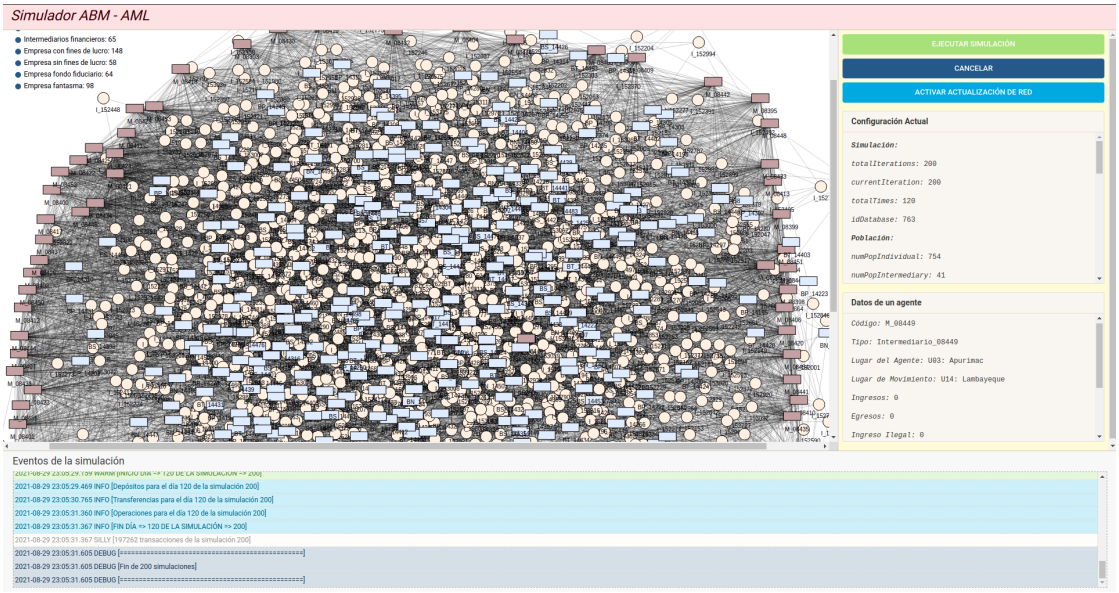


Figura 6. Muestra de una red con más de 1500 agentes y 9000 conexiones.

- *rangeHighPropensityFraud* **[0.829617586]**: El parámetro tiene correlación positiva con el Outcome. Es decir, el parámetro del valor alto de predisposición al fraude que elige el agente de manera aleatoria, influye de manera positiva en las salidas. Este parámetro es el que presenta mayor correlación y está sujeta a factores que no son susceptibles de medición, dado que es parte de la persona, ambiente, región, genética, psicología, estrés, economía, etc. Ver la figura 12.

- En la figura 8, de PRCC se puede visualizar la correlación del Outcome y los parámetros de entrada.

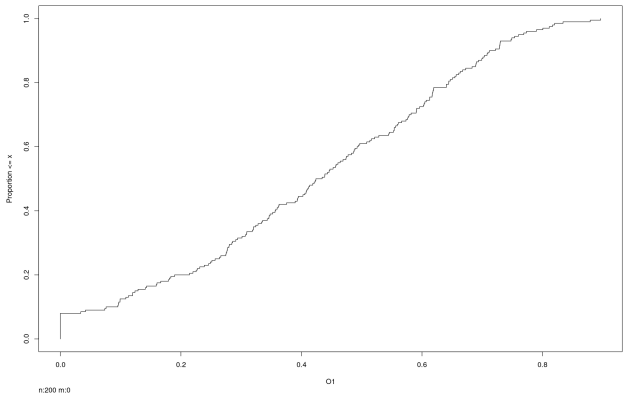


Figura 7. Distribución acumulada del Outcome.
Fuente: pse, PRCC.

Parámetro	PRCC
rangeMaxTimesWatchList	-0.227546195
rangeMaxTimesCleanWatchList	0.315445893
rangeMaxAmountTransaction	0.658932723
rangePropensityFraud	0.114238044
rangeLinkedIntermediary	0.011139089
rangeMaxLinkedNoIntermediary	0.104609436
rangeMaxLinkedIndBusInter	0.135285659
rangeExecuteDeposit	0.186759631
rangeExecuteTransfer	0.143579939
rangeExecuteWithdrawal	-0.020701911
rangeNewLinkTransact	-0.079030242
rangePopulation	0.068174119
rangeIndividual	0.009458012
rangeBusiness	-0.090407205
rangeIntermediary	-0.032441369
rangeNoProfitBusiness	0.027471308
rangeProfitBusiness	-0.073413330
rangeTrustBusiness	0.104533005
rangeShellBusiness	0.006335857
rangeHighPropensityFraud	0.829617586
rangeWatchList	-0.090226869

Cuadro 2.
TABLA DE VALORES DE CORRELACIÓN DESPUÉS DE APLICAR EL ANÁLISIS DE SENSIBILIDAD.

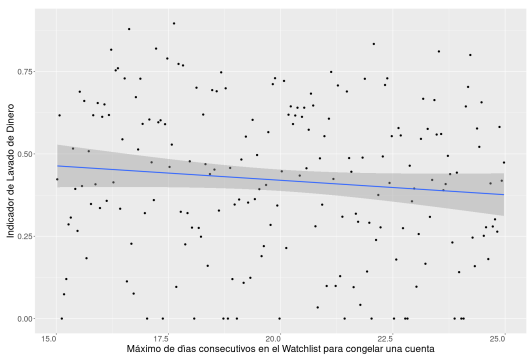


Figura 9. Correlación de rangeMaxTimesWatchList y el Outcome.

Orden	Descripción	Nombre	Valores LHS	Distribución	Valores Final	Pre-procesado
1	Número de experimentos	*totalIterations*	∅	∅	200	No
2	Número de días por simulación	*totalTimes*	∅	∅	120	No
3	Población total	Solo para LHS	$[200, 2000]$	*Uniforme*	De la muestra	No
4	Población de individuos	*numPopIndividual*	$[0.5, 0.8]$	*Uniforme*	∅	Sí
5	Población de intermediarios	*numPopIntermediary*	$[10, 74]$	*Uniforme*	∅	Sí
6	Población de empresas	Solo para LHS	$[0.1, 0.4]$	*Uniforme*	De la muestra	No
7	Población de empresa con fines de lucro	*numPopProfitBusiness*	$[0.1, 0.8]$	*Uniforme*	∅	Sí
8	Población de empresa sin fines de lucro	*numPopNoProfitBusiness*	$[0.1, 0.8]$	*Uniforme*	∅	Sí
9	Población de empresa de fondo fiduciario	*numPopTrustBusiness*	$[0.1, 0.8]$	*Uniforme*	∅	Sí
10	Población de empresa/negocio fantasma	*numPopShellBusiness*	$[0.1, 0.8]$	*Uniforme*	∅	Sí
11	Población de individuos con alta predisposición al fraude	*numPopHighPropensityFraud*	$[0.1, 0.8]$	*Uniforme*	∅	Sí
12	Porcentaje de población de individuos y empresas a incluir en el Watchlist	*perPopWatchList*	$[0, 0.3]$	*Uniforme*	∅	Sí
13	Máximo de días consecutivos en el Watchlist para congelar la cuenta de un individuo o empresa	*maxTimesWatchList*	$[15, 20]$	*Uniforme*	De la muestra	No
14	Máximo de días consecutivos que no esta en Watchlist para reiniciar el conteo	*maxTimesCleanWatchList*	$[2, 10]$	*Uniforme*	De la muestra	No
15	Predisposición de fraude a asignar a *numPopHighPropensityFraud*	*maxPropensityFraud*	$[0.6, 0.9]$	*Uniforme*	De la muestra	No
16	Alto valor de la predisposición al fraude para los enlaces extras	*maxHighPropensityFraud*	∅	∅	*min* $(maxPropenFraud + 0.2, 0.9)$	Sí
17	Porcentaje de enlaces entre intermediarios	*perLinkedIntermediary*	$[0.4, 0.8]$	*Uniforme*	De la muestra	No
18	Número de enlaces entre individuos y empresas	*numMaxLinkedNoIntermediary*	$[1, 16]$	*Uniforme*	De la muestra	No
19	Número de enlaces de individuos/empresas con los intermediarios	*numMaxLinkedIndBusInter*	$[2, 6]$	*Uniforme*	De la muestra	No
20	Porcentaje de individuos y empresas que realizan transacciones de tipo depósito	*perExecuteDeposit*	$[0, 0.3]$	*Uniforme*	De la muestra	No
21	Porcentaje de individuos y empresas que realizan transacciones de tipo transferencia	*perExecuteTransfer*	$[0.3, 0.7]$	*Uniforme*	De la muestra	No
22	Porcentaje de individuos y empresas que realizan transacciones de tipo retiro	*perExecuteWithdrawal*	$[0.3, 0.7]$	*Uniforme*	De la muestra	No
23	Umbral para determinar si es una operación sospechosa	*amountSuspiciousOperation*	∅	∅	9000	No
24	Valor máximo para ser seleccionado de manera aleatoria en una transacción	*rangeAmountTransaction*	$[500, 110000]$	*Uniforme*	[0, De la muestra]	Sí
25	Porcentaje para seleccionar de manera aleatoria un agente que no es vecino para una transacción	*rangeNewLinkTransact*	$[0, 0.2]$	*Uniforme*	De la muestra	No

Cuadro 1.
TABLA DE PARÁMETROS PRINCIPALES

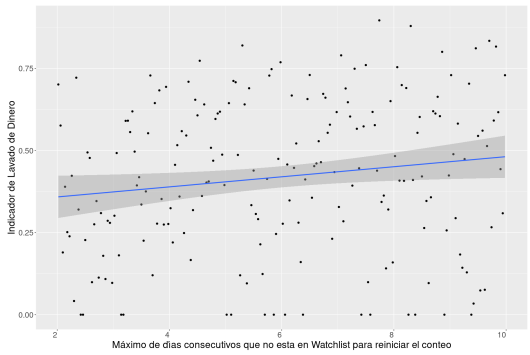


Figura 10. Correlación de rangeMaxTimesCleanWatchList y el Outcome.

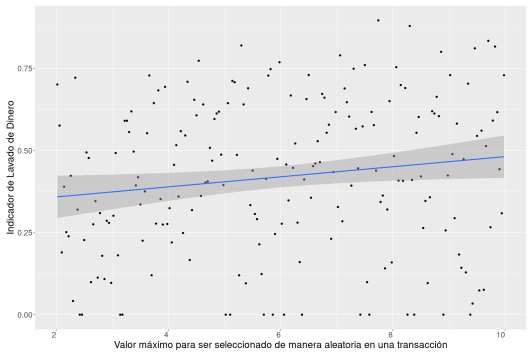


Figura 11. Correlación de rangeMaxAmountTransaction y el Outcome.

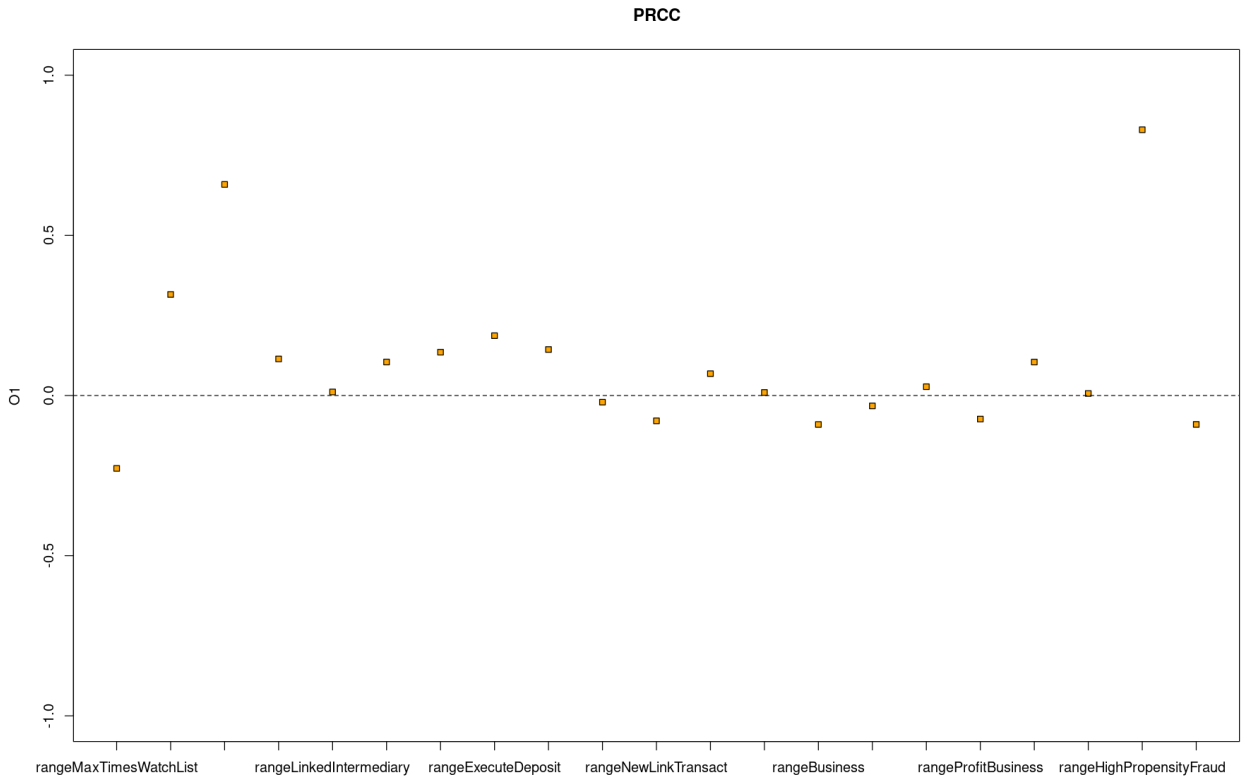


Figura 8. Partial Rank Correlation Coefficient.

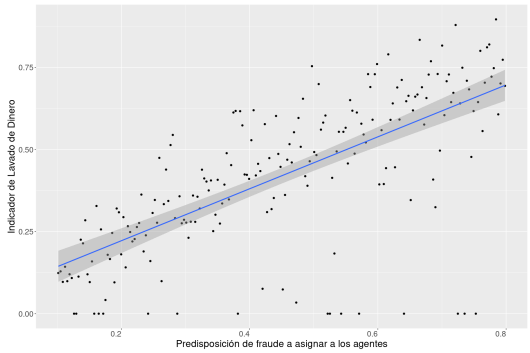


Figura 12. Correlación de rangeHighPropensityFraud y el Outcome.

6. Conclusiones

- AML es un problema complejo, en el desarrollo de este trabajo solo se consideró las principales variables. Sin embargo faltó incluir información de sectores económicos, predisposición al fraude por el lugar donde vive agente, entre mucho otros.
- En el presente trabajo se muestra que modelar AML bajo el enfoque de ABM, puede adelantar el análisis de sensibilidad de variables que son difíciles de medir con la data histórica o aún no son considerados bajo estudio. Esto sucede dado que la información generada ha sido sometida a la aleatoriedad de la incertidumbre sistémica que presenta AML.
- Contar con una herramienta, como el desarrollado en este proyecto, es de uso necesario cuando se quiere explorar más allá de los datos con los que se cuenta. Este tipo de enfoque sirven para generar datos sintéticos que ayuden a mejorar el rendimiento de los modelos de Machine Learning/IA.