

# Algoritmos Genéticos para estrategias de marketing en un modelo de comportamiento de consumo

Juan Francisco Robles Fuentes<sup>\*</sup>, Manuel Chica Serrano<sup>†</sup>, Óscar Cordón García<sup>\*</sup>

<sup>\*</sup> Dpto. Ciencias de la Computación e Inteligencia Artificial. Universidad de Granada.

<sup>†</sup> Dpto. Ciencias de la Computación - IN3. Universitat Oberta de Catalunya.

juanfranrobles@correo.ugr.es, manuel.chica.serrano@gmail.com,  
ocordon@decsai.ugr.es

**Abstract.** En este artículo utilizamos la metodología de modelado basada en agentes para simular procesos de compra en entornos de mercado. Extendemos un modelo basado en agentes incorporando conocimiento a los consumidores para hacerlo más próximo a la realidad. Así, los consumidores no parten de un conocimiento completo del entorno sino que deben adquirirlo por medio de procesos de boca a boca dentro de su red social. Adicionalmente, hacemos uso de algoritmos genéticos para generar estrategias de marketing viral basadas en el análisis de distintas métricas de las redes sociales. Finalmente, mostramos cómo las estrategias basadas en WOM se ralentizan debido al filtro añadido en el conocimiento de los agentes y cómo los algoritmos genéticos son capaces de encontrar buenas soluciones para focalizarse en los miembros más influyentes del mercado de acuerdo a la información contenida en la red social.

**Keywords.** Modelado basado en Agentes, Marketing Viral, Algoritmos Genéticos, Redes Sociales.

## 1 Introducción

El Marketing Viral (MV) se ha convertido en una de las piedras angulares del mundo del marketing [1] debido a su fuerza para alentar la adopción de productos mediante iteraciones consumidor-consumidor dentro de su red social [2].

El modelado del MV [15] es posible gracias al uso de redes sociales artificiales (RSAS) [3] y al uso de procesos de difusión de información [4] para el modelado de mecanismos de boca a boca<sup>1</sup> (WOM) [10], [11]. La introducción del WOM proporciona un mayor realismo a la difusión ya que permite a los agentes intercambiar conocimiento de forma realista [5]. Para ser capaces de modelar estos procesos necesitamos un sistema para simular interacciones entre consumidores a nivel de red como las simulaciones basadas en el modelado con agentes (MBA) [6], [7].

---

<sup>1</sup> Boca a boca: El boca a boca, Word of Mouth (WOM) en inglés, hace referencia al intercambio de opiniones entre individuos sobre un determinado producto, tendencia u opinión.

Con el uso del MBA es posible crear resultados a nivel de mercado permitiendo la incorporación de reglas de comportamiento simples y descartando fórmulas que simulen el comportamiento general del sistema [9].

En este artículo utilizaremos las técnicas comentadas anteriormente para simular procesos de compra (Sección 2) sobre distintas topologías de red social [8] como las redes libres de escala [13] y las redes de mundos pequeños [14] y estudiar cómo varía el comportamiento social en cada una de ellas (sección 4). También haremos uso de algoritmos genéticos [16] (sección 3) para generar estrategias automáticas basadas en MV [17] y detección de individuos influyentes para optimizar los costes y beneficios de campañas de venta simuladas. Para finalizar, expondremos nuestras conclusiones y trabajos futuros sobre este tema en la sección 5.

## 2 Nuestro modelo CONSUMAT extendido

### 2.1 Descripción del modelo CONSUMAT original

El modelo CONSUMAT es un MBA centrado en el comportamiento del consumidor desarrollado por [12]. Consiste en un simulador de mercados formado por  $N$  agentes y  $M$  productos en el que los agentes tienen un conocimiento completo de las opiniones de sus contactos a la hora de tomar decisiones [20] y comprar productos.

Las reglas principales de comportamiento de los agentes del modelo se resumen en los siguientes puntos:

#### 1. Utilidad ( $U_{ij}$ )

Refleja la utilidad total esperada (tanto individual como colectiva) de un agente  $i$  cuando consume un producto  $j$ .

#### 2. Incertidumbre ( $Unc_{ij}$ )

Refleja la convicción que muestran los agentes a haber tomado la decisión correcta al elegir un determinado producto.

#### 3. Procesos cognitivos y toma de decisiones

Los agentes pueden tomar diferentes decisiones dependiendo de su nivel de satisfacción e incertidumbre. Los umbrales que controlan estos comportamientos son  $U_{\min}$  (el nivel mínimo de satisfacción) y  $Unc_{\max}$  (el nivel de tolerancia a la incertidumbre), ambos estáticos para cada agente durante la simulación.

Cuando todos los agentes han tomado sus decisiones se actualizan los niveles de incertidumbre y satisfacción esperada. De esta forma, podemos distinguir entre el umbral esperado, que hace referencia a lo que los agentes experimentan mientras consumen un producto que está siendo comprado en su entorno, y el experimentado, que es fijo para cada agente.

Las siguientes cuatro heurísticas representan las acciones de los agentes y utilizan los umbrales mencionados anteriormente:

- Repetición (agente satisfecho y convencido de su elección [ $U_i \geq U \min \mid Unct_i \leq Unct$ ]): Los agentes siguen comprando el producto que compraron en su última compra.
- Deliberación (agentes no satisfechos pero convencidos de su elección [ $U_i < U \min \mid Unct_i \leq Unct$ ]): Los agentes comprueban cuál es la utilidad esperada de adquirir cada producto y utilizan una función logística para elegir su nueva compra.
- Imitación (agentes satisfechos pero no convencidos de su elección [ $U_i \geq U \min \mid Unct_i > Unct$ ]): Los agentes evalúan los productos que están siendo consumidos por su entorno. Aquellos productos consumidos por un mayor número de contactos serán más propensos a ser escogidos en la siguiente toma de decisiones.
- Comparación Social (agentes insatisfechos y no convencidos de su elección [ $U_i < U \min \mid Unct_i > Unct$ ]): Los agentes evalúan los productos consumidos en su entorno sin tener en cuenta aquellos que no han sido consumidos por lo que el conjunto de posibles decisiones se reduce.

## 2.2 Incluyendo conocimiento en el modelo CONSUMAT original

Para obtener un modelo de consumo más realista incorporaremos un filtro adicional al comportamiento de los consumidores basándonos en la racionalidad limitada de la conducta humana con el que tendrán limitaciones en el conocimiento de los productos del modelo. De esta forma, el modelo incorporará dos nuevas variables: la probabilidad de difusión, que permite a los agentes compartir parte de sus conocimientos en su RS y la pérdida de conocimiento con la que olvidarán parte de la información que tenían sobre los productos del mercado.

Las decisiones de los consumidores dependen ahora del conocimiento que pierden y adquieren haciendo uso del WOM dentro de su RS.

Utilizamos un mecanismo basado en probabilidades para modelar la adopción y pérdida de conocimiento:

- Probabilidad de difusión: cada agente  $i$  tiene una probabilidad de difusión dada por  $a_{ip} \in [0,1]$  para cada producto del mercado. Si este valor no supera un umbral de difusión delimitado por  $A_j$  el agente hablará sobre el producto  $j$  con cada agente  $i$  que pertenezca a su vecindario. Si el valor sobrepasa  $A_p$ , el agente no difundirá su conocimiento sobre el producto  $p$  en sus vecindarios.
- Pérdida de conocimiento: después del proceso de difusión, todos los agentes pueden olvidar parte del conocimiento adquirido sobre los productos ofertados en el entorno de compra. Para ello se establece un valor de pérdida de conocimiento  $d_{ij} \in [0,1]$  para todos los agentes y productos. Si este valor no supera el umbral de

pérdida de conocimiento  $D_j$ , el agente  $i$  olvidará la información que tenga sobre el producto  $j$ . En otro caso, el agente continuará manteniendo la información de dicho producto. Así, para valores altos de  $D_j$  los agentes serán más propensos a perder la información que adquirieron sobre los productos en la fase de difusión.

### 3 Marketing Viral basado en Algoritmos Genéticos

#### 3.1 Planteamiento del algoritmo

El objetivo de nuestro AG es encontrar el número óptimo de agentes influyentes en cada RS a partir de una serie de medidas que permitan maximizar la difusión de un producto para obtener unas buenas cifras en ventas y un coste reducido de la campaña.

Para comenzar con el planteamiento del algoritmo necesitamos determinar una manera de ponderar o medir cómo es de influyente cada uno de los individuos que forman las RSS. Para ello seguimos una adaptación de la aproximación definida por [24]. Así, valoramos el grado de influencia de cada agente combinando las siguientes tres medidas:

- Grado: el número de vecinos de un agente dado normalizado por el mayor grado de la red,  $grado(i) = \frac{grado(i)}{\max(grado)}$ . Un mayor grado de los nodos permite una mayor influencia de los mismos ya que son capaces de difundir información a un mayor número de individuos.
- Grado a 2 pasos: Es una medida relacionada con el número de nodos alcanzables recorriendo como máximo dos enlaces desde un nodo origen,  $2pasos(i) = \frac{2pasos(i)}{\max(2pasos)}$ . Es una extensión de la medida de grado.
- Coeficiente de agrupamiento (CA): es una medida del grado de agrupación de los nodos de un grafo. En la mayoría de las redes del mundo real, los nodos tienden a crear grupos muy conectados entre sí que se caracterizan por una gran densidad de enlaces entre ellos [13].

Cada cromosoma del AG definirá una combinación de estas tres medidas para formar una solución a la detección de agentes influyentes de acuerdo a las características de distintas topologías de RS. Las principales variables de decisión de nuestro algoritmo son el número de personas influyentes del modelo junto con los pesos asociados a la combinación lineal de las tres medidas definidas anteriormente.

Así, se define la siguiente codificación para cada cromosoma de la población utilizada en el AG:

$W_d$	$W_{2s}$	$W_{cc}$	$F_s$
-------	----------	----------	-------

Donde:

- $W_d \in [0,1]$  representa el peso asociado al grado de los nodos.
- $W_{2s} \in [0,1]$  representa el peso para el número de consumidores alcanzables recorriendo dos enlaces desde un nodo origen.
- $W_{cc} \in [0,1]$  representa el peso asociado al coeficiente de agrupamiento de los nodos.
- $F_s \in [1, Max_{agentes}]$  representa el número de agentes que actúan como influyentes en la simulación.

Utilizando los valores anteriores definimos la siguiente relación lineal para ponderar la importancia total de cada agente:

$$W_{agente_i} = W_d grado_i + W_{2s} 2pasos_i + W_{cc} cc_i$$

Haciendo uso de la fórmula anterior podemos generar los  $F_s$  primeros agentes de la población como influyentes y recompensarlos.

### 3.2 Etapas del algoritmo genético

La primera inicialización de la población se realiza de forma aleatoria y se evalúa simulando CONSUMAT con los valores dados por el cromosoma.

Para obtener el fitness del cromosoma simulamos el MBA recompensando a los  $F_s$  agentes más influyentes basándonos en los valores de  $W_d$ ,  $W_{2s}$  y  $W_{cc}$  codificados en cada cromosoma y calculando el número de ventas conseguidas al final de la ejecución y el coste de la campaña. Utilizaremos la siguiente expresión para calcular los beneficios obtenidos en la campaña (NPV), tanto para los nuevos consumidores del producto difundido como para los agentes difusores:

$$NPV = \sum_{t=1}^{\infty} a(t) p \lambda^t$$

Donde  $a(t)$  representa el número de compradores en el instante  $t$ ,  $p$  es el beneficio obtenido de la adopción de los nuevos productos (un precio ficticio utilizamos donde asignamos 1 para los nuevos consumidores y 1/10 para los agentes escogidos como difusores) y  $\lambda$  es el factor de descuento aplicado (0.8 tal y como se utiliza en [17]).

Así, el valor de fitness para cada cromosoma será el NPV obtenido con los nuevos compradores menos el NPV de los agentes influyentes.

$$Fitness = NPV_{nuevosCompradores} - NPV_{difusores}$$

Para la generación de la población de hijos seguimos una estrategia de tipo estacionario en la que  $k$  padres ( $k = 2$ ) participan en un torneo aleatorio. Esto significa que seleccionamos dos padres en cada iteración y compiten, siendo el ganador aquel padre con mayor valor de fitness.

Tras identificar a los dos miembros de la población  $p'$ , cruzamos los cromosomas con probabilidad  $p_c \in [0,1]$  utilizando un operador de cruce  $BLX - \alpha$  [18].

Una vez cruzados, cada nuevo cromosoma tendrá una probabilidad  $p_m \in [0,1]$  de mutar uno de sus genes de forma aleatoria dentro del rango de valores de dicho gen.

Por último, tras obtener los nuevos cromosomas que forman la población  $p'$ , éstos competirán con la población  $p$  de padres para reemplazar a dos individuos de la misma. La estrategia de reemplazo se basa en comparar el fitness de los cromosomas de la población de hijos con los dos peores de la población de padres. Si los hijos superan a los padres los reemplazan en la nueva población.

## 4 Experimentos

### 4.1 Inicialización de los experimentos

La tabla I muestran los parámetros para las simulaciones en el CONSUMAT original y en el extendido así como los utilizados en el AG.

**Tabla 1.** Parámetros para el modelo CONSUMAT original y extendido. Parámetros del AG.

<b>Parámetros del modelo original</b>	
<b>Probabilidad de Compra</b>	0.1
$\beta$	0.6
$U_{\min}$	0.4
$U_{\text{net}}$	0.2
<b>Ejecuciones Monte Carlo</b>	15
<b>Parámetros del modelo extendido</b>	
<b>Probabilidad de compra</b>	1.0
<b>Probabilidad de difusión</b>	0.4
<b>Prob. Pérdida de información</b>	0.1
<b>Parámetros utilizados en el AG</b>	
<b>Número de ejecuciones</b>	5
<b>Población</b>	30
<b>Probabilidad de cruce (<math>P_c</math>)</b>	0.7
<b>Probabilidad de mutación (<math>P_m</math>)</b>	0.1

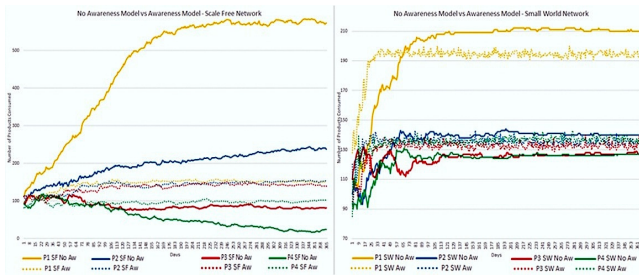
La probabilidad de compra determina el número de agentes que actúan como consumidores (hablan y compran) en cada etapa del modelo. Utilizamos los mismos parámetros para cada tipo de RS utilizada en los experimentos (1000 agentes, un grado medio de 10 para la red y 365 días de compra). Para la red libre de escala utilizamos un valor de 3 para el número de posibles conexiones para cada nodo durante la creación de la red [19]. Para la red de mundos pequeños escogemos 0.2 como valor para la probabilidad de reconexión.

## 4.2 Comparación entre el CONSUMAT original y el extendido.

Se consideran dos modelos de RS (libre de escala y mundos pequeños) para comparar el modelo CONUMAT original con el extendido con los parámetros descritos anteriormente. Se analizará el comportamiento de los agentes en la compra de productos y en su toma de decisiones en ambos modelos.

### Red Libre de Escala.

La figura 1 (izqda.) muestra la evolución de las ventas en el modelo sobre una red Libre de Escala. En este caso, los consumidores tienden a sesgarse por la compra de un único producto que termina liderando el mercado. Cuando el conocimiento es limitado, el comportamiento cambia y los productos consumidos varían (se consume un conjunto de los mismos).



**Fig. 1.** Evolución de ventas para los 4 productos más vendidos en redes libres de escala (izqda.) y de mundos pequeños (dcha.) en CONSUMAT original (línea sólida) y extendido (línea de puntos)

Cuando se limita el conocimiento la difusión de información se produce de una forma más gradual siendo los nodos de mayor grado los que ejercen una mayor influencia en los procesos de WOM. Aquí, el tiempo necesario para difundir un producto es mayor que en el modelo simple y las ventas se reparten de una forma más equitativa.

La figura 2 muestra la evolución del uso de heurísticas en los dos modelos. Se observa que los agentes comienzan a utilizar de forma masiva heurísticas que implican un cambio en el producto comprado para terminar repitiendo su última compra. En nuestro modelo los agentes requieren una mayor interacción con su entorno para satisfacer sus necesidades y como puede verse en la imagen 2 (dcha.), la heurística de deliberación adquiere un mayor protagonismo frente a otras como imitación o comparación social.

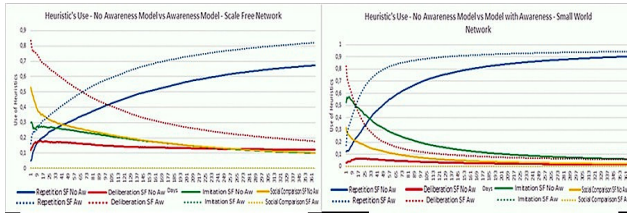


Fig. 2. Evolución del uso de heurísticas en redes Libres de Escala (izqda.) y de mundos pequeños (dcha.) para CONSUMAT original (línea sólida) y extendido (línea de puntos)

### Experimentos sobre la red de Mundos Pequeños

En la figura 1 (dcha.) se observa que el comportamiento de los dos modelos es muy similar. El producto más difundido en la RS también es el más vendido. Sin embargo, en este caso la difusión de información es más lenta debido a que esta topología de red tiene una distribución de grados menor que la libre de escala.

Con respecto al uso de heurísticas, podemos ver en la figura 2 (izqda.) cómo los agentes pasan de utilizar imitación, comparación social y repetición en el modelo original a deliberación y repetición en el extendido.

### 4.3 Aplicación de AGS para generar estrategias de marketing viral en el modelo CONSUMAT extendido

En esta sección mostramos los resultados de la aplicación de AGS para optimizar campañas de marketing.

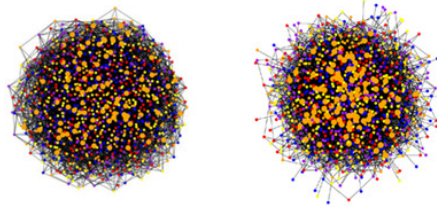
Tabla 2. Resultados del AG en redes Libres de escala y de mundos pequeños.

Tipo de Red Social	Valores del cromosoma				Fitness	
	Wd	W2s	Wcc	Fs	Beneficios	Ventas
Libre de escala	0.5292	0.6703	0.6168	82	2678.651	621
Mundos pequeños	0.355	0.629	0.122	99	718.4988	948

Los resultados obtenidos con el AG, tabla 2, muestran que no siempre son los agentes con grado alto los más útiles para las campañas de difusión sino que es la topología de la RS y las características de los consumidores de la misma los que influyen de forma más notable en la identificación de estos agentes.

Como puede verse en la figura 3, en la mayoría de los casos, los productos comprados por los hubs (nodos de mayor grado de la red) son los que terminan liderando el mercado en detrimento de los productos comprados por consumidores de menor grado.





**Fig. 3.** Estrategias de MV encontrada por el AG en la red libre de escala (izqda.) y en la red de mundos pequeños (dcha.). Los nodos de mayor tamaño representan a los difusores mientras que los de menor tamaño son consumidores normales. Cada color representa un producto distinto.

En la figura 3 podemos observar que las estrategias generadas son muy diversas optando unas veces por ponderar más a nodos de grado alto y otras por centrarse más en otras medidas o combinarlas entre sí.

## 5 Conclusiones

Limitar el conocimiento en CONSUMAT nos ha proporcionado un comportamiento de consumidor más cercano a la realidad. Los patrones de consumo utilizados por los agentes resultan ser más variados y su uso se produce de forma pausada y gradual.

Las características de cada topología de RS son percibidas de una forma más notable en nuestro modelo extendido ya que medidas como el grado o el coeficiente de agrupamiento cobran una mayor importancia durante los procesos de difusión de información.

Por último, el uso de AGS para generar estrategias de MV resulta ser muy útil para identificar diferentes criterios con los que identificar individuos importantes en distintas topologías de RS.

Futuros estudios podrían enfocarse en la aplicación de estrategias de optimización multi-objetivo para extender la generación de campañas de MV [19] y en la aplicación de técnicas de computación inteligente para calibrar MBA a partir de datos históricos.

## 6 Agradecimientos

Este trabajo ha sido apoyado por el Ministerio Español de Economía y Competitividad bajo el proyecto NEWSOCO (ref. TIN2015-67661-P), incluido en el Fondo Europeo de Desarrollo Regional (FEDER).

## 7 Bibliografía

1. M. Trusov, R. Bucklin and K. Pauwels, "Effects of word-of-mouth versus traditional marketing: Findings from an internet social networking site," *Journal of Marketing* 73, pp. 90-102, 2009.
2. B. Libai, R. Bolton, M. Bügel, K. De Ruyter, O. Götz, H. Risselada and A. Stephen, "Customer-to-customer interactions: broadening the scope of word of mouth research," *Journal of Service Research* 13, p. 267-282, 2010.
3. D. Watts and P. Dodds, "Influentials, networks, and public opinion formation," *Journal of Consumer Research* 34, p. 441-458, 2007.
4. E. Rogers, Diffusion of innovations, 5 ed., S. a. Schuster, Ed., New York: Free Press, 2003.
5. D. Duzevic, A. Anev, P. Funes and P. Gaudiano, "The Effects of Word-of-Mouth: An Agent-Based Simulation of Interpersonal Influence in Social Networks," *Word of Mouth Research Symposium*, 2007.
6. W. Rand and R. Rust, "Agent-based modeling in marketing: Guidelines for rigor," *International Journal of Research in Marketing* 28, p. 181-193, 2011.
7. J. Epstein, Generative social science: Studies in agent-based computational modeling, Princeton University Press, 2006.
8. M. Newman, A. Barabási and D. Watts, "The structure and dynamics of networks," *Princeton University Press*, 2006.
9. E. Bonabeau, "Agent-based Modeling: Methods and Techniques for Simulating Human Systems," *Annual Review of Sociology*, pp. 7280-7287, 2002.
10. J. Goldenberg, B. Libai and E. Muller, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing letters* 12, p. 211-223, 2001.
11. B. Libai, E. Muller and R. Peres, "Decomposing the value of word-of-mouth seeding programs: Acceleration versus expansion", *Journal of Marketing Research* 50, p. 161-176, 2013.
12. M. A. Janssen and W. Jager, "Simulating Market Dynamics: Interactions between Consumer Psychology and Social Networks," *Artificial Life*, pp. 344-355, 2003.
13. D. J. Watts and S. H. Strogatz, "Collective dynamics of 'smallworld' networks," *Nature* 393, p. 440-442, 1998.
14. R. Albert and A. Barabási, "Emergence of scaling in random networks," *Science* 286, pp. 509-512, 1999.
15. J. Leskovec, L. Adamic and B. Huberman, "The Dynamics of Viral Marketing," *ACM Transactions on the Web*, 2007.
16. D. Goldberg and J. Holland, "Genetic algorithms and machine learning," *Machine Learning* 3, pp. 95-99, 1988.
17. F. Stonedahl, W. Rand and U. Wilensky, "Evolving Viral Marketing Strategies," *GECCO*, 10, pp. 7-11, 2010.
18. F. Herrera, M. Lozano and J. Verdegay, "Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis," *Artificial Intelligence Review* 12, pp. 265-319, 1998.
19. C. Coello, G. Lamont and D. Van Veldhuizen, Evolutionary Algorithms for Solving Multi-objective Problems, 2, Ed., Springer, 2007.
20. J. Bettman, E. Johnson and J. Payne, "Consumer decision making," *Handbook of consumer behavior* 44, pp. 50-84, 1991.