

## SISTEMA MULTIAGENTE PARA ENLACE DE INTERESES

Erick Muñiz Torres

José Gpe. Escamilla de los Santos

{emuniz, jescamil}@campus.ruv.itesm.mx

Instituto Tecnológico y de Estudios Superiores de Monterrey

Av. Eugenio Garza Sada 2501 sur Edificio CETEC torre sur

64849 Monterrey N.L. Méx.

Tel. (52 - 8)358 - 2000 Ext. 6616

Fax (52 - 8) 328 - 4055

### RESUMEN

La motivación para la realización de este trabajo, es el lograr el fortalecimiento de la vinculación y difusión de los trabajos desarrollados por los investigadores, de manera que se conozcan sus investigaciones y resultados de forma automática y en el momento en que ellos desean compartirlos, sin necesidad de esperar a que sean publicados en algún congreso o revista especializada. Se presenta en este artículo una breve introducción a los sistemas de filtrado de información, el desarrollo del trabajo, los resultados obtenidos, las conclusiones y el trabajo futuro.

### 1. INTRODUCCIÓN

En la actualidad la sociedad se encuentra en una transición hacia una sociedad de información, en la que, la continua digitalización y el crecimiento de las redes electrónicas en los negocios, gobierno, escuelas, etc. forman las bases de dicha sociedad [Brennen, Zarnekow, Wittig, 1998]. El establecimiento de dichas redes permiten obtener y proveer información digitalizada entre los usuarios [Brennen, Zarnekow, Wittig, 1998].

Los retos fundamentales de la búsqueda de información en Internet, son los

mismos que buscar en el catálogo de una biblioteca. El usuario indica el tema que le interesa y se busca que éste se encuentre dentro de un catálogo con la descripción de los documentos. Los documentos que más se parezcan al tema del usuario son los que se le presentan [Williams, 1996].

Internet ha tenido resultados satisfactorios para apoyar el trabajo en grupo ya que tiene el potencial necesario para una gran colaboración social.

Dentro de los esfuerzos que se han realizado para fortalecer la vinculación y colaboración entre grupos de trabajo por medio de Internet se pueden mencionar entre otros a los sistemas multiagentes de los cuales destacan los filtradores sociales de información [Shardanand, Maes, 1995].- También llamados sistemas sociales descubridores de recursos. Facilitan la distribución de información útil de personas que son buenas para encontrar cosas interesantes en la red hacia personas que no son tan buenas buscando información.

Los sistemas de filtrado sociales utilizan las opiniones de otras personas para seleccionar la información más valiosa de las colecciones existentes.

Sin embargo, este tipo de agentes requieren del trabajo de otras personas y éstas desafortunadamente no hacen el trabajo altruista extra para diseminar la información

que tienen especialmente cuando dichas personas son expertas con bastantes ocupaciones más [Starr, Ackerman, Pazzani, 1996].

## 2. DESARROLLO

Como solución se propone un sistema de agentes que cumpla con las siguientes características:

- Efectúe la búsqueda y vinculación entre personas.
- Permita la colaboración social para que los agentes con intereses afines puedan tener acceso a la información que los otros agentes han ido almacenando, esto con la finalidad de realizar la divulgación de artículos de investigación.
- El aprendizaje del perfil de usuario se ejecutará exclusivamente en los artículos que el usuario desee compartir tanto a los demás usuarios como al agente, pudiendo conservar la información que se desee mantener privada aún para el agente.
- El agente deberá obtener las áreas de interés del usuario con la mínima intervención del dueño para completar su actividad.
- El agente de usuario será capaz de poder manejar más de un área de interés del usuario, esto con el fin de no tener varias copias del agente para cada uno de los intereses que se tengan.
- El sistema deberá asegurar que los artículos que un agente desee compartir lleguen a todos los posibles interesados en ellos aún y cuando los interesados no estuvieran registrados en el sistema al momento de que se compartieron los artículos.

Los agentes que componen el Sistema Multiagente de Enlace de Intereses se dividen en agentes de servicio al usuario, los cuales son los encargados de interactuar directamente con el usuario y se encuentran almacenados dentro de la computadora del usuario y los agentes de filtrado de información, los cuales se encuentran en los servidores, estos son los encargados de hacer el enlace de investigadores afines y de filtrarles la información que le puede ser de interés a los distintos usuarios.

La arquitectura del sistema multiagente se compone de cuatro agentes, de los cuales dos serán para servicio al usuario y dos para filtrar los artículos.

Para el servicio al usuario se proponen los siguientes agentes [Muñiz, Escamilla, García, 1998]:

- Agente de Usuario. Este agente es el encargado de la interacción entre el usuario y el ambiente, además de coordinar a los demás agentes de servicio al usuario.
- Agente de Aprendizaje. Este agente aprenderá las áreas de interés del usuario revisando los artículos y las páginas Web a compartir. Cada vez que el usuario comparta un nuevo artículo o un nuevo apuntador(bookmark), el Agente de Usuario activará al Agente de Aprendizaje para que realice su tarea. Una vez que el Agente de Usuario confirma el interés, el Agente de Aprendizaje envía el interés al Agente de Intereses.

Para el filtrado de artículos se proponen dos agentes [Muñiz, Escamilla, García, 1998]:

- Agente de Artículos. Se encarga de almacenar y distribuir los artículos que

los distintos usuarios han querido compartir.

- Agente de Intereses. Este agente es encargado de almacenar y distribuir los intereses y las direcciones de los distintos usuarios.

La base de conocimientos del sistema fue modelada utilizando el modelo entidad - relación y construida utilizando el modelo relacional de bases de datos.

El modelado del comportamiento de los agentes se efectuó utilizando Redes de Petri. Se utiliza este modelo porque se aprovechan las siguientes propiedades de las Redes de Petri [Brena, 1998]:

- No - determinismo. En un momento dado puede disparar cualquiera de las transiciones cuyos lugares de entrada tienen todos marca.
- No - simultaneidad. Solamente puede disparar una transición a la vez.
- Paralelismo o Concurrencia. La ejecución de dos o más partes de una Red de Petri puede darse en forma traslapada o concurrente.

Estas propiedades permiten que el agente pueda ejecutar una acción en cualquier momento y que realice varias acciones en paralelo, sin embargo, las acciones no se desarrollen simultáneamente.

El sistema está implementado en el lenguaje de programación JAVA y se utilizó el paquete de clases JATLite para desarrollar los métodos de comunicación [JATLite, 1997].

Se tomó la decisión de utilizar este paquete debido a las siguientes ventajas proporcionadas por el ruteador de mensajes que este paquete provee:

- Se almacenan los mensajes en una cola por si el agente receptor no se encuentra conectado o no puede atender el mensaje por estar realizando otras acciones. Con esto, los mensajes no pierden la secuencia de cómo fueron recibidos [JATLite, 1997].
- Todos los mensajes son almacenados en un archivo. Cuando el agente receptor lo desea puede recoger y borrar los mensajes que ya atendió [JATLite, 1997].

La coordinación de cada uno de los agentes está implementada dentro del mismo nivel de comportamiento del agente, esto porque al modelar el comportamiento se modeló también la coordinación para desarrollar las acciones.

El algoritmo de coordinación básicamente consiste en que una vez que se registra el agente al ruteador de mensajes, la red de Petri es inicializada y el método para realizar el escaneo del puerto es activado. Una vez realizado esto la red de Petri se encarga de realizar la coordinación de acuerdo al mensaje que se recibe.

Para la coordinación entre los agentes que componen el sistema se desarrolló un protocolo propio de coordinación, el cual tiene las siguientes características:

- La conversación se inicia con el simple envío de la tarea a ejecutar.
- Si es una tarea de consultas, el agente emisor da por hecho que la conversación se estableció.
- Si el agente receptor no envía una respuesta, el agente emisor supone que su consulta no tenía respuesta.
- Si la tarea es de insertar datos a la base de conocimientos de un agente, el agente emisor espera confirmación por parte del

agente receptor de que la tarea ha sido realizada.

- Si el receptor es el Agente de Usuario y lo que recibe es un artículo por parte del Agente de Artículos, el Agente de Usuario esta obligado a confirmar de su recepción para que dicho artículo no le vuelva a ser enviado, si esta confirmación no se realiza, podrá recibir indefinidamente el mismo artículo.

El protocolo desarrollado permitió tener coordinación entre los agentes al establecer una conversación con un número bajo de mensajes para iniciar la comunicación, pues, se da por hecho que la conversación se establece en la mayoría de los casos, y sólo se espera hasta que el agente receptor indique que la tarea ha sido realizada cuando es una tarea que requiere de confirmación para que no se vuelva a ejecutar.

En lo que corresponde a la obtención de los intereses del usuario se decidió utilizar un método de aprendizaje que pudiera extraer las frases claves de un documento y dependiendo de la frecuencia de una frase decidir si era área de interés del usuario.

Se encontró una librería llamada Extractor [Turney, 1997, Turney, 1999], la cual, realiza la conversión de html a texto y extrae frases significativas del documento simplificando dos pasos del proceso. Se desarrollaron pruebas para verificar su funcionamiento y se decidió utilizarla debido a la efectividad de la herramienta.

Una vez que el Agente de Aprendizaje ha recibido la orden de aprender por parte del agente de usuario se inicia el siguiente algoritmo:

- Abre el archivo bookmark.htm del cual se indicó su ruta completa al configurar

el usuario sus agentes de servicio al usuario. El archivo bookmark es generado por el navegador Netscape Navigator y en él se encuentran almacenadas todas las ligas de interés del usuario.

- Buscar carpeta “**Compartir**”.
- Mientras no se encuentre fin de carpeta hacer.
- Obtener dirección de Internet.
- Obtener contenido de la página de Web.
- Guardar contenido en un archivo de trabajo (**temporal.tmp**).
- Aplicar algoritmo de extracción de frases significativas (extrae cinco frases significativas).
- Guardar frases significativas en archivo de trabajo (**thinking.tmp**).
- Fin mientras.
- Abrir archivo **thinking.tmp**.
- Aplicar algoritmo de extracción de frases significativas (extrae tres frases significativas).
- Si el peso de la frase significativa es mayor que 30 guardar la frase en **Result.dat**

### 3. RESULTADOS

Con la finalidad de obtener resultados que pudieran proporcionar información acerca de la funcionalidad tanto de la identificación de los intereses del usuario como de la arquitectura propuestos, se realizaron una serie de experimentos, los cuales, se dividieron en dos áreas: obtención de intereses y funcionamiento general del sistema.

En lo correspondiente a obtención de intereses estas son las condiciones bajo las cuales se realizaron los experimentos para ver las reacciones del agente a las alteraciones:

- Las páginas se almacenaron en la carpeta Compartir en un rango entre 1 a 200 direcciones.
- Se usaron páginas de Internet tomadas al azar.
- Se utilizaron relacionadas con diversos tipos de intereses.
- El tamaño de los artículos fue variable, desde un párrafo hasta un capítulo de un libro.

Se realizaron un total de 49 experimentos modificando el rango de páginas que el Agente de Aprendizaje utilizaría para obtener los intereses del usuario, los experimentos consistieron en medir el tiempo requerido para detectar los intereses y el porcentaje de pertinencia de los resultados obtenidos con relación a los intereses reales del usuario.

Los resultados que se obtuvieron fueron los siguientes:

- El tiempo que tarda el agente en detectar los intereses está relacionado directamente con el número de páginas que tiene que explorar, de tal forma que a mayor número de páginas el tiempo se incrementa.
- Cuando el número de páginas es muy pequeño el agente no es capaz de detectar exactamente el interés que se desearía encontrar, esto debido al algoritmo que utiliza para su aprendizaje. Cuando se realiza el segundo procesamiento a las frases significativas de los documentos las frecuencias de las mismas son muy bajas y no logra detectar el interés.
- Conforme aumenta el número de páginas el porcentaje de pertinencia se incrementa, sin embargo, llega el momento en que el incremento en el porcentaje se detiene, los problemas

observados principalmente son las ligas rotas, páginas desarrolladas en un lenguaje distinto al HTML, ligas hacia páginas con otro tipo de archivos, etc.

- En el rango de 5 a 20 direcciones de artículos, el agente logró detectar los intereses con mayor claridad, siendo este rango el que alcanzó mayor grado de pertinencia a los intereses del usuario.
- A partir de 20 páginas en adelante no se observaron cambios significativos en cuanto a lo que el agente de aprendizaje detecta como área de interés, sin embargo, el tiempo de aprendizaje se hace mucho mayor ya que el agente tiene que establecer una conexión con cada uno de los sitios que se le están presentando, es recomendable utilizar sólo hasta 20 ligas para que el proceso no sea tan largo y asegurar que el interés obtenido será lo más aproximado al del usuario.
- Por último destacar el hecho que el agente está habilitado para detectar máximo tres intereses en cada sesión de aprendizaje, esto ocasiona dos connotaciones, cuando sólo se tiene uno o dos intereses el agente de cualquier manera presentará tres opciones dejándole la autoridad al usuario de aceptar, eliminar o corregir los intereses. Por otra parte, si se desea que el agente detecte más de tres intereses se recomienda agregar o cambiar las ligas de interés de la carpeta compartir y volver a iniciar el proceso de aprendizaje para que el nuevo interés sea detectado.

En las figuras 1 y 2 se encuentran concentrados los resultados que se obtuvieron durante el desarrollo de los experimentos.

En cuanto al funcionamiento del sistema se estableció el siguiente experimento:

- Se ejecutó un Agente de Artículos y un Agente de Intereses.
- Se crearon ocho perfiles de usuario, esto es, se definieron ocho archivos con diversas ligas de interés dentro de una carpeta “**Compartir**”.
- Para cada uno de estos perfiles se ejecutaron sus respectivos agentes de servicio al usuario, esto es, su Agente de Usuario y su Agente de Aprendizaje.
- Después de llevar a cabo todos los procesos que cada uno de ellos puede realizar se integraron nuevos usuarios al sistema creando nuevos perfiles de usuarios y realizaron los mismos procesos que los agentes que se habían definido con anterioridad.

Los resultados que se obtuvieron de este experimento se enlistan a continuación:

- Los agentes de aprendizaje obtuvieron las áreas de interés del usuario con la mínima intervención del usuario.
- El aprendizaje se realiza tomando en cuenta exclusivamente lo que el usuario definió como material para compartir.
- El Agente de Usuario fue capaz de manejar todos los intereses que el usuario tuvo durante el tiempo del experimento.
- El agente de artículo probó su capacidad para divulgar los artículos hacia todos los posibles usuario interesados, sin repetir el envío una vez que el Agente de Usuario confirmó haber recibido dicha liga de interés.
- Esta prueba también validó el que los usuarios que pertenecen a un grupo de interés reciban los todos los artículos que se han enviado para su divulgación al Agente de Artículos aún y cuando los interesados no se encuentren registrados en el sistema al momento en el que se pusieron disponibles los artículos.

- El proceso que sigue el Agente de Aprendizaje para detectar los intereses del usuario es muy ágil, ya que sólo necesita un mínimo de siete ligas de interés proporcionadas por el usuario.
- El Agente de Usuario no requiere de información acerca de los demás agentes o de agentes con intereses similares, lo único que requiere es enviar su interés al Agente de Intereses y éste junto con el agente de artículos se encargan de conseguir artículos que sean de su interés y de divulgar los artículos que él desee a los posibles interesados en ellos.
- Por último se logró el objetivo propuesto, el sistema multiagente para enlace de intereses basado en la colaboración social permitió la divulgación de artículos tanto al enviar lo que se desea compartir, así como, recibiendo artículos que pudieran ser de interés para el usuario. Además se alcanzó la vinculación de usuarios con intereses similares, para lograr esto, el Agente de Usuario obtuvo del Agente de Intereses el nombre de los investigadores con áreas de interés afines, el centro de investigación donde laboran y su correo electrónico para que el usuario pueda establecer comunicación personal con quien desee.

#### 4. CONCLUSIONES

Se define la arquitectura del sistema multiagente de esta manera, debido a que se requiere de un sistema flexible y modular, de tal forma que si se desea agregar una nueva capacidad a este sistema, tal como el filtrado de audio, vídeo, otro tipo de formato de archivo; o cualquier otro medio, esta arquitectura pueda crecer simplemente agregando otro agente con una base de conocimientos especializada para almacenar

dicho tipo de medio sin tener que realizar otro tipo de modificaciones al sistema.

Por otra parte, si se requiere aprender las áreas de interés del usuario, a partir de un grupo de noticias, correo electrónico, o un sistema de conferencias sincrónico, sólo se necesita agregar un nuevo Agente de Aprendizaje capacitado para descubrir los intereses del usuario a partir de dicha información, de modo que no se altere el funcionamiento actual de los agentes de servicio al usuario.

Se definió el uso de Agentes de Artículos, con la finalidad de que estos mismos guarden un registro de los artículos que los distintos usuarios han querido compartir para que, cuando nuevos usuarios se incorporen al sistema puedan tener acceso a todo lo que se compartió anteriormente.

Esta característica implica que el sistema se convierta en un ambiente distribuido, en el cual, ciertos agentes tienen que comunicarse con otro agente que cumpla las funciones de un servidor con los problemas que ocasiona el que deje de funcionar uno de los agentes de filtrado.

Durante el estudio realizado se decidió desarrollar un protocolo propio de coordinación entre agentes, el cual, permite establecer las conversaciones entre los agentes utilizando un mínimo de mensajes para iniciar la comunicación. Asimismo, dependiendo de la acción realizada se envía un mensaje para finalizar la conversación o simplemente se termina sin necesidad de notificarlo.

La creación de este protocolo tiene la ventaja de que las conversaciones entre los agentes sean más cortas, permitiendo a éstos atender un mayor número de agentes, así también, el número de mensajes que

transitan por la red disminuye considerablemente.

Se utilizan las ligas que el usuario va almacenando como información para detectar los intereses del usuario debido a que, durante el desarrollo de sus investigaciones o en el tiempo en que navega por Internet el usuario va guardando las ligas de sitios de Internet que le son de su agrado, pues piensa que pueden serle de utilidad más adelante.

La investigación desarrollada permitió comprobar que es factible intentar obtener las áreas de interés del usuario a partir de esta información, creando así un método para obtener el perfil del usuario, en el cual, no se requiere que el usuario tenga que interactuar con el Agente de Aprendizaje contestando cuestionarios o evaluando páginas de Internet, que harían que el proceso de aprendizaje se convirtiera en una tarea tediosa para el usuario.

## **5. TRABAJO FUTURO**

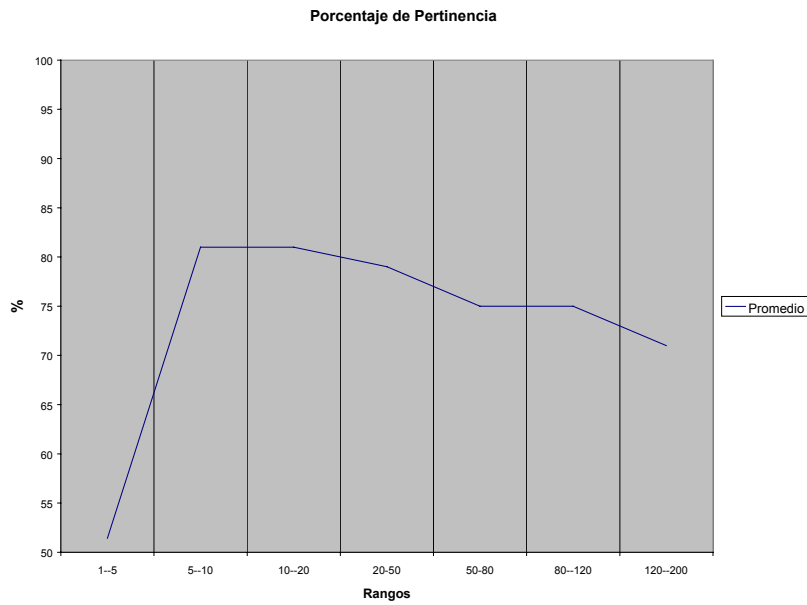
Se proponen las siguientes adecuaciones como mejoras al desempeño del sistema desarrollado:

- Diseño de una interfaz más amigable al usuario, de tal manera, que se presente la información de una forma más sencilla de entender, esto es, mayor comunicación con el usuario, de tal forma que éste sepa en que parte del proceso se encuentra el agente y que acciones está realizando, además de recibir retroalimentación de las acciones que el usuario realiza.
- Mejorar el método de aprendizaje, en este rubro, se requiere que el agente encargado de detectar los intereses tenga un método de aprendizaje supervisado,

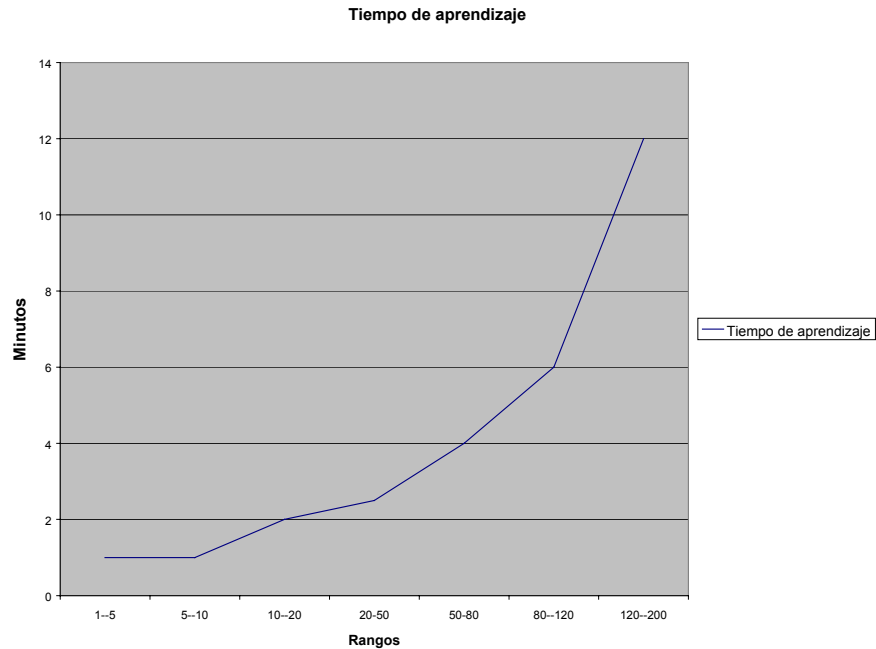
de tal forma, que pueda mejorar su desempeño, tomando en cuenta la retroalimentación del usuario y las ejecuciones anteriores.

- Notificación de cambios en páginas de interés. Esto con la finalidad de que se efectúen monitoreos frecuentes a las

ligas que el agente ha recibido, con el propósito de que si existe algún nuevo cambio se le vuelva a notificar para su lectura posterior.



**Figura 1.** Porcentaje de pertinencia de los intereses detectados al usuario



**Figura 2.** Tiempo que tarda el Agente de Aprendizaje en obtener los intereses del usuario

## REFERENCIAS

Brena R. (1998). *Lenguajes Formales y Automatas*.  
 Brenner W., Zarnekow R., Wittig H. (1998). *Intelligent Software Agents*. Editorial Springer. Alemania.  
 JATLite (1997) *JATLite*. [online] <http://java.stanford.edu/>  
 Starr, Ackerman, Pazzani (1996). *Do I Care: A Collaborative Web Agent*. proceedings CHI'96 ACM.  
 Muñiz E., Escamilla J., García M. (1998) *Arquitectura Distribuida de Agentes para Enlace de Intereses*. XX Congreso Electro. Chihuahua, Chih.  
 Shardanand, Maes P. (1995) *Social Information Filtering: Algorithms for Automating "Word of Mouth"*. CHI 1995

Turney, P.(1997). *Extraction of Keyphrases from Text: Evaluation of Four Algorithms*. NRC/ERB-1051  
 Turney.P.(1999) *Learning to Extract Keyphrases from Text*. NRC/ERB-1057:  
 Willams J. (1996). *Bots and Other Internet Beasties*. Sams Net Publishing