

Arquitectura Multi Agente Extensible Para La Administración Del Correo Electrónico

Darío Juan José Aznar

darioaznar@frre.utn.edu.ar

Grupo de Investigación Sobre Inteligencia Artificial
Universidad Tecnológica Nacional
Facultad Regional Resistencia
Chaco - Argentina

Resúmen. Los sistemas multi agentes son una herramienta importante para construir sistemas complejos que interactúen en ambientes abiertos, como Internet, colaborando para lograr objetivos de diseño. La mayoría de las aplicaciones modernas, en el área de las tecnologías de la información, funcionan en ambientes abiertos, donde determinar si la información disponible es útil o no, es una tarea no trivial. El presente trabajo analiza una arquitectura multi agente, cuya finalidad es facilitar la administración del correo electrónico no deseado, a través de la colaboración entre agentes que trabajen de forma autónoma. En la primera parte se hace referencia al problema y su importancia. En la segunda, se expone el concepto de agente de software inteligente y se describen los tipos de agentes empleados. En la tercera se presenta una arquitectura inicial y otra extendida para solucionar el problema junto con sus características relevantes.

1. Introducción

El presente trabajo tiene como objeto de estudio la problemática del correo electrónico no deseado (SPAM) de la cual no escapa prácticamente ninguna organización con acceso a Internet.

Nuestro país se encuentra entre los principales proveedores de SPAM en el mundo, lo que hace imperativo el contar con herramientas adecuadas para combatir este problema y disminuir los efectos del mismo (para más información del problema argentino ver [1]).

Por ello, la realización de un sistema capaz de detectar, eliminar y realizar otras funciones asociadas a la administración del correo en forma autónoma dentro de una organización, es un área de estudio y desarrollo muy necesaria en estos momentos.

La basta cantidad de información que se requiere para decidir autónomamente qué es correo deseado o no deseado para un determinado usuario hace que sea prácticamente imposible implementar esta solución en sistemas del tipo convencional. Para brindar este tipo de soluciones una aproximación posible es la realización de sistemas de información basados en agentes que sean multi agente y adaptativos [2]. En este trabajo se plantea utilizar una arquitectura multi agente capaz de sacar beneficios de las cualidades de distintos tipos de agentes, como ser la rapidez de los agentes reactivos, la adaptabilidad de los agentes de estado mental.

2. Agentes Utilizados

Para desarrollar sistemas que se encarguen de realizar tareas específicas en representación de sus usuarios sin que deba haber una intervención explícita del mismo, es necesaria la utilización de entidades de software colaborativas denominadas *agentes inteligentes*.

En el estado actual de la Inteligencia Artificial no hay una definición única de agente de software inteligente, pero la mayoría de las definiciones coinciden en que son programas que actúan en representación de sus usuarios humanos realizando tareas no triviales. Las características más consensuadas de los agentes de software es que son: *autónomos*, *colaborativos* y *adaptativos* [2].

A continuación se describen brevemente los tres tipos de agentes utilizados en la arquitectura.

2.1 Modelo BDI

Este tipo de agentes tiene sus raíces en la comprensión del *razonamiento práctico*, el cual envuelve dos procesos: una **deliberación** (decidir qué metas se desean alcanzar), y un **razonamiento para alcanzar la meta** (decidir cómo alcanzar dichas metas). Esto significa observar qué opciones se encuentran disponibles en un determinado instante y luego elegir entre ellas y *comprometerse* a una de tales opciones. Estas opciones elegidas se convierten en **intenciones**, las que determinan las acciones del agente. Si las intenciones no son reconsideradas con la frecuencia suficiente, el agente puede seguir intentando alcanzar un objetivo inalcanzable o inválido; pero si son reconsideradas constantemente, el agente puede fallar en dedicar suficiente trabajo para alcanzar cualquier meta.

El modelo BDI reduce las nociones abstractas de deseos e intenciones a los conceptos de metas y planes. Las creencias (*beliefs*) representan la información que un agente tiene sobre el mundo que habita y sobre su propio estado interno. Las actitudes motivacionales de agentes son capturadas en forma de *deseos* o menos precisamente *metas* (goals). La meta es un concepto central en la arquitectura BDI, representando un cierto estado objetivo que el agente está tratando de alcanzar (en el caso de la detección y eliminación de SPAM, el objetivo principal de un agente de filtrado del correo es el de mantener la casilla de correo del usuario sin correo no deseado y sin eliminaciones equivocadas de correo). Las metas representan explícitamente los estados a ser alcanzado y son las razones por las que el agente realiza sus acciones. El concepto de meta permite modelar agentes que no sean solamente reactivos: aquellos agentes que persiguen sus propios objetivos exhiben un comportamiento proactivo. Los planes, que son actitudes deliberadas, son los medios por los cuales los agentes alcanzan sus metas. Un plan no es solo una secuencia de acciones básicas sino que además puede incluir submetas. Otros planes se ejecutan para alcanzar dichas submetas, formando una jerarquía de planes [3].

2.2 Agentes Reactivos

Los agentes reactivos disponen de varias modalidades sensoriales para percibir sus mundos y de distintas formas de actuar sobre ellos. Los más complejos tienen la capacidad de recordar propiedades y almacenar modelos internos de sus mundos, pero en todos los casos las acciones decididas son función del estado actual y de los estados anteriores de sus mundos. También pueden disponer de procesos perceptivos y motores muy complejos. [4]

Tenemos como casos más sencillos de agentes reactivos [5] a:

- **Agentes de reflejo simple.**
- **Agentes bien informados de todo lo que pasa.**

2.3 Agentes Lógicos

Una concepción tradicional en la construcción de sistemas inteligentes artificiales, sugiere que el comportamiento inteligente puede ser generado mediante la incorporación de una representación simbólica del dominio de aplicación y del comportamiento deseado del sistema, además del mecanismo de manipulación correspondiente de esa representación[6].

Un mecanismo de representación simbólica válido es el de las fórmulas lógicas y el mecanismo de manipulación es el de la deducción lógica.

Si se dispone de una teoría de agencia, se puede describir el comportamiento del agente en cuanto a la definición de sus objetivos parciales para lograr sus objetivos de diseño. Esta teoría puede ser considerada una especificación del comportamiento del agente.

Una aproximación es la implementación de agentes basados en lógica o agentes deliberativos[6] en los que el estado interno es una colección de fórmulas de predicados de primer orden. Esta colección de fórmulas es el conjunto de creencias que tiene el agente acerca del estado del ambiente. Obviamente, estas creencias pueden, por diversas razones, no estar acorde a la realidad del ambiente en un momento dado y por lo tanto generar comportamientos equívocos del agente.

En la selección de la mejor acción a llevar a cabo según las percepciones del ambiente, se genera un espacio de búsqueda que a veces es necesario acotar. En este sentido, se pueden utilizar técnicas que permitan mejorar la eficiencia en dicha selección. Una alternativa es utilizar técnicas de planificación para poder determinar las acciones a realizar para alcanzar un estado objetivo.

3. Arquitectura Utilizada

El diseño del sistema debe ser realizado de manera tal que permita satisfacer de manera eficiente no sólo el objetivo principal (la detección y filtrado del SPAM), sino también permitir que se agregue con el tiempo mayor funcionalidad al sistema con el menor grado de dificultad posible.

Para alcanzar estas metas se plantea para el diseño arquitectónico como una adaptación de RETSINA[2], la cual es una arquitectura reutilizable para sistemas multi agente en donde los agentes separan sus conocimientos sobre tareas especializadas y se coordinan para recabar y filtrar información en respuesta a un problema iniciado por el usuario [2].

De esta manera se posibilita la ampliación del sistema con el agregado de nuevos agentes que cumplan con tareas particulares y su correspondiente interacción con el resto de los agentes previamente existentes.

El sistema se centra en torno de tres tipos de agentes (en función de sus tareas):

- Agentes de *interfase*: están directamente relacionados con los objetivos del usuario humano. Estos agentes son los que proveen al usuario de la respuesta obtenida por el sistema.
- Agentes de *tareas*: quienes están relacionados con los procesos de solución de tareas arbitrarias.
- Agentes de *información*: están relacionados con fuentes de información necesarias.

La arquitectura cuenta con las siguientes características principales:

- Los agentes pueden distribuirse en muchas máquinas.
- Los agentes cooperan entre sí comunicándose a través de un mismo lenguaje, en este caso el lenguaje de comunicación KQML. (Para mas información sobre KQML ver [7])
- Los agentes publican su funcionalidad utilizando “agentes del medio” (middle agents), quienes se encargan de relacionar las necesidades de unos agentes con las capacidades de otros.

3.1 Arquitectura básica

La arquitectura que se desarrolla a continuación está planteada para realizar el filtrado inteligente del correo entrante. Cuenta con un agente de interfaz que se encarga de adaptar las opciones del sistema a las preferencias de cada usuario (por ej. eliminar directamente el correo no deseado o enviarlo a una carpeta especial). Esta adaptación no necesariamente se realiza por órdenes explícitas que el usuario da al agente de interfase, sino que este agente se encuentra constantemente activo y aprende cómo interactúa el usuario con el sistema con el objetivo de facilitarle el uso del mismo.

El filtrado del SPAM es realizado por uno o más agentes de tareas (*agente/s de filtrado*), el cual toma la decisión de filtrado tanto por el remitente como por el contenido del mail. Para filtrar por el contenido, el agente de filtrado recurre a la colaboración de otro agente de tareas que analiza los cuerpos de los correos de acuerdo a lo que aprende de las preferencias del usuario, en búsqueda de frases o palabras específicas. Toda interacción entre agentes se realiza a través del lenguaje KQML[7]. Para determinar el tipo de remitente, el agente de filtrado se comunica con agentes de información los cuales buscan en la red datos sobre fuentes de SPAM conocidas. Gracias a las características anteriormente nombradas estos agentes tienen la facultad de desplazarse por la red, ya sea interna o por Internet en búsqueda de la información solicitada.

Cada uno de los agentes puede clonarse tanto en la misma máquina como en otra con el objetivo de balancear la carga de trabajo. Debido a esta movilidad inherente de los agentes, los mismos no pueden saber a priori dónde encontrar cierta funcionalidad que requieran, por lo que se emplea para ello a los agentes del medio: agentes bien conocidos y que se encargan de publicar los servicios disponibles que brinda cada una de las entidades en una arquitectura multi agente.

3.2 Arquitectura ampliada

Las características de esta arquitectura (Figura 1) permite ampliar el alcance del sistema creando nuevos agentes con funciones específicas que colaboren con los agentes de la arquitectura básica.

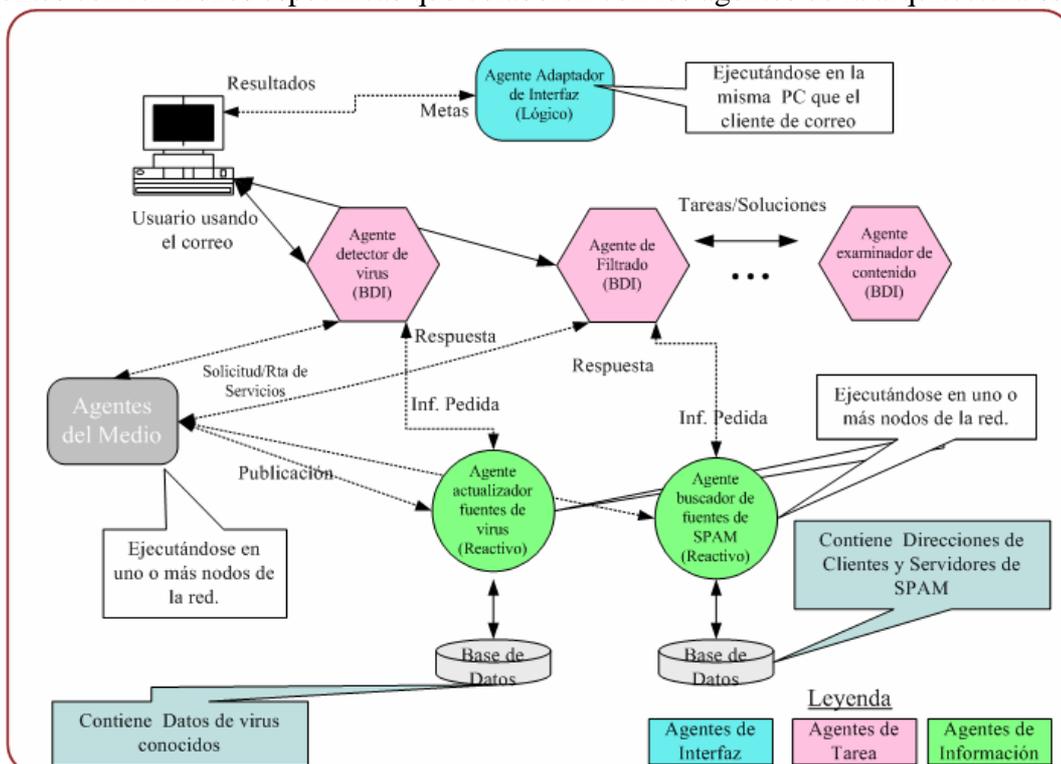


Fig. 1. Arquitectura Ampliada

Uno de los principales orígenes del correo basura está dado en los virus que al ejecutarse instalan servidores de SPAM en las máquinas de usuarios desprevenidos. Como solución a este nuevo inconveniente, se agrega al sistema un agente encargado de controlar los archivos anexados de acuerdo a la información de virus existentes entregada por otro agente de información el cual actualiza una base de datos de virus conocidos. El comportamiento del agente *actualizador de virus* es conceptualmente idéntico al del agente de información que busca las fuentes de SPAM: tiene la aptitud para desplazarse por una red con el objetivo de recabar y filtrar la información que se le solicita.

De acuerdo a lo que determine el agente de interfaz (el cual decide en base a lo que aprende de las preferencias del usuario) se puede enviar un mensaje automático al remitente con el objeto de ponerlo en aviso sobre la existencia de un SPAMMER en su máquina.

Opcionalmente, se puede agregar más funcionalidad de administración automática al sistema de acuerdo a lo requerido por una situación determinada.

Referencias

- [1] Andrea Ferrari. **El ataque de la basura argentina.** <http://www.pagina12web.com.ar/diario/sociedad/3-29283.html>, Ultimo acceso: 31/03/2005.
- [2] Katia P. Sycara. **Levels of Adaptivity in Systems of Coordinatin Information Agents** – 1998
- [3] L. Braubach, W. Lamersdorf, A. Pokahr. **Jadex:Implementing a BDI-Infrastructure for JADE Agents**
- [4] Nilsson, N. - **“Inteligencia Artificial - Una nueva síntesis”** - Ed. Mc Graw Hill. 2000. Pág. 11.
- [5] Russell S., Norvig P. - **“Inteligencia Artificial, Un Enfoque Moderno”**. Ed. Prentice Hall. 1996. Pág. 33 a 55
- [6] Wooldridge M. J., y Jennings, N. R., **Intelligent Agents: Theory and Practice.** The Knowledge Engineering Review 1995, Vol 10 (Nº 2) pp115-152
- [7] T.Finin, R. Fritzson, D. McKay and R. McEntire. **Kqml as an agent communication lenguaje.** In proceedings of the Tirad Int Conference on Information and Knowledge Management, 1994.