

Los agentes de la compañía fractal aprenden a seleccionar los socios con quienes establecer relaciones cliente-servidor.

M.M. Canavesio^(a) E.C. Martinez^(b)

^(a) CIDISI Centro de Investigación y Desarrollo en Ingeniería en Sistemas de Información. UTN Facultad Regional Santa Fe, Santa Fe. Argentina.
mcanaves@frsf.utn.edu.ar

^(b)INGAR Instituto de Desarrollo y Diseño. CONICET, Santa Fe. Argentina.
ecmarti@santafe-conicet.gob.ar

Abstract. El actual entorno de negocios enfrenta a las empresas a profundos cambios y para sobrellevarlo forman alianzas estratégicas entre ellas. [1] proponen un modelo de la compañía fractal basada en proyectos para la integración virtual y temporal entre diferentes empresas en el logro de objetivos de negocios. El éxito de la aplicación del modelo radica en el establecimiento de relaciones cliente-servidor entre los gestores de los proyectos. Este trabajo analiza la incorporación de aprendizaje en los agentes de gestión, cuando seleccionan socios con quienes establecer relaciones cliente-servidor más beneficiosas. Para ello propone la aplicación del aprendizaje por refuerzo.

Keywords. Compañía fractal. Gestión por proyectos. Aprendizaje por refuerzo. Agentes inteligentes.

1 Introducción

Las empresas enfrentan a profundos cambios en su entorno de negocios, que las conduce a formar alianzas estratégicas o redes con sus pares para así satisfacer demandas de un mercado, cada vez más exigente. Para que las empresas que participan de estas redes o alianzas alcancen los beneficios y ventajas competitivas esperadas, se requiere desarrollar un modelo de empresa integrada que identifique y defina concretamente la estructura, procesos, información y relaciones entre las empresas que las componen. Atendiendo a esta problemática, [1] proponen un modelo de la compañía fractal basada en proyectos para la integración virtual y temporal entre empresas. En este modelo, el proyecto es la unidad fractal de gestión, auto-gobernada, interdependiente y temporal que otorga a la red de empresas flexibilidad, rapidez de respuesta y agilidad para adaptarse a los cambios de su

entorno de negocios. La unidad fractal de gestión se concibe constituida por un gestor de la misma y un objeto gestionado. Este último puede ser un fin (meta) o un medio (recurso) necesario para el logro de un dado fin. De allí que los roles que los gestores de proyecto pueden asumir son el de gestor de fines o gestor de medios. Los actores que desempeñan tales roles se vinculan entre sí e interactúan a través de relaciones cliente-servidor. Este concepto de relación cliente-servidor entre los gestores, es fundamentalmente importante para el modelo de empresa integrada, por lo que la selección tanto del mejor servidor para cada cliente y también el mejor cliente para cada servidor, es clave en la obtención del mayor beneficio de la integración entre empresas. Por ello, es necesario dotar a los gestores de proyecto con capacidad de aprendizaje, que individualmente les permita definir y modificar su comportamiento sobre la base de sus preferencias, percepciones, habilidades e interacción con otros gestores de la compañía fractal.

Este trabajo analiza la incorporación de aprendizaje en los roles: gestor de fines y de medios, que asumen los gestores de proyectos en el marco de la compañía fractal cuando seleccionan los socios con quienes establecer relaciones cliente-servidor más beneficiosas. Para ello se propone la aplicación del aprendizaje por refuerzo [2].

2 La compañía fractal basada en proyectos

La idea de la compañía fractal [3] es un modelo de empresa conceptual, que a través de unidades autónomas, descentralizadas e interdependientes, denominadas fractales, otorga a las empresas mayor flexibilidad y agilidad para adaptarse a los cambios en su entorno de negocios. Un fractal es definido como una estructura que describe un patrón idéntico que se replica a sí mismo a distintos niveles de abstracción de una manera recursiva (Fig. 1). En el modelo de empresa fractal propuesto por [1], la unidad fractal de gestión es un proyecto. Dentro de la red de empresas, cada proyecto es una entidad auto-gestionada, interdependiente y temporal, en la que se combinan distintos tipos de habilidades, conocimientos y recursos para lograr una meta concreta (Ej. completar una orden, diseñar un nuevo producto, satisfacer un requerimiento de recursos, etc.).

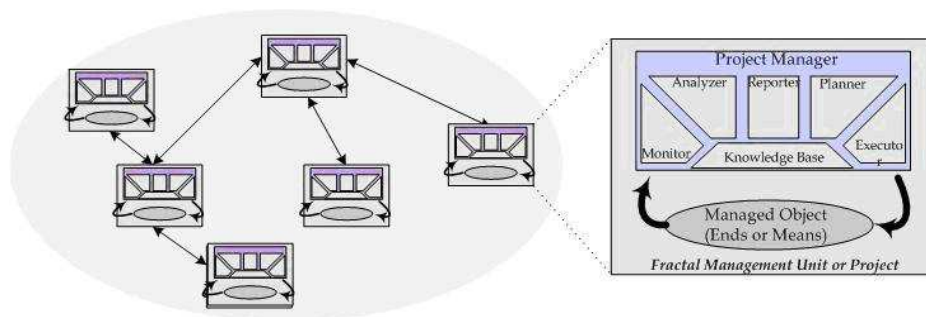


Fig. 1. Estructura del fractal de gestión.

La unidad fractal propuesta se compone de un gestor de proyecto que la gestiona y un objeto que es gestionado por éste. Como en el modelo, tanto los fines o metas como los medios o recursos son gestionados a través de proyectos, el gestor de un proyecto asumirá el rol de *gestor de fines* o *gestor de medios*, respectivamente. Los actores que desempeñan tales roles se vinculan mediante relaciones cliente-servidor, establecidas a través de algún mecanismo de negociación entre los agentes. Así, la compañía fractal es vista como un conjunto de relaciones temporales cliente-servidor, a través de las cuales los gestores de proyecto interactúan para diversificar su portafolio de productos, acceder a una mayor variedad de recursos, reducir costos, tiempo e incertidumbre.

3 La compañía fractal como un sistema multi-agente

La compañía fractal es un sistema social, intervienen actores o agentes inteligentes, que desempeñan roles e interactúan, a través de relaciones cliente-servidor, para lograr sus objetivos en un contexto del proyecto como unidad de gestión. Un agente se define como una entidad física o virtual que en un dado entorno, es capaz de actuar de manea autónoma con el fin de alcanzar sus objetivos [4]. Además, un agente posee capacidad de aprendizaje, le permite capitalizar experiencia, modificar su política de actuación y la representación interna del entorno para adaptar su comportamiento a los cambios del mismo. El modelo de la compañía fractal propuesto basa su estructura organizacional y de gestión en el concepto de las relaciones cliente-servidor, establecidas entre gestores de proyecto, donde el gestor de proyecto cliente negocia la utilización de un dado recurso o la delegación del logro de una sub-meta a otro gestor proyecto servidor. Por ello, la selección del mejor servidor en cada oportunidad es clave para la obtención del mayor beneficio de la integración entre las empresas. Este proceso de selección involucra información incompleta e incierta con respecto a los servidores candidatos, sus habilidades y experiencias. Además, algunas de las características consideradas (performance, calidad, confiabilidad, etc.) con subjetivas o cualitativas dificultando aún más la definición o valoración concreta de las mismas [5,6].

El modelo compañía fractal propuesto es un mercado de coincidencias [10] donde existen clientes que anuncian requerimientos por la provisión de recursos a diversos proveedores, y a la vez, existen proveedores de recursos que seleccionan a un conjunto de clientes a quienes desean proveerle sus recursos. Por ello, es necesario dotar a los gestores de proyecto, en sus diferentes roles, con capacidad de aprendizaje, que individualmente les permita definir y modificar su comportamiento sobre la base de sus preferencias, percepciones, habilidades, nivel de conocimiento, comunicación e interacción con otros gestores en la compañía fractal. Este proceso de aprendizaje les permitirá mejorar constantemente sus desempeños y reducir la incertidumbre al aprender sobre las preferencias, creencias y estrategias de los otros agentes, para así competir eficazmente con ellos.

3.1 Los agentes aprenden a establecer relaciones cliente-servidor

Distintos aspectos del aprendizaje involucrado en la decisión que los agentes realizan al establecer relaciones cliente-servidor han sido abordados por [7,8,9]. Este trabajo implementa algoritmos de aprendizaje por refuerzo [2] que permiten analizar cómo, en un mercado de coincidencias como el de la compañía fractal basada en proyectos, los agentes clientes identifican servidores eficaces y de mínimo costo, mientras los servidores aprenden a detectar clientes, cuyos contratos les otorguen el mayor beneficio económico.

3.1.1. Aprendizaje por refuerzo

Este tipo de aprendizaje es un enfoque computacional para entender y automatizar el aprendizaje orientado al logro de metas y toma de decisiones en una secuencia [2]. Un agente aprende por prueba y error cual acción ejecutar mientras interactúa con su entorno. En cada episodio, el agente selecciona una acción posible en el actual estado y la ejecuta, causando que el entorno se mueva al siguiente estado. El agente recibe una recompensa que refleja el valor de la acción tomada. El objetivo del agente es maximizar la suma de las recompensas acumuladas desde un estado inicial hasta que alcanza el estado final. Inicialmente, el agente desconoce el curso de acción a tomar en función del contexto. A través de la interacción, el agente descubre qué acciones tienen mayor recompensa tras un análisis retrospectivo de los resultados (aciertos y errores) que ha obtenido. La implementación de agentes que aprenden por refuerzo se lleva a cabo utilizando una estructura compuesta por los siguientes elementos [2]: *Política* define el objeto de optimización y mejora el conocimiento disponible por el agente. *Función recompensa* define el objetivo que se espera satisfacer al final de cada episodio. *Función valor o utilidad* proporciona una medida de la efectividad de una dada política. El *modelo del entorno* imita el comportamiento del mismo.

3.1.2 El gestor de proyecto aprende a seleccionar el mejor servidor para una tarea

Un plan de proyecto se integra de cientos o miles de tareas que requieren de diferentes recursos y habilidades. El gestor de proyecto dispone de un centro de potenciales proveedores, algunos calificados como confiables y capaces de proveer recursos con el nivel de calidad requerido y otros que no reúnen tales condiciones. A priori el gestor de proyecto desconoce esta calificación de los potenciales proveedores por lo que, deberá aprender a distinguirlos.

Para ello, el gestor de proyecto utiliza un parámetro denominado *factor de credibilidad*, que considera la capacidad de un proveedor de cumplir en tiempo y forma sus contratos. Inicialmente, todos los gestores de recursos son considerados confiables en igual grado. Luego de cada interacción, el gestor califica el desempeño del servidor y actualiza el factor de credibilidad del mismo para una particular tarea. Por ello, mientras el gestor de proyecto tenga poca experiencia sobre la credibilidad de los proveedores de recursos, incurrirá en elevados costos, por elegir servidores

económicos e ineficaces o eficaces pero onerosos. Cuando el agente ha adquirido conocimiento respecto a quienes son los servidores confiables, se reduce la búsqueda, la incertidumbre y los costos asociados debido a que las negociaciones se circunscriben sólo a este grupo de proveedores. Para no perder la oportunidad de considerar nuevos y potenciales mejores proveedores de recursos para la tarea, que se incorporen a la compañía fractal, el gestor de proyecto posee la capacidad de explorar el centro de potenciales proveedores y probar así nuevas soluciones. Este mecanismo de exploración debe ser empleado estratégicamente.

Para implementar este aprendizaje en los agentes se emplea el algoritmo Q-Learning con una política de selección de la acción ϵ -greedy, que le permitirá explorar la posible incorporación de nuevos y potenciales proveedores de recursos para una dada tarea, en el marco de la compañía fractal.

```

Inicializar  $Q(a_i, t_a, p)$  arbitrariamente
Repetir N episodios
Elegir  $a_i, t_a, p$  utilizando la política  $\epsilon$ -greedy
Aplicar la acción  $a_i, t_a, p$ , observar  $r, B$  para la acción
seleccionada
Actualizar  $Q_{t+1}(a_i, t_a, p) = Q_t(a_i, t_a, p) + \alpha [r_{t+1} - Q_t(a_i, t_a, p) + B_{t+1}(a_i, t_a, p)]$ 
FinRepetir

```

Fig. 2. Algoritmo Q-learning para implementar el aprendizaje

Cada episodio en el algoritmo de la Fig. 2, representa una instancia de negociación para establecer una relación cliente-servidor entre gestores de proyecto. La acción que el agente lleva a cabo es la elección de un dado servidor para la provisión de recursos en una tarea particular del plan de proyecto. Luego de cada negociación, el valor-Q de la acción seleccionada se actualiza según la Eq (2)

$$Q_{t+1}(a_{i,t_a,p}) = Q_t(a_{i,t_a,p}) + \alpha [r_{t+1} - Q_t(a_{i,t_a,p}) + B_{t+1}(a_{i,t_a,p})], \quad (2)$$

Donde,

- $a_{i,t_a,p}$ acción del gestor i de elegir el proveedor p que provee recursos a la tarea t_a
- $Q_t(a_{i,t_a,p})$ valor que en el tiempo t ha tenido la acción para el gestor i de elegir al proveedor p para proveer a la tarea t_a
- r_t , la función recompensa es el costo de contrato (C) por la provisión de recursos para la tarea t_a que se obtienen de una distribución normal. $r_{t+1} = (-1) * C_{(a_i,t_a)}$, el costo se considera como un valor negativo, por ello se multiplica por -1 .
- $B_{t+1}(a_{i,t_a,p})$, *factor de credibilidad* del gestor i respecto a que el proveedor p posea la capacidad para satisfacer efiazmente la tarea t_a (Eq. 3).

$$B_{t+1}(a_{i,t_a,p}) = \begin{cases} B_t(a_{i,t_a,p}), & \text{proveedor capaz} \\ B_t(a_{i,t_a,p}) - 25, & \text{proveedor incapaz} \end{cases} \quad (3)$$

Parámetros considerados en el algoritmo de aprendizaje: valores-Q inicializados en cero, velocidad de aprendizaje $\alpha = 0,1$, factor de confiabilidad $B = 100$ créditos

(para todos los servidores), política de selección de la acción con $\epsilon = 0,1$. Así, el algoritmo explora las ofertas presentadas por los potenciales proveedores (con $B > 0$) y la de los nuevos proveedores, con una probabilidad de 1 cada 10 y explota la acción que tiene el máximo valor-Q con probabilidad 0,9. El experimento se ejemplifica para la tarea *ta* que requiere el recurso *RI*, para el cual en la compañía fractal existen 10 potenciales proveedores con diferentes capacidades.

El primer experimento considera el aprendizaje del gestor de proyecto en distinguir servidores confiables y de mínimo costo con quien establecer relaciones cliente-servidor. Para ello, se consideraron 1000 interacciones asociadas con la selección de proveedores de recursos para la tarea y se efectuaron cinco repeticiones. Luego, se obtuvo el porcentaje de veces que en promedio fue elegido cada proveedor durante los experimentos, resultando la gráfica que muestra la Fig. 3. Se observa que cuando el gestor de proyecto no posee suficiente conocimiento sobre los potenciales servidor, explora muchas alternativas (buenas, regulares, malas). Con la acumulación de experiencia en la selección de proveedores, se comienza a notar su preferencia por contratar sólo aquellos que son confiables y de mínimo costo. Así, los proveedores 1, 2, 3, 5, 7 y 8 rápidamente son descartados por tratarse de proveedores no confiables; luego el gestor de proyecto aprende quién es el proveedor confiable de mínimo costo, infiriendo después de varias interacciones, que se trata del proveedor 4, descartando así a los proveedores 9 y 10.

El siguiente experimento, supone que luego de las primeras 100 interacciones, cinco nuevos gestores de recursos para la tarea considerada, ingresan a la compañía fractal, de ellos dos son confiables, uno de ellos con costos semejante al del actual mejor proveedor. Como se observa en la Fig. 4, los proveedores 13 y 14 entran a competir fuertemente con los antiguos servidores confiables mientras que los nuevos servidores no confiables son rápidamente descartados. También el proveedor 14 que si bien es oneroso, ofrece mejores precios que los antiguos mejores proveedores 9 y 10.

3.1.3. El agente gestor de recursos aprende a aceptar los mejores contratos

Los servidores se diferencian en término de costos de prestación, confiabilidad, dispersión geográfica y también por capacidades y habilidades que le permiten proveer recursos sólo a determinadas tareas de los proyectos.

En un mercado de coincidencias, los proveedores de recursos deben aprender qué contratos de provisión aceptar para incrementar beneficios y credibilidad ante sus clientes. Durante su aprendizaje, un gestor de recursos novato desconoce su potencial y por lo tanto aceptará todos los contratos adjudicados para la provisión de recursos sin discriminar las tareas, incurriendo en elevados costos adicionales que reducirán considerablemente beneficios y credibilidad, debido a los incumplimientos. Esta situación se irá revirtiendo a través de sucesivas interacciones que le permitirán al servidor reconocer a que tareas es capaz de proveerle recursos.

Nuevamente, el algoritmo Q-Learning se implementa con una política de selección de la acción ϵ -greedy. En este caso, cada episodio representa una interacción en la que el servidor debe decidir cuál contrato aceptar para la provisión de recursos a una dada tarea, suponiendo que involucran a un recurso no divisible

cuya provisión se requiere para un mismo período de tiempo. En cada interacción, el valor-Q se actualiza según la Eq (5)

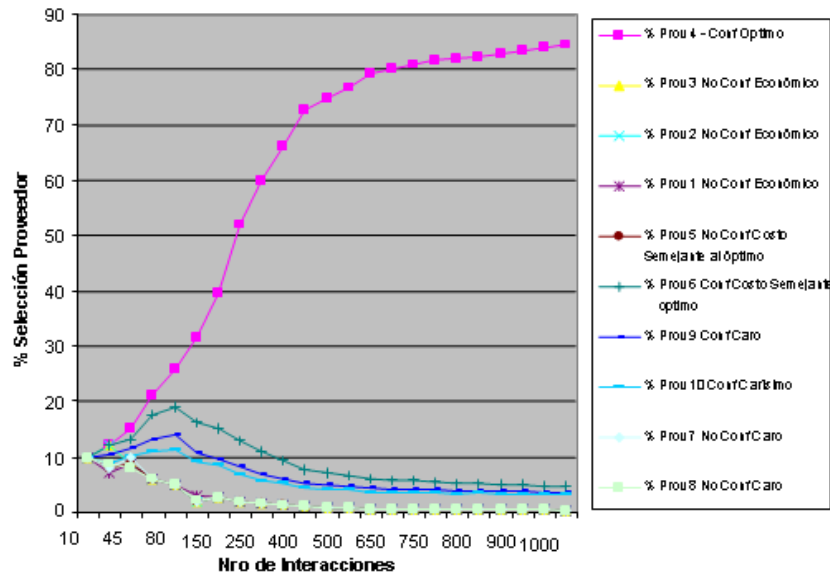


Fig. 3. Curva aprendizaje del cliente en discriminar servidores confiables.

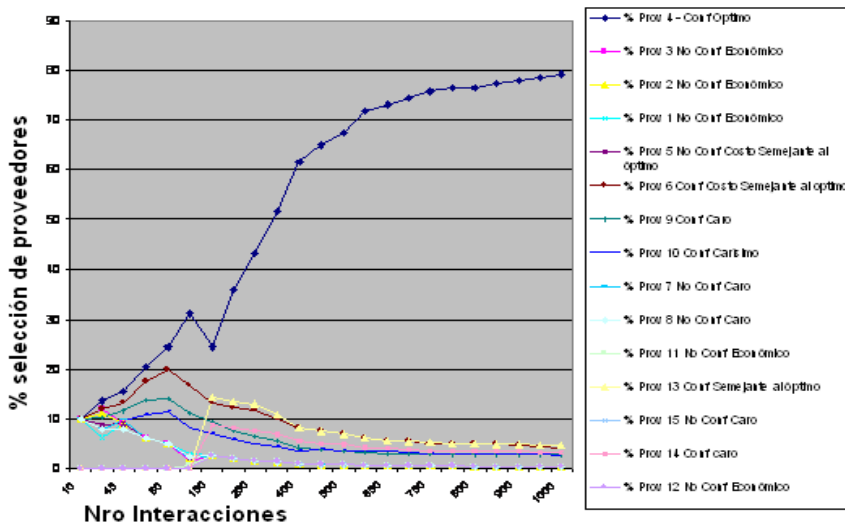


Fig. 4. Curva aprendizaje del cliente cuando ingresan nuevos potenciales servidores.

$$Q_{t+1}(a_{j,ta,r}) = Q_t(a_{j,ta,r}) + \alpha[rt+1 - Q_t(a_{j,ta,r})] \quad (5)$$

Donde,

- $a_{j,ta,r}$ es la acción del agente servidor j de elegir proveer el recurso r a la tarea ta .

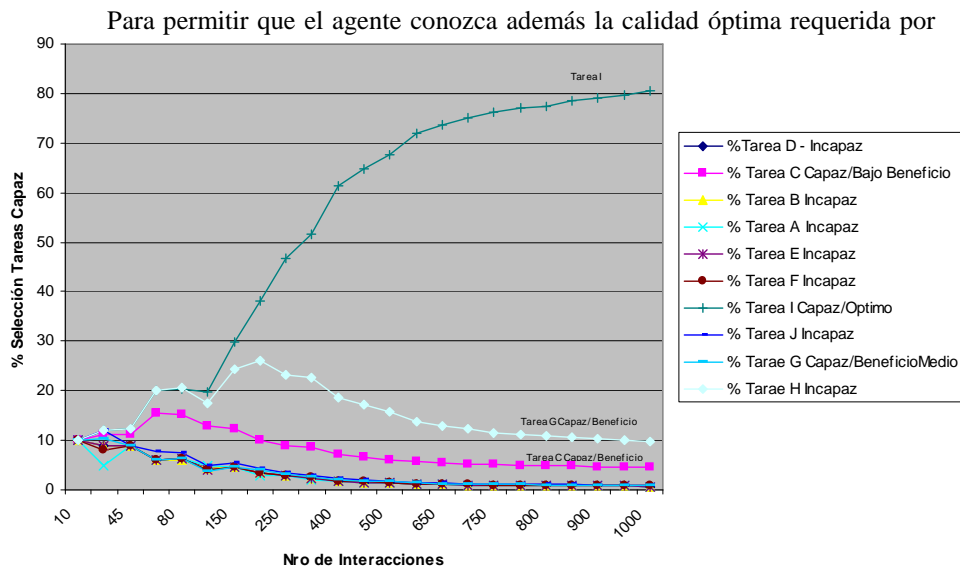
- $Q_t(a_{j,ta,r})$ es el valor que en la interacción t ha tenido la acción para el servidor j de elegir proveer el recurso r a la tarea ta .
- r_t es la función recompensa que se compone del beneficio esperado por el contrato de provisión del recurso (B) más una penalidad (P) o factor de incapacidad del agente para proveer del recurso r a la tarea ta .

$$r = B(a_{j,ta,r}) + P(a_{j,ta,r})$$

$$P(a_{j,ta,r}) = \begin{cases} 0 & \text{si el agente } j \text{ es capaz de ser proveedor de la tarea } ta \\ \text{Crece 25} & \text{si el agente } j \text{ no es capaz para ser proveedor de la tarea } ta \end{cases}$$

El algoritmo Q-learning para el agente servidor, posee los siguientes parámetros: los valores-Q inicializados con un valor muy elevado respecto a los valores de los beneficios esperados. Velocidad de aprendizaje $\alpha = 0,1$, política de selección de acción con $\epsilon = 0,1$, se inicializa $P = 0$ para todo agente. En los experimentos se consideró que existen en la compañía fractal 10 diferentes tareas para las cuales el gestor puede proveer su recurso R1. Cada una de estas tareas tiene asociado un beneficio esperado que surge de una distribución normal. El algoritmo le permitirá al agente reconocer las tareas para las cuales es capaz de proveer recursos y cuál de ella le otorga el mayor beneficio. De la Fig. 5, se observa que el agente es capaz de proveer el recurso a las tareas identificadas como C,I,G de las cuales la tarea I le reditúa el mayor beneficio.

Fig. 5. Curva de aprendizaje agente servidor en identificar tareas para las que es capaz.



cada tarea a la que potencialmente pretende proveer de recursos, el algoritmo fue modificado según la propuesta de [9], por cuanto no sólo se penaliza la acción elegida en una interacción, sino que si el gestor de recursos eligió proveer a un nivel de calidad mayor (o menor) al óptimo requerido, también se penalizan todos los niveles de calidad por sobre (o por debajo) del elegido porque cualquiera de ellos implicaría

una pérdida de ganancias o gastos adicionales. Así planteado, la acción $a_{j,ta,r}$ representa el nivel de calidad elegido por el agente j para proveer el recurso r a la tarea ta . La función recompensa r_i responde a la EQ (6), donde $B(a_{j,ta,r})$ es el beneficio esperado al proveer con el nivel de calidad a , el recurso r a la tarea ta . Este valor surge de una distribución normal.

$$r_i = \begin{cases} B(a_{j,ta,r}) & \text{si } a_{ta} = a_{ta}^* \\ 0 & \text{si } a_{ta} > a_{ta}^* \\ B(a_{i,ta,r}) * (1 - \beta) & \text{si } a_{ta} < a_{ta}^* \text{ donde } \beta = a_{ta}^* - a_{ta} \end{cases} \quad (6)$$

$Q_t(a_{j,ta,r})$ es el valor que en la interacción t ha tenido la acción para el agente j de elegir el nivel de calidad a para proveer del recurso r a la tarea ta . La actualización del valor-Q para cada interacción t depende del nivel de calidad seleccionado (a_{ta}^e), según el caso de acuerdo a ecuaciones (7), (8) y (9).

$$Q_{t+1}(a_{j,ta,r}) = Q_t(a_{j,ta,r}) + \alpha[r_{t+1} - Q_t(a_{j,ta,r})]; a^e = a^* \quad (7)$$

$$\forall k \leq e \quad Q_{t+1}(a_{kj,ta,r}) = Q_t(a_{kj,ta,r}) + \alpha[r_{t+1} - Q_t(a_{kj,ta,r})]; a^e < a^* \quad (8)$$

$$\forall k > e \quad Q_{t+1}(a_{kj,ta,r}) = Q_t(a_{kj,ta,r}) + \alpha[r_{t+1} - Q_t(a_{kj,ta,r})]; a^e > a^* \quad (9)$$

Los experimentos se han corrido con los siguientes parámetros: valores-Q inicializados con valores muy altos, $\alpha = 0.1$ y $\epsilon = 0.1$. Para el ejemplo previo, se obtuvo la curva de aprendizaje en la que al principio los agentes desconocen el nivel de calidad de los recursos requerido por las tareas y su propia capacidad, por lo que se generan fluctuaciones. Con el tiempo, el agente aprende a identificar cuales tareas puede satisfacer con el nivel de calidad requerido.

4 Conclusiones

El trabajo presenta brevemente el modelo de la compañía fractal basada en proyectos para la integración virtual y temporal entre diferentes empresas para el logro de un dado objetivo de negocios. En este modelo, el establecimientos de relaciones cliente-servidor entre gestores de proyectos es fundamental.

La compañía fractal propuesta es un mercado de coincidencias donde los gestores clientes requieren recurso para las tareas de sus proyectos y los servidores de recursos seleccionan el conjunto de clientes a quienes desean proveerles sus recursos.

Las simulaciones realizadas sobre la incorporación de aprendizaje en los agentes de gestión para establecer relaciones cliente-servidor, mostró que los clientes aprenden a requerir recursos sólo a proveedores confiables, capaces y de mínimo costo. Por otro lado, el servidor aprende a identificar a qué tareas es capaz de proveer un dado recurso para maximizar sus beneficios y credibilidad ante sus clientes.

Referencias

1. Canavesio, MM; Martinez, EC. Enterprise modeling of a project-oriented fractal company for SMEs networking. *Computers in Industry*. Elsevier (2007) Nro 58, Pp 974-813
2. Sutton, RS; Barton,AG. Reinforcement Learning. An introduction. The MIT Press (1998)
3. Warnecke, HJ. The fractal company: a revolution in corporate culture. Springer-Verlag (1993).
4. Russell, A; Norving,P. Inteligencia artificial. Un enfoque moderno. Pearson Educacion SA. (2004).
5. Zeydan,M; Colpan,C; Cobanoglu,C. A combined methodology for supplier selection and performance evaluation. *Expert Systems with Applications*. (2011). Nro 38, Pp 2741-2751.
6. Huang,B.; Gao, Ch.; Chen,L. Partner selection in a virtual enterprise under uncertain information about candidates. *Expert Systems with Applications*. (2011). Nro 38, Pp. 11305-11310.
7. Hu,J.; Wellman,MP. Learning about other agents in a dynamic multiagent system. *Journal of Cognitive Systems Research*. (2001). Vol 2, Pp. 67-79.
8. Ren,Z; Anumba,CJ. Learning in multi-agent systems: a case study of construction claims negotiation. *Advanced Engineering Informatics*. (2002). Vol 16, Pp. 265-275.
9. Valluri,A.; Croson, DC. Agent learning in supplier selection models. *Decision Support Systems*. (2003). Vol 39. Pp. 219-240.
10. Sotomayor,M. Implementation in the many-to-many matching market. *Games and economic behaviour*. (2004). Nro 46, Pp. 199-212.