

# Confianza y Reputación de Agentes en Sistemas Multi-agente para Entornos Dinámicos

Federico Joaquín    Luciano H. Tamargo    Alejandro J. García

Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET),  
Departamento de Ciencias e Ingeniería de la Computación,  
Universidad Nacional del Sur, Bahía Blanca, Argentina.

San Andrés 800, (B8000CPB).

Tel: (0291) 459-5135 / Fax: (0291) 459-5136

e-mail: { federico.joaquin, lt, ajg }@cs.uns.edu.ar

## RESUMEN

Esta línea de investigación se enfoca en el área de confianza y reputación de agentes en sistemas multi-agente. Su objetivo general es el análisis, desarrollo y formalización de la dinámica de la confianza y reputación de los agentes, a partir de la interacción con sus pares en el marco de un sistema multi-agente (SMA). Esto involucra el desarrollo y formalización de técnicas de representación y actualización del grado de confianza y de reputación de un agente, y también, la integración de estas técnicas con mecanismos de razonamiento automático y dinámica de creencias. En particular, se espera avanzar en el estudio y desarrollo de formalismos para aplicaciones de naturaleza dinámica y distribuida, que combinen mecanismos de confianza, dinámica de creencias y argumentación.

**Palabras claves:** *Confianza y reputación, Sistemas Multi-Agente, Dinámica del conocimiento.*

## CONTEXTO

Esta línea de investigación se realizará dentro del ámbito del Instituto de Ciencias e Ingeniería de la Computación (ICIC) de doble dependencia entre CONICET y la Universidad Nacional del Sur, y dentro del Departamento de Ciencias e Ingeniería de la Computación de dicha

Universidad. La investigación está asociada con los proyectos de investigación: “Formalismos para el tratamiento de confianza y reputación en sistemas multi-agente” (PGI 24/ZN32) y “Argumentación y dinámica de creencias para mejorar las capacidades de razonamiento y representación de conocimiento en sistemas multi-agente” (PGI 24/N035) ambos financiados por la Universidad Nacional del Sur. Además, el primer autor posee una beca doctoral financiada por CONICET.

## INTRODUCCIÓN

En [SS05], Sabater y Sierra sostienen que la importancia de la confianza y reputación en sociedades humanas está fuera de discusión, por lo cual no es sorprendente que varias disciplinas, cada una desde una perspectiva diferente, haya estudiado y utilizado ambos conceptos. La investigación científica en el área de mecanismos computacionales de confianza y reputación en sociedades virtuales, es una disciplina reciente orientada a incrementar la fiabilidad y performance de comunidades electrónicas. En artículos recientes como [Chan10, Bhu10, Jos08, Jos07, SS05, Del03, McKChe02, AM02, eBAY02, Mont02, Mui02, ARH00, GrSlo00, BK01] podemos notar que en ciencias de la computación hay dos elementos que han contribuido sustancialmente a incrementar el interés en confianza y reputación:

el paradigma de SMAs y la evolución creciente del comercio electrónico.

El estudio de confianza y reputación tiene muchas aplicaciones en tecnologías de Comunicación e Información. Estos sistemas han sido reconocidos como factores claves para el éxito de la adopción del comercio electrónico. Los mismos son usados por agentes de software inteligentes como un mecanismo para buscar compañeros confiables y como un incentivo en la toma de decisiones acerca de si se tiene en cuenta un contrato. La reputación es usada en el mercado electrónico como un mecanismo para evitar fraudes y estafas [AM02, eBAY02, Del03]. Los e-markets no son el único campo de aplicación; por ejemplo, en [TGFS14, TGFS12, BK01] usan la confianza para mejorar la performance de mecanismos de revisión de creencias. Otra importante área de aplicación en tecnología de agentes es el trabajo de equipo y cooperación [Mont02]. Además, entre los formalismos de argumentación que se proponen sobre sistemas multi-agente, también hay algunos que usan la información de confianza y reputación para razonar [PSM11, Tang12]. Las investigaciones en confianza y reputación se requieren tanto para ambientes colaborativos como competitivos [LLRJ12, C11], y considerando tanto valores de confianza como valores de desconfianza [HCJ10].

En la actualidad, existen pocos trabajos que estudien una visión general de confianza y reputación desde el punto de vista de ciencias de la computación. Por ejemplo, Dellarocas en [Del03] presenta una visión general de mecanismos de reputación online que son usados en sitios web comerciales. En el área de confianza, Grandison [GrSlo00] examina varias definiciones de confianza que existen en la literatura y proveen una definición de confianza para aplicaciones de Internet. Hay también algunas propuestas que establecen una tipología

de reputación [Mui02] y confianza [McKChe02].

El estudio y desarrollo de agentes inteligentes, y la investigación en el área de SMA, han demandado en los últimos años el estudio y desarrollo de formalismos de representación y mantenimiento de conocimiento. En particular esto ha dado un fuerte impulso al área de Revisión de Creencias [Han99, HFCF01, FKS02] donde se estudia formalmente parte de estos temas. Además, en los últimos años se han producido importantes avances en un área relativamente joven que se enfoca en la Revisión de Creencias en SMA, entre cuyos artículos se pueden destacar [DGB97, Can98, LW99, LW01, TGFS12, TGFS14, TGFS16].

En esta línea de investigación se pretende realizar un aporte en el estudio y desarrollo de nuevos formalismos, técnicas y métodos de representación y actualización del grado de confianza y de reputación de un agente, y su integración con mecanismos de razonamiento automático. Este aporte es de importancia para mejorar las capacidades de los agentes que sean utilizados en aplicaciones concretas que requieran el uso de agentes inteligentes y sistemas multi-agente. Este tipo de aplicaciones son requeridas en la actualidad para la resolución de problemas de mucha complejidad en ciencias de la computación y en otras disciplinas.

### **LINEAS DE INVESTIGACIÓN Y DESARROLLO**

Esta línea de investigación, toma como punto de partida varias publicaciones vinculadas a la misma. Por ejemplo, en el libro [FC10], se realiza un análisis sistemático y acabado de los sistemas computacionales de confianza; en particular, sus autores distinguen lo desarrollado como la visión global del naciente campo de la modelización de confianza y los modelos computacionales de confianza. Se discute la

confianza y el concepto aliado de la reputación de una gama de diferentes orígenes. En particular, se plantean como objetivo poder sistematizar una teoría general de la 'confianza' para proporcionar un modelo orgánico de este fenómeno muy complejo y dinámico en el desarrollo cognitivo, afectivo, y de niveles sociales (interactivos y colectivos).

En el artículo [TGFS12], publicado en *Journal of The Knowledge Engineering Review*, se ha desarrollado una primera propuesta para modelar la dinámica de unabase de creencias de agentes basada en la credibilidad de la información que es recibida de sus pares informantes. En particular, se han definido operadores de revisión priorizados y no priorizados, donde la credibilidad de la información es usada para aceptar o rechazar la información recibida.

En [Tam12], se ha publicado una primera aproximación para utilizar la credibilidad de la información recibida en una extensión de un sistema de argumentación estructurada.

En [TGFS14], publicado en *Artificial Intelligence Journal*, se propuso operadores de expansión, contracción y revisión de órdenes de credibilidad de informantes. En esta propuesta un agente que tiene un orden de credibilidad sobre sus informantes, puede modificar dinámicamente ese orden ante la presencia de nueva información recibida de sus pares.

En [TGFS16], publicado recientemente en la revista *Knowledge and Information Systems*, se realiza una propuesta para compartir información entre agentes, donde cada agente puede tener su propio orden de credibilidad entre sus informantes.

## RESULTADOS OBTENIDOS/ESPERADOS

La implementación de aplicaciones que integren adecuadamente mecanismos de confianza y

reputación en sistemas multi-agente, requiere del estudio y desarrollo de nuevos formalismos teóricos. En esta línea de investigación, se consideran varios objetivos específicos para tal fin:

- Proponer un formalismo de actualización de valores de confianza y reputación usando la experiencia generada a partir de la interacción con otros agentes del sistema. Se estudiará en particular cómo un cambio en las creencias de un agente puede afectar los valores de confianza que el mismo mantiene de sus informantes.
- Analizar y definir un formalismo de confianza y reputación multi-contexto. La confianza y la reputación son propiedades dependientes del contexto. Un modelo de confianza y reputación que administra un único contexto, es diseñado para asociar un único valor de confianza o reputación a cada miembro del SMA sin tomar en consideración el contexto. Sin embargo, un modelo de multi-contexto permite la posibilidad de tratar con varios contextos, manteniendo diferentes valores de confianza y reputación asociados a cada contexto particular de un agente.
- Formalizar un modelo de mantenimiento de confianza que, además de considerar los elementos indicados en los objetivos específicos anteriores, permita integrar a ellos el concepto de la opinión de la mayoría. Dado que muchas propuestas de la literatura consideran únicamente a la opinión de la mayoría como método de evaluación, se planea integrar esta propuesta a las propuestas basadas en un orden de desconfianza entre los agentes.
- Evaluar la relación entre mecanismos de mantenimiento de valores de confianza simbólicos y numéricos, y desarrollar una integración de ambas propuestas.

- Analizar los formalismos mencionados tanto en ambientes colaborativos como competitivos, y considerar tanto valores de confianza como valores de desconfianza

## FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación forma parte de los trabajos que se realizarán para la concreción de un Doctorado en Ciencias de la Computación.

### BIBLIOGRAFÍA

- [AM02] Amazon (2002). Amazon Auctions. <http://auctions.amazon.com>.
- [ARH00] A. Abdul-Rahman and S. Hailes. Supporting trust in virtual communities. In *Proceedings of the Hawaii's International Conference on Systems Sciences (Maui, Hawaii, 2000)*.
- [BK01] Barber, K. S. and Kim, J. (2001). Belief Revision Process based on Trust: Simulation Experiment. In: *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada*. pp.112.
- [Bhu10] Touhid Bhuiyan, Yue Xu, Audun Josang. A Review of Trust in Online Social Networks to Explore New Research Agenda. *International Conference on Internet Computing*, pp. 123-128, 2010.
- [C11] Cholvy L (2011) How strong can an agent believe reported information? 11th European conference, ECSQARU 2011, Belfast, UK, June 29–July 1, 2011, pp 386–397
- [Can98] Cantwell, J. 1998. Resolving conflicting information. *Journal of Logic, Language and Information* 7(2):191–220.
- [Chan10] Changyu Dong, Naranker Dulay. Shinren: Non-monotonic Trust Management for Distributed Systems. *IFIPTM, 2010*. pp. 125-140.
- [Del03] Dellarocas, C. 2003. The digitalization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms. *Management Science*.
- [DGB97] Dragoni, A.F.; Giorgini, P.; and Baffetti, M. 1997. Distributed belief revision vs. belief revision in a multi-agent environment: First results of a simulation experiment. In *MAAMAW*, 45–62.
- [eBAY02] eBay (2002). eBay. <http://www.eBay.com>.
- [FC10] R. Falcone. C. Castelfranchi. Trust theory: A socio-cognitive and computational model, 2010. Wiley Ed. ISBN-13: 978-0470028759
- [FKS02] M. A. Falappa, G. Kern-Isberner, G. R. Simari. Explanations, Belief Revision and Defeasible Reasoning. *Artificial Intelligence Journal*, 141: pp. 1-28, 2002.
- [GrSlo00] Grandison, T. & Sloman, M. (2000). A survey of trust in Internet application, *IEEE, Communications Surveys, Fourth Quarter, 2000*.
- [Han99] Hansson, S. O. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer 1999.
- [HCJ10] Harwood WT, Clark JA, Jacob JL (2010) Networks of trust and distrust: towards logical reputation systems. In: Gabbay DM, van der Torre L (eds) *Logics in Security*, Copenhagen, Denmark.
- [HFCF01] S. O. Hansson, E. L. Fermé, J. Cantwell, M. A. Falappa. Credibility Limited Revision. *The Journal of Symbolic Logic*, Volume 66, Number 4: pp. 1581-1596, 2001.
- [Jos07] A. Josang, R. Ismail, C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, Volume 43, Number 2: pp. 618-644, 2007.

- [Jos08] Audun Josang, Touhid Bhuiyan, Yue Xu, Clive Cox. Combining Trust and Reputation Management for Web-Based Services. *TrustBus*. 2008. pp. 90-99.
- [McKChe02] McKnight, D. H. & Chervany, N. L. (2002). Notions of Reputation in Multi-Agent Systems: A Review. In: Proceedings of the 34th Hawaii International Conference on System Sciences.
- [Mont02] Montaner, M., Lopez, B. & de la Rosa, J. (2002). Developing Trust in Recommender Agents. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multi agent Systems (AAMAS-02), Bologna, Italy. pp.304-305.
- [LLRJ12] Luke Teacy WT, Luck M, Rogers A, Jennings NR (2012) An efficient and versatile approach to trust and reputation using hierarchical bayesian modelling. *Artif Intell* 193:149–185
- [LW99] Liu, W., and Williams, M. 1999. A framework for multi-agent belief revision, part i: The role of ontology. In *Australian Joint Conference on A. I.*, 168–179.
- [LW01] Liu, W., and Williams, M. 2001. A framework for multi-agent belief revision. *Studia Logica* 67(2):291–312.
- [Mui02] Mui, L., Halberstadt, A. & Mohtashemi, M. (2002). Notions of Reputation in Multi-Agent Systems: A Review. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02), Bologna, Italy. pp.280-287.
- [PSM11] Parsons S., Sklar E., McBurney P. (2011) Using argumentation to reason with and about trust. In: *Argumentation in multi-agent systems—8th international workshop, ArgMAS 2011, Taipei, May 3, 2011, Revised Selected Papers*. pp 194–212.
- [SS05] Sabater, J and Sierra C. 2005. Review on Computational Trust and Reputation Models. *Artificial Intelligence Review* 24:33-60.
- [Tang12] Tang Y, Cai K, McBurney P, Sklar E, Parsons S (2012) Using argumentation to reason about trust and belief. *J Log Comput* 22(5):979–1018.
- [TGFS12] Tamargo, L. H., Garcia, A. J., Falappa, M. A., and Simari, G. R. Modeling knowledge dynamics in multi-agent systems based on informants. *Journal of The Knowledge Engineering Review (KER)* 27 (1) (2012) 87-114.
- [Tam12] Luciano H. Tamargo, Sebastián Gottifredi, Alejandro J. García, Marcelo A. Falappa and Guillermo R. Simari. Deliberative DeLP agents with multiple informants. *Journal Inteligencia Artificial v 15 nro 49* pp 13-30. *Asociación Española de Inteligencia Artificial (AEPIA)*. ISSN 1137-3601 (2012).
- [TGFS14] Luciano H. Tamargo, Alejandro J. García, Marcelo A. Falappa and Guillermo R. Simari. On the revision of informant credibility orders. *Artificial Intelligence Journal (AIJ)*. Vol. 212 (2014), pp: 36-58. ISSN: 0004-3702. ELSEVIER SCIENCE BV. doi:10.1016/J.ARTINT.2014.03.006.
- [TGFS16] Luciano H. Tamargo, Sebastián Gottifredi, Alejandro J. García and Guillermo R. Simari. Sharing beliefs among agents with different degrees of credibility. *Knowledge and Information Systems (KAIS)*. Vol. 47. Num. 43. (2016), pp: 1-33. SPRINGER. ISSN: 0219-1377.