

Crenças de Grupo como Instrumento de Formação da Reputação

Tiago L. Schmitz¹, Jomi F. Hübner¹ e Carine G. Webber²

¹ Departamento de Engenharia de Automação e Sistemas - Universidade Federal de Santa Catarina
PO Box 476, Florianópolis, SC, CEP 88040-900, Brasil
tlschmitz, jomi@das.ufsc.br

² Centro de Computação e Tecnologia da Informação - Universidade de Caxias do Sul
R. Francisco Getúlio Vargas, 1130 - Caxias do Sul, RS, CEP 95070-560, Brasil
cgwebber@ucs.br

Resumo

A qualidade das interações entre agentes é um fator determinante na eficiência de um SMA. Uma das ferramentas utilizadas para melhorar essas interações é o uso do conceito de reputação. Os modelos cognitivos de reputação são modelos recentes e dentre eles o ForTrust[6], apesar de sua excelente fundamentação teórica, que faz uso de alguns conceitos que ainda não possuem um modelo de implementação. Nesta direção, o presente artigo apresenta um modelo de implementação das crenças de grupo usando o conceito de artefatos, que são estruturas de um ambiente, passivas e orientadas a funções, que podem ser construídas, destruídas e operadas pelos agentes. Propõe-se a construção de um artefato para representar as crenças de um grupo de agentes usando-se como domínio de aplicação a edição colaborativa de documentos da Wikipédia.

Palavras chave: Crenças de Grupo, Reputação, ForTrust, Modelo Cognitivo, Artefato

1. INTRODUÇÃO

Em um sistema multi-agentes (SMA) o comportamento coletivo de agentes autônomos é utilizado como forma de resolver problemas. A metáfora de sociedade traz consigo conceitos de outras áreas, tais como: Sociologia, Psicologia e Economia. Um dos conceitos mais importantes herdados é a interação entre os participantes, função elementar para o funcionamento de um SMA. A qualidade dessas interações é decisiva na concepção da solução, pois um comportamento desleal de um participante ou uma falha na comunicação tem potencial para afastar o sistema da melhor solução. Por exemplo, suponha que um agente A solicita a um agente B um cálculo. O agente B processa a requisição e por um ato delibe-

rado pela sua concorrência com A ou por simples falta de competência em realizar tal cálculo, ele retorna um valor inconsistente. O agente A utiliza tal resposta e o sistema tende a se afastar do resultado desejado.

A questão que permeia este problema é a avaliação das informações e a escolha de parceiros em um ambiente não confiável. Dois conceitos são socialmente utilizados para responder a esta questão: confiança e reputação. A confiança é uma relação entre dois agentes, ou seja, o agente A considera B confiável. A reputação é uma relação de vários indivíduos para com outro, por exemplo, um grupo I acredita na reputação de B para desempenhar uma determinada tarefa. Estes conceitos aplicados a um SMA tornam a ecologia das relações saudável, uma vez que reforçam os comportamentos positivos e desmotivam os negativos [3]. Em SMA abertos, onde agentes entram e saem da sociedade, com frequência é difícil construir uma rede de confiança uma vez que as interações podem ser limitadas a um curto espaço de tempo. O uso da reputação, foco desse trabalho, atribui um grau de aceitação a um indivíduo do grupo para desempenhar uma tarefa sem que necessariamente o agente requisitante tenha interagido com o agente alvo anteriormente. Para tanto utiliza as experiências dos demais participantes do grupo.

O presente trabalho utiliza o modelo ForTrust proposto em [6]. Este modelo de reputação utiliza o conceito de crença de grupo (CG). Uma crença de grupo é um conhecimento comum a um grupo de agentes e representa algo que o grupo acredita. Enquanto as crenças individuais estão claramente representadas no estado mental dos agentes, não é evidente a localização das crenças de todo um grupo de agentes.

O trabalho apresentado neste artigo se encontra em fase preliminar e apresenta uma proposta inicial de implementação de crenças de grupo para a construção de uma

arquitetura de implementação de um modelo cognitivo de reputação.

A seção a seguir apresenta brevemente algumas características e modelos de reputação. A terceira seção apresenta o projeto ForTrust. A quarta seção apresenta o uso de artefatos para a composição das crenças de grupo propondo uma solução para implementar o modelo de reputação ForTrust. A quinta seção apresenta um cenário de aplicação para a reputação baseada em crença de grupos. Por fim, a última seção descreverá o estado atual do trabalho e planos futuros.

2. MODELOS DE REPUTAÇÃO

Existem na área de SMA diversos modelos de reputação. Contudo há uma característica comum a todos os modelos que é a questão temporal. Pode-se dizer que a reputação baseia-se na manutenção e análise de um histórico de comportamento dos participantes, auxiliando a decidir em quem confiar, encorajando comportamentos leais e dissuadindo comportamentos desonestos [4]. As principais propriedades dos modelos de reputação são paradigma, visibilidade, granularidade e fonte de informação [9].

O paradigma trata do modo como é feita a manipulação das informações, dividindo-se em cognitivo e teoria dos jogos. O cognitivo baseia-se em crenças, desejos e intenções. Define a confiança em alguém baseado em um estado mental [9]. Exemplos de modelos nesta categoria são: ForTrust [6], Castelfranchi e Falcone [2]. Baseado na Teoria dos Jogos, que usa funções e métodos probabilísticos [9], tem-se por exemplo: ReGreT [8], Carter et al. [1], schillo et al [10].

A visibilidade trata do escopo da reputação, podendo ser subjetiva ou global. Na visibilidade subjetiva cada agente tem uma visão própria da reputação de seus pares [9], sendo um exemplo é o modelo ReGreT [8]. Na visibilidade global, os agentes compartilham de uma única base de reputação [9], por exemplo: ForTrust [6] e Carter et Al. [1].

A granularidade trata de como a reputação é aplicada ao agente, dividindo-se em duas subcategorias: a livre de contexto e a contextual. Livre de contexto é aquela que o agente possui apenas uma reputação independente dos papéis que desempenha no grupo [9], por exemplo, os modelos Schillo et al. [10] e eBay [5]. Já na contextual, o agente possui uma reputação para cada papel que desempenha no grupo [9], por exemplo, os modelos ForTrust [6] e ReGreT [8].

A fonte de informação trata da origem das informações utilizadas para estabelecer a reputação. Um modelo pode se valer de uma ou mais fontes de informação. Os principais tipos de fontes são: a interação direta, o teste-

munho e o preconceito. Na interação direta as informações são obtidas através das interações de dois agentes; no testemunho as informações são obtidas através de um terceiro agente que fornece informações sobre o agente alvo e no preconceito a reputação é atribuída a um agente com base na sua classe de agente [9].

Dentre os modelos citados nesta seção o ForTrust é o único que não possui uma implementação. Uma das razões para isso está na sua modelagem cognitiva, que faz uso de um conceito que não possui ainda uma arquitetura de implementação. Este fato motiva a busca por tal arquitetura e a verificação de seu potencial, sendo portanto explorado na próxima seção.

3. PROJETO FORTRUST

O projeto ForTrust trabalha com um modelo cognitivo de confiança e reputação. Baseado no modelo de confiança de Castelfranchi e Falcone [2], sua constituição não se resume a uma mera probabilidade atualizada a cada interação dos agentes. A confiança é construída através de uma série de crenças que quando reunidas levam a um estado interno que permite ou não confiar em um determinado agente [6].

No modelo ForTrust, as crenças que fundamentam uma relação de confiança são: objetivo, capacidade, competência e intenção. Sendo assim, para que um agente A confie em um agente B para fazer uma tarefa J é necessário que: A tenha um objetivo X; A acredite que B é capaz de fazer J; A acredite que B tem competência para atingir X fazendo J e A acredite que B tem a intenção de fazer J. O agente A, tendo essas crenças, pode confiar em B [6]. Tomando como exemplo um time de futebol. O técnico tem o objetivo de levar a equipe à vitória, para tanto deve escalar 11 jogadores que acredita serem capazes de levar o time a atingir este objetivo. Ele escolherá para cada posição um jogador que ele acredite competente para a posição; seja capaz de desempenhar o papel satisfatoriamente e tenha a intenção de desempenhar a função.

Para a reputação o caso é similar, contudo faz uso do conceito de crenças de grupo. Uma CG é um conjunto de crenças partilhadas por um grupo, logo, não é mais um agente A que tem uma crença sobre um agente B, mas um grupo I que tem uma crença sobre um agente B [6]. Sendo assim, para que I crie uma reputação para um agente B para fazer uma tarefa J é necessário que: I tenha um objetivo X; I acredite que B é capaz de fazer J; I acredite que B tem competência para atingir X fazendo J; I acredite que B tem a intenção de fazer J. O grupo I tendo essas crenças pode estabelecer uma reputação para B [6]. Considere o seguinte exemplo: na escolha de um representante de uma agremiação os sócios nem sempre conhecem diretamente os candidatos a representante. O grupo de sócios

acaba por escolher o representante através de sua reputação, sendo assim, o grupo acredita que: o candidato pode exercer a função de representante; tem competência para representar satisfatoriamente o grupo e tem a intenção de representar o grupo.

O fato do grupo compartilhar essas crenças aumenta o grau de complexidade da concepção da arquitetura do sistema, pois não temos hoje na proposta de como as CGs são instituídas e concretizadas em sistemas computacionais. Através de um estudo preliminar observa-se que algumas técnicas como consenso, votação, recomendação, entre outras são factíveis para a implementação das CGs. Contudo não são técnicas triviais e não é sabida qual técnica é mais adequada.

4. SUPORTE ÀS CRENÇAS DE GRUPO USANDO ARTEFATOS

O meta-modelo de agentes e artefatos proposto por [11] enriquece o ambiente dos agentes com artefatos que são ali colocados para auxiliá-los na realização e coordenação das suas atividades. Artefatos são estruturas do ambiente passivas e orientadas a funções que podem ser construídas, destruídas e operadas pelos agentes. Por exemplo, uma calculadora seria um artefato enquanto o indivíduo que a manipula seria o agente, ao invés de o indivíduo realizar o cálculo ele requisita que a calculadora execute-o.

Artefatos possuem duas interfaces, A primeira, permite ao agente disparar operações do artefato. A segunda interface permite apenas a observação do artefato [11]. Estas interfaces podem ter grau de visibilidade distintos. Desta maneira, um determinado agente pode visualizar apenas as informações ou disparar operações do artefato dependendo de suas permissões.

Uma proposição inicial para implementação das crenças de grupo é a criação de um artefato CG. Sua função é *manter* um conjunto de crenças entendidas pelo grupo como verdade. Nesta estrutura os agentes pertencentes ao grupo anunciam por meio dos artefatos as proposições que eles entendam relevante para o grupo. Estas proposições não são necessariamente as crenças do agente, mas o que eles desejam informar ao grupo. Os agentes têm liberdade para anunciar qualquer proposição ao grupo, porém apenas em um ambiente sincero esses anúncios correspondem garantidamente as crenças do agente. O artefato ao receber os anúncios dispara uma tarefa que irá sintetizá-los procurando agregar o que vem a ser a crença do grupo.

A figura 1 representa um artefato de CG. Onde (1) é um display que fornece uma lista das crenças do grupo; (2) é um display que fornece um histórico dos anúncios com seus respectivos agentes; (3) é um botão que o agente

pode usar caso ele acredite que o agente alvo satisfaz a crença representada pelo artefato passando por parâmetro Agente alvo (A) e uma atividade (F); por fim, (4) é a negação de (3), o agente pode usar essa operação caso ele não acredite que A satisfaz a crença representada pelo artefato. A maneira como o artefato fará a análise dos anúncios



Figura 1. Exemplo de artefato de CG

é um processo investigado no estágio atual deste trabalho, e incorpora a busca por identificar algoritmos que façam as CG serem as mais adequadas para o grupo. Algumas possibilidades abordadas são: votação, consenso e testemunho. Na *votação*, o grupo seria conclamado a expressar seu voto periodicamente sobre o agente alvo e sobre a sua credibilidade perante a crença que o artefato representa. A proposição com maioria simples dos votos uma crença grupo. A regra que conclama os agentes a votar é descrita no regimento do grupo. O artefato de grupo é responsável por avisar os agentes da necessidade de votar sob pena de exclusão do grupo. No *consenso*, o artefato de CG recebe um novo anúncio e consulta o grupo mediando uma audiência conciliatória para determinar a validade deste anúncio para o grupo. No *testemunho* quando o artefato recebe uma proposição é recalculada a crença de grupo da relação agente alvo e papel proposta com base nas proposições anteriores e no nível hierárquico dos agentes. A proposição de um agente de maior grau tem um peso maior na decisão do que a dos agentes de grau inferior.

A criação do artefato de reputação baseado no projeto ForTrust é um modelo que agrega vários artefatos de CG, onde as crenças que ganham relevância são competência, objetivo, intenção e capacidade. A figura 2 apresenta um arquitetura em estágio embrionário proposta neste trabalho. Nela são observados quatro artefatos CGs cada qual é a representação de uma categoria de crença (competência, objetivo, intenção e capacidade). Estes quatro artefatos recebem anúncios dos agentes e por testemunho os sintetizam em CGs. Uma vez sintetizada a CG, a mesma é informada ao artefato reputação, que sintetizará a reputação baseada no modelo ForTrust. Por fim o artefato de grupo faz a regulamentação dos indivíduos conforme o regimento do grupo e disponibiliza uma lista com os agentes pertencentes ao grupo. Para exemplificar o funcionamento estabeleceremos o uso do testemunho para sintetização das crenças do grupo. Um agente A acredita que B não tem competência para realizar uma tarefa X; A informa ao artefato competência que “Ñ acredita (B,X)”; no Artefato competência é disparado o processo que analisa

todas as proposições recebidas que relaciona o agente B com a tarefa X; É constatado que não há nenhuma proposição positiva de B competente em X; A proposição é ratificada como uma crença do grupo disparando um processo de no artefato de reputação que através do modelo ForTrust classificará o agente B com uma reputação inapto a realizar X.

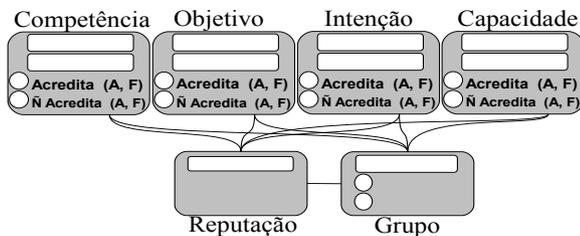


Figura 2. Exemplo de arquitetura de implementação do modelo de reputação com artefato de CG

5. CENÁRIO DE APLICAÇÃO

A Wikipédia, enciclopédia livre, permite a seus usuários a livre criação e edição de artigos contanto que a sua manipulação esteja em acordo com as políticas de uso da ferramenta. O grande domínio abrangido pela enciclopédia, o volume de inserções e alterações executadas diariamente torna complexa a tarefa de validar as novas informações. Esta característica torna o ambiente propício ao surgimento de atos de vandalismo maliciosos ou não, imprecisão de informações, propagandas, entre outros. Para analisar as alterações e para evitar estas desconformidades, a Wikipédia implementa diversos dispositivos. Dentre eles é previsto o uso de uma entidade denominada patrulheiro. Os patrulheiros são usuários normais que assumem uma função de analisar, voluntariamente, a lista de alterações em busca destas desconformidades [7].

Como auxílio aos patrulheiros é proposto o uso de um SMA para a criação de um sistema de reputação dentro da comunidade da Wikipédia. O uso desse tipo de sistema permite estabelecer um grau de confiança entre o patrulheiro e o usuário contribuidor, mesmo que um determinado patrulheiro nunca tenha analisado a contribuição de um referido usuário. A reputação perante a comunidade permite ao patrulheiro inferir um grau de confiança no usuário, facilitando assim a análise das suas alterações. O patrulheiro pode assim se focar nos usuários com reputação duvidosa e analisar suas contribuições [7].

6. PLANOS FUTUROS

No presente momento é feito um estudo mais aprofundado do uso de artefatos e suas capacidades a fim de

garantir a viabilidade de implementação das crenças de grupo. O próximo passo é a definição das regras para a incorporação das crenças do indivíduo no artefato de crenças de grupo. Uma vez definidos estes requisitos, será iniciada a modelagem e a implementação da arquitetura. A avaliação será realizada através de uma série de experimentos analisando o tempo de detecção de danos em relação ao tempo decorrido.

Referências

- [1] J. Carter, E. Bitting, and A. Ghorbani. Reputation formalization for an information-sharing multi-agent system. *Computational Intelligence*, 1:515–534, 2002.
- [2] C. Castelfranchi and R. Falcone. Principles of trust for MAS: Cognitive anatomy, social importance and quantification. *ICMAS*, 1:72–79, 1998.
- [3] C. Castelfranchi, R. Falcone, B. Firozabadi, and Y. Tan. Special issue on trust, deception and fraud in agent societies. *Applied Artificial Intelligence Journal*, 1:763–768, 2000.
- [4] S. J. Cesare. Uma ontologia funcional de reputação para agentes. Master's thesis, Escola politécnica da Universidade de São Paulo, 2005.
- [5] eBay. eBay. <http://www.eBay.com/>, 2009.
- [6] A. Herzig, E. Lorini, J. F. Hübner, and L. Vercouter. A logic of trust and reputation. *Logic Journal of the IGPL*, 2009.
- [7] Y. Krupa, L. Vercouter, J. F. Hübner, and A. Herzig. Trust based evaluation of wikipedia's contributors. In *Proc. of 10th Annual International Workshop*, volume 5881 of LNCS, pages 148–161, 2009.
- [8] J. Sabater and C. Sierra. Regret: A reputation model for gregarious societies. *Proc. of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, 1:61–69, 2001.
- [9] J. Sabater and C. Sierra. Review on computational trust and reputation models. *Kluwers Academic Publishers*, 1:27, 2005.
- [10] M. Schillo, P. Funk, and M. Rovatsos. using trust for detecting deceitful agents in artificial societies. *Applied Artificial Intelligence*, 1, 2000.
- [11] M. Viroli, T. Holvoet, A. Ricci, K. Schelfhout, and F. Zambonelli. Infrastructures for the environment of multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 14(1):49–60, 2007.