

1 Introdução

Segundo [17], **confiança** é a segurança, certeza daquele que tem fé na probidade de alguém e **reputação** é um conceito atribuído a uma pessoa por parte da sociedade em que vive, estritamente ligado ao comportamento e utilizado para medir o grau de confiança.

Sistemas de reputação coletam, distribuem e agregam informações sobre o comportamento anterior dos seus participantes. Estes sistemas ajudam os participantes envolvidos a decidir em quem podem confiar e a evitar a participação daqueles que não são confiáveis, desonestos [21]. Sistemas baseados na tecnologia de agentes podem utilizar os conceitos de confiança e reputação através de mecanismos que têm o objetivo de inibir o mau comportamento dos agentes nas sociedades virtuais. Em tais sistemas, o mau comportamento de um agente de software corresponde a não execução de uma ação considerada obrigatória ou a execução de uma ação considerada proibida. O mau comportamento, quando detectado, pode causar a diminuição da reputação do agente e por tanto a diminuição da confiança dos demais agentes a este agente. Em sistemas multi-agentes abertos temos sociedades de agentes heterogêneos, implementados por diferentes desenvolvedores, com objetivos semelhantes ou não. Surge então a importância da existência de tais mecanismos para identificar agentes que não se comportam adequadamente.

Em alguns sistemas são utilizadas normas para definir o que os agentes de software devem fazer e o que não podem fazer, seja restringindo efetivamente as suas ações ou então permitindo violações de normas que poderão ser punidas de alguma forma. Mecanismos de reputação têm um papel importante neste cenário, com o objetivo de avaliar os agentes, permitindo que todos os agentes que habitam o sistema possam identificar aqueles que não estão se comportando como deveriam, agentes defeituosos ou mal intencionados, e desta forma inibir a proliferação de violações que podem levar o sistema ao caos [10].

1.1. Definição do problema

Em sistemas multi-agentes abertos os agentes de software devem saber identificar em quais agentes podem confiar. Diversos modelos de reputação foram propostos nos últimos anos, (centralizados e descentralizados) para dar suporte aos agentes ao selecionar parceiros. Tais modelos têm como objetivo selecionar os melhores agentes naquele contexto e evitar aqueles que podem trazer algum prejuízo. Os modelos atuais não conseguem informar aos agentes de software que participam de um sistema aberto se um agente desconhecido é confiável. Este é o problema que queremos resolver.

Modelos de reputação centralizados tais como eBay [7], têm a intenção de informar ao comprador a atuação do vendedor em negociações anteriores. No eBay os agentes são pessoas. Seu sistema de reputação reúne todas as avaliações enviadas pelos agentes participantes, porém não existe nenhum mecanismo para inibir o envio de avaliações falsas.

Nos modelos de reputação descentralizados os próprios agentes de software estão dotados de um mecanismo de reputação capaz de fazer avaliações dos agentes com quem interagiram. Tais modelos não funcionam de forma eficaz em sistemas multi-agentes abertos de larga escala, porque os agentes podem ser desconhecidos (existem agentes com os quais o agente em questão nunca interagiu) ou então o número de interações diretas anteriores pode ser insuficiente para que as avaliações das reputações possam ser consideradas consistentes com o comportamento do agente. Estes dois pontos estão relacionados com a necessidade de ter interagido com um agente para saber se este agente é confiável ou não.

Alguns pesquisadores exploraram este problema, entretanto as soluções adotadas não resolveram o problema em definitivo. Algumas abordagens tais como REGRET [22, 23, 24], propõem o uso de testemunhas que já interagiram anteriormente com um determinado agente para repassar o resultado desta interação, ou seja, a reputação deste agente do ponto de vista do outro agente que já interagiu com ele. Porém estas abordagens possuem três desvantagens principais. A primeira está relacionada à confiabilidade do agente testemunha que está fornecendo a reputação de outro agente. Testemunhas podem mentir. Um

agente de software mentir significa fornecer informações incorretas. A segunda está relacionada ao fato dos agentes terem interesses próprios e não estarem dispostos a compartilhar suas informações sem ter alguma contra-partida [20]. A terceira está relacionada com o tamanho do sistema. Se considerarmos sistemas de larga escala, um agente estará procurando por testemunhas dentro de um universo grande e o tempo / custo desta procura pode ser limitado.

O modelo FIRE [13,14] propõe o uso de reputação certificada onde os próprios agentes de software são capazes de fornecer avaliações sobre suas atuações anteriores para os agentes com quem deseja interagir como se fossem cartas de referência. Estas avaliações são feitas e certificadas pelos agentes com quem interagiu. O principal problema desta abordagem está no fato de que os agentes podem mascarar a sua verdadeira reputação fornecendo aos agentes uma reputação superestimada, calculada por um algoritmo que considera apenas as melhores avaliações que receberam anteriormente de outros agentes.

Outra desvantagem encontrada nestas abordagens está relacionada à quão antigas são as informações utilizadas para cálculo da reputação. Informações antigas podem ter como resultado uma reputação que não corresponde ao comportamento atual do agente. Uma carta de referência sobre uma interação que ocorreu há muito tempo pode não refletir o comportamento recente de um agente.

1.2. Proposta

Este trabalho propõe um modelo de reputação híbrido no qual é utilizada uma abordagem descentralizada combinada com uma outra (semi-) centralizada. Em nossa abordagem os agentes de software são capazes de avaliar as atuações dos agentes com os quais interagiram e calcular as reputações de acordo com as suas próprias experiências, de acordo com o seu ponto de vista sobre o desempenho do agente, caracterizando assim a abordagem descentralizada. Além disso, os agentes de software também possuem a sua disposição organizações com a habilidade de avaliar e armazenar as reputações dos agentes que estão atuando nestas organizações. As organizações avaliam o comportamento dos agentes de acordo com violações de normas aplicadas aos agentes e estabelecidas pelas

organizações. Violações de normas que se aplicam a um agente geram alterações na reputação deste agente.

Estamos assumindo que sistemas multi-agentes de larga escala são compostos por uma hierarquia de grupos ou organizações onde agentes atuam, i.e., desempenham papéis. As organizações são entidades capazes de fornecerem reputações confiáveis e imparciais dos agentes com base nos testemunhos de violações de normas enviados por outros agentes. A hierarquia de organizações estabelece a abordagem (semi-) centralizada do nosso modelo de reputação porque cada suborganização centraliza um conjunto de reputações dos agentes que desempenham papéis nesta organização e cada superorganização grupa as reputações que estão distribuídas nas suas suborganizações. As reputações fornecidas pelas organizações são importantes porque:

- (i) Agentes podem obter informações sobre outros agentes com quem eles nunca interagiram;
- (ii) Agentes podem atualizar a reputação dos parceiros com quem eles não interagem há muito tempo;
- (iii) A fonte de informação utilizada para calcular as reputações é confiável, uma vez que os testemunhos são previamente julgados;
- (iv) As reputações não estão influenciadas pelo ponto de vista de um agente. A reputação é calculada da mesma maneira para todos os agentes, de forma imparcial, dependendo apenas da norma violada;
- (v) São oferecidas diferentes reputações para um mesmo agente que levam em consideração o contexto da violação. Pode se considerar no cálculo da reputação apenas violações de uma norma específica (independentemente do papel desempenhado pelo agente), violações de normas enquanto o agente desempenhava um determinado papel ou então se considerar todas as violações de normas do agente. Além disso, as reputações oferecidas podem ser do ponto de vista local (i.e., do ponto de vista de uma organização), e também do ponto de vista global, dentro de uma hierarquia organizacional.

1.3. Principais Contribuições

As principais contribuições deste trabalho são:

1. Proposta de um modelo de reputação híbrido, reunindo os benefícios de abordagens descentralizadas e centralizadas encontradas na literatura e detalhadas na seção 2. O mecanismo descentralizado do modelo híbrido pode ser implementado utilizando modelos já existentes na literatura [13, 14], [22, 23, 24], [1] e [31, 32], enquanto que o mecanismo (semi-) centralizado é a principal contribuição do modelo proposto e está apresentado nesta dissertação.
2. Implementação do mecanismo centralizado do modelo proposto através da criação do *framework* chamado REPORT (REPUtation-ORganization-Testimony) para o cálculo de reputações de agentes de software baseado em testemunhos.
3. Avaliação da nossa proposta através de um experimento que compara o seu desempenho com o sistema REGRET, um dos sistemas de confiança e reputação mais citados na literatura.

1.4. Organização da dissertação

A dissertação está estruturada da seguinte forma:

- No Capítulo 2 são apresentados os trabalhos relacionados destacando as principais características dos trabalhos encontrados na literatura sobre confiança e reputação;
- No Capítulo 3 é apresentada a nossa proposta, o modelo híbrido, combinando abordagens de cálculo de reputações descentralizadas e centralizadas;
- No Capítulo 4 é detalhado o *framework* REPORT para o cálculo de reputações baseado em testemunhos, representando a abordagem centralizada do modelo híbrido proposto;

- No Capítulo 5 é apresentado o estudo de caso *Consolidador de Cargas*, exemplificando o uso do *framework* em sistemas multi-agentes abertos;
- No Capítulo 6 é apresentado um experimento comparando o modelo híbrido com o sistema REGRET;
- No Capítulo 7 são apresentados os trabalhos futuros e a conclusão desta dissertação.