SIMULAÇÃO E ANÁLISE DO MERCADO GERADO POR SPAMMERS E POTENCIAIS CONSUMIDORES

CÉSAR FERNANDES TEIXEIRA

SIMULAÇÃO E ANÁLISE DO MERCADO GERADO POR SPAMMERS E POTENCIAIS CONSUMIDORES

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Orientador: Virgílio Augusto Fernandes de Almeida

Belo Horizonte

de 2010

Teixeira, César Fernandes
T266s Simulação e Análise do mercado gerado por spammers e potenciais consumidores / César Fernandes Teixeira. — Belo Horizonte, 2010
xxii, 83 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de Minas Gerais
Orientador: Virgílio Augusto Fernandes de Almeida

1. Modelagem Econômica - Teses. 2. Simulação (computação) - Teses. 3. Spam - Teses. I. Orientador II. Título.

CDU 519.6*86(043)

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Simulação e análise do mercado gerado por spammers e potenciais consumidores

CESAR FERNANDES TEIXEIRA

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. VIRGÍLIO AUGUSTO FERNANDES ALMEIDA - Orientador
Departamento de Ciência da Computação - UFMG

PROFA. CRISTINA DUARTE MURTA
Centro Federal de Educação Tecnológica - CEFET - MG

PROF. DORGIVAL OLAVO GUEDES NETO
Departamento de Ciência da Computação - UFMG

PROFA. JUSSARA MARQUES DE ALMEIDA
Departamento de Ciência da Computação - UFMG

Belo Horizonte, 18 de dezembro de 2009.

Agradecimentos

Agradeço primeiramente à minha família, principalmente à minha mãe e minhas irmãs. O apoio e descontração de vocês foi fundamental para mim. Muito obrigado por não me internarem pelas inúmeras vezes em que andei pela casa falando sozinho coisas incompreensíveis (até para mim) sobre a dissertação.

Agradeço ao meu orientador, professor Virgílio Almeida, não somente pela ótima orientação mas pela diversão que foi trabalhar com você. Com certeza pude aprender lições valiosas durante o mestrado que sem dúvida nenhuma não poderia ter aprendido em outro local ou com outra pessoa.

Agradeço aos meus amigos de laboratório, essas pessoas divertidas e inteligentes com quem tive o prazer de trabalhar e conviver. Agradeço aos amigos de mais longa data Fabrício, Matheus, Adriano Veloso e Djim. Agradeço também aos que não estavam mais no laboratório na época da defesa: Fabiano, Fernando e Gustavo Siqueira. Por fim, agradeço aos novos amigos: Rauber, Emanuel, Tatiana, Marisa, Tiago e Gabriel.

Agradeço aos meus amigos dos tempos de graduação, com quem pude conviver também durante o mestrado. Agradeço principalmente a Luciana Fujii, Douglas e David.

Agradeço aos meus amigos fora da computação, que sempre com bom humor e simpatia me lembravam que há vida fora da computação :-) Agradeço principalmente a Carlos Eduardo, Horácio, Danilo e Frederico.

Agradeço ao grupo de dança de salão do ICB. Com certeza a 1:15 por semana de aula fizeram uma diferença enorme, tanto em relaxamento quanto diversão. Agradeço principalmente à Karin, Alessandra, Vera, Cláudia, Ana Paula, Dulce, Daniel, Ronaldo e Roberto.

Resumo

Spam é um problema recorrente que gera perdas financeiras grandes para toda a sociedade. Porém, o spam é baseado em uma relação de consumo. A existência, ou possibilidade de existência, de uma demanda gera uma oferta via spam.

Portanto, entender esse mercado gerado entre spammers e consumidores é fundamental. Afinal, através dessa compreensão é possível entender o efeito de diversos fatores fundamentais, o resultado e deficiências de técnicas anti-spam já estabelecidas e propor alternativas.

Este trabalho realiza uma modelagem, baseada em agentes e em conceitos e modelos econômicos , do mercado por trás do spam. Além disso, as análises realizadas fornecem contribuições na compreensão do spam e de seu mercado e na inferência das conseqüências, nesse mercado, de estratégias anti-spam comuns. Como objetivo final, sugestões sobre melhorias e novas estratégias são propostas.

Palavras-chave: Spam, Modelagem econômica, Simulação baseada em agentes.

Abstract

Spam is an important problem that incurs in huge financial losses. However, spam is based on consumption. The existence, or the possibility of existence, of a demand generates a supply offered through spam.

Therefore, understanding this market generated by spammers and potential consumers is imperative. This is due to the fact that understanding this market can bring valuable information about the effects of fundamental factors, the consequences and deficiencies of the available anti-spam techniques, and to propose alternatives.

This work creates a model, based on agents and in economics concepts and models, of the market behind spam. Moreover, the analysis done contributes to understanding spam and its market, to inferring the consequences, in this market, of the usual anti-spam techniques. As a final goal, suggestions about improvements and new strategies are proposed.

Keywords: Spam, Economic Models, Agent-based simulation.

Lista de Figuras

Lista de Tabelas

Sumário

Capítulo 1

Introdução

1.1 Motivação

Spam é um problema recorrente no mundo da Internet. Apesar de atualmente existirem filtros poderosos capazes de impedir que a grande maioria dessas mensagens alcance a caixa de entrada do usuário, ainda assim esse é um problema relevante. Atualmente, estima-se que 80% [51] dos *emails* gerados sejam *spam*. Há estatísticas mais atuais que afirmam que esse volume atingiu a casa dos 90% [35]. Esse volume extra gera desperdícios, afinal, banda e recursos computacionais precisam ser alocados para tratar essas mensagens. Além disso, o mesmo estudo ([51]) analisa o *spam* sob a perspectiva de energia consumida e conclui que a energia gasta anualmente devido ao *spam* poderia abastecer 2,4 milhões de casas americanas. Ademais, cada mensagem *spam* corresponde à emissão de 0,3 g de CO_2, o que, somando-se todo o volume de *spam* mundial, equivaleria a dar 1,6 milhão de voltas, usando um automóvel, ao redor da Terra. Reconhecendo esse problema, há diversas estratégias de combate ao spam.

De longe a mais popular é a utilização de filtros. Atualmente, é impensável que um provedor de *email* não ofereça um filtro anti-spam. O objetivo do filtro é impedir que *spam* alcance a caixa de entrada dos usuários. Assim, o princípio básico é obter as características dos *spams* que os diferem de mensagens regulares e utilizar essas características na distinção do que é ou não *spam*. Obviamente ele pode, inadvertidamente, excluir mensagens legítimas que se pareçam com *spam*, bem como entregar spams que se pareçam com mensagens legítimas. Entretanto, considera-se que atualmente os filtros são capazes de realizar um bom trabalho.

Infelizmente, os filtros não têm se mostrado capazes de reduzir o volume de *spam* gerado. Algumas análises consideram que, inclusive, a existência de filtros agravou esse volume. A justificativa é que spammers, ao perceberem que suas mensagens não

chegavam à caixa de entrada, aumentaram o volume enviado. Com a utilização de *botnets* [1], esse volume pôde ser incrementado enormemente a baixos custos.

Em uma tentativa de resolver também o problema do volume excessivo de *spam*, estratégias econômicas foram propostas [45; 44; 62; 29; 66; 49]. A idéia central que permeia as abordagens econômicas é aumentar o custo de envio de *spam*. Porém, como a priori é impossível sobretaxar apenas *spammers*, esse aumento no custo é aplicado a todas as mensagens geradas. Claramente o impacto em se aumentar o custo de envio é proibitivamente maior para *spammers* que para usuários legítimos, já que os últimos geram um volume infinitamente menor de emails. Ainda, há variações dessa abordagem que incluem uma em que o pagamento é a realização de cálculos pelo computador do remetente [45] e outra em que o pagamento é cobrado somente se o destinatário considerar a mensagem como sendo spam [29].

Finalmente, outra abordagem é o rastreamento e processo de *spammers*. Nessa linha, entidades se esforçam em entender a atividade de *spammers*, bem como em obter dados que possam, no futuro, identificá-los e acusá-los criminalmente. Porém, a atividade de envio de *spam* é uma atividade internacional por natureza, afinal um *spammer* de um país pode enviar mensagens a partir de um segundo país com o objetivo de atingir pessoas de um terceiro. Portanto, descobrir a fonte do spam, bem como processar os culpados, é um desafio. Vale mencionar que investiga-se que, atualmente, a maior parte dos *spams* de farmácias canadenses sejam originários de *spammers* da Rússia e Ucrânia [21; 54].

Grande parte da dificuldade em se combater *spam* vem da própria dificuldade em se definir o que é *spam*. Geralmente, *spam* é definido como mensagem não-solicitada. Porém, esse conceito não pode ser aplicado na prática, afinal, provavelmente todo primeiro contato via *email* seria considerado spam. Outra definição é a de mensagens indesejadas. Mas esse conceito, apesar de ser o mais popular de *spam*, é abstrato demais e extremamente dependente do destinatário para fornecer uma definição precisa. Em uma tentativa de resolver esse impasse, existe o ato americano CAN-SPAM [4]. Nele, *spam* é definido como toda mensagem não-solicitada, enviada em massa, sem que haja uma maneira explícita de excluir o destinatário da lista de remetentes (*opt-out*). Essa definição, apesar de ser precisa e de permitir a punição de criminosos, não é apreciada por diversos usuários. A principal justificativa é a de que a lei permite que todas as empresas enviem uma mensagem, para todos os usuários de *email* do mundo, sem que a mesma seja considerada spam. Neste trabalho adotaremos a definição de *spam*

[1] Botnets são conjuntos de máquinas de usuários legítimos infectadas. Essas máquinas são controladas remotamente por pessoas mal-intecionadas que causaram a infecção. Geralmente são utilizadas em ataques ou no envio de *spam*.

apresentada pelo CAN-SPAM.

Um fato importante sobre *spam*, como definido pelo ato CAN-SPAM e usado ao longo desta dissertação, é que ele depende de um mercado. Por mais que a maior parte dos destinatários considere *spam* inútil e se pergunte se há alguém que compre, somente há *spam* porque alguém compra ou porque alguém acredita que há quem compraria o que é anunciado. Portanto, se há demanda ou expectativa de demanda, há pessoas dispostas a suprir essa demanda. Dessa forma, *spam* nada mais é que uma ferramenta barata pela qual pessoas capazes de prover uma oferta de um produto tentam alcançar a demanda. Obviamente, como achar essa demanda é complicado, *spam* é enviado a uma quantidade muito grande de pessoas desinteressadas.

Assim sendo, uma abordagem para o problema é estudar esse mercado. Afinal, sem ele não haveria *spam*. Compreendê-lo pode ser fundamental para o entendimento dos fatores que influenciam esse mercado. Por exemplo, aumentar o custo do *spam* enviado é uma solução que trás alto ganho, como sugerem as estratégias econômicas? Ou a melhor abordagem é investir constantemente em aperfeiçoar o filtro? Como a fraude, ou seja, a probabilidade de um *spammer* não entregar corretamente o produto comprado, interfere nesse sistema? A partir desse entendimento, pode-se avaliar o real impacto de estratégias anti-spam, onde elas falham, onde acertam e como a sua existência influencia o mercado. Ainda, pode-se, através desse entendimento, sugerir novas abordagens ou estratégias complementares, de maneira a criar um cenário mais favorável que desestimule a geração de *spam*.

1.2 Objetivos

O objetivo desta dissertação é fornecer subsídios para a criação de novas técnicas anti-spam e aperfeiçoamento das já existentes. Outro objetivo é o estudo de fenômenos e cenários que podem facilitar o combate ao *spam*, bem como de situações que precisam ser evitadas, já que viabilizam e tornam o *spam* uma atividade altamente lucrativa. Para tal, o mercado formado por *spammers* e potenciais consumidores é estudado. Focar no mercado permite uma visão mais abrangente do problema, gerando análises mais completas e robustas. Para realizar tal abordagem, esta dissertação modela e simula esse mercado. Assim, através do estudo de cenários diferentes, é possível uma melhor compreensão do sistema.

1.3 Contribuições

Esta dissertação apresentou diversos resultados interessantes:

1. Modelo e simulador baseado em agentes para a representação do mercado composto por *spammers* e usuários.

2. Análises qualitativas do efeito de fatores e combinações de fatores em cenários de interesse. Essas análises permitem a compreensão do mercado, entendimento das conseqüências e impactos de várias ações, bem como o planejamento e melhoria de estratégias anti-spam.

3. O efeito da confiança foi entendido. Confiança se refere a quanto o destinatário confia que receberá o item corretamente ao comprá-lo de um dado *spammer*. Foram estudadas duas fontes principais para a formação de confiança em um *spammer* desconhecido: experiência pessoal com a compra de produtos de *spammers* e dados externos sobre o *spammer* em questão. A utilização de experiência externa se mostrou bastante eficaz no combate ao *spam* em situações em que há altas taxas de fraudes facilmente identificáveis.

4. O efeito da probabilidade de fraude por parte dos *spammers*, bem como da capacidade dos usuários em reconhecer essas fraudes, foi estudado. A probabilidade de fraude se refere as chances de um dado *spammer*, do qual foi realizada uma compra, não entregar o produto corretamente. Devido à possibilidade de os usuários inferirem a confiança em um dado *spammer* antes de uma compra, o aumento na taxa de fraudes pode ajudar no combate ao *spam*. Porém, em situações em que essas fraudes são dificilmente reconhecidas, esse ganho é menor.

5. Compreensão do efeito, no mercado, da melhoria na qualidade do filtro (redução nos falsos-negativo [2]). Nem sempre é positivo para o sistema melhorar o filtro, da mesma forma que nem sempre é negativo piorá-lo.

6. Análise, em diferentes contextos, dos custos dos *spammers* e da relevância desses custos na redução do lucro. Por exemplo, aumentar o custo por mensagem enviada perde o seu efeito à medida que a popularidade do produto vendido aumenta.

7. O uso da caixa de *spam* foi estudado. Assim, pode-se inferir como a utilização da caixa de *spam* na busca de fornecedores impacta o sistema e a eficácia dos

[2] taxa de erros que o filtro comete, ao classificar mensagens *spam* como legítimas

filtros. Como foi percebido, o uso da caixa de *spam* anula o ganho obtido com a melhora no filtro, quando esse ganho ocorre.

8. Estudo do impacto do interesse do usuário no produto ofertado via spam. O grau de interesse exerce um papel fundamental no lucro que *spammers* podem obter, portanto, menor interesse significa melhor combate ao *spam*.

9. Investigação da influência, no mercado, do preço cobrado por empresas que não realizam *spam*. Como uma investigação complementar, estudou-se o impacto causado pelo grau de conhecimento do usuário acerca dos preços praticados fora do mundo dos *spams*. À medida que o grau de informação aumenta, o lucro de *spammers* tende a diminuir. Além disso, com o aumento no preço externo, devido, por exemplo, à dificuldade em se obter o produto de fornecedores não-spammers, maior o lucro dos *spammers* e maior o ganho dos usuários durante uma compra bem-sucedida.

10. Aplicação dos resultados no estudo do *spam* de farmácias canadenses, um dos *spams* mais populares atualmente [54].

1.4 Organização da Dissertação

Esta dissertação é organizada da seguinte forma. Inicialmente, um capítulo contendo a introdução é apresentado. Nele os objetivos, contribuições e motivação para o trabalho são descritos.

A seguir há o capítulo de referencial teórico. Nele, as principais estratégias e técnicas para se abordar, estudar e remediar o problema do *spam* são estudadas. Além disso, apresenta descrição de estudos econômicos, teóricos ou aplicados em ciência da computação, bem como de modelos e simulações baseadas em agentes.

O capítulo 3 apresenta o modelo e a simulação realizadas no trabalho. Detalha as equações utilizadas, seu embasamento, o ambiente de programação do simulador e seu funcionamento.

O capítulo 4 apresenta os parâmetros utilizados para a simulação e seu embasamento. Além disso, apresenta os resultados obtidos através da simulação de cenários de interesse.

Por fim, o caítulo 5 traz os comentários finais e possíveis direções futuras de trabalho.

Capítulo 2

Referencial Teórico

2.1 Teorias econômicas

Atualmente, diversas teorias econômicas têm sido utilizadas no estudo de *spam* e de comportamentos maliciosos, como o *phishing* [1]. Essas teorias serão explicadas aqui, bem como sua relação com o trabalho executado.

Akerlof [28] apresenta o conceito de *Mercado de Limões*. Ele modela um mercado com consumidores e vendedores em que há assimetria de informação. Mais especificamente, os vendedores sabem exatamente a qualidade do produto vendido, enquanto os consumidores somente podem estimar a qualidade média. Assim, como o comprador não pode reconhecer a qualidade do produto, estará disposto a pagar, no máximo, o preço justo para um produto de qualidade média. Portanto, o comprador estará disposto a pagar o valor justo para um produto de qualidade média, tanto para produtos de alta qualidade (que valem mais que o valor pago) quanto para produtos de qualidade inferior (que valem menos que o preço pago). Se o usuário pagar somente o valor correspondente a um produto de qualidade média, isso afastará do mercado vendedores com produtos de qualidade superior à média, afinal, estariam recebendo um valor inferior ao justo para o seu produto. Dessa maneira, a qualidade média dos produtos ofertados iria diminuir, decrescendo também o preço que os consumidores estariam disposto a pagar, já que no modelo de Akerlof, o consumidor sempre sabe a qualidade média. Diversas interações nesse mercado levariam o preço a tender a zero, eliminando completamente o mesmo.

O trabalho de Herley [41] se baseia no conceito de *Mercado de limões* e em

[1] *Phishing* são mensagens não-solicitadas com o objetivo de realizar uma fraude. Geralmente, o remetente finge ser outra pessoa ou entidade, com o objetivo de enganar o destinatário e criar credibilidade suficiente para que o mesmo informe dados sigilosos ou realize pagamentos.

dados sobre os preços de identidades e dados de cartões de crédito comercializados por *phishers*[2]. Os preços praticados nesse mercado são muito inferiores ao valor de um item de alta qualidade (especialmente dados de cartões de crédito, que podem render fraudes de milhares de dólares por cartão). A partir disso, os autores inferem que o mercado contém uma fração muito grande de vendedores com itens com qualidade baixa, ou, até mesmo, vendedores que não possuem qualquer informação a ser vendida. Assim, aplicando diretamente a teoria de Akerlof, o mercado de *phishing* possui a tendência de desaparecer.

Em relação a *spam*, temos também uma assimetria de informação, já que o *spammer* sabe qual a qualidade de seu produto, enquanto o comprador não tem esse dado. Porém, a teoria de Akerlof se baseia no fato de que o comprador pode estimar a qualidade média do mercado. E isso não é verdade nesse mercado. Um comprador não sabe a probabilidade de ser enganado por um *spammer*. Então, incorporamos em nosso modelo esse conceito de *Mercado de Limões* e comportamento fraudulento, porém, o comprador não sabe a qualidade média. Ele tem uma expectativa inicial e, através de informações sobre experiências de outros consumidores e de experiência própria, ele estima essa qualidade média. No caso específico do simulador, essa qualidade média é chamada de confiança.

O trabalho de Coase [33] propõe uma nova maneira de se compreender e avaliar o impacto de situações de conflito econômico. Por exemplo, imaginemos duas empresas, uma fábrica e uma fazenda de trigo. Suponhamos que a fábrica gere poluição, que afeta negativamente a fazenda. Geralmente, a decisão mais corrente é taxar a fábrica para que o fazendeiro seja recompensado. Porém, o trabalho de Coase propõe uma outra abordagem. É necessário analisar o ganho para a sociedade no caso de não haver a taxa e compará-lo ao caso em que há a taxa. O ganho para a sociedade é a soma do ganho obtido devido aos produtos manufaturados pela fábrica mais o ganho com o que foi produzido pela fazenda em cada cenário. Assim, nem sempre é proveitoso realizar a taxação.

Esse trabalho tem aplicação direta no caso em questão. O envio de spam gera um gasto extra para toda a sociedade. Porém, uma mensagem spam pode ser valiosa para alguma pessoa. A partir dessa idéia, o simulador a ser proposto incorpora uma métrica que avalia o ganho, para a sociedade, com o *spam*. Na avaliação dessa métrica foram utilizados valores médios para a perda que cada usuário tem ao receber *spam*. O ganho foi atribuído como a economia financeira que o usuário interessado faz ao comprar um produto via spam ao invés de adquirí-lo de outra maneira. Obviamente, o ótimo a ser

[2] *Phishers* são indivíduos ou grupos que realizam *phishing*.

alcançado é que esse valor seja o máximo possível.

Outro conceito fundamental é o de *Tragédia dos Comuns* [39]. Se existir uma área comum da qual todos possam tirar proveito e cujos danos serão divididos entre todos, essa área será explorada ao máximo. Por exemplo, consideremos um rio com peixes e pescadores. Cada pescador, se aumentar o número de redes, conseguirá pescar mais e aumentar seu lucro. Porém, isso causará um impacto ambiental, que será dividido por todos, já que todos os pescadores, não importando a sua taxa de pesca, terão menor quantidade de pesca no futuro. Nesse caso, é esperado que cada pescador aumente o número de redes, já que o ganho com esse aumento supera a perda que ele terá (uma fração do impacto total) ao adicionar essa rede.

Esse conceito é utilizado para explicar porque a taxa de *spams* tem crescido [58]. A idéia é que os recursos de Internet e infraestrutura para envio de *emails* é um bem comum. Então, faz sentido que seja explorado ao máximo, pelos *spammers*. Também em [40], o problema do *phishing* é modelado usando esse conceito. A idéia é que os recursos financeiros disponíveis a serem roubados constituem o bem comum. Então, caso os *phishers* explorem em excesso esse bem comum, ele tenderá a acabar, atingindo uma situação em que cada *phisher* obterá o mínimo possível de sua atividade. E, como descrito anteriormente, essa exploração em excesso é esperada.

No simulador proposto, esse conceito foi usado para justificar a suposição de que existem *spammers* fraudulentos. Assim, como há assimetria de informação, é esperado que surjam pessoas que ajam de maneira desonesta. Afinal, existe um bem comum (recursos de potenciais compradores) que pode ser facilmente explorado através de atividade fraudulenta. Ademais, o peso de uma fraude é dividido por todos os *spammers*, quer sejam fraudulentos ou não. Porém, não podemos inferir qual a fração de comportamento fraudulento.

2.2 Técnicas Anti-spam

Devido ao fato de o *spam* ter se tornado um dos grandes problemas que assolam a utilização da Internet, diversas técnicas foram criadas e desenvolvidas com a intenção de contê-lo. Esta seção tem por objetivo expor, brevemente, as técnicas mais usuais e difundidas de combate ao *spam*.

2.2.1 Restrições ao envio do *spam*

Para desonerar a rede, diminuindo o tráfego devido ao *spam*, mecanismos foram criados para dificultar o envio de mensagens não solicitadas.

Uma das maneiras mais usuais de envio de *spam* é através de máquinas zumbis. Essas máquinas, de usuários comuns, são infectadas e passam a ser controladas externamente, enviando *spams* sem que o dono da máquina tome conhecimento disso. Para combater essa atividade, ou pelo menos abrandá-la, Xie [53] sugere uma maneira de se identificar tais máquinas. Ao identificá-las seria possível bloquear o tráfego gerado por elas, impedindo o envio de *spam*.

O estudo de Goodman [46] trata da situação em que o *spammer* utiliza contas gratuitas em provedores de *email* legítimos para o envio de *spam*. Os autores demonstram que a estratégia adotada de restringir o limite máximo diário de mensagens (o Hotmail restringe esse valor a 100 emails por dia) não gera bons resultados. A justificativa para isso é que seria fácil criar novas contas e utilizá-las até o limite definido.

Aquele trabalho avalia também a utilização de *Human Interactive Proofs - HIPs*(prova de que a interação é feita por um humano). Um exemplo comum de HIP é o CAPTCHA, que requisita que usuários identifiquem palavras ou conjuntos de caracteres para provar que não se trata de um processo automatizado. O trabalho mostra que os HIPs usados durante o cadastro da conta de *email* não geram qualquer resultado contra o *spam*, já que o custo econômico de se realizar um cadastro manual e, em seguida, utilizar a conta para o envio de *spam*, é irrisório. Os autores também avaliam a estratégia de se utilizar um HIP a cada *email* enviado. Essa abordagem é narrada como bastante desconfortável para o usuário comum, então avaliam a idéia de se obrigar a realização de um HIP a cada n mensagens. Caso o usuário responda x HIPs corretamente, não seria mais necessária qualquer ação especial do usuário ao enviar emails. No artigo, essa estratégia surte tanto efeito quanto a realização de HIPs a cada *email* enviado.

Por fim, o artigo associa a sua estratégia de HIPs ao esquema de reclamações. A cada *email* recebido, seria oferecido ao destinatário uma interface simples para indicar caso o *email* seja *spam*. Assim, o servidor teria condições de identificar mais rapidamente o *spammer* e cancelar a conta, obrigando-o novamente a criá-la manualmente.

Essas estratégias buscam dificultar o envio de *emails* por parte dos spammers. Entretanto, como avaliado em [56], *spammers* também contratam servidores, em diversas partes do mundo, que se dedicam ao envio de *spam*. Além disso, o combate do envio de *emails* por parte de servidores gratuitos precisa da adesão de todos os servidores, pois o *spammer* pode simplesmente migrar de um que tenha adotado a proteção, para outro que não a adote. Concluindo, o combate a máquinas zumbis requer a análise do tráfego gerado pelas máquinas, o que é potencialmente bastante oneroso.

2.2.2 Blacklists e Whitelists

Heckerman e Schwartz [45; 62] apresentam uma visão geral sobre a criação de listas, especialmente para auxiliar o trabalho dos filtros. Uma das possibilidades difundidas é a utilização de *whitelists,* que listam os usuários ou *hosts* nos quais se confia como não-spammer. Assim, qualquer mensagem de alguém nessa lista é recebida diretamente, sem qualquer verificação pelo filtro, evitando-se os falsos positivos.

Outra solução pode ser listar os usuários considerados *spammers*, em uma *blacklist.* Essa técnica é facilmente contornada por emails forjados. Diversos trabalhos [45; 44; 62] sugerem estratégias para identificar o emissor. Algumas das técnicas envolvem assinaturas criptográficas e outras alterações no protocolo SMTP para que seja necessária a identificação do remetente.

Além disso, existem também *blackholes,* que são listas de servidores considerados fontes de *spam* e dos quais não se aceitam quaisquer mensagens. Serjantov [64] há um estudo econômico detalhado que apresenta diretrizes para a decisão de incluir um servidor suspeito na lista. O servidor é tratado como suspeito quando há mensagens legítimas e *spams* sendo recebidos a partir desse servidor. O dilema, nesse caso, é decidir incluir um servidor suspeito e perder as mensagens legítimas ou não incluí-lo e continuar a receber *spam.*

Existem tentativas visando a criação de listas de identificação de servidores para se evitar que *emails* sejam enviados forjando-se o servidor remetente. Essa estratégia tenta assegurar que uma mensagem, dita originária do servidor X, realmente partiu do servidor X. Para tal, foram criados vários padrões [31; 44] tais como DomainKeys [9], proposto pela Yahoo e o SenderID [23], pela Microsoft.

2.2.3 Filtros anti-spam

A técnica mais usual de combate ao *spam* é a utilização de filtros, que impedem que os *spams* alcancem a caixa de entrada do usuário. O grande questionamento dessas técnicas, além, claro, de sua eficácia em barrar mensagens que sejam *spam*, é o de se evitar falsos positivos. Os falsos positivos são as mensagens, que embora não sejam *spam*, são classificadas como tal erroneamente, levando o usuário a perder um email legítimo potencialmente importante. Infelizmente, os índices atualmente calculados para falsos positivos pressupõem que toda mensagem comercial seja *spam*, ignorando a possibilidade de o usuário desejar recebê-la.

Heckerman e Schwartz [45; 62] apresentam diversas técnicas para filtragem de *spam.* Uma das técnicas é comparar o conteúdo de uma mensagem recebida com uma lista de *spams* conhecidos, de forma a determinar se o novo *email* é ou não um *spam.*

Evidentemente, essa técnica não funciona contra novos *spams*, nem contra variações aleatórias no corpo dos *spams* já conhecidos, dependendo da função de computação utilizada para gerar essas variações.

Outra técnica é a utilização de algoritmos de aprendizagem de máquina, mais especificamente através da utilização do *naive Bayes*. A idéia é construir uma rede de palavras e, para cada palavra, interpolar através de uma base de treinamento qual a probabilidade de uma mensagem ser *spam* dada a presença daquela palavra no corpo do texto. A abordagem atualmente é uma das mais populares, porém os *spams* mais atuais e complexos tentam utilizar menos palavras perigosas (*free, money, etc.*) e mais palavras consideradas boas (*hi, oi, etc.*) para tentar enganar os filtros. Alguns filtros mais recentes têm buscado montar redes utilizando, também, combinações de palavras para melhorar a precisão. Por exemplo, em muitos *spams* a palavra *agora* pode estar presente, mas ela apresenta perigo apenas se estiver acompanhada da palavra *compre*, na sentença *compre agora*.

Ainda, alguns filtros possuem modelos de compressão, um para *spam* e outro para não-spam. Uma mensagem é comprimida com os dois modelos; se essa mensagem for melhor comprimida com o modelo de *spam*, então ela é classificada como *spam*; caso comprima melhor com o não-spam, então é dita não-spam. Os modelos de compressão podem ser extremamente complexos, levando a filtros eficientes, algumas vezes mais eficientes que os baseados em *naive Bayes* [45].

Em [42] é apresentada uma estratégia combinada de diversos algoritmos de *machine learning* para a detecção de *spam*. A eficácia foi de 99% com uma taxa de 0.025% de falsos positivos.

Os *spammers*, para combater essas estratégias, têm buscado cada vez mais que suas mensagens se pareçam com mensagens legítimas. Assim, os filtros precisarão ser cada vez melhores e mais treinados para barrar e distingüir entre *spam* e não-spam. Afinal, à medida que os spams se parecerem mais com mensagens legítimas, maior a probabilidade de que uma ferramenta automática tanto deixe um *spam* passar como legítima quanto barre um *email* legítimo. Além disso, a utilização de filtros consome muitos recursos dos servidores de email, já que o processo de detecção de *spam* é complexo e custoso.

2.2.4 Abordagem econômica

Vários trabalhos focam nas relações econômicas para inibir o *spam*. Afinal, caso o *spammer* tenha mais custo com o envio de mensagens que a receita que poderia ter com a venda de seus produtos, não teria lucro, inviabilizando sua atividade.

Os trabalhos [45; 44; 62], descrevem e analisam as abordagens que buscam adicionar um custo fixo, muito baixo, por *email* enviado. Esse custo seria calculado de maneira a trazer pouco ônus ao usuário legítimo, que envia poucas mensagens, e um custo alto para um *spammer*, que envia milhões de emails.

Além da utilização de dinheiro e depósito direto, é também sugerida uma alternativa [44; 45]. O custo seria computacional, ou seja, ao invés de pagar financeiramente por mensagem enviada, o computador do usuário resolveria uma tarefa, que iria requerer um certo tempo de processamento. Assim, um usuário legítimo, enviando poucas mensagens, não teria qualquer problema em fornecer esse tempo de processamento para o envio do email. Por outro lado, o *spammer*, ao enviar milhões de emails, no menor tempo possível, teria praticamente todo o tempo de processamento da máquina ocupado pela resolução dessas tarefas.

Os trabalhos [29; 66] apresentam um arcabouço para inviabilizar economicamente o envio de *spams*. Cada usuário de email determinaria seu custo de atenção: um valor monetário que determinaria um montante que compensaria as perdas desse usuário ao abrir um email que seja *spam*. Um usuário A deseja enviar um *email* para B. Se A conhecer B (se A estiver na *whitelist* de B) então a mensagem é recebida diretamente. Caso contrário, B envia para A um desafio, contendo o seu custo de atenção. A, então, pode ignorar esse desafio e não ter sua mensagem entregue, ou A pode depositar um título monetário com o valor de atenção de B, em nome de B e enviar novamente o email. Assim, B receberia o *email* de A. Ao abrí-lo, poderia decidir se o email é ou não *spam*. Se decidir que é *spam*, receberá o dinheiro correspondente ao título depositado por A. Caso contrário, A receberá seu dinheiro de volta.

Schechter [61] aborda a questão de risco de ataques remotos. Entretanto, sua modelagem de lucro com um ataque, que leva em conta o risco de ser descoberto, aperfeiçoa as fórmulas econômicas mais usuais para modelagem do *spam*, que não incluem esse custo devido ao risco.

Krishnamurthy [49] apresenta uma abordagem que utiliza selos, que são comprados, no envio de emails. O usuário A deseja enviar um email para B. A tem um conjunto de selos já previamente comprados de uma entidade unificada confiável. A, então, anexa um selo em sua mensagem e envia a B. B avalia a mensagem. Se considerar que é um spam, notifica isso. Caso contrário, A recebe seu selo de volta. Assim, se A tiver muitas mensagens taxadas como *spam*, rapidamente ficará sem selos e precisará comprar novos selos. Se não for um *spammer*, nenhuma mensagem será taxada como *spam* e nunca precisará comprar novos selos. O artigo também sugere como implantar a estratégia, fornecendo mecanismos para que *emails* possam ser enviados sem o selo, seguindo o caminho usual atual.

Yang [56] apresenta estimativas do custo de envio de *spam*. Esse trabalho inclui a compra de listas de emails, bem como do aluguel de servidores para o envio nessa estimativa. Além disso, apresenta também uma estimativa de qual a fração, dentre todas os *emails* enviados, precisam ser respondidos para que o *spammer* tenha lucro. Por fim, afirma, de maneira pouco embasada, que o *spam* está fadado a um fim breve e que os *spammers* têm migrado para atividades mais lucrativas, como *phishing* .

A maior crítica contra essas abordagens é a dificuldade de implantação. Como argumentado em [56; 45; 62], para que o sistema se torne robusto, ele deve ser abrangente o suficiente para que um usuário possa optar por não receber *email* da maneira usual, sem arriscar perder *emails* importantes. Esses sistemas impõem modificações drásticas na maneira como os *emails* são enviados. Novas organizações seguras precisariam ser criadas, bem como o próprio protocolo SMTP precisaria ser modificado. Isso é largamente apresentado como inviável na prática.

Um sistema chamado CentMail [7] encontra-se em fase de implantação. O sistema realizará o controle da emissão e pagamento de selos para o envio de mensagens. O sistema funciona da mesma maneira que o proposto por Krishnamurthy [49], porém, ao invés de o pagamento por selo ser realizado à empresa certificadora, ele será feito na forma de doação a uma instuição de caridade. Assim, espera-se incentivar a adoção inicial, já que o pagamento representa uma doação e a divulgação da entidade que recebeu a doação no *email* enviado.

2.3 Ferramentas de Marketing Eletrônico

Além dos diversos filtros anti-spam existentes nos diversos servidores de email, existem também ferramentas comerciais que procuram atacar o problema do *spam*.

O GoodMail [14] é um sistema para facilitar o marketing direcionado. A empresa aceita apenas cadastros de empresas que enviem emails para usuários que se cadastraram ativamente para receberem o email. Além disso, a empresa anexa ao *email* enviado por seu esquema uma opção de reclamação contra o *email* quando o usuário o considerar spam. Se uma empresa atingir uma taxa alta de reclamações, seus emails não poderão mais ser enviados pelo GoodMail. Em troca, o sistema oferece acordos com diversas empresas de email, que colocam os *emails* enviados pelo GoodMail com ícones diferentes na caixa de *email* do usuário e as mensagens não passam pelo filtro anti-spam, garantindo entrega inclusive de *emails* compostos apenas por imagens. A empresa afirma que a utilização de seus serviços gera um aumento no lucro de 20 a 30% para as organizações participantes.

O Boxbe [3] é um provedor de *email* que exige o pagamento para que uma mensagem seja recebida, caso o remetente não esteja na *whitelist* do destinatário. Parte desse pagamento iria diretamente ao usuário. Além disso, existe um cadastro em que o usuário informa suas preferências. O Boxbe também aceita acordos com *spammers*. Caso desejem, podem ter acesso a informações anônimas das preferências de cada usuário e determinar para quais deseja enviar sua propaganda, pagando também pelo envio.

Tem surgido o conceito de *Admediary* [38]. *Admediary* é uma empresa que se responsabiliza em aproximar consumidores interessados e os vendedores capazes de oferecer os produtos nos quais os consumidores se interessam. Assim, usuários se cadastram voluntariamente junto ao mediador e informam seus interesses. O mediador, então, é pago por anunciantes para entregar suas ofertas aos usuários que possuem interesse. Esse modelo de negócios foi implementado em [8; 10; 12; 15; 16]. Alguns desses serviços oferecem pagamento ao usuário por receber emails. Obviamente, essa é uma técnica questionável, já que, apesar de aumentar a base de usuários cadastrados, pode diminuir a habilidade de encontrar usuários interessados, ou seja, pode dificultar o *targeting*.

2.4 Caracterização do *spam* e modelagem através de redes sociais

Existe uma diferença grande entre o comportamento gerado por um endereço de *email* utilizado por um *spammer* e o usado por um usuário legítimo. O endereço de *email* utilizado por um usuário legítimo tem uma pessoa real por trás, com círculos de amizade e padrões de comportamento humanos, enquanto o *email* de um *spammer* tem por trás um *script* de envio de mensagens.

Buscando separar essas diferenças, Gomes [37], realizou uma análise de diversas métricas que diferenciam entre *spam* e *email* legítimo. Por exemplo, existe uma diferença no tamanho médio das mensagens enviadas, no número médio de *emails* no campo *to* e *cc*. Os autores também analisam o período da semana e a hora do dia, e descobre que os *emails* legítimos são mais prováveis de serem enviados de segunda a sexta e no horário comercial, enquanto o *spam* permanece praticamente constante durante os diversos períodos. Além disso, existem claras relações sociais de amizade nos emails legítimos. Exemplificando, se um usuário legítimo A envia uma mensagem para B, e B já enviou mensagem para *C,* existe uma probabilidade grande de que A conheça e se comunique com C. O mesmo não ocorre no caso de *spam*. Assim, se A é um *spammer* e se comunica com B e C, a probabilidade de que B e C se comuniquem

é baixa.

Calais [32] utiliza técnicas de mineração de dados para separar mensagens *spam* em campanhas. Após essa separação, avalia diversas informações e compara as várias campanhas obtidas.

Plice [55] analisa o problema do *spam* segundo a óptica de recursos comunitários. Assim, por exemplo, a rede é um recurso comunitário, que sofre abuso do *spammer*. Além disso, a atenção do usuário também é considerada recurso comunitário, sendo a atenção o tempo que é dedicado à leitura de emails. Dessa forma, os autores apresentam que se um usuário recebe mais *emails* do que pode ler, esse recurso (atenção) foi sobrecarregado. O texto analisa *emails* comerciais e conclui que não há crescimento na taxa de envio de *emails* desse tipo, ou seja, afirma que o problema do *spam* não tem apresentado uma tendência de piorar. Por fim, conclui que os *spams* apresentam sim características temporais, assim, os *spams* também são mais frequentes durante os dias de semana que durante os fins de semana, o que o contrapõe diretamente à conclusão de Gomes [37]. Importante mencionar que ambos utilizam bases de dados de universidades, sendo que [37] utilizou a base, com duração de oito dias, de *emails* da UFMG e [55] utilizou a base de *emails* da Alpha State University com duração de 366 dias. Talvez essa diferença possa ser explicada pelos tamanhos das bases investigadas.

Em [63], é apresentada uma ferramenta baseada em redes sociais que busca reduzir o número de falso positivos em filtros anti-spam. Para tal, utiliza um esquema de certificação. Um usuário A certifica um usuário B como não *spammer*, primariamente por conhecê-lo. Então, se B quiser enviar *email* para algum dos amigos de A (pessoas que certificam que A não seja *spammer*), ele poderá fazê-lo, sem que seu *email* sequer passe pelo filtro.

A grande questão que se coloca através dessas abordagens é a necessidade de se ter um conhecimento grande da rede para se realizar boas inferências. Por exemplo, um servidor apenas com conhecimento local, pode ter conhecimento de A e B em seu servidor. Mas se A envia pela primeira vez um email para B e eles compartilham um amigo em comum, C, em outro servidor, então não seria possível inferir que ambos são legítimos e que a comunicação pode ocorrer.

2.5 Investigações e Meios Legais

Outra linha de combate ao *spam* diz respeito à busca, investigação e processo criminal. Nessa linha, os atos dos *spammers* são monitorados, de maneira a se identificar a origem dos mesmos. Essa linha enfrenta grandes problemas. O primeiro é a dificuldade em

se identificar o culpado já que, obviamente, os *spammers* se esforçam em manter sua identidade em segredo. Por exemplo, Alex Polyakov [2], responsável por diversos crimes virtuais, bem como pela disseminação do *spam* "My Canadian Pharmacy", é conhecido apenas por seu pseudônimo. Apesar de sua ação constante, ainda não existem dados sobre sua real identidade. O segundo é o caráter internacional das operações. Um *spam* recebido em um país pode ter sido proveniente de um segundo país e enviado por um cidadão de um terceiro país. Isso torna não só a descoberta dos responsáveis mais complicada, como também impõe barreiras à punição dos criminosos.

Apesar disso, avanços foram realizados. Muitas informações sobre *spammers* e *spams* conhecidos estão disponíveis [26; 21; 1; 6]. Inclusive, há organizações como a *spam*haus [25] que busca ativamente informações sobre *spammers*, na tentativa de, posteriormente, identificá-los e processá-los. As investigações apresentam o modo como as organizações de *spammers* operam, a ligação entre *spammers* e organizações e os atos criminosos realizados. Esses avanços culminaram em casos famosos, como a punição de *spammers* notórios como Leo Kuvayev e a organização por trás do SanCash [19; 22]. Também culminou na desativação da McColo [50], servidor associado a geração de um alto volume de spam.

Outro avanço importante nessa linha foi o ato americano CAN-SPAM [4]. Esse ato estipula a definição do que é ou não *spam*, facilitando o julgamento e criminalização do *spam*. O ato considera *spam* como sendo toda mensagem não-solicitada que não contenha uma forma de o destinatário se descadastrar da lista de envio e nunca mais receber mensagens do remetente. Esse ato recebe críticas por legalizar o envio de mensagens não-solicitada permitindo que qualquer empresa possa enviar uma mensagem para qualquer endereço eletrônico. Basicamente, ao invés de o usuário demonstrar interesse em receber uma mensagem, ele poderá primeiro receber uma mensagem, perder seu tempo com ela, para depois demonstrar desinteresse e não mais receber mensagens do remetente.

2.6 Modelagem de *spam*

A maior parte dos estudos relacionados a *spam* tem como objetivo detectá-lo e eliminá-lo. Poucos estudos e trabalhos foram realizados no sentido de se modelar o problema, com o objetivo maior de se compreender a situação e avaliar estratégias. Esses estudos são relevantes já que as estratégias de detecção e combate ao *spam* se baseiam em informação antiga. Por exemplo, um filtro é proposto com uma taxa de acerto em relação ao *spam* gerado. Nada se pode afirmar sobre a taxa de acerto desse mesmo

filtro após sua implantação. Afinal, *spammers* modificarão seus *spams* com o objetivo de enganar esse novo filtro. Portanto, compreender a situação e o impacto de diversas estratégias pode ser valioso na decisão de qual a melhor estratégia a se utilizar, quer seja um filtro ou não.

Os trabalhos [57; 58; 55] caminharam nesse sentido. Neles são propostos modelos de simulação para o *spam*. São modelados destinatários e *spammers*. Os destinatários podem receber tanto *spam* quanto *email* legítimo e irá ler uma certa taxa de *emails* por dia. Caso a taxa de chegada de *emails* seja muito alta, ele sofrerá de sobrecarga de informação e deletará, sem qualquer consideração, uma certa quantidade de *emails* recebidos. Os modelos também apresentam o conceito de falso-negativo, não como a taxa de *spam* que o filtro permite passar, mas como a taxa de *emails* inúteis que o filtro entrega. Outra característica é que à medida que os *spammers* aumentam seu lucro, o número de *spams* gerados também aumenta na modelagem. Além disso, os destinatários têm uma probabilidade de compra, caso leiam um *spam*. Essa probabilidade é independente e fixa para cada *spam* recebido. Por fim, o modelo trata o problema através de valores médios e fórmulas fechadas.

Esta dissertação se baseia nesses trabalhos mas, ao invés de seguir um modelo de valores médios, cada *spammer* e cada destinatário é modelado. Assim, é possível modelar melhor o processo de compra do usuário e incluir efeitos como aprendizagem, confiança e fraude. Também, a taxa de falsos-negativos é modelada como sendo a probabilidade de uma mensagem spam ser entregue, não importando se ela será relevante ou não. Afinal, os filtros, na grande maioria das vezes, não fazem essa distinção, distinguindo tão somente entre *spam* e não-spam. Outra diferença importante é em relação à probabilidade de compra. Nos modelos mencionados, para cada mensagem recebida, uma certa fração delas resultará em vendas. Em nosso modelo, esse valor não é fixo. Dessa maneira, podemos modelar concorrência, confiança e o recebimento de múltiplos *spams* para o mesmo produto e fornecedor, que afetam essa fração. Portanto, ao invés de modelar as vendas como uma fração das mensagens, preferimos modelar o interesse do usuário como uma probabilidade. Daí, cada usuário tem uma probabilidade de se interessar por um produto; se ele irá comprar ou não é uma decisão em que outros fatores, que podem variar ao longo do tempo, interferem.

2.7 Modelagem Baseada em Agentes

Modelagem baseada em agentes tem se tornado popular em modelagens econômicas. Agentes são as entidades que possuem comportamentos e que participam do modelo.

A grande vantagem dessa abordagem é que, ao invés de se modelar o sistema como um todo, é possível atribuir comportamentos relativamente simples para os agentes e a interação entre agentes dá origem a fenômenos mais complexos e difíceis de se modelar.

Ao se realizar uma modelagem por agentes, a primeira entidade a ser criada é o agente. Diversos tipos de agentes podem coexistir. Cada classe de agente apresenta seus próprios comportamentos, que, inclusive, podem ser respostas a fenômenos externos ou ações de outros agentes. A forma como os agentes se relacionam também é modelada. Assim, é possivel a modelagem de aprendizado dos agentes e outros comportamentos, cuja modelagem analítica seria extremamente complicada.

Há dois estudos muito importantes em relação a modelagem baseada em agentes. O primeiro [65] apresenta uma revisão acerca dos modelos e problemas abordados em economia, utilizando essa técnica. O segundo [52] apresenta um modelo para *spam*. Porém, esse trabalho foca nas relações sociais e de amizade entre os diversos usuários de email. Assim, o estudo parte da base de dados dos emails dos funcionários da Enron e, a partir daí, modela, usando agentes, a maneira como as relações entre os diversos usuários são estabelecidas. O relacionamento é definido como o envio de email.

Capítulo 3

Modelo e Simulador

3.1 Modelo

3.1.1 Visão Geral

Em linhas gerais, o modelo pode ser descrito da seguinte maneira. Os passos a seguir mostram a ordem das ações modeladas.

1. Os *spammers* enviam mensagens *spam* para todos os usuários do sistema.

2. Essas mensagens chegam ao servidor do destinatário. Elas podem ser bloqueadas automaticamente pelo Mail Transfer Agent (MTA)[1]. No mundo real, esse bloqueio ocorre devido a erros na mensagem, *blacklists* e outros indicadores fortes que indicam que uma mensagem é *spam*. No modelo, a situação é modelada como uma probabilidade de a mensagem *spam* ser sumariamente recusada pelo MTA. *Spams* rejeitados pelo MTA sequer serão considerados pelo filtro e nunca serão entregues sequer à caixa de *spam*.

3. Mensagens não rejeitadas pelo MTA são encaminhadas ao filtro. Seguindo uma probabilidade de falsos-negativos, o filtro pode entregar mensagens *spam* à caixa de entrada do usuário. As outras mensagens são encaminhadas à caixa de spam.

4. O usuário, então, irá acessar sua caixa de *email*. Ele pode possuir interesse no produto comercializado via *spam*. Esse interesse é determinado por uma distribuição de probabilidade, sendo que, cada usuário possui chances de, a cada leitura, ter interesse ou não no produto.

[1]Mail Transfer Agent é um processo ou aplicação que transfere uma mensagem recebida de um computador para outro.

5. Caso o usuário não tenha interesse, todas as mensagens são apagadas.

6. Caso o usuário possua interesse, ele irá acessar sua caixa de entrada para escolher os fornecedores. Dessa caixa, ele seleciona, aleatoriamente, quais *spams* irá ler, até um limite determinado.

7. Se não houver *spams* suficientes para completar o limite, o usuário poderá acessar sua caixa de *spam* para completar o limite. Ele somente irá acessá-la em experimentos específicos, como será detalhado na seção de resultados.

8. De posse dos *spams* que serão lidos, e de informações sobre o preço de um fornecedor externo (*não-spammer*) ele decidirá de quem comprar a mercadoria. Caso compre do fornecedor externo, nenhum *spammer* será pago.

9. Se a compra for realizada de um *spammer* o usuário realiza o pagamento e espera pela mercadoria, que poderá ser entregue corretamente ou não. Sempre que um produto é entregue incorretamente há uma probabilidade de o usuário perceber que foi enganado e de aprender e influenciar outros com seu aprendizado. O processo de compra será explicado em 3.1.2. A maneira como o usuário escolhe o fornecedor será explicada em 3.1.2.2. A forma como o usuário aprende e influencia outros será descrita em 3.1.2.3.

3.1.2 Processo de compra do usuário

3.1.2.1 Etapas do processo de compra

O artigo [60] descreve as etapas de um consumidor no seu processo de compra. A primeira etapa é a de detecção de necessidade, na qual o consumidor percebe suas necessidades. Em seguida, na próxima etapa, o potencial consumidor descobre qual o produto que atenderia a essa necessidade. Após isso, precisa pesquisar e decidir de qual fornecedor irá comprar o produto escolhido. Por fim, faz um balanço em relação a sua compra e a avalia como bem-sucedida ou não.

O modelo desta dissertação considera as duas primeiras etapas como uma probabilidade. Cada usuário possui uma probabilidade de diagnosticar uma dada necessidade e de apontar o produto vendido via *spam* como sendo adequado a atender seus objetivos. Importante ressaltar que há apenas um tipo de produto e que todos os *spammers* do modelo vendem exatamente o mesmo produto.

A etapa seguinte é a de escolha do fornecedor. Para tal, o usuário possui informação sobre o preço do produto em uma empresa confiável que não realiza *spam*. Assim,

ele pode julgar que é melhor comprar nesta empresa, não gerando qualquer receita a qualquer *spammer*.

No processo de decisão, o usuário calcula a utilidade (como descrita na seção 3.1.2.2), calculada para cada *spam* lido e para o fornecedor não-spammer. Para escolher os *spams* lidos, o usuário acessa a sua caixa de entrada e seleciona, aleatoriamente, um número definido de *spams*. Então, calcula a utilidade para cada um dos fornecedores presentes nessa amostra de *spams*. Caso não haja na caixa de entrada mensagens suficientes para atingir o número de *spams* lidos, o usuário pode recorrer à sua caixa de *spam*, ou simplesmente ignorar esse limite e ler todas as mensagens da caixa de entrada. A decisão de recorrer ou não à caixa de *spam* depende de uma variável no modelo (o impacto da leitura ou não da caixa de *spam* será avaliada na parte experimental).

Por fim, o usuário adquire o produto do fornecedor de maior utilidade e, após isso, realiza o balanço acerca da transação realizada. Caso compre do fornecedor não-spammer, nenhum processo é realizado, já que esse fornecedor é considerado sempre honesto e a transação sempre bem-sucedida. Caso compre de um *spammer*, a transação pode ser mal-sucedida, interferindo na confiança do usuário (vide seção 3.1.2.3). Vale antecipar a informação de que o usuário se vale de seus conhecimentos anteriores e da confiança no cálculo da utilidade, e, consequentemente, na escolha do fornecedor.

3.1.2.2 Cálculo da Utilidade

A função de utilidade foi baseada em [34]. A diferença é o acréscimo do fator $d \times \#spams$ que representa o quanto o usuário pondera negativamente um fornecedor à medida que recebe mais *spams* em sua caixa de entrada. *spams* na caixa de *spam* não são computados. Basicamente, esse fator considera a sobrecarga de informação do usuário, já que, quanto maior o número de *spams*, menor será a atenção que o usuário irá destinar aos mesmos, aumentando a probabilidade de descartá-los sem qualquer investigação. Além disso, usuários geralmente não gostam de receber *spam*, portanto, quanto maior o número de *spams* mais o usuário irá desejar punir todos os *spammers*, e, analogamente, recompensar um não-spammer (no caso deste modelo existe um não-spammer).

A função é dada a seguir, para cada fornecedor i (*spammers* cuja mensagem foi lida e para o *não-spammer*):

$$U_i = a \times Confianca_i + b \times \ln(Preco_i) + c \times Confianca_i \times \ln(Preco_i) + d \times \#spams$$

A fórmula é composta por parcelas que contribuem para a formação da utilidade.

$a \times Confianca_i$ representa a influência da confiança no dado fornecedor, cuja utilidade está sendo calculada. A confiança varia de 0 a 1, sendo que 1 significa completa confiança e 0 representa desconfiança total.

$b \times \ln(Preco_i)$ representa a importância do preço para o comprador. Em suma, representa o quanto diminuir o preço é atrativo ao usuário. O logaritmo é utilizado por questões de escala em relação ao preço do produto.

$c \times Confianca_i \times \ln(Preco_i)$ representa a sensibilidade do usuário em relação a variações no preço. A idéia é o quanto o usuário estaria disposto a pagar, a mais, por uma marca ou fornecedor conhecido e considerado confiável em detrimento de outro fornecedor, de menor preço e menos conhecido. Como pode ser percebido, essa parcela difere da ponderação simples da confiança, afinal, também considera o preço do produto, sinalizando a porcentagem a mais, no preço, que o usuário está disposto a pagar em um produto de fornecedor confiável e conhecido.

Por fim, como foi antecipado, $d \times \#spams$ representa o quanto a utilidade de um *spammer* decresce a medida que o usuário é sobrecarregado com *spams* em sua caixa de entrada.

Os parâmetros a,b,c,d são os pesos para cada parcela que forma a utilidade. Dado que o objetivo não é conseguir valores absolutos de utilidade, e sim relativos, já que o ponto é escolher o fornecedor de maior utilidade, os parâmetros a,b,c,d foram escolhidos arbitrariamente. Durante a fase experimental, a alteração destes parâmetros, ou seja, a maneira como cada parcela da fórmula é ponderada, será avaliada.

3.1.2.3 Cálculo da confiança

O cálculo de confiança precisa ser feito em dois momentos distintos. O primeiro ocorre quando o usuário já realizou pelo menos uma compra com o dado *spammer*. O outro momento ocorre quando o usuário precisa inferir o quanto confia no *spammer* no momento da primeira interação. No caso do vendedor não-spammer, a confiança é sempre 1 (o que significa confiança máxima).

Quando o usuário já realizou alguma transação com um dado *spammer*, a confiança do usuário nesse *spammer* é dada pela equação:

$$Confianca_i = Confianca_{i,anterior} + (1 - Confianca_{i,anterior}) \times \left[1 - \left(\frac{1}{\#transacoes_i + 1}\right)\right]$$

A fórmula é baseada na tese de doutorado de Klos [48]. A diferença é que a fórmula proposta na tese aborda um problema em que somente haverá uma segunda

transação caso a primeira tenha sido bem sucedida. Então, a fórmula proposta somente inicia o cálculo da confiança a partir da segunda interação. No nosso caso, logo após a primeira transação, a confiança já precisa ser calculada, afinal, esse valor irá ser decisivo para o usuário escolher se haverá outra transação com o mesmo anunciante.

Explicando a fórmula. $\#transacoes_i$ representa o número de transações, consecutivas, consideradas pelo usuário como bem-sucedidas com o vendedor i. Portanto, se o usuário considerar uma transação como mal-sucedida, o $\#transacoes_i$ retorna a zero. $Confianca_{i,anterior}$ representa a confiança do usuário no *spammer* antes de iniciar a sequência de transações bem-sucedidas. Assim, caso somente tenham ocorrido transações bem-sucedidas, esse valor é igual à confiança do usuário antes de qualquer contato. Caso a confiança tenha sido quebrada, esse valor corresponde à confiança após a traição.

Porém, a fórmula acima apenas atua enquanto o usuário realiza compras que considera bem-sucedidas com o *spammer*. Quando o usuário é traído, um fator é multiplicado à confiança atual e o valor de $\#transacoes_i$ retorna a zero. Assim, a nova confiança é uma fração da confiança original antes da traição. Vale mencionar que essa nova confiança, após a aplicação do fator, será a nova $Confianca_{i,anterior}$, usada nas equações caso o usuário decida se relacionar novamente com o *spammer*. Assim, em caso de quebra a confiança:

$$Confianca_i = Confianca_{i,anterior} = Confianca_i * fator$$

Caso não tenha havido qualquer contato, o usuário precisa decidir o seu grau de confiança no *spammer*. Para tal, o usuário utiliza o conhecimento adquirido por ele nas compras anteriores de outros *spammers*. Além disso, o usuário também se vale de dados sobre o número de compras bem-sucedidas e mal-sucedidas de todos os outros usuários em relação a esse *spammer*. A utilização de informações obtidas por outros usuários equivale a pesquisar sobre o dado *spammer*, já que o usuário não o conhece. A fórmula gerada para tal é a seguinte:

$$Confianca_i = 1 + g \times (Compras_{Sucesso} - Compras_{Fracasso}) + h \times (ComprasExt_{i,Sucesso} - ComprasExt_{i,Fracasso})$$

Na fórmula anterior, o valor da confiança inicial é igual a 1, ou seja, o usuário confia cegamente no fornecedor. g e h correspondem a pesos dados a importância de cada uma das fontes de informação. g é o peso para o comportamento observado pelo usuário e h é o peso para informações obtidas de opiniões dos outros usuários. O pri-

meiro parêntesis representa o ganho do usuário em relação às compras de *spammers*. Caso o ganho seja positivo, ou seja, ele acredita ter obtido maior número de transações bem-sucedidas que mal-sucedidas, a confiança aumentará. O segundo parêntesis corresponde à opinião que a comunidade tem do dado *spammer*. Basicamente, todas as compras feitas desse *spammer* são contabilizadas e o ganho é calculado. Importante frisar que é o usuário que comprou que avalia a compra como bem-sucedida ou não, portanto, a probabilidade de detectar uma fraude é crucial na confiabilidade da fórmula. A probabilidade de detectar uma fraude será explicada em 3.1.3.

A equação acima utiliza valores absolutos em relação ao número de compras bem-sucedidas ou não. Assim, o valor da equação pode ser maior que 1 ou menor que 0. Nos casos em que é maior que 1 o valor é transformado em 1; nos casos menor que 0 os valores são considerados iguais a 0.

3.1.3 Comportamento Fraudulento

O modelo prevê a possibilidade de avaliar a fraude por parte de *spammers*. O produto ofertado pode não ser entregue após o pagamento, ou, quem sabe, pode ser entregue um produto que não atenda às especificações originais. Por exemplo, em casos de medicamentos, ao invés da dosagem correta, pode ser entregue um placebo.

Para modelar esse comportamento, foi adicionada a probabilidade de um *spammer* não entregar o produto adequado. Quando o produto correto não é entregue, o *spammer*, no modelo, não possui qualquer gasto com o mesmo. Assim, nesse caso, toda a receita corresponde a lucro. Como a fraude é modelada como uma probabilidade, o mesmo *spammer* pode, aleatoriamente, em alguns casos entregar o produto correto, e, em outros, não fazê-lo. Essa decisão de modelar como uma probabilidade se baseia no artigo [43]. Nele, o autor fez quatro encomendas de Viagra ao conhecido *spam* da Canadian Pharmacy [6]. Para todas as encomendas, o autor recebeu itens correspondentes. E, em apenas uma das entregas, o produto não continha o princípio ativo correto do medicamento. Portanto, um mesmo *spammer*, ou uma mesma organização de *spammers*, pode, em alguns momentos entregar o produto correto, e, em outros casos, não.

Outra questão fundamental no modelo é o comprador perceber que foi enganado. Em alguns casos, essa percepção é óbvia. Por exemplo, em uma compra de celular, o comprador pode saber, facilmente, se o celular entregue possui algum defeito e se o produto combinado foi realmente entregue. Porém, a fraude em alguns produtos é mais difícil de ser detectada. Um caso clássico é a venda de medicamentos. O comprador dificilmente pedirá a um laboratório para analisar se um medicamento entregue possui

o princípio ativo correto. Além disso, o tempo entre a compra e a detecção da fraude pode ser grande, ainda mais em casos de medicamentos de uso contínuo. Inclusive, o médico responsável pode recomendar novo coquetel de medicamentos ao invés de perceber que os medicamentos receitados continuariam a surtir efeito se não fossem falsificados. Adicionalmente a isso, temos o clássico efeito placebo, em que um medicamento sem qualquer efeito químico pode contribuir para uma melhora no quadro clínico do paciente.

Para abordar essa questão, foi adicionada a probabilidade de um comprador perceber a fraude. Caso ele não perceba, considerará a transação como bem-sucedida. Assim, votará incorretamente e atribuirá uma confiança superior à justa a esse *spammer*. Portanto, acabará por interferir no processo de formação da confiança de outros usuários, bem como no seu próprio processo de formação de confiança. Veja seção 3.1.2.3 para um detalhamento do processo de formação de confiança.

3.1.4 Filtragem de *spam*

Inicialmente, importante frisar que o usuário já conhece os fornecedores *não-spammers* por algum outro meio, por exemplo, é a farmácia ao lado de sua casa. Portanto, no modelo e simulador não há qualquer *email* proveniente do fornecedor não *spammer*.

A filtragem de *spam* do modelo segue duas etapas básicas. Na primeira, uma fração das mensagens é rejeitada pelo MTA (Mail Transfer Agent). Todos os *spammers* têm a mesma fração de mensagens rejeitadas.

Em seguida, as mensagens aceitas serão entregues ao filtro propriamente dito. O filtro é modelado como uma probabilidade. Essa probabilidade corresponde à taxa de falsos-negativo, ou seja, corresponde à probabilidade de uma mensagem *spam* ser entregue à caixa de entrada. O modelo não inclui mensagens legítimas, portanto, a taxa de falsos-positivo não foi modelada. Assim, cada mensagem *spam* entregue ao MTA tem a dada probabilidade de ser entregue à caixa de entrada. As mensagens não entregues pelo filtro, serão direcionadas automaticamente para a caixa de *spam*.

3.2 Premissas e Simplificações

A principal premissa do modelo e simulador é a de que o mercado formado por *spammers* e compradores segue regras básicas e comuns a outros mercados. Por exemplo, o usuário segue as mesmas etapas do processo de compra, se vale do cálculo e otimização da utilidade na escolha do fornecedor e usa memória e confiança para ajudar na ponderação da utilidade.

As principais simplificações serão descritas abaixo:

- Simulação de apenas um produto, vendido por todos os *spammers*.
- Cada *spammer* tem a mesma probabilidade de ser fraudulento.
- Não é considerado qualquer impacto do *spam* filtrado na utilidade calculada pelo usuário.
- A taxa de *spams* enviados por usuário é fixa e constante.
- O modelo contempla iterações com períodos de um mês. Ou seja, a cada mês, o usuário tem uma probabilidade de ter interesse nesse mês. Não importa em qual dia ou hora. Se tiver, pode ler qualquer mensagem enviada no período, sem qualquer ordem específica. A caixa de *spam* permanece constante por um mês, não tendo informação sobre o dia de chegada do email.
- Parâmetros das equações não variam por agente, exceto quando se tratarem de memória e aprendizado.
- No sistema modelado não há diferenciação entre o *spammer* e quem disponibiliza o produto (*sponsor*). São encarados como um agente único.
- Cada endereço de *email* corresponde a um usuário.
- O modelo foi construído para modelar comportamentos. Como não há estudos completos acerca do comportamento de usuários e *spammers*, muitos valores para os fatores do modelo precisaram ser escolhidos de maneira arbitrária. Uma descrição dos parâmetros e valores encontra-se em 4.1.

3.3 Simulador

3.3.1 Ambiente de implementação utilizado

A simulação implementada foi baseada em agentes. Para a implementação da simulação, o ambiente Repast [20] foi utilizado. O ambiente fornece suporte à linguagem Java [18]. Ainda, provê integração simples com a IDE Eclipse [13], facilitando o processo de configuração e codificação do simulador. Também disponibiliza tutoriais para o aprendizado sobre o ambiente, recursos, bibliotecas, funções disponíveis e sobre simulação baseada em agentes.

Dois trabalhos [59; 36] avaliaram diversos ambientes para simulação baseadas em agentes. Ambos concluíram que o Repast é um ambiente adequado para simulação de modelos de ciências sociais. Como modelamos o comportamento do mercado e das relações de compra e venda, justamente um ambiente voltado para ciências sociais é o mais adequado. Ainda, avaliaram que o Repast possui um ótimo desempenho e ótimos materiais de estudo, que ajudam no processo de aprendizagem e de construção dos modelos.

O ambiente trabalha com o conceito de proto-agentes, comportamentos, contexto e projeções. Os proto-agentes são os agentes do sistema, ou seja, são todos os seres modelados. Obviamente, podem ser criadas diversas classes de agentes que interajam entre si, bem como diversas instâncias de cada uma das classes de agentes.

O sistema implementa um *clock*. Esse *clock* é disparado sempre que todas as ações designadas a ocorrer na execução anterior terminam. A cada vez que o *clock* é disparado temos uma execução. Assim, o *clock* permite executar ações síncronas, que ocorrem com uma certa periodicidade. Essa periodicidade pode ser diferente de um. Logo, é possível agendar ações que ocorrem a cada x execuções (ou x eventos do *clock*).

Os comportamentos são as possíveis ações dos agentes. Cada classe de agentes implementa seu próprio conjunto de comportamentos. Esses comportamentos podem ser acionados por alterações no meio, como por exemplo alterações em variáveis do ambiente, ações de outros agentes, ou temporalmente com a utilização do *clock*. Assim, por exemplo, a ação de um agente pode acionar um comportamento de outro agente, que representa a reação deste.

Contextos são os elementos do simulador que armazenam conjuntos de agentes. Esses agentes podem ser da mesma classe ou não. Podemos imaginar contextos como países, por exemplo. O Brasil, que seria um contexto, armazena os agentes brasileiros. O Irã, outro contexto, armazenaria outra classe de agentes, os iranianos. Obviamente, podem existir iranianos vivendo no Brasil e brasileiros vivendo no Irã, portanto, há agentes brasileiros no Irã e agentes iranianos no Brasil.

As projeções modelam o relacionamento entre os agentes e sua disposição física no universo. Por exemplo, é possivel modelar relações de amizade utilizando grafos ou relações de vizinhança utilizando uma distribuição em grade.

3.3.2 Descrição da simulação

O simulador implementado utilizou o ambiente Repast [20]. Na seção 3.3.1 há uma descrição desse ambiente. O simulador foi programado em turnos. A cada turno, que corresponde a 2 ciclos do *clock*, duas etapas são realizadas. Essas etapas são realizadas

seqüencialmente e em ciclos diferentes. A necessidade de ciclos diferentes se deve à necessidade de sincronização, afinal, a primeira etapa, obrigatoriamente, deve ocorrer antes da segunda.

A primeira etapa é o envio de *spams*. Essa etapa precede todas as outras. Nela, cada *spammer* envia um número definido de *spams* por usuário. O cálculo desse número será explicado na seção 4.1. Toda mensagem enviada é filtrada, para que se decida caso ela será entregue à caixa de entrada, à caixa de *spam*, ou será completamente eliminada. Para maiores detalhes sobre o processo de filtragem, vide 3.1.4. Vale lembrar que a decisão de quais mensagens serão entregues à caixa de *spam* e quais serão entregues à caixa de entrada é não-determinística e depende de uma probabilidade.

A segunda etapa é a de compra. Para maiores detalhes, vide seção 3.1.2. Os usuários que terão interesse no produto são definidos. A definição de quais usuários terão interesse é não-determinística e todos os usuários possuem a mesma probabilidade de terem interesse. Assim, não há classes de usuários com maior interesse que outros. Além disso, o interesse é um processo independente, logo, o interesse ou não no produto em um dado instante não interfere em outro instante. Em seguida, para os usuários interessados, *spams*, aleatoriamente escolhidos, são investigados e, finalmente, decide-se qual será o fornecedor. Após a venda, se o produto for comprado de um *spammer*, esse *spammer* receberá dinheiro para o seu fundo de recursos. Em seguida, o *spammer* decide se entregará o produto ou não. E, finalmente, a confiança é recalculada.

Importante mencionar que, por questões de simplificação na execução dos experimentos, bem como no cálculo de alguns parâmetros retirados de dados reais, um turno foi modelado como sendo um mês. Assim, as probabilidades são adaptadas a esse período. Por exemplo, um usuário tem a probabilidade definida, de, em algum momento do mês, ter interesse em comprar o produto.

Uma limitação dessa abordagem discreta é que somente *spams* de um mês são considerados. Um usuário, ao decidir comprar um produto, somente utilizará informações do mês corrente. Assim, não há distinção alguma se o usuário descobriu seu interesse no início, meio ou final do mês. Obviamente, essa simplificação permitiu que não fosse necessário armazenar e controlar *spams* com a granularidade diária, permitindo, ainda, que não houvesse qualquer preocupação com a duração do processo de busca de fornecedor nem em relação a quais e de quais dias os *spams* seriam lidos, permitindo uma amostragem estocástica.

3.3.3 Métricas reportadas

O simulador foi projetado para reportar três métricas principais. A primeira é o lucro dos *spammers*. Essa métrica é a soma dos recursos disponíveis a cada *spammer* no final da simulação. Vale mencionar que os recursos, de cada *spammer*, equivalem a soma da receita obtida com a venda de produtos menos os custos (com os *spams* enviados e com a aquisição dos produtos vendidos e entregues honestamente).

A segunda métrica é o número de *spams* gerados. O sistema é capaz de reportar o número total de *spams* gerados, bem como o número de *spams* entregues à caixa de entrada dos usuários. Novamente, esses valores são a soma para todos os *spammers* e todos os usuários. Como o total de *spams* enviados é fixo, essa métrica é reportada somente por sanidade. Também, pelo mesmo motivo de sanidade, o número total de *spams* entregues à caixa de entrada também foi reportado.

A terceira métrica reporta o saldo social do *spam*. A métrica foi inspirada no artigo [33]. Para maior detalhamento do artigo, veja a seção 2.1. A idéia por trás da métrica é reportar o impacto financeiro do *spam* na sociedade. Assim, precisamos contabilizar o quanto pessoas interessadas, que compram o produto, economizaram ao poderem pesquisar preços e comprar de *spammers*. Precisamos contabilizar o quanto pessoas perdem caso o produto seja entregue incorretamente. A perda equivale ao preço pago. Também, é importante incluir o custo que usuários têm com mensagens *spam*. Esse custo depende do destino do *spam*. Caso ele seja entregue à caixa de entrada, o custo é maior que para a caixa de *spam*. Também, é necessário adicionar os custos com infra-estrutura para tratar *spams*.

O cálculo dessa métrica é feito de acordo com a seguinte fórmula:

$$Saldo = \sum_{m=1}^{m=\#sucesso}(Preco_{externo} - PrecoPago_{sucesso,m}) - k \times \#spams_{inbox} - q \times \#spamsTotal - \sum_{n=1}^{n=\#fraude}(PrecoPago_{fraude,n})$$

Assim, contabiliza-se a economia, em relação ao preço de um anunciante externo não-spammer (definido na seção 3.1.2), feita pelo usuário em cada transação bem-sucedida. Nesse caso, o conceito de bem-sucedida engloba apenas produtos entregues corretamente. Ou seja, mesmo que o usuário não perceba a fraude, nessa métrica ela é contabilizada como fraude. Também contabiliza-se o gasto de tempo do usuário ao tratar as mensagens recebidas na caixa de entrada. Contabiliza o gasto com infra-estrutura e atenção do usuário ao lidar com mensagens spam filtradas. Por fim, considera a perda com fraudes que ocorreram.

Os preços serão melhor explicados na seção 4.1. Os parâmetros k e q foram

calculados através de dados de [17; 30; 51]. Em [17] diz-se que *spam* custará 130 bilhões em 2009 para as empresas. As estatísticas incluem apenas empresas, seus funcionários e respectivos endereços de email. Porém, consideram custos com infra-estrutura e falsos-positivo. Ademais, o trabalho [17] também estima os custos para mensagens entregues à caixa de entrada do usuário. Assim, foi possível obter o custo médio para o destinatário por *spam* enviado e o custo médio por *spam* enviado e recebido na caixa de entrada. Apesar de o estudo focar em empresas, nesta dissertação os valores serão usados para qualquer endereço de email, profissional ou não. Também não há estatísticas sobre o custo de mensagens filtradas, os estudos reportam apenas custo médio (filtradas e não-filtradas) e o custo de cada *spam* recebido na caixa de entrada. Então, para mensagens filtradas (q) foram utilizados os valores médios de custos.

Importante deixar claro que o saldo é uma métrica puramente financeira. Nela, contabiliza-se somente valores financeiros perdidos ou ganhos com a ação do *spam*. Questões e valores de caráter qualitativo ou filosófico não foram incluídos. Por exemplo, o efeito nocivo que a ausência de um medicamento com o princípio ativo correto pode fazer não é analisado. A razão para isso é que esse valor não pode ser calculado. E, como o sistema também pode entregar remédios corretos, ou entregar placebo a hipocondríacos, o ganho também pode ser maior que o valor economizado na compra.

Outra questão não abordada pela métrica são impostos, pirataria e outras possíveis perdas para a sociedade devido à venda de produtos via *spam*. Devido ao fato de *spam* ser uma atividade internacional cujos rastros são difíceis de seguir, a determinação de quais países receberiam impostos e os valores seria um problema enorme. A questão da pirataria também é uma questão delicada. Afinal, grande parte da venda a partir de *spams* são medicamentos, e vários países adotam leis de genéricos. Assim, para evitar esses problemas, foi-se estudado apenas o sistema *spammers* e usuários de email, excluíndo-se, da análise a existência de qualquer governo ou questões legais.

A métrica de saldo apresenta uma visão agregada do problema. Isso acaba por não trazer uma visão específica do ganho com *spam*, nem de sua perda, mas do valor resultante dessas duas métricas. Como o problema é uma questão social, analisar todos os participantes conjuntamente pode ser melhor que uma visão focando determinados jogadores. Por exemplo, se, em uma situação o ganho aumentar, e não analisarmos o problema de maneira agregada, podemos ser levados a concluir que esta situação é melhor. Porém, a situação pode também aumentar a perda, tornando-a pior ainda. Em algumas situações, em que for interessante para compreender o resultado, uma visão não-agregada (somente o ganho) será apresentada, em todas as outras, o foco será no saldo total.

Capítulo 4

Resultados Experimentais

4.1 Parâmetros do modelo

O simulador possui diversos parâmetros. A tabela abaixo descreve cada um desses parâmetros e apresenta quais os valores o parâmetro pode assumir durante os experimentos. Além disso, apresenta também os parâmetros que foram escolhidos como constantes durante todos os experimentos. O valor em negrito, na coluna valores apresenta o valor atribuído ao parâmetro nos experimentos nos quais o mesmo não é variado.

Parâmetro	**Descrição**	**Valores nos Experimentos**
Probabilidade de detecção da fraude	Se, em uma compra, o usuário for enganado, ele tem uma dada probabilidade de perceber que foi enganado. Esse valor varia de [0,1].	{0,01; 0,1; **0,5**; 0,7; 0,99}

Probabilidade de fraude	Probabilidade de os *spammers* agirem de forma fraudulenta. A cada venda, o *spammer* responsável pela venda possui essa probabilidade de não entregar a mercadoria corretamente. Esse valor varia de [0,1].	{0,01; 0,1; **0,5**; 0,7; 0,99}
Taxa de falsos-negativo	Fração de mensagens *spam*, que foram entregues pelo MTA, e que serão enviadas à caixa de entrada do usuário. As outras mensagens *spam* entregues pelo MTA e filtradas serão enviadas à caixa de *spam*. Esse valor varia de [0,1].	{0,0; 0,00001; **0,0001**; 0,001; 0,01} O valor em negrito está na mesma ordem de grandeza da média simples dos valores estimados em [47] para diversos provedores de email.

Probabilidade de interesse	Probabilidade, a cada mês, de cada usuário decidir comprar o produto. Essa probabilidade corresponde às etapas de compra que englobam a detecção da necessidade e escolha do produto que irá atender às necessidades. A subseção 3.1.2 explica melhor esse valor. Importante mencionar que o usuário poderá comprar de um vendedor não-spammer. Esse valor varia de [0,1].	{0,00005; 0,0001; **0,0005**; 0,001; 0,01} O valor em negrito resulta em uma taxa de compra uma ordem de grandeza maior que a obtida em [47]. O valor foi escolhido como uma ordem de grandeza maior, já que o artigo apresenta a taxa de compra, que, por natureza, é menor que a probabilidade de interesse. Afinal, a taxa de compra é obtida a partir da probabilidade de interesse e do processo de escolha do fornecedor (esse processo é modelado neste simulador).

Peso na confiança de informação externa	Esse valor corresponde a quanto a informação externa influencia na formação da confiança de um dado usuário em um *spammer* do qual nunca comprou anteriormente. As opiniões externas são exclusivamente sobre esse dado *spammer*. Ou seja, representa qual fração de opiniões externas chega ao usuário e qual o impacto dessa informação recebida. Essas opiniões externas correspondem a pedaços de informações que o usuário pode obter através de pesquisas na Internet. Por exemplo, ao buscar informações sobre o *spammer* X, ele pode ter acesso a blogs, foruns ou outras fontes de informação que irão permitir que o usuário forme uma opinião inicial sobre o *spammer*. Esse valor é um número real positivo. Vide a subseção 3.1.2.3.	{0,01; **0,1**; 1,0}

Peso na confiança de experiência própria	Esse valor corresponde a quanto a experiência com outros *spammers* influencia na confiança durante a realização de uma primeira compra com um *spammer* não-conhecido. É um valor real positivo. Vide a subseção 3.1.2.3.	{0,01; **0,1**; 1,0}
Peso, na utilidade, da confiança	Peso que a confiança possui no cálculo do valor da utilidade. Vide 3.1.2.2. É um valor real positivo.	{5,0; **10,0**; 15,0}
Peso, na utilidade, do preço	Peso que o preço possui no cálculo do valor da utilidade. Vide 3.1.2.2. É um valor real negativo.	Constante: -1,0
Peso, na utilidade, da sensibilidade em relação ao preço	Peso que a sensibilidade em relação ao preço possui no cálculo do valor da utilidade. Sensibilidade em relação ao preço corresponde ao valor a mais, que um usuário estaria disposto a pagar por um produto, dado que confia no fornecedor. Vide 3.1.2.2. É um valor real positivo.	Constante: 1,2
Peso, na utilidade, do número médio de *spams* na caixa de entrada.	Peso que o número médio de *spams* na caixa de entrada possui no cálculo do valor da utilidade. Vide 3.1.2.2. É um valor real negativo.	Constante: -0,5

Fator de decréscimo da confiança quando enganado	O quanto decai a confiança de um usuário em um dado *spammer* quando este percebe que foi enganado por esse *spammer*. Vide 3.1.2.3. É um valor entre [0,1]	Constante: 0,7

Número de *spams* por usuário	Número que cada *spammer* gera de mensagens para cada usuário. Esse valor é definido para cada *spammer*. É um número natural.	*spammer* modelando o CanadianPharmacy: 1205 spams por mês por destinatário. Outras farmácias: 240 spams por mês por destinatário. O valor foi calculado a partir do volume de *spam*, em um mês, proveniente do CanadianPharmacy [54] e do número de emails no mundo [30]. Para as outras farmácias o volume de *spam*, como não está disponível, foi estimado como 1 quinto do volume restante (não enviados pela Canadian Pharmacy) dos *spams* de conteúdo relacionado a farmácias. Nos experimentos, para simular os efeitos de um redução ou aumento no volume de *spam* gerado, também foram usadas frações do valor originalmente obtido: {20%; 60%; **100%**; 150%; 200%}

Preço externo do produto	Preço, cobrado pelo fornecedor externo, na venda do produto em questão. É um número real.	Os preços utilizados foram obtidos de 3 farmácias para o produto Viagra 30 comprimidos de 100mg. O primeiro preço ($448,97) foi o da Drugstore [11] uma farmácia americana que vende sob prescrição. O segundo ($266,72) foi o do CanadaDrugs [5], uma farmácia canadense legítima que vende com receita. O terceiro (**$301,00**) foi obtido da US Online Rx, que é reconhecida e vende sem receita.

Preço do produto ofertado via *spam*	Preço, cobrado pelo *spammer*, na venda do produto. Cada *spammer* possui seu próprio preço. É um número real.	Os valores foram obtidos, a partir de consulta aos sites modelados anunciados via *spam*. Para o CanadianPharmacy foi obtido o preço de $89,99. Para os outros, o preço foi de $105,00. O preço dos outros é idêntico, já que são sites da mesma organização, como afirma *spam*haus [25]. Nos experimentos, também foram usadas frações do valor originalmente obtido: {20%; 60%; **100%**; 150%; 200%}

Custo da mercadoria vendida	Custo da mercadoria. Basicamente, quanto o fornecedor gasta para produzir ou comprar a mercadoria que irá vender ou revender ao usuário. O valor é definido como sendo o mesmo para todos os *spammers*. É um número real positivo.	{5,0; 15,0; **25,0**; 37,5; 50,0}. Esses valores foram escolhidos arbitrariamente, afinal, não há informações disponíveis sobre o custo do medicamento para o *spammer*. O maior valor foi escolhido baseado no preço de venda (foi escolhido como aproximadamente 50
Número de mensagens lidas	Número máximo de mensagens *spam* que cada usuário está disposto a ler para a escolha do fornecedor. É um número natural.	Valor definido como uma constante igual a 10.

Utiliza a caixa de *spam*	Um valor booleano, que define se os usuários também acessarão a caixa de *spam* na busca de fornecedores, quando o limite de mensagens lidas não houver sido estourado. Corresponde a combinação de conhecimento do usuário de que há fornecedores na sua caixa de *spam* e de sua disposição em procurá-los nela.	Varia em verdadeiro e **falso**.

Número de *spammers* no sistema	Número de *spammers* existentes no sistema. É um valor natural.	Foram modeladas 5 *spammers*. O número foi escolhido como sendo as farmácias recebidas como *spam* no mês de junho/09. De acordo com a *spam*trackers [24] e *spam*haus [25], as organizações por trás desses *spams* são as duas maiores. O primeiro *spammer* modelado foi o CanadianPharmacy. Os próximos 4 *spammers* modelados fazem parte da Bulker.biz: my canadian pharmacy, International Legal Rx, Canadian Health&Care Mall, cvs pharmacy.
Número de usuários no sistema	Número de usuários existentes no sistema. É um valor natural.	O valor foi mantido constante, como sendo 100.000. A escolha desse valor se deve à capacidade das máquinas usadas nos experimentos.

Custo por *spam*	Custo pago pelos *spammers* a cada mensagem *spam* enviada. É um número real positivo.	Foi obtido de [27] para o valor cobrado por *botnets* para o envio de *spam* ($5E-6). O valor foi variado, como uma fração desse valor obtido. {20%; 60%; **100%**; 150%; 200%}
Taxa de falsos-negativo MTA	Fração de *spams* aceitos pelo MTA, e que serão, posteriormente, filtrados. É um número entre [0,1]	O valor foi obtido do artigo *spam*alytics [47]. Sendo igual a 23,8%

4.2 Procedimento experimental e Análises

Em todos os gráficos apresentados, os valores de todos os pontos representam a média de 40 execuções. Para cada gráfico foi plotado um correspondente, apresentando o intervalo com 90% de confiança. Assim, visualmente, é possível perceber quais pontos são realmente diferenciáveis com 90% de confiança. Esses gráficos foram suprimidos desta dissertação, afinal, as barras de erro acabam por poluir os gráficos, dificultando sua visualização. Seu papel fundamental foi ajudar a analisar quais conclusões, obtidas a partir do gráfico sem as barras de erro (apresentados nesta dissertação), poderiam ser tiradas e quais não se justificavam.

Todas as análises apresentadas neste capítulo embasaram-se nos gráficos de erros. Portanto, somente análises que puderam ser comprovadas também nesses gráficos de erros foram apresentadas. Porém, para não tornar a leitura do texto massante, somente quando se fizer relevante, haverá comentários sobre os intervalos de confiança.

4.3 Projeto variando um fator

4.3.1 Alternativas dos *spammers* para aumentar do lucro

O principal objetivo desta subseção é analisar as alternativas que spammers podem adotar para aumentar seus lucros. A partir dessa análise, é possível inferir qual a melhor abordagem para eles. Assim, técnicas anti-spam precisam focar em impedir que os *spammers* possam adotar essa abordagem, com o objetivo de reduzir seus lucros e viabilidade do envio de *spam*.

Os gráficos 4.1 avaliam o impacto de várias alternativas no lucro final dos *spammers*. Essas alternativas contemplam a análise de ações que, teoricamente, podem ser tomadas por *spammers* para aumentar seus lucros. A primeira alternativa é o decréscimo no custo pago por mercadoria entregue, o que equivale a procurar novos fornecedores de insumos (gráfico 4.1(a)). A segunda representa uma redução no custo para o envio de *spam*, o que representa a utilização de *botnets* e de novos sistemas de parceria para o barateamento dos custos (gráfico 4.1(b)). A terceira é a variação no preço cobrado por *spammers*, o que sinaliza uma variação no valor cobrado pelos *spammers* para a venda de suas mercadorias (gráfico 4.1(c)). Por fim, a variação no volume de mensagens enviadas por usuário foi estudada (gráfico 4.1(d)).

No gráfico 4.1(a) pode-se perceber que o aumento no custo dos insumos decresce o lucro dos *spammers* para uma probabilidade fixa de fraude. Claramente, a medida que a probabilidade de fraude aumenta, menor é o impacto desse preço. Isso é algo esperado, afinal, o aumento no custo da mercadoria aumenta o valor que os *spammers* precisam pagar por um produto entregue honestamente. Além disso, aumentar de \$5 para \$50 o custo para se obter a mercadoria, ou seja, aumentar em nove vezes, decresceu o lucro em 6,27%.

Ainda no gráfico 4.1(a) pode-se perceber um fenômeno estranho. Quando o custo da mercadoria foi 25, o lucro foi maior que para o custo 15. Isso se deve ao fato de que, neste gráfico, somente os pontos correspondentes ao primeiro e ao último custo puderam ser diferenciados com 90% de confiança. Porém, como esses dois pontos são diferenciáveis, o cálculo acima está correto.

No gráfico (b) o aumento no custo para se enviar *spam* tem um grande impacto no lucro obtido. Ao se aumentar o custo em nove vezes, o lucro cai 28,44 vezes.

Em (c), ao se aumentar o preço da mercadoria vendida, o lucro aumenta. Obviamente, o preço não pode ser maior que o praticado no mercado não-spammer. Ademais, se os valores forem próximos, a taxa de fraude por parte dos *spammers* precisa ser baixa, ou a probabilidade de detecção baixa, já que a utilidade, caso contrário, rapidamente

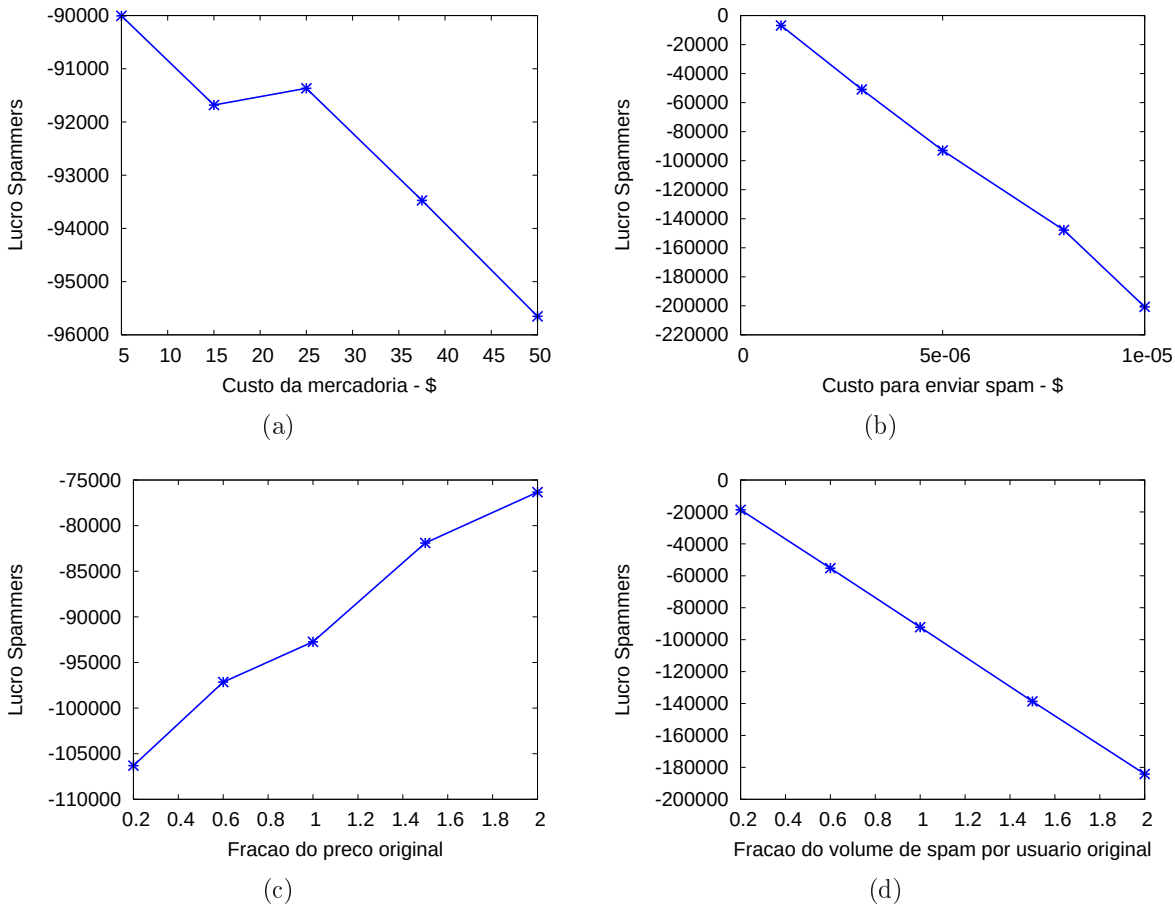


Figura 4.1. Gráficos demonstrando a influência do modo de envio de *spams* e do custo da mercadoria vendida no ganho dos *spammers*. Para efeitos de comparação, os valores no eixo x correspondem à mesma fração do valor final. Por exemplo, o primeiro ponto sempre representa 10% do valor do último ponto. (a) O gráfico no canto superior esquerdo ilustra o impacto do custo para a produção ou aquisição da mercadoria por parte do *spammer*. (b) O gráfico no canto superior direito mostra o impacto da variação do custo para o envio de um *spam*. (c) O gráfico no canto inferior esquerdo demonstra a conseqüência, no lucro, do preço cobrado por produto vendido. Os valores do eixo x representam a fração do preço cobrado em relação ao preço real. (d) O gráfico no canto inferior direito mostra o que ocorre ao se variar o número de mensagens enviadas por usuário. Os valores no eixo x representam o fator pelo qual o número de mensagens geradas de acordo com dados reais foi multiplicado em cada experimento.

seria menor que a utilidade externa. Portanto, o aumento no preço possui um limite. Nos casos estudados para a venda de Viagra, esse limite é alto, já que a diferença de preços, entre o praticado fora do mercado *spam* e dentro, é de $201,11. Esse valor é alto, comparado ao preço cobrado pela Canadian Pharmacy, que é $99,89. Pode-se perceber que ao aumentar o preço em nove vezes o lucro aumentou em 28,2%.

Em (d), os valores no eixo x representam a fração de mensagens enviadas. Assim, por exemplo, para o Canadian Pharmacy, que envia 1205 mensagens por usuário nos experimentos, o ponto 0,2 implica que o mesmo enviará apenas 20% das mensagens, ou seja, 241.

Ao se decrescer o volume de *spam* por usuário, os *spammers* conseguiram aumentar o lucro. Isso demonstra que o envio de múltiplas mensagens por usuário nem sempre é vantajoso para os *spammers*. O aumento em nove vezes do volume de *spam* por usuário gerou um prejuízo 8,87 vezes maior. Em parte, esse resultado se deve ao fato de que *spammers* tiveram prejuízo por mensagem enviada. Assim, o aumento no número de mensagens claramente aumenta o prejuízo. Portanto, não importa o volume de mensagens enviadas, o *spammer* sempre terá prejuízo. Afinal, o custo por *spam* enviado é maior que o lucro amortizado esperado por *spam*. Logo, na situação apresentada, o ideal é que os *spammers* não enviem *spam*.

Esses resultados, obviamente, podem ser influenciados pelo fato de os *spammers* apresentarem prejuízo. Por exemplo, o efeito do preço cobrado, ou o preço dos insumos pode ser maior no caso em que há uma alta procura.

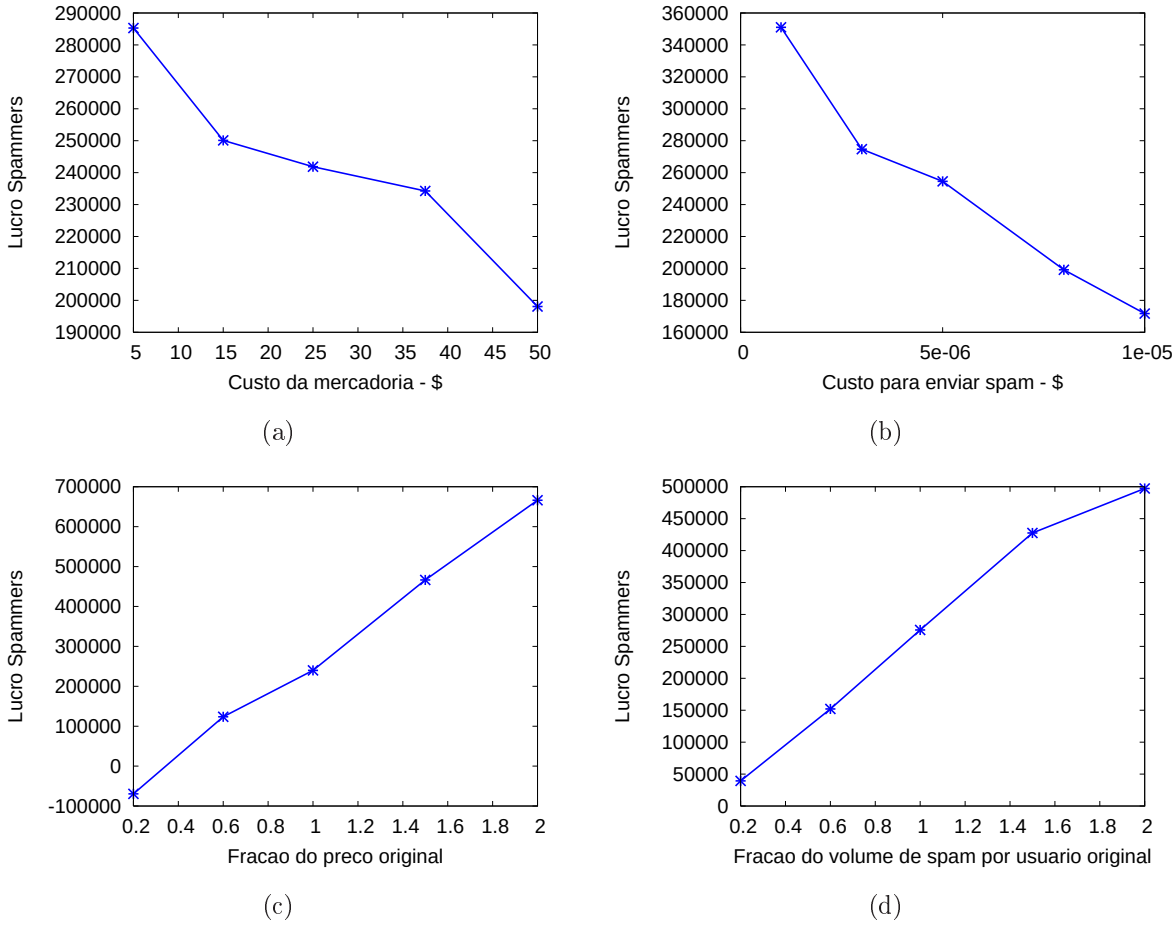


Figura 4.2. Gráficos demonstrando a influência do modo de envio de *spams* e do custo da mercadoria vendida no ganho dos *spammers*. Porém, agora a probabilidade de interesse é alta (0,01). Para efeitos de comparação, os valores no eixo x correspondem à mesma fração do valor final. Por exemplo, o primeiro ponto sempre representa 10% do valor do último ponto. (a) O gráfico no canto superior esquerdo ilustra o impacto do custo para a produção ou aquisição da mercadoria por parte do *spammer*. (b) O gráfico no canto superior direito mostra o impacto da variação do custo para o envio de um *spam*. (c) O gráfico no canto inferior esquerdo demonstra a conseqüência do preço cobrado por produto vendido no lucro. Os valores do eixo x representam a fração do preço cobrado em relação ao preço real. (d) O gráfico no canto inferior direito mostra o que ocorre ao se variar o número de mensagens enviadas por usuário. Os valores no eixo x representam o fator pelo qual o número de mensagens geradas de acordo com dados reais foi multiplicado em cada experimento.

Os gráficos 4.2 apresentam a mesma análise do conjunto de gráficos 4.1. Entretanto, para verificar a influência do lucro nos resultados, a probabilidade de interesse

foi escolhida como sendo um valor alto (0,01), para que haja lucro na grande maioria dos experimentos.

O gráfico 4.2(a) apresenta os resultados para a variação no custo da mercadoria (preço que o *spammer* precisa pagar para adquirí-la ou produzí-la). Ao se aumentar o preço em nove vezes, o lucro cai em 30%. Esse valor é superior ao obtido na situação em que poucas vendas eram realizadas. O que demonstra que o custo da mercadoria pode se tornar importante, especialmente em casos em que muitas vendas são realizadas.

Ainda em relação ao gráfico 4.2(a), somente o primeiro e o último ponto puderam ser diferenciados, entre si, com 90% de confiança.

O gráfico 4.2(b) analisa o impacto da variação no custo para o envio de *spam*. Ao se aumentar o custo em nove vezes, o lucro caiu em 51%. Esse valor é significativamente menor ao obtido no gráfico 4.1(b). Assim, caso o sistema dê lucro ao spammer e muitas vendas sejam realizadas, o impacto do preço por mensagem enviada se torna cada vez menor.

O gráfico 4.2(c) verifica o impacto do preço cobrado pelo *spammer* no lucro. Ao aumentar o preço, o lucro subiu em 10,58 vezes. Esse impacto é maior que na situação com menor interesse.

O gráfico 4.2(d) apresenta o impacto do número de mensagens geradas por usuário. Ao se aumentar o número de mensagens o lucro aumentou em 11,65 vezes. Esse resultado difere do obtido em uma situação de prejuízo. Na situação anterior, o aumento resultava em aumento no prejuízo. Em uma situação com lucro, aumentar o número de mensagens por usuário pode gerar maior lucro, obviamente mantendo o usuário não saturado. Afinal, caso o usuário esteja saturado, irá rejeitar todos os produtos vendidos por *spammers*.

Em suma, pode-se retirar algumas conclusões principais dos conjuntos de gráficos 4.1e 4.2:

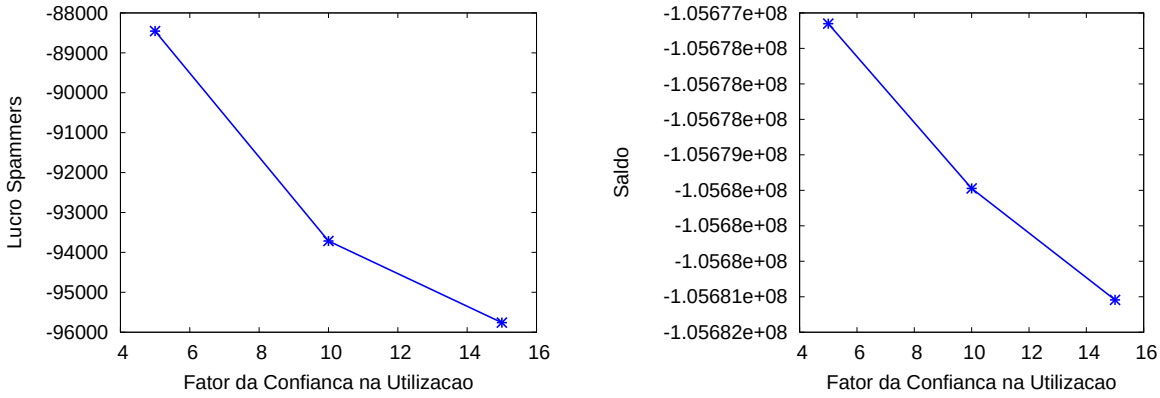
1. A primeira conclusão a ser percebida é que *spam* pode não ser tão lucrativo quanto esperado. Em todos os gráficos do conjunto 4.1, os *spammers* tiveram prejuízo. Esse Nos experimentos seguintes serão analisados vários fatores mantidos constantes, como a taxa de filtragem e a probabilidade de interesse, dessa maneira, será possível perceber o quanto esses parâmetros interferiram nos prejuízos.

2. A segunda conclusão é a de que o custo de envio é importante no lucro. Isso corrobora com as abordagens econômicas, que investem em aumentar esse custo para combater *spam*. Porém, caso a probabilidade de interesse e compra sejam altas, a estratégia econômica de aumentar o custo de envio pode não ser muito útil.

Afinal, nessa situação, o melhor é investir em diminuir o preço das mercadorias vendidas por *spammers*, ou, de aumentar seus custos com insumos.

3. Também é possível perceber que nem sempre é vantajoso enviar múltiplas mensagens por usuário, já que esse fator foi o de segundo maior impacto no conjunto de gráficos 4.1. Isso pode ser justificado pelo fato de que múltiplas mensagens significam, em um dado período, excesso de informação, e não, necessariamente, múltiplas compras pelo mesmo destinatário. Entretanto, essa análise vale apenas para a situação com poucas vendas. Em situações em que o interesse é grande, aumentar, até certo ponto, o número de mensagens por usuário é sim vantajoso.

4.3.2 Variação no peso da confiança na utilidade e variação dos pesos na formação de opinião sobre um *spammer* não-conhecido



(a) (b)

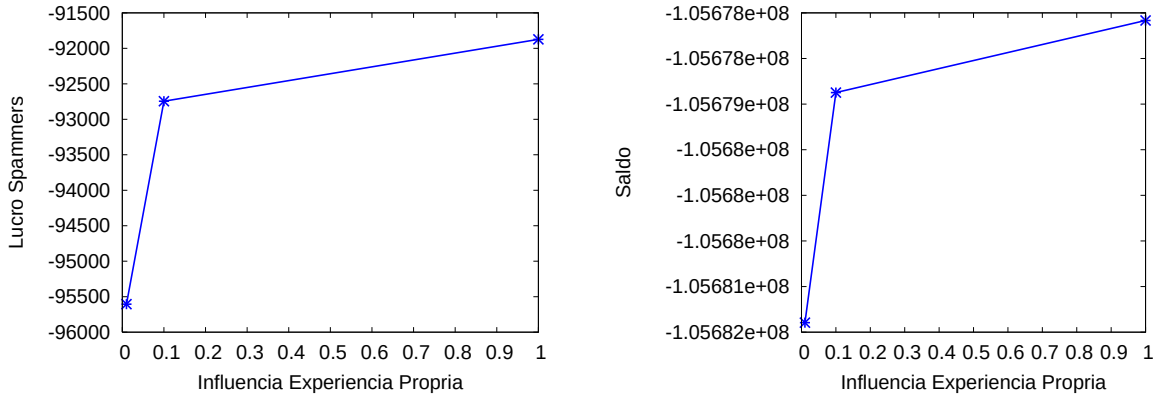
Figura 4.3. Variação no peso da confiança na utilidade. O primeiro gráfico (a) representa a influência no lucro e o segundo (b) a influência no saldo.

Como pode ser percebido nos gráficos 4.3(a) (b), o aumento da influência da confiança na utilidade decresce tanto o lucro de *spammers* quanto o saldo do sistema. Como os *spammers* podem agir de maneira maliciosa, aumentar a importância da confiança decresce o número de compras de *spammers*, diminuindo os lucros.

No gráfico 4.3(b), o saldo também diminuiu. Vide seção 3.3.3 para uma explicação sobre a métrica saldo. Isso devido ao fato de que a diferença de preços entre o vendedor externo e os *spammers* é alta. Portanto, no sistema atual é melhor se os usuários forem enganados e comprarem, que se não comprarem. Importante mencionar que o sistema

apresentava uma probabilidade de fraude de 50%. Portanto, fazendo-se as contas com os dados, um roubo equivale a uma perda de $99,89 enquanto um transação bem-sucedida equivale a um ganho de $201,11, portanto, com 50% de fraude, o sistema tem maior ganho caso a confiança atue pouco e sempre sejam realizadas compras a partir de *spammers*.

Obviamente, como o lucro e o saldo, ambos, diminuíram, e o saldo é negativo, o sistema precisa ser encarado a longo prazo. Então, mesmo que o saldo diminua com o aumento da confiança, esse pode ser um pequeno preço a ser pago para se desestimular o envio de *spams* (diminuir o lucro), e, em um futuro, aumentar o saldo, tornando-o positivo.

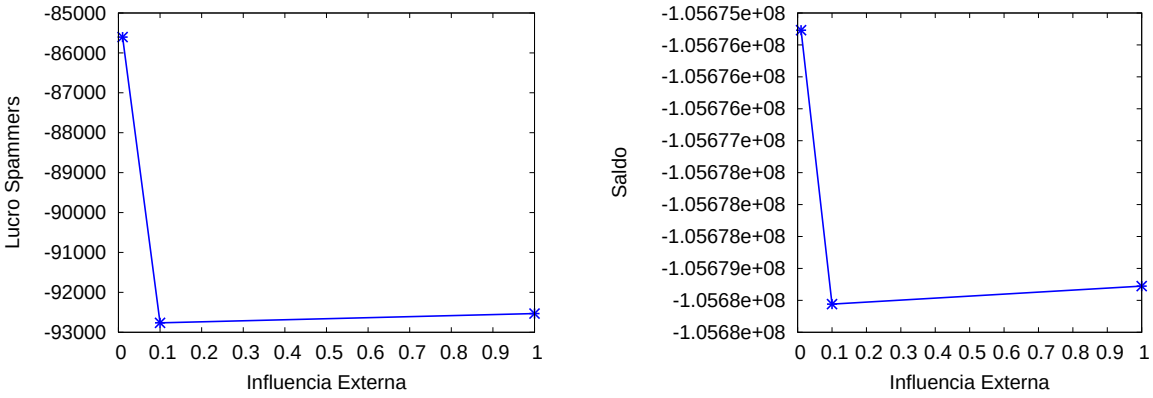


(a) (b)

Figura 4.4. Análise do impacto do peso da experiência pessoal na determinação da confiança em um *spammer* não conhecido. Avalia esse impacto para o lucro (a) e o saldo do grupo (b).

Os resultados na figura 4.4 mostram que confiar cada vez mais na experiência própria (vide 4.1), no momento de determinar a confiança para um *spammer* desconhecido, não somente aumenta o lucro como pode ser benéfico aos usuários. Isso se deve aos preços externo e *spam*, bem como à probabilidade de fraude, tornando vantajoso aumentar o número de compras. E esse aumento no número de compras é o resultado de se utilizar mais a própria experiência e ignorar o conhecimento externo.

Importante mencionar que, nos gráficos 4.4, somente o primeiro e o último ponto puderam ser diferenciados entre si, com 90% de confiança. Ainda, se verificarmos os intervalos de confiança, é possível perceber que a diferença entre esses pontos pode ser muito pequena.



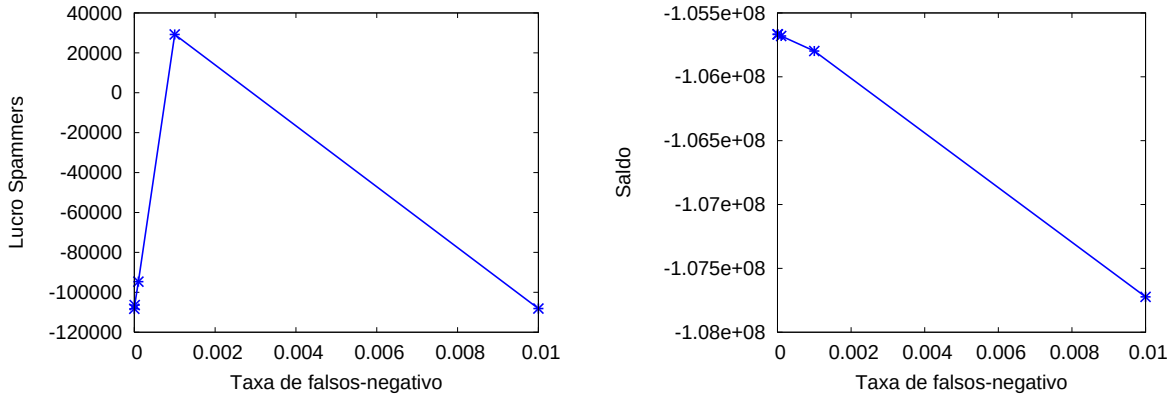
(a) (b)

Figura 4.5. Análise do impacto do peso da experiência externa na determinação da confiança em um *spammer* não conhecido. Avalia esse impacto para o lucro (a) e o saldo do grupo (b).

Como mostra a figura 4.5, aumentar a confiança na experiência externa aumenta a quantidade de informação obtida, diminuindo o total de compras realizadas. Mesmo havendo uma probabilidade de 50% de uma fraude não ser detectada, a quantidade de informação externa supera o uso apenas da experiência pessoal. Isso diminui a taxa de primeiras compras, também diminuindo o ganho do sistema.

Os últimos e penúltimos valores não puderam ser diferenciados com confiança de 90%. Isso sugere uma tendência à estabilização do gráfico, ou seja, a partir desse ponto, não há impacto significativo em se continuar a aumentar a influência externa. Esse efeito será visto novamente nos gráficos 4.12 e 4.13.

4.3.3 Qualidade do filtro



(a) (b)

Figura 4.6. Variação da taxa de falsos-negativo do filtro e seu impacto no sistema. A medida que essa taxa aumenta, pior é a qualidade do filtro. O primeiro gráfico representa a influência no lucro (a) e o segundo a influência no saldo (b).

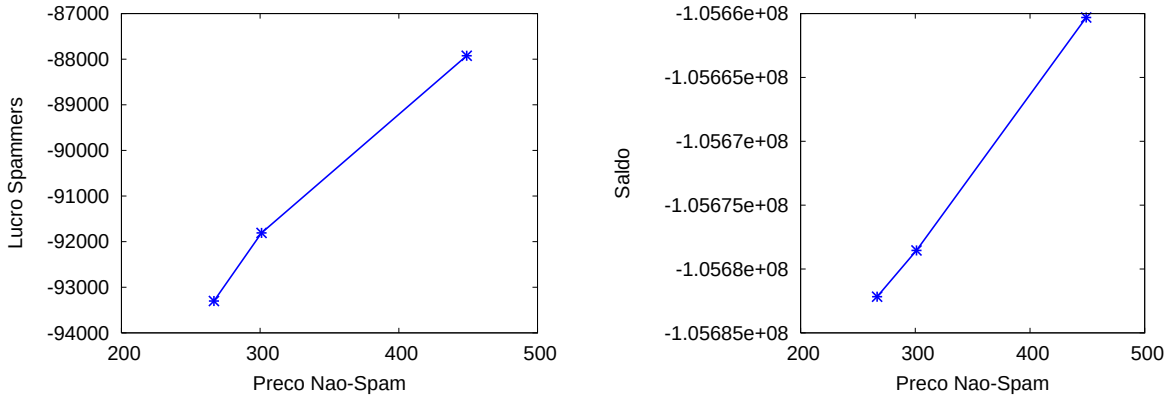
O gráfico 4.6(a) demonstra que a taxa de falsos-negativo interfere no sistema. Vale lembrar que a taxa de falsos-negativo representa o percentual de mensagens *spams* classificadas como legítimas e entregues à caixa de entrada. Portanto, aumentar essa taxa representa que o filtro se torna cada vez pior, deixando um maior número de *spams* serem entregues.

Diferentemente do esperado, a piora do filtro nem sempre traz aumento no lucro de *spammers*, bem como a melhora no filtro nem sempre reduz o lucro. A idéia por trás disso é que para uma taxa baixa do filtro, não chegam mensagens suficientes e os usuários não podem pesquisar fornecedores *spam* (vale mencionar que neste experimento a caixa de *spam* não é utilizada). Daí, quando o filtro piora, até certo ponto, o lucro aumenta devido ao aumento no número de mensagens entregues. A partir desse ponto, o lucro volta a cair, devido ao excesso de *spams* e à rejeição aos produtos que oferecem. Vide seção 3.1.2.2, que apresenta o efeito da rejeição devido ao excesso de *spam* na utilidade. Dessa maneira, melhorar o filtro pode ser prejudicial ao sistema como um todo. Ainda, o saldo, no gráfico 4.6(b), cai a medida que o filtro piora. Isso se deve ao fato de que um *spam* entregue à caixa de entrada é mais prejudicial ao sistema que um filtrado.

Nas análises anteriores, considerou-se o número de mensagens como fixo e a qualidade do filtro como variável, para demonstrar o impacto apenas da melhora do filtro em um cenário estável. Porém, *spammers* podem regular o número de *spams* por usuário de acordo com a taxa de filtragem. Assim, a qualidade alta do filtro poderá obrigar

o aumento no número de mensagens enviadas para compensar. Além disso, a piora no filtro pode reduzir a taxa de *spams* gerados. A idéia para esse último caso é simples. A entrega de uma quantidade excessiva de *spams* é negativa para *spammers*. Afinal, usuários tendem a rejeitar produtos vendidos via *spam* caso o volume de *spam* recebido seja alto. Portanto, como o filtro entrega uma quantidade excessiva de *spams*, a única resposta possível, por parte dos *spammers*, para aumentar o lucro, é reduzir o número de *spams* gerados. Assim, o volume de *spam* entregue seria reduzido e a rejeição a produtos vendidos por *spammers* diminuiria.

4.3.4 Influência do preço cobrado por não-spammers

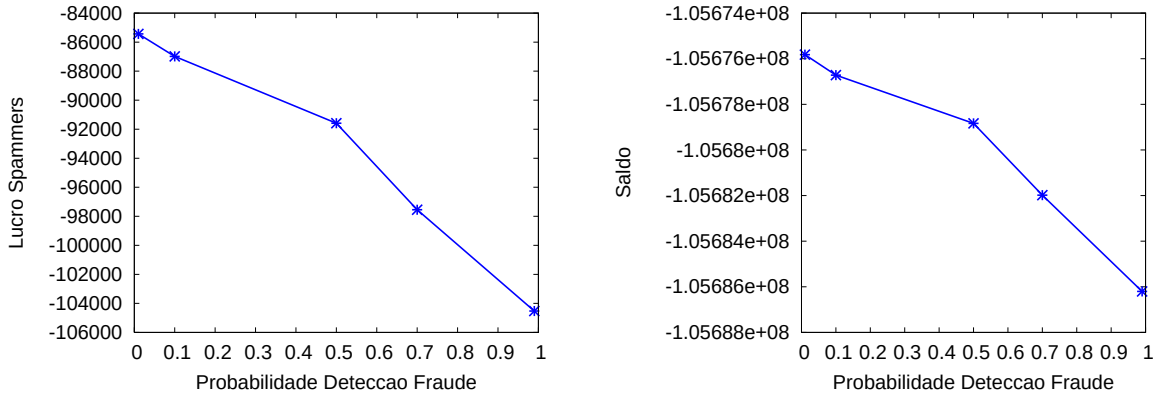


(a) (b)

Figura 4.7. Impacto do preço externo (cobrado pelo fornecedor não-spammer) no sistema. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Nos gráficos da figura 4.7, o aumento no preço cobrado externamente aumenta tanto o lucro dos *spammers* como o saldo geral do sistema. Esse resultado é esperado. O mais interessante é ponderar que o preço externo depende do propósito da compra e do conhecimento do usuário. Se todos os usuários possuírem prescrição e puderem comprar em farmácias americanas, o saldo e o lucro são os maiores, já que essa situação corresponde ao preço mais alto do experimento. Porém, se não possuírem prescrição, o lucro e o saldo são médios, pois essa situação corresponde ao segundo maior preço. Entretanto, se os usuários possuírem prescrição e forem bem informados, o lucro e o saldo diminuem, afinal, o usuário pode comprar de uma fonte mais barata.

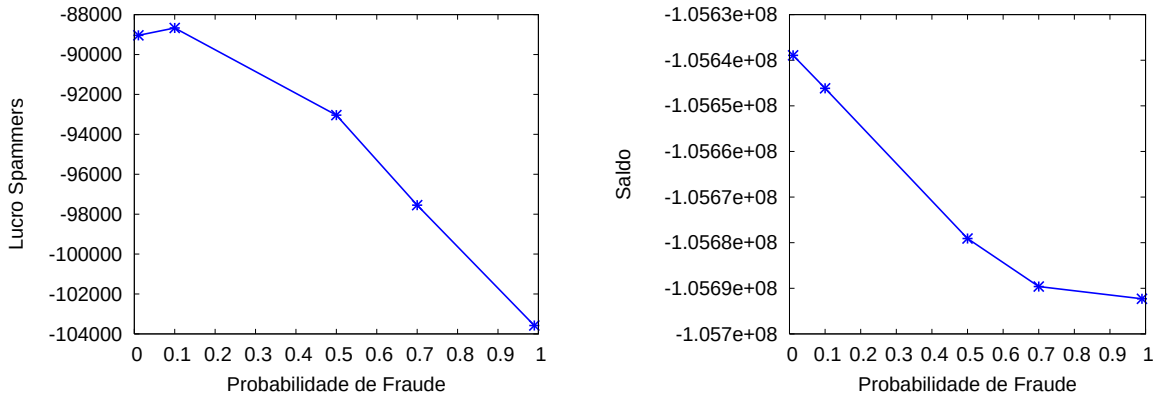
4.3.5 Efeito da fraude e capacidade de detecção de fraudes



(a) (b)

Figura 4.8. Impacto da capacidade de detectar uma fraude por parte dos compradores no sistema. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Os gráficos 4.8 mostram que o aumento na probabilidade de detecção decresce o lucro dos *spammers*. Dependendo da probabilidade de fraude, da diferença de preços entre vendedor externo e *spammer*, aumentar a probabilidade de detecção pode ser benéfico ou não ao sistema. Novamente, no caso não é, afinal, a diferença de preços compensa a taxa de fraudes. Vale ressaltar que a análise e o cálculo do saldo são puramente financeiros (ver seção 3.3.3).



(a) (b)

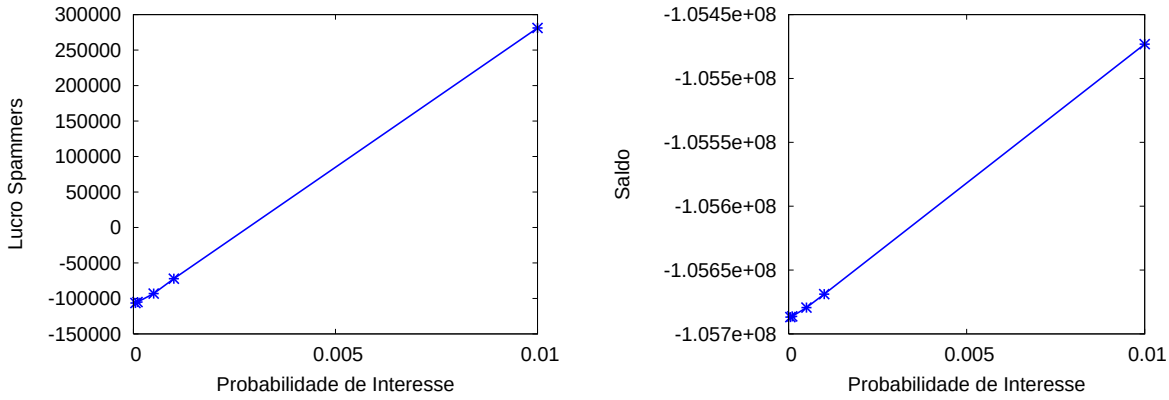
Figura 4.9. Impacto da probabilidade de fraude no sistema. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Como visto em 4.9(a), o aumento na taxa de fraudes ocasiona uma redução nos lucros. Isso se deve ao fato de que, mesmo que uma fraude represente maior receita, a existência de confiança e de mecanismos para calculá-la antes de realizar uma primeira compra protegem o sistema.

Importante mencionar que o segundo ponto apresenta um lucro maior que o primeiro. Isso sugere a existência de um leve aumento no lucro, para uma situação em que o aumento nas fraudes é pequeno o suficiente para aumentar os lucros e passar despercebido pelos usuários. Porém, esse comportamento não pôde ser verificado com 90% de confiança. Ainda, ele não mais se revelou nos gráficos 4.12 e 4.14.

Como percebido e esperado, o saldo, 4.9(a), cai a medida que a probabilidade de fraude aumenta, já que o número de pessoas enganadas aumenta. Porém, o sistema se torna cada vez mais resistente, afinal, o saldo varia pouco de 0,7 para 0,99, demonstrando que, em ambos os casos, o número de vendas é muito próximo, ou seja, os usuários já não estão comprando produtos de *spammers* como faziam com probabilidades de fraude menores. Além disso, a partir de 0,7 um novo fenômeno começa a ocorrer. A redução no número de vendas, devido à confiança, passa a ser benéfica ao sistema. Afinal, para 0,7, a expectativa de ganho por compra, considerando-se os preços externos, os preços ofertados por *spams* e a probabilidade de fraude, se torna negativa.

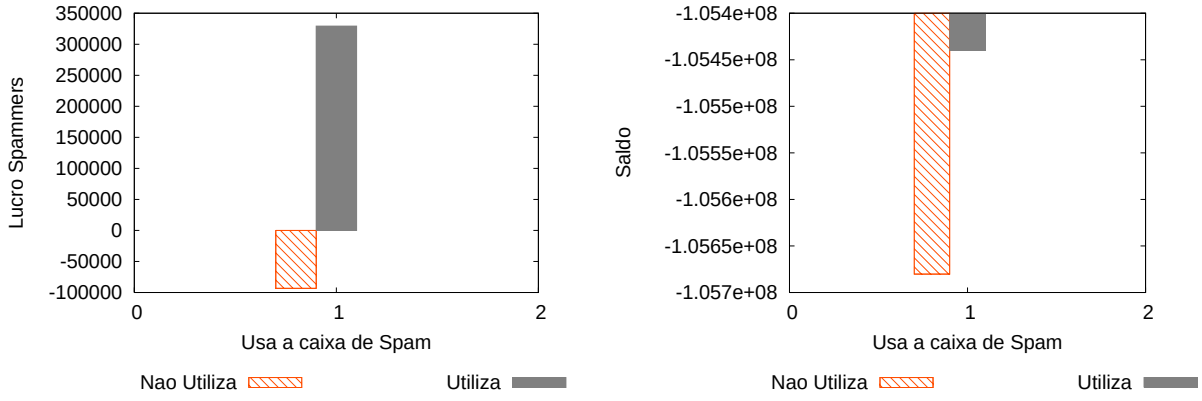
4.3.6 Interesse do usuário e utilização da caixa de *spam*



(a) (b)

Figura 4.10. Impacto da probabilidade de interesse no sistema. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

De acordo com a figura 4.10, o aumento na probabilidade de interesse tem um impacto positivo tanto no lucro quanto no saldo. Vale mencionar que, inclusive, o lucro foi positivo. Porém, esse efeito pode ser anulado com o aumento na probabilidade de fraude, afinal, a confiança seria comprometida, reduzindo o número de vendas realizadas por *spammers*.



(a) (b)

Figura 4.11. Impacto da utilização da caixa de *spam* na busca por fornecedores. O valor 1 no eixo x corresponde a não utilização da caixa de *spam*. O valor 2 corresponde ao seu uso. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Nos gráficos 4.11, o valor 1 corresponde à não utilização da caixa de *spam*, enquanto o valor 2 corresponde à utilizá-la. Pode-se perceber que utilizar a caixa de *spam* aumenta tanto o lucro quanto o saldo, especialmente nos experimentos realizados, nos quais a taxa de mensagens entregues na caixa de entrada é baixa. Inclusive, isso é um fato nos dias de hoje. A taxa de *spam* de produtos farmacêuticos entregues é muito pequena. Portanto, é necessário avaliar se os usuários utilizam a caixa de *spam* em suas busca e, caso utilizem, uma estratégia de combate a *spam* deve ser idealizada com o objetivo de reduzir o acesso à caixa de *spam*.

4.3.7 Sumário dos resultados

- O custo por *spam* enviado é importante para *spammers*. Assim, sua diminuição obteve um aumento significativo no lucro. Dessa maneira, as estratégias econômicas que focam em aumentar esse custo seguem um caminho adequado. Por outro lado, *spammers* têm adotado estratégias para diminuir esses custos, tais como a utilização de *botnets* ou terceirização dos serviços de envio de mensagens. O sistema de terceirização adotado segue um esquema de parceria, no qual quem envia

as mensagens recebe uma porcentagem dos lucros pelas vendas obtidas através das mensagens que enviou. Logo, não haveria custo de envio para mensagens que não gerarem retorno. Porém, em um cenário no qual haja probabilidades de compra e interesse altas, o custo por mensagem reduz seu impacto. Nesses cenários, os custos com os insumos ou o valor final cobrado exercem maior influência. Portanto, estratégias econômicas podem não ser tão efetivas em cenários com alta probabilidade de interesse e procura.

- Nem sempre é vantajoso financeiramente para a sociedade depender da confiança. Apesar de ela proteger o usuário de fraudes, em alguns casos, dependendo do preço dos produtos vendidos via *spam* e via não-spam, as fraudes são compensadas. Importante mencionar que isso segue uma visão apenas financeira, e a perda com fraudes equivale ao custo do produto.

- Pode haver casos em que melhorar o filtro gera aumento no lucro de spammers. Piorar o filtro pode reduzir o lucro de *spammers*.

- O efeito (diminuição no lucro) de se piorar o filtro pode ser compensado, pelos *spammers*, através da redução no volume de *spams* gerados. Assim, podem haver casos em que o sistema como um todo se beneficiaria mais da piora do filtro que de sua melhora, considerando essa diminuição no volume de mensagens.

- Em vários experimentos o saldo e o lucro cresceram nas mesmas circunstâncias. Isso delimita decisões mais complicadas. Afinal, é necessário escolher melhorar o saldo, ou piorá-lo com o intuito de diminuir o lucro dos *spammers* apostando em um ganho a longo prazo devido à falência dos mesmos.

- Na maior parte dos experimentos o lucro foi negativo. Isso mostra que o mercado de *spam* pode não ser tão lucrativo como diz o senso comum. Inclusive, o artigo [47] chega a essa mesma conclusão. Para que haja lucro, é necessário altas taxas de interesse, uma combinação ideal entre mensagens enviadas e taxas de filtragem, além, claro, de baixos custos para o envio de mensagens. Além disso, a utilização da caixa de *spam* na busca de fornecedores apresentou um aumento no lucro dos *spammers*. Assim, é importante verificar se os usuários recorrem à caixa de *spam* para a busca de fornecedores *spammers*, e, se o fizerem, são necessárias estratégias para reduzir esse acesso.

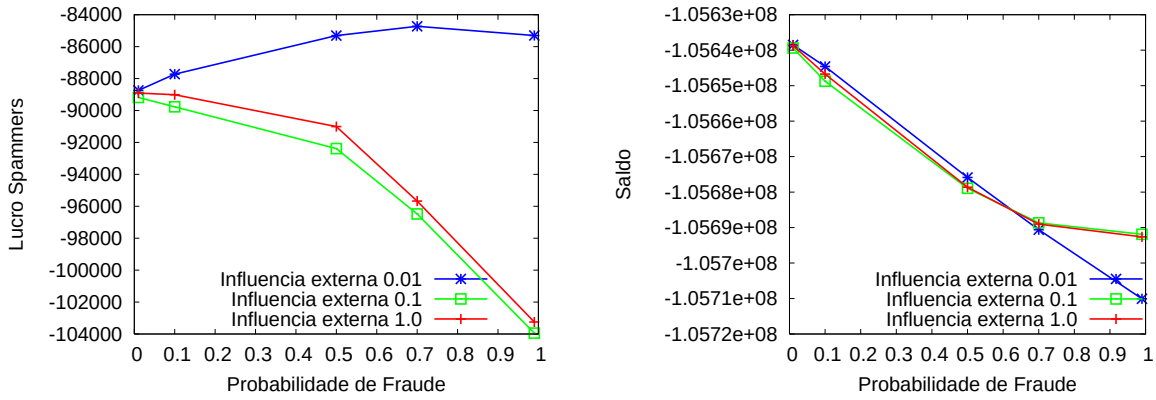
- A existência de fraudes é esperada, devido à assimetria de informação e a tragédia dos comuns. Porém, o aumento na probabilidade de fraude acaba por diminuir o lucro de *spammers*, devido ao efeito da confiança. Entretanto, ao se diminuir a

capacidade dos usuários de detectarem uma fraude, esse efeito é perdido, e menor capacidade significa maior lucro.

4.4 Verificação de hipóteses

Todos os pontos em todos os gráficos correspondem a uma média de 40 execuções. Todas as análises em que um aumento ou decréscimo é mencionado se basearam em comportamentos que puderam ser observados com 90% de confiança.

4.4.1 Hipótese 1: À medida que a probabilidade de fraude aumenta, o aumento na influência externa melhora o sistema.



(a) (b)

Figura 4.12. Os gráficos apresentam o efeito do aumento da influência externa no impacto da probabilidade de fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

No gráfico em 4.12(a), para o sistema com baixa influência externa, a medida que a probabilidade de fraude aumenta, o lucro dos spammers aumenta, justamente por não haver comunicação que previna a existência de primeiras compras. Esse aumento ocorre até o ponto 0,7. No ponto 0,99 há uma queda no lucro, demonstrando que, a partir daí, há informação suficiente para que a confiança atue no sistema e evite primeiras compras.

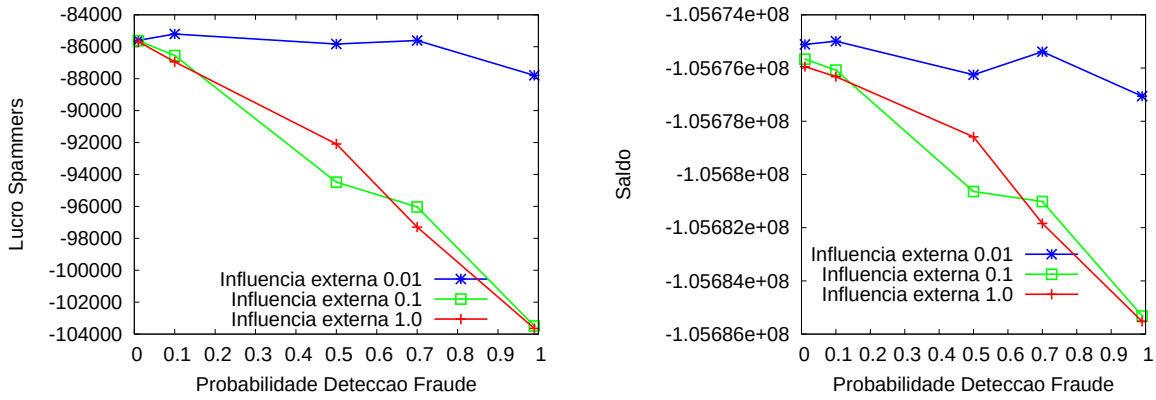
Porém, a medida que a influência externa aumenta, o lucro cai a medida que a probabilidade de fraude aumenta. Isso demonstra que, em um sistema com uma

probabilidade razoável de detecção de fraudes (nesse experimento 50%), confiar na opinião externa pode reduzir o lucro de *spammers*. Também é possível observar que as curvas para influência externa 0,1 e 1,0 apresentaram comportamentos próximos, não sendo possível diferenciá-las com 90% de confiança. Isso sugere que, a partir de certo ponto, aumentar a influência externa e fontes de informação não surte efeito, ou tem o efeito bem reduzido.

Em relação ao saldo (4.12(b)), todas as curvas apresentam um comportamento semelhante. Porém, para probabilidades de fraudes altas, o sistema acaba se beneficiando do aumento na influência externa. Isso demonstra que o ganho do aumento da influência externa somente é sentido, no saldo, para taxas altas de fraudes. Isso se deve ao fato de que, para baixas taxas de fraude, as fraudes interferem moderadamente no saldo, portanto, evitá-las tem menor impacto que evitar o próprio *spam*. Entretanto, ao se aumentar a probabilidade de fraude, sua influência no saldo aumenta, também aumentando o efeito benéfico da influência externa.

Em suma, a hipótese é parcialmente válida. O lucro diminui com o aumento da influência da experiência externa, porém, somente até certo ponto, a partir do qual aumentar essa influência não resulta em ganho. Em relação ao saldo, a hipótese somente é válida para altas taxas de fraudes, afinal, para baixas taxas, o tratamento de *spams* domina completamente o valor do saldo. Além disso, novamente, o ganho ocorre somente até certo ponto, após isso, aumentar a influência externa não melhora o saldo.

4.4.2 Hipótese 2: À medida que a habilidade em detectar fraude diminui, o aumento na influência externa se torna menos eficaz em melhorar o sistema.



(a) (b)

Figura 4.13. Os gráficos apresentam o efeito do aumento da influência externa no impacto da probabilidade de detecção fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Pode-se perceber, pelo gráfico 4.13(a), que o impacto no lucro da probabilidade de detecção de fraudes depende da influência da opinião externa. Quando a influência é pequena, o aumento na probabilidade de detecção tem um impacto muito menor no lucro que para taxas de influência maiores. Também, para influências mais altas, os valores obtidos para o lucro são próximos, e não puderam ser diferenciados com 90% de confiança. Assim, a influência externa tem uma influência decisiva no impacto da probabilidade de detecção no lucro, porém, a medida que a influência aumenta, esse impacto tende a se manter constante.

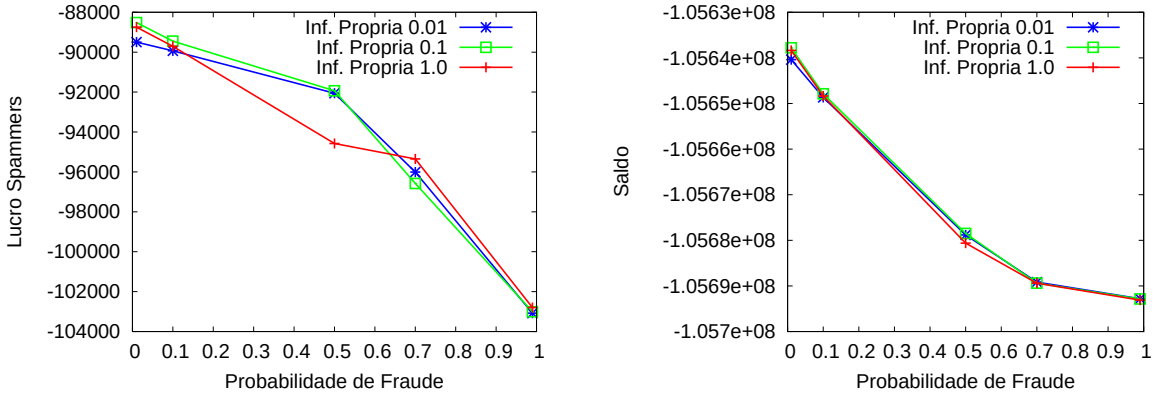
Além disso, o gráfico 4.13(a) apresenta um comportamento inusitado. Para a curva 0,1, os valores para 0,5 e 0,7 não puderam ser diferenciados com 90% de confiança. Assim, não podemos concluir que 0,7 apresenta lucro maior que 0,5.

O saldo, de acordo com o gráfico em 4.13(b), decresce juntamente com o lucro dos *spammers*. Novamente, isso se deve a relação de preços e de probabilidade de fraude, assim, é melhor para o sistema que pessoas sejam enganadas. Pois, caso contrário, o sistema tende mais rapidamente a uma situação em que não há compras de *spammers*.

Em suma, a hipótese é verdadeira. A medida que a probabilidade tendeu a zero, menor foi a diferença de lucros e saldos entre os valores das curvas de diferentes

influências externas. Esse efeito é mais perceptível ao se observar que a curva 0,1 se aproxima das demais à medida que a probabilidade tendeu a zero.

4.4.3 Hipótese 3: À medida que a probabilidade de fraude aumenta, o aumento na influência de experiências pessoais melhora o sistema.



(a) (b)

Figura 4.14. Os gráficos apresentam o efeito do aumento da influência pessoal no impacto da probabilidade de fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

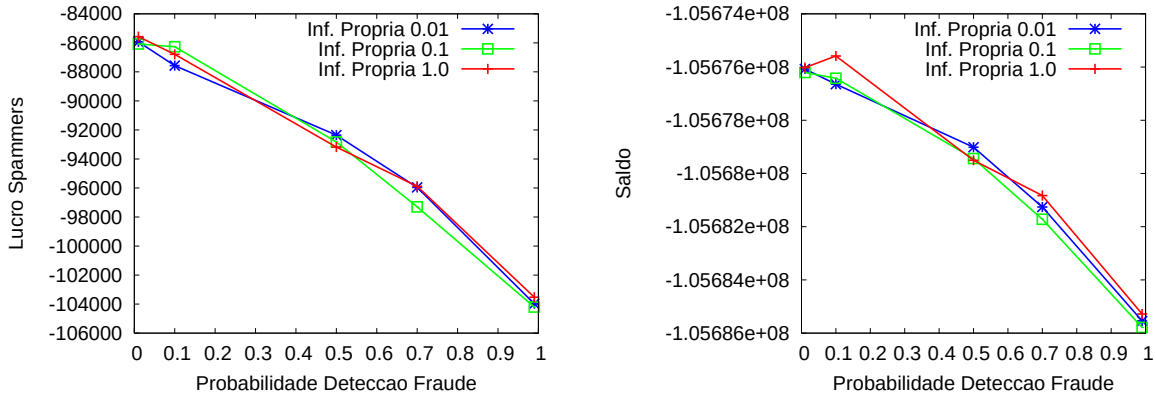
Os gráficos 4.14(a) demonstram que a alteração na influência própria tem pouca influência no impacto da probabilidade de fraude, tanto no lucro quanto no saldo. Isso provavelmente se deve ao fato de que, com a existência de uma influência externa média (0,1), usuários não devem ser enganados vezes suficientes para que sua opinião própria passe a ser determinante. Em relação ao saldo, 4.14(b), novamente os valores são dominados pela influência externa, demonstrando que a opinião pessoal exerce pequeno efeito no sistema.

Importante mencionar que, nos gráficos 4.14, nenhum das curvas pôde ser diferenciada de nenhuma das outras, com 90% de confiança.

Em suma, a hipótese é falsa. Os valores sugerem que a influência pessoal não tem, ou tem um impacto pequeno no efeito da probabilidade de fraude no sistema. Ou seja, o uso da informação externa têm maior impacto, já que traz maior quantidade de informação que apenas experiências pessoais. E o uso dessa informação extra contribui mais na redução dos lucros. Obviamente, o fato de haver um fornecedor não-spammer

tem um papel decisivo, posi, caso contrário, o usuário não poderia obter o produto de outra maneira.

4.4.4 Hipótese 4: À medida que a habilidade em detectar fraude diminui, o aumento na influência de experiências pessoais se torna menos eficaz em melhorar o sistema.



(a) (b)

Figura 4.15. Os gráficos apresentam o efeito do aumento da influência pessoal no impacto da probabilidade de detecção fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Novamente, como apresentado nos gráficos 4.15, a hipótese não é verdadeira. A influência própria não apresentou efeito significativo sobre o impacto da detecção de fraude no saldo e no lucro. Importante mencionar que, para probabilidades de detecção baixas, a influência externa perde seu efeito sobre o impacto da detecção de fraudes em ambos os casos. Assim, para probabilidades de detecção de fraudes baixas, o sistema é completamente vulnerável e nenhuma forma de opinião pode ajudar. Já para probabilidades mais altas, o lucro dos *spammers* decai mais para influências maiores da opinião externa, sem apresentar efeitos significativos devido à opinião pessoal. Em relação ao saldo, novamente os valores são dominados pela influência externa, demonstrando que a opinião pessoal exerce pequeno efeito no sistema.

Em relação ao gráfico 4.15(b), o ponto 0,1 para a curva 1,0 apresenta um comportamento inesperado. Esse comportamento não pôde ser verificado com 90% de confiança.

4.4.5 Hipótese 5: Na ausência de informações externas ou quando elas são desconsideradas, o sistema é vulnerável a fraudes

Os resultados das hipóteses 3 e 4 mostram que o uso de experiência pessoal tem um impacto no sistema muito inferior ao do uso de informação externa. Porém, isso induz a pergunta: em um ambiente em que não haja troca de informações, ou seja, em que as informações externas sejam inexistentes ou ignoradas, como o sistema se comportaria? Assim, como o sistema se comportaria se o usuário, no cálculo de sua confiança, pudesse se basear apenas em experiências pessoais?

Os gráficos 4.16(a),(b) possuem um comportamento não usual. Porém, esses gráficos foram gerados com 70 execuções e, nenhum dos pontos pôde ser diferenciado com 90% de confiança.

O gráfico 4.16(a) mostra que o uso da experiência pessoal tem pequeno impacto no lucro, se comparado com a influência externa, como pode ser percebido pela escala do gráfico. Ainda, é possível perceber que com a ausência de influência externa, mesmo no caso em que a influência pessoal foi igual a 1,0 (uma compra fraudulenta a mais que compras honestas representa que nunca mais o usuário comprará de *spammers*) o lucro foi positivo.

Os gráficos 4.16(c),(d) mostram que o uso de experiência pessoal, apenas, não melhora o sistema. Afinal, aumentar o valor da influência pessoal não alterou o grau de proteção do sistema contra fraudes.

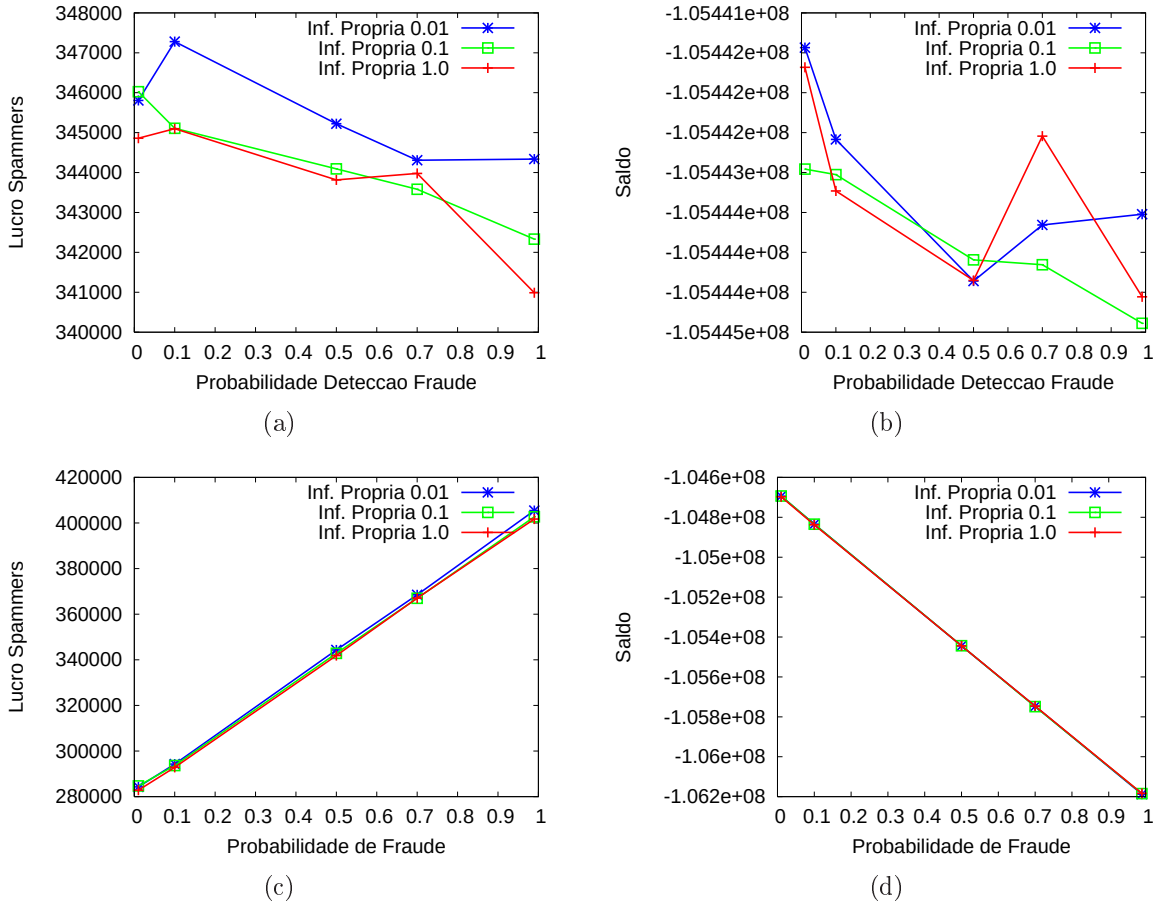
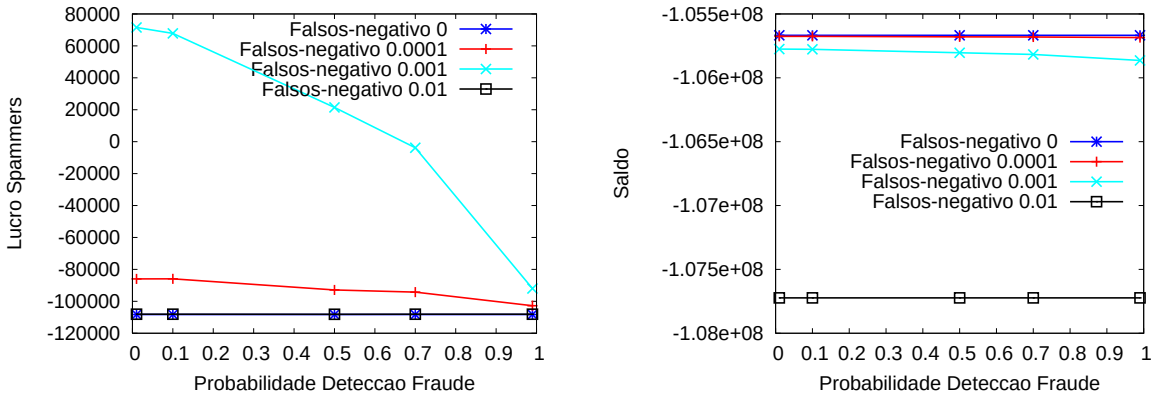


Figura 4.16. Os gráficos apresentam o impacto da influência pessoal quando a influência externa é igual a zero. Assim, a situação em que não há informação externa, ou essa é ignorada, é estudada. O gráfico (a), no canto superior esquerdo, apresenta os resultados do impacto da influência própria e da probabilidade de detecção de fraude no lucro. O gráfico (b), no canto superior direito, apresenta os mesmos resultados, porem, para o saldo.O gráfico (c), no canto inferior esquerdo, apresenta os resultados do impacto da influência própria e da probabilidade de fraude no lucro. O gráfico (d), no canto inferior direito, apresenta os mesmos resultados, porem, para o saldo.

4.4.6 Hipótese 6: À medida que o filtro se torna mais preciso, e a probabilidade de detecção de fraude diminui, mais o sistema se beneficia com filtro.



(a) (b)

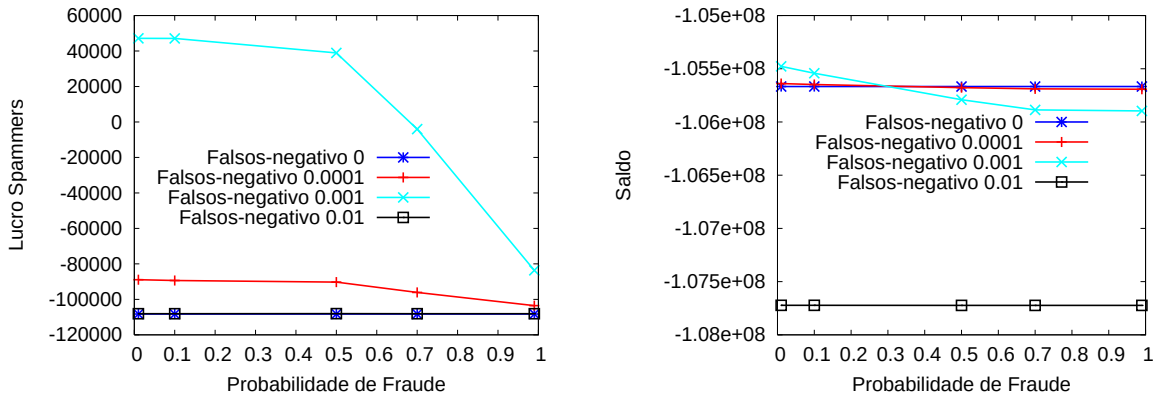
Figura 4.17. Os gráficos apresentam o efeito da qualidade do filtro no impacto da probabilidade de detecção fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Como explícito nos gráficos 4.17, para taxas de falsos-negativo baixas ou altas, o sistema é robusto à variação na probabilidade de detecção de fraudes. Assim, em relação ao lucro (4.17(a)), o sistema se torna melhor (menor lucro) com a melhora do filtro. Para taxas intermediárias, o filtro perde sua influência no impacto da probabilidade de detecção de fraude no lucro. Afinal, ele entrega quantidades não-excessivas de *spam*, portanto, a ação do filtro diminui e o sistema passa a depender mais de outros fatores.

O saldo do sistema, 4.17(b), depende claramente da taxa de falsos-negativo. Para taxas altas, o saldo é menor que para taxas baixas. Isso se deve ao fato de que *spams* entregues na caixa custam mais à sociedade que *spams* filtrados. E isso acaba por dominar o saldo.

Em suma, somente filtros muito bons ou ruins influenciam o impacto da probabilidade de detecção de fraudes no sistema. Filtros intermediários não são capazes de proteger o sistema contra fraudes. Inclusive, alguns podem contribuir para que fraudes ocorram, ao entregar um número ideal de *spams* (entrega *spam*, mas não um número grande de mensagens, capaz de despertar rejeição).

4.4.7 Hipótese 7: À medida que o filtro se torna mais preciso, e a probabilidade de fraude aumenta, mais o sistema se beneficia com filtro.



(a) (b)

Figura 4.18. Os gráficos apresentam o efeito do aumento da qualidade do filtro no impacto da probabilidade de fraude. O impacto é avaliado tanto para o lucro obtido por *spammers* quanto para o saldo. O primeiro gráfico representa o impacto no lucro (a) e o segundo no saldo (b).

Os gráficos 4.18 apresentam um efeito semelhante ao gráfico relativo à probabilidade de detecção de fraude. Portanto, um aumento na qualidade do filtro tem efeitos positivos somente para filtros de alta qualidade. O mesmo efeito também pode ser obtido por filtros ruins.

O gráfico 4.18(b) assemelha-se ao gráfico relativo à probabilidade de detecção de fraudes e o saldo. Porém, pode-se notar que a curva relativa a 0,001 apresenta um comportamento inesperado. Esse comportamento realmente ocorreu, com 90% de confiança. A justificativa para isso é que esse ponto é o de maior número de vendas. Inicialmente, para um volume de vendas mais alto, o sistema se beneficia do ganho social com essas vendas. Porém, basta que haja um aumento na probabilidade de fraude para que esse efeito seja perdido.

4.4.8 Hipótese 8: O efeito do filtro diminui se os usuários utilizarem a caixa de *spam* na pesquisa por fornecedores.

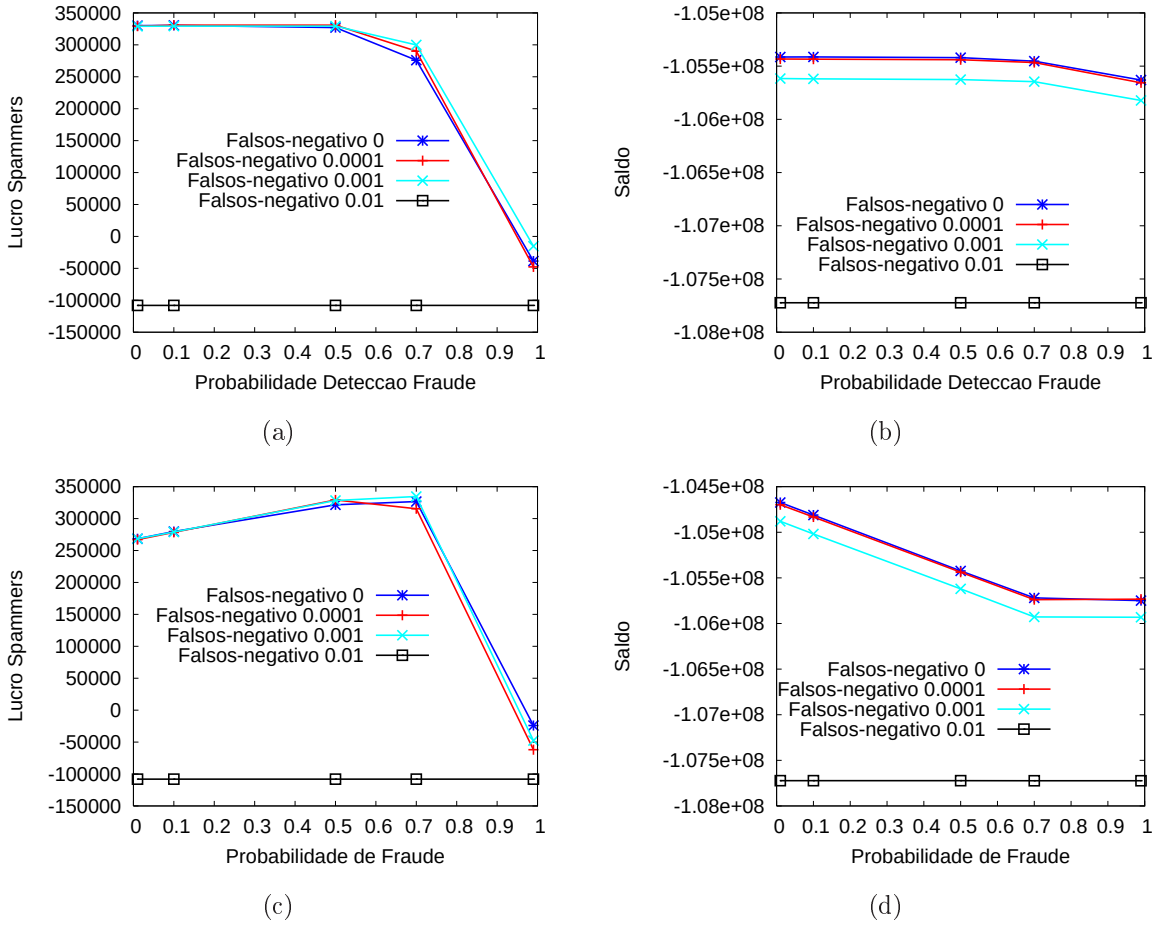


Figura 4.19. Os gráficos demonstram o efeito que o uso da caixa de *spam* tem na eficácia do filtro em lidar com fraudes. O gráfico no topo esquerdo (a), apresenta a análise para a detecção de fraude e seu impacto no lucro para filtros de qualidades diferentes e com o uso da caixa de *spam*. O gráfico no topo direito (b) apresenta os mesmos resultados que (a), porém, para o saldo. O gráfico (c) no canto inferior esquerdo apresenta os resultados para a probabilidade de fraude sob diversas qualidades de filtro e seu impacto no lucro. O gráfico no canto inferior direito (d) apresenta a mesma análise, entretanto, em relação ao saldo.

Em relação ao lucro, pode-se perceber, pelos gráficos 4.19(a),(c), claramente que a melhora do filtro perde o efeito. Afinal, para filtros de alta qualidade, que anteriormente mantinham o lucro em níveis baixos apesar do aumento da fraude ou diminuição da detecção, agora apresentam um comportamento semelhante ao dos filtros de qualidade intermediária. O filtro de má qualidade apresentou os melhores resultados, afinal, ele

se baseia na rejeição em relação ao *spam*, e não em evitar que mensagens cheguem ao destinatário.

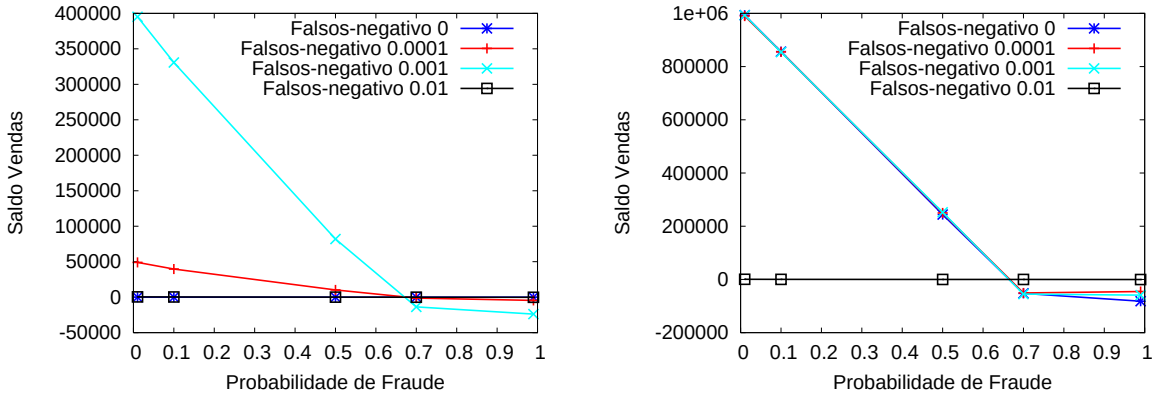
Ainda nos gráficos 4.19(a),(c), é possivel perceber que o lucro aumenta, se comparado à situação sem o uso da caixa de *spam*. Também, é possivel perceber que os usuários se tornam mais vulneráveis a fraudes e menos protegidos pelo aumento da detecção de fraudes. Afinal, o ponto de inflexão, no qual o lucro começa a cair, em ambos os gráficos, ocorreu para um ponto superior, se comparado à situação sem a utilização da caixa de *spam*.

O efeito do filtro no impacto da probabilidade de detecção de fraudes no saldo (gráfico 4.19(b)) demonstra que o sistema, para uma mesma taxa de falsos-negativo, apresenta um ganho maior se mais pessoas forem enganadas, já que, como dito anteriormente, a análise é puramente financeira e o par probabilidade de fraude e diferença entre preço externo e do *spammer* acaba compensando as fraudes. Também é possível perceber que o pior filtro apresentou o pior saldo, já que, além de diminuir o número de vendas de *spammers*, ainda aumentou o volume de *spam* na caixa de entrada. Outro efeito a ser percebido é que, para probabilidades de detecção altas (0,99 no gráfico), há uma queda no saldo. Isso demonstra a influência da confiança em reduzir o número de vendas. Esse comportamento também é observado no gráfico 4.17 correspondente. Porém, nesse gráfico o efeito somente ocorre para o filtro de 0,001, ou seja, para o pior filtro intermediário. Entretanto, no gráfico utilizando a caixa de *spam*, esse fenômeno ocorre para todos os filtros intermediários e para o filtro de alta qualidade.

O gráfico 4.19(d) mostra que a adição da caixa de *spam* diminuiu a eficácia do filtro contra um aumento na probabilidade de fraude. Inicialmente, o saldo aumenta com uso da caixa de *spam*, para probabilidades de fraude baixas e intermediárias. Isso se deve ao fato de que temos o melhor dos dois mundos, os filtros de alta qualidade evitam que spam chegue à caixa de entrada e os usuários interessados podem comprar de *spammers* com baixa probabilidade de fraude. Porém, para taxas altas de fraude, o saldo cai para níveis próximos do caso sem a caixa de *spam*. Se repararmos bem, o comportamento é semelhante ao apresentado para o filtro 0,001 nos gráficos 4.18, para a situação correspondente sem o uso da caixa de *spam*. Ou seja, os filtros de qualidade alta e intermediária, com o uso da caixa de *spam*, se comportaram de maneira semelhante ao pior filtro de qualidade intermediária no conjunto de gráficos sem o uso da caixa de *spam*. Novamente, o pior filtro apresentou comportamento estável e saldo baixo.

Para visualizar melhor o efeito do uso da caixa de *spam* no saldo, os gráficos 4.20, com os valores correspondentes ao saldo das vendas. O saldo das vendas inclui somente o saldo com a compra de produtos vendidos por *spammers* e a taxa de fraude, exclui,

portanto, o custo com o tratamento de mensagens *spam*. Assim, esse gráfico corresponde aos anteriores de saldo sem incluir o efeito do filtro no número de mensagens entregues.



(a) (b)

Figura 4.20. Os gráficos apresentam o saldo obtido exclusivamente com a venda de produtos de *spammers*. Nestes gráficos, a perda que a sociedade tem com o tratamento de *spam* não é considerada. Somente contabiliza-se a parcela resultante de economia com a compra de produtos entregues corretamente e a perda com fraudes. Isso ajuda a compreender os resultados do gráfico 4.19 em relação ao saldo. O gráfico (a) na esquerda, apresenta os resultados sem a utilização da caixa de *spam*. O gráfico (b) na direita apresenta os resultados para o caso utilizando a caixa de *spam*.

Pode-se observar que, pelos gráficos 4.20, para o caso sem o uso da caixa de *spam* (a), o filtro de taxa 0,001 sofreu o maior impacto no ganho devido ao *spam*, já que ele entrega uma quantidade ideal de mensagens. Porém, ao se adicionar a caixa de *spam* (b), todos os filtros apresentaram o mesmo ganho, exceto o filtro que se baseia na rejeição. Afinal, em nenhum dos outros filtros o usuário recebe uma quantidade de mensagens excessiva, a ponto de rejeitar completamente produtos de *spammers*.

Outro comportamento que pode ser percebido nos gráficos 4.20 é o de que, a partir de 0,7, a compra de *spammers* deixa de ser benéfica. Inicialmente, com probabilidade de fraude de 0,5, mesmo que algumas pessoas sejam enganadas, a economia com a compra de produtos de *spammers* compensa a perda com fraudes. Porém, para probabilidades de 0,7 ou maiores, a economia deixa de compensar, e a sociedade passa a perder, na média, com cada venda realizada por *spammers*.

Em suma, o efeito da adição da caixa de *spam* influenciou todos os resultados. Nos gráficos relativos ao lucro, anulou o efeito de filtros mais poderosos. Nos casos em que o

saldo foi analisado, os gráficos para os filtros de alta qualidade e qualidade intermediária se comportaram de maneira semelhante ao pior filtro de qualidade intermediária da situação sem o uso da caixa de *spam*. Novamente, isso demonstra que o efeito do filtro é reduzido ou eliminado com a adição do uso da caixa de *spam*.

4.5 Aplicação dos resultados: Análise do *spam* de farmácias canadenses

A venda de produtos de supostas farmácias canadenses é um das formas de *spam* mais populares atualmente. De acordo com [54], 75% do *spam* mundial é devido a esse tipo de *spam*. Mas por que esse spam é tão popular? Vários dos resultados apresentados anteriormente ajudam a responder essa pergunta.

O primeiro fator é que é difícil diferenciar medicamentos falsos de verdadeiros. Além disso, a existência de efeito placebo, no qual há resultados apesar de o medicamento não conter qualquer princípio ativo, também é um fator que contribui. Vide o efeito da probabilidade de detecção de fraudes na figura 4.8.

A existência de laboratórios clandestinos de medicamentos em países mais pobres. Isso permite que medicamentos com o princípio ativo sejam entregues a preços bem mais baixos. Assim, parte dos medicamentos vendidos pode conter o princípio ativo. Vale lembrar que no estudo [43], a maior parte dos medicamentos (Viagra) entregues continham o princípio ativo. Além disso, o acesso a esses laboratórios não é trivial. É necessária a formação de relativamente grandes e estruturadas organizações criminosas (chamadas *sponsors*), que sejam capazes de encontrar esses laboratórios, comprar seus produtos e revendê-los. Isso diminui a concorrência nesse negócio. Vide o impacto do custo dos insumos nas figuras 4.1 e 4.2. O impacto da probabilidade de fraude pode ser visto na figura 4.9.

O custo para o envio de spams *é importantíssimo.* A existência de *botnets* e outros recursos para o envio de *spam* facilitam enormemente o envio de *spam*. Também, esquemas de parcerias como o *spam*it [26], diminuem ainda mais o custo. Já que, agora, uma terceira parte é responsável pelos custos de envio de *spams*, e recebem somente pelas vendas. Ou seja, agora criou-se uma nova camada nesse ambiente. Existem as organizações por trás do produto e de seu comércio (*sponsor*) e os indivíduos por trás do envio de mensagens (os *spammers* propriamente dito). Normalmente consideraremos as duas entidades como uma só, mas a sua divisão reduz custos fixos. Outro custo que tem sido alvo dos *spammers* é o com a hospedagem e registro de suas lojas virtuais. O esquema Glavmed [43; 26] é um exemplo, em que pessoas ofereceriam espaço em suas

páginas para a venda de medicamentos, ficando com uma parcela dos lucros. Vide o efeito desses custos nas figuras 4.1 e 4.2.

O efeito do filtro, que é capaz de barrar a grande maioria dos spams. Daí, resulta a pergunta: as compras são realizadas a partir de mensagens que chegaram à caixa de entrada ou também incluem mensagens na caixa de *spam*? Como esse *spam* é extremamente popular nas caixas de *spam*, é difícil de se acreditar que, caso o usuário tenha interesse, ele não saberá da existência dessas mensagens na sua caixa de *spam*. Vide o efeito do filtro na figura 4.6, 4.18 e do uso da caixa de *spam* nas figuras 4.11 e 4.19.

O sistema americano também apresenta uma influência grande nesse esquema. Afinal, os preços de seus medicamentos são mais altos que em outros locais e a grande maioria dos medicamentos precisa ser vendida com prescrição. Então, surgiu a prática de se comprar medicamentos em outros países próximos, como Canadá e México. *spammers* se aproveitaram da divulgação dessa prática, e do fato de que nem todo americano poderia realizar viagens periódicas a esses países, para supostamente oferecer o serviço. Supostamente devido ao fato de que geralmente *spammers* entregam medicamentos originários da China ou Índia. Portanto, seriam capazes de oferecer produtos mais baratos, confiáveis (afinal os medicamentos supostamente viriam de uma farmácia canadense) e sem necessidade de prescrição. Então, essa combinação de fatores criou uma taxa alta de interesse entre americanos. Veja o impacto dos preços na figura 4.7e o da taxa de interesse nas figuras 4.10 e 4.2.

A compra de medicamentos de farmácias canadenses com registro nos EUA e com prescrição é legal nos EUA. Porém, a compra de medicamentos sem prescrição não é. A maior parte dos *spams* explora esse fato ao apresentar certificados falsos [6], tanto da legalidade da farmácia no Canadá quanto nos Estados Unidos. O detalhe de ser ilegal vender sem prescrição normalmente não é informado. A idéia é criar confiança suficiente no possível comprador, para que ele compre o medicamento que deseja comprar, quer tenha prescrição ou não. Todos os experimentos consideraram que, sem qualquer informação além do próprio site do *spammer* nem experiência prévia, o usuário considera o site legítimo e confiável.

A troca de informações entre compradores não é difundida. A questão legal não é exatamente clara. Então, para um usuário divulgar que realizou uma compra e foi enganado, há um risco de processo criminal (porte de substância controlada). Além disso, admitir publicamente a compra de certos tipos de medicamentos, bem como admitir que foi enganado, não são muito atrativos. Assim, o sistema se beneficia pouco da influência externa em evitar que compras sejam realizadas. O impacto do uso de informação externa pode ser visto nas figuras 4.5,4.12 e 4.13. O impacto do uso de

informação pessoal pode ser visto nas figuras 4.14, 4.15 e 4.16.

O sistema pode não apresentar lucro ou somente uma parte da cadeia de spam *(*sponsor*) apresentaria lucro.* Porém, a expectativa de lucro atuaria da mesma maneira que a existência de lucro, especialmente para as camadas terceirizadas como o envio de *spam* e a hospedagem de lojas virtuais. Então, essas camadas poderiam se renovar constantemente, enquanto a principal (*sponsor*) se manteria constante e lucrativa. Esse seria um fenômeno semelhante ao levantado no artigo [41], relativo a outra atividade criminosa, o *phishing*. Isso justificaria o fato de que a maior parte dos gráficos apresentou prejuízo para o custo de envio igual a $0,000005.

Portanto, como se pode perceber, o *spam* de farmácias canadenses explora grande parte dos principais fatores do sistema de maneira bastante eficaz. Assim, podemos ver que esse tipo de *spam* perdura até hoje por utilizar esses fatores em seu saldo, tornando o sistema mais vantajoso e mais difícil de ser combatido.

Capítulo 5

Conclusão

Esta dissertação realizou a modelagem do mercado por trás do *spam*. Esse mercado é formado por *spammers* interessados em vender seus produtos e os usuários, que podem ter interesse em adquirir esses produtos ou não. O modelo se baseou na premissa de que o mercado formado por spammers e usuários é semelhante a outros mercados e, portanto, segue regras comuns a esses últimos.

Esses dois participantes foram modelados como agentes. Seus comportamentos e ações também foram modelados. O modelo também inclui a possibilidade de fraude na venda de um produto via *spam* e a capacidade dos usuários em detectar uma fraude. Ainda, o efeito da confiança foi analisada nesse sistema. Também, o uso ou não da caixa de *spam* na busca de fornecedores foi modelado.

Devido à insuficiência de dados na literatura sobre o mercado e o comportamento de usuários e *spammers*, algumas análises, especialmente as de caráter quantitativo, não puderam ser realizadas. Assim, por exemplo, é impossível, com o grau de informação atual, inferir o lucro que os *spammers* teriam ao se melhorar o filtro em 20%. Portanto, devido a essa limitação, as análises se basearam no comportamento qualitativo dos fatores, sugerindo comportamentos interessantes e interações entre fatores.

O modelo possui uma quantidade grande de fatores fundamentais. Isso dificulta a análise e o estudo completo do modelo. Então, algumas interações entre fatores, bem como fatores, não foram analisados. Os fatores e interações analisados foram escolhidos como os de maior interesse e de aplicação mais prática. A maior parte das interações não analisadas resulta apenas em alterações quantitativas, sem nenhuma alteração no comportamento dos fatores participantes. Por exemplo, analisar a influência do aumento no interesse na qualidade do filtro resultaria em maiores lucros, não em mudanças significativas na forma da curva obtida. Essas mudanças são, em muitos casos previsíveis e, como o modelo não pode apresentar valores quantitativos de ma-

neira precisa, essas alterações são de interesse bastante limitado. Além disso, durante as análises, fez-se questão de discutir a influência de outros parâmetros no resultado, quando esse impacto é claro e de interesse.

Várias simplificações foram necessárias para tornar o sistema tratável. Essas simplificações se basearam em simplificações correntes em modelos econômicos. Porém, pela falta de dados reais para o modelo, não é possivel inferir o grau de impacto dessas simplificações.

A partir do estudo de vários cenários, envolvendo configurações diferentes de fatores, diversas conclusões puderam ser alcançadas:

1. A obtenção de lucro por parte dos *spammers* pode não ser tão fácil quanto se espera. A maior parte dos cenários gerados obtiveram prejuízo. Inclusive, isso induz a conclusão de que há necessidade de condições especiais do mercado para que haja lucro. Como por exemplo, uma taxa de filtragem ótima, alta probabilidade de interesse por parte dos usuários no produto vendido, pouca fraude ou pouca capacidade de detecção de fraudes. Esse resultado talvez contradiga o crescimento recente no volume de *spam*, apontado por [35]. Porém, o artigo [41] pode fornecer uma explicação. Não necessariamente o crescimento se deve, obrigatoriamente, ao aumento no lucro. Um aumento no volume pode significar também um aumento na expectativa de lucro, o que pode não corresponder a um lucro de fato. Importante ressaltar que os valores negativos representam que o lucro obtido foi inferior ao valor cobrado por donos de *botnets* para o envio de *spam*. Nada impede que, ao receber valores inferiores a esse valor cobrado, todas as partes envolvidas no *spam* obtenham lucro. Portanto, esta dissertação não afirma que a atividade não seja lucrativa, apenas demonstra que o lucro pode não ser tão alto e que, maiores lucros ocorrem em cenários específicos e não indiscriminadamente.

2. A existência de fraudes torna a utilização de experiência externa útil na redução dos lucros de *spammers*, porém, é necessário que os compradores sejam capazes de identificar corretamente as fraudes. Portanto, a troca de experiências entre usuários pode ser valiosa nesse sistema, especialmente se houver uma taxa de fraude alta e se forem fornecidos mecanismos para facilitar a detecção das mesmas. Além disso, a informação externa não precisa ser, rigorosamente, proveniente de opiniões de usuários. Organizações anti-spam podem promover uma análise dos produtos vendidos via *spam*, relatando qual a qualidade desses itens. E, se essa taxa de fraudes for alta, a confiança irá agir de maneira a reduzir o lucro dos *spammers*.

3. A melhora na qualidade de um filtro não necessariamente reduz o lucro de *spammers*. Em alguns casos, é possível que essa melhora gere aumento no lucro. Nem sempre piorar o filtro aumenta o lucro dos *spammers*. Devido ao fato de existir rejeição a *spam* caso o volume de *spam* na caixa de entrada seja grande, piorar o filtro pode trazer uma redução significativa no lucro. Portanto, uma estratégia de combate ao *spam* pode ser, simplesmente, piorar o filtro.

4. Obviamente, em relação ao filtro, é necessário avaliar também as respostas dos *spammers*. Melhorar um filtro pode sim resultar em um maior volume de *spam* gerado. Afinal, *spammers* podem aumentar o volume de *spam* gerado para compensar essa melhora no filtro. De maneira análoga, piorar o filtro pode decrementar o volume de *spam* gerado. A idéia é que, ao se piorar o filtro, mais *spam* chegaria à caixa de entrada. Devido a esse aumento no volume de *spam* recebido, a rejeição também aumentaria, diminuindo o número de vendas. E a resposta dos *spammers* a esse fenômeno seria a redução no volume de *spam* gerado, com o objetivo de entregar um menor número de *spams* à caixa de entrada e reduzir a rejeição. Dessa forma, apesar de o filtro pior resultar em um saldo pior, a resposta dos *spammers* pode melhorar esse saldo, inclusive para níveis mais altos que os melhores filtros.

5. Uma conjectura interessante que pode ser retirada dos estudos e análises anteriores é que, talvez, o filtro tenha agravado o problema do *spam* comercial. Possivelmente, se em momento nenhum da história filtros tivessem sido criados, o próprio mercado iria se auto-regular em um número tolerável de *spams* por usuário. Claramente, se um usuário recebesse um número maior que o tolerável, sua rejeição iria atuar e produtos não seriam comprados.

6. Os usuários podem utilizar sua caixa de *spam* na busca de fornecedores. Vale lembrar que o conceito de *spam*, para o usuário, é algo subjetivo. Dado que o usuário tem interesse no produto, ele pode muito bem considerar que uma dada mensagem não-solicitada enviada em massa é desejável e, portanto, não a classificaria como *spam*. Assim, o uso da caixa de *spam* acaba por anular o efeito positivo de se melhorar o filtro, quando esse efeito positivo ocorre. Por exemplo, em alguns casos, aumentar a qualidade do filtro poderia gerar uma redução nos lucros de *spammers*, porém, com o filtro, essa redução já não mais ocorreria. Somente o filtro ruim foi resistente a esse efeito do uso da caixa de *spam*, afinal, ele se baseia na rejeição dos usuários a produtos vendidos por *spammers*, e não na filtragem propriamente dita.

7. O custo por mensagem enviada foi um dos fatores de maior impacto no lucro dos *spammers*. Ou seja, as estratégias econômicas seguem sim um bom princípio. Porém, se a probabilidade de interesse e a de compra forem suficientemente altas, o custo por mensagem perde seu impacto. Nesse cenário, é melhor aumentar os custos dos insumos ou reduzir os preços cobrados por *spammers*.

8. A análise do saldo, mostrou que, financeiramente, pode ser melhor permitir uma certa taxa de fraude em troca da manutenção do mercado. Ou seja, em alguns casos, especialmente quando não colocarem a vida de pessoas em risco e quando a diferença de preços entre *spammers* e fornecedores externos for grande, o sistema tolera fraudes. Assim, nesses casos, incentivar a troca de informações e a confiança, apesar de reduzir o lucro, pode decrescer o saldo. Nos experimentos, uma taxa de 0,5 de fraude foi tolerada, enquanto uma taxa de 0,7 tornou melhor evitar toda e qualquer venda.

9. Em vários experimentos, lucro e saldo aumentaram sob as mesmas circunstâncias. Isso acaba por delimitar uma escolha. Sacrificar o saldo para reduzir o lucro (pensando na redução do *spam* a longo prazo), ou incentivar o lucro para também incentivar o saldo.

10. Os saldos, em todos os gráficos, foram negativos. Isso demonstra que a sociedade está melhor sem *spam*. Portanto, pode ser tolerável sacrificar o saldo momentaneamente, se isso trouxer redução nos lucros e falências de *spammers*.

O fundamental a ser feito, como trabalho futuro, é coletar dados. Entender o comportamento, bem como os valores dos parâmetros do modelo. Assim, seria possível refinar o modelo e apresentar resultados quantitativos, que permitiriam planejamentos mais apurados.

Outro ponto que merece maior investigação é o ganho que o mundo teria com a piora dos filtros. Apesar de parecer contra-intuitivo, essa estratégia merece ser melhor investigada. Afinal, ela tem o potencial de reduzir o total de *spam* gerado para níveis toleráveis, reduzindo custos e o volume total de *spams* gerados.

Por fim, estratégias de *targeting,* ou seja, de entregar mensagens spam somente a pessoas interessadas poderia ser uma boa alternativa. Especialmente se esse mecanismo fosse utilizado por *spammers* na escolha dos destinatários de suas mensagens. Obviamente, questões como o aumento no custo devido ao *targeting* e a redução no volume de mensagens precisam ser ponderadas.

Referências Bibliográficas

[1] The 10 worst rokso spammers. *Spamhaus*.

[2] Alex polyakov. *Spam Trackers Wiki*.

[3] Boxbe. *www.boxbe.com*.

[4] Can-spam act. *http://www.ftc.gov/bcp/edu/pubs/business/ecommerce/bus61.shtm*.

[5] Canadadrugs. *www.canadadrugs.com*.

[6] Canadian pharmacy. *Spam Trackers Wiki*.

[7] Centmail. *http://centmail.net*.

[8] coolsavings. *www.coolsavings.com*.

[9] Domainkeys identified mail. *http://www.dkim.org/*.

[10] Doubleclick.com. *www.doubleclick.com*.

[11] Drugstore. *www.drugstore.com*.

[12] E-mail pays u. *www.e-mailpaysu.com/*.

[13] Eclipse. *http://www.eclipse.org*.

[14] Goodmail. *www.goodmail.com*.

[15] Hits4pay. *http://hits4pay.com/*.

[16] Inboxdollars. *www.inboxdollars.com*.

[17] Industry statistics. *Ferris Research*.

[18] Java. *http://java.sun.com*.

[19] Leo kuvayev. *Wikipedia*.

[20] Repast. *http://repast.sourceforge.net*.

[21] The rokso list. *Spamhaus*.

[22] Sancash. *Spam Trackers Wiki*.

[23] Senderid. *http://www.microsoft.com/mscorp/safety/technologies/senderid/default.mspx*.

[24] Spam trackers. *http://www.spamtrackers.eu*.

[25] Spamhaus. *http://www.spamhaus.org*.

[26] Spamit. *Spam Trackers Wiki*.

[27] Sex, drugs and software lead spam purchase growth. *M86 Security*, August 2008.

[28] George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

[29] Marshall Van Alstyne. Curing spam: Rights, signals & screens. *The Economists' Voice*, 4(2), 2007.

[30] Mark Brownlow. Email and webmail statistics. *Email Marketing Campaign*.

[31] Robert Buderi. Microsoft declares war on spam. *Technology Review*, feb 2005.

[32] Pedro Calais, Douglas Pires, Dorgival Olavo Guedes Neto, Wagner Meira Jr., Cristine Hoepers, and Klaus Steding-Jessen. A campaign-based characterization of spamming strategies. In *CEAS*, 2008.

[33] R. H. Coase. The problem of social cost. *Journal of Law and Economics*, 3:1–44, 1960.

[34] Swait J. Erdem, T. and J. Louviere. The impact of brand credibility on consumer price sensitivity. *International Journal of Research in Marketing*, 19(1), 2002.

[35] David M. Ewalt. Nine out of ten emails are spam. *Forbes*, Maio 2009.

[36] Bill Gates, Nathan Myhrvold, and Peter Rinearson. *The Road Ahead*. Viking, 1995.

[37] Luiz Henrique Gomes, Cristiano Cazita, Jussara M. Almeida, Virgílio Almeida, and Wagner Meira, Jr. Characterizing a spam traffic. In *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, pages 356–369, New York, NY, USA, 2004. ACM.

[38] Ram D. Gopal, Arvind K. Tripathi, and Zhiping D. Walter. Economics of first-contact email advertising. *Decis. Support Syst.*, 42(3):1366–1382, 2006.

[39] Garrett Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, December 1968.

[40] C. Herley and D. Florencio. A profitless endeavor: Phishing as a tragedy of the commons. *NSPW*, 2008.

[41] C. Herley and D. Florencio. Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. *WEIS*, 2009.

[42] Shlomo Hershkop. *Behavior-based email analysis with application to spam detection.* PhD thesis, New York, NY, USA, 2006. Adviser-Salvatore J. Stolfo.

[43] Mark Iype. Online pharma spammers capitalize on canada's health-care reputation. *Ottawa Citizen*, agosto 2009.

[44] Eric S. Johansson and Keith Dawson. A better way to squelch spam? *Technology Review*, mar 2004.

[45] David Heckerman Joshua Goodman, Gordon V. Cormack. Spam and the ongoing battle for the inbox. *Communications of the ACM*, 50(3):25 – 33, feb 2007.

[46] Robert Rounthwaite Joshua Goodman. Stopping outgoing spam. *ACM Conference on Electronic Commerce*, pages 30 – 34, mai 2004.

[47] Chris Kanich, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M. Voelker, Vern Paxson, and Stefan Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the 15th ACM Conference on Computer and Communications Security*, pages 3–14, Alexandria, Virginia, USA, October 2008.

[48] Tomas Benjamin Klos. *Agent-based Computational Transaction Cost Economics.* PhD thesis, 2000. Adviser- B. Nooteboom.

[49] B. Krishnamurthy and E. Blackmond. Shred: Spam harassment reduction via economic disincentives.

[50] Robert Lemos. Mccolo takedown nets massive drop in spam. *Security Focus*, 2008.

[51] Mcafee. The carbon footprint of email spam report. *Mcafee.*

[52] Fabian Menges, Bud Mishra, and Giuseppe Narzisi. Modeling and simulation of e-mail social networks: a new stochastic agent-based approach. In *WSC '08: Proceedings of the 40th Conference on Winter Simulation*, pages 2792–2800. Winter Simulation Conference, 2008.

[53] Haining Wang Mengjun Xie, Heng Yin. An effective defense against email spam laundering. *Conference on Computer and Communications Security*, pages 179–190, 2006.

[54] Ellen Messmer. Experts link flood of canadian pharmacy spam to russian botnet criminals. *Network World*, 2009.

[55] Robert K. Plice Nigel Melville, Aaron Stevens and Oleg V. Pavlov. Unsolicited commercial e-mail: Empirical analysis of a digital commons. *International Journal of Electronic Commerce*, 10(4):143, 2006.

[56] W. Yang P. Judge, D. Alperovitch. Understanding and reversing the profit model of spam. *Workshop on Economics of Information Security*, jun 2005.

[57] Oleg V. Pavlov, Nigel Melville, and Robert K. Plice. Toward a sustainable email marketing infrastructure. *Journal of Business Research*, 61(11):1191–1199, November 2008.

[58] Robert K. Plice, Nigel P. Melville, and Oleg V. Pavlov. Toward an information-compatible anti-spam strategy. *Commun. ACM*, 52(5):128–130, 2009.

[59] Steven F. Railsback, Steven L. Lytinen, and Stephen K. Jackson. Agent-based simulation platforms: Review and development recommendations. *SIMULATION*, 82(9):609–623, September 2006.

[60] Pattie Maes Robert, Robert H. Guttman, and Ros G. Moukas. Agents that buy and sell: Transforming commerce as we know it, 1999.

[61] Stuart E. Schechter. Toward econometric models of the security risk from remote attack. *IEEE Security & Privacy*, 3(1):40 – 44, feb 2005.

[62] Evan I. Schwartz. Spam wars. *Technology Review*, jul 2003.

[63] Michael J. Freedman Brad Karp David Mazières Haifeng Yu Scott Garriss, Michael Kaminsky. Re: Reliable email. *3rd Symposium on Networked Systems Design and Implementation*, mai 2006.

[64] Andrei Serjantov and Richard Clayton. Modelling incentives for email blocking strategies. *Workshop on the Economics of Information Security*, jun 2005.

[65] Leigh Tesfatsion. Agent-based computational economics: A constructive approach to economic theory. In Leigh Tesfatsion and Kenneth L. Judd, editors, *Handbook of Computational Economics*, volume 2, chapter 16, pages 831–880. Elsevier, 1 edition, 2006.

[66] Rick Wash Theodore Loder, Marshall Van Alstyne. Advances in economic analysis & policy. *The Berkeley Electronic Press*, 6(1), 2006.