Anonimato na Internet

Edelberto Franco Silva[1]

[1]Instituto de Computação – Universidade Federal Fluminense (UFF)
Disciplina: Tópicos Avançados em Redes e Sistemas Distribuídos e Paralelos II
(Segurança em Redes de Computadores)
2012.1

{esilva}@ic.uff.br

Resumo. *O presente trabalho tem por objetivo apresentar as mais difundidas técnicas de anonimato utilizadas atualmente na Internet. Apesar deste tema ser difundido desde o início da década passada, ainda há muitos trabalhos relevantes que têm gerado propostas inovadoras. O intuito, portanto, deste trabalho é introduzir o leitor nesta área e consolidar uma base teórica para um estudo continuado.*

1. Introdução

Desnecessária se faz comprovar a atual importância da Internet na sociedade. Sabe-se que hoje há uma forte dependência dos usuários pelos serviços oferecido através da Internet, assim como uma dependência infraestrutural com relação ao protocolo de comunicação base dessa rede, o IP (*Internet Protocol*). Como o único dado que *a priori* se tem em relação à informação de localização do usuário é seu endereço IP, uma característica intrínseca com relação à confiança desta informação é gerada uma vez que se sabe que com a escassez de endereços IP técnicas de reutilização e compartilhamento de IPs torna muito difícil a real localização de um usuário.

O anonimato na Internet pode ocorrer por diversos fatores; Um usuário pode esconder sua real origem utilizando desde técnicas básicas de *spoof* até outras muito mais refinadas, como as de roteamento oculto realizado por vários saltos, conforme será apresentado neste trabalho.

Deve-se, antes de qualquer embasamento teórico, discutir a importância ou não do anonimato na grande rede, ou quando esta pode ser utilizada visando uma boa ou uma má finalidade. Sendo assim, é importante expor o questionamento: o anonimato é bom? Quando? Por quê? E as respostas para essas perguntam dependem; Dependem exclusivamente da finalidade a que o anonimato se presta. Como ilustração, pode-se visualizar o anonimato de forma positiva para ocultar uma fonte e impedir a análise de tráfego de um cidadão cujo país fiscaliza e pune qualquer ideologia contrária ao seu governo. Por outro lado, o anonimato pode se tornar ruim, com uma finalidade indesejável socialmente, quando utilizado por um criminoso que visa o mal de terceiros, seja ele um terrorista trocando informações sobre um ataque ou um pedófilo exibindo suas imagens.

Neste trabalho serão apresentadas as principais técnicas de anonimato utilizadas atualmente na Internet. Para facilitar a compreensão deste ambiente, criou-se uma taxonomia que será apresentada na Seção 2. Após apresentada tal taxonomia, poder-se-á discutir com mais detalhes as técnicas abordadas neste trabalho, como serão abordadas na Seção 3, passando pelas técnicas baseadas em *proxy* e aquelas baseadas em redes P2P

(*Peer-to-Peer*), além daquelas mais clássicas (e simples). Por fim, finalizar-se-á na Seção 4.

2. Taxonomia

Como forma de facilitar a compreensão do presente trabalho e facilitar o acompanhamento durante o desenvolvimento do texto, foi criada uma taxonomia para as técnicas de anonimato na Internet mais utilizados atualmente. Acredita-se que a partir desta taxonomia fique mais clara a compreensão das técnicas existentes, desde às mais simples às mais sofisticadas de anonimato na Internet.

A Figura 1 classifica em duas grandes áreas o anonimato na Internet, onde, técnicas utilizadas em aplicações P2P e outras consideradas mais simples do ponto de vista teórico são agrupadas distintamente daquelas baseadas em *proxy*. Técnicas baseadas em *proxy*, são, na realidade, técnicas que se utilizam de roteamento de múltiplos saltos (*proxies* em cascata - aninhados) para ocultar a fonte. São consideradas melhores técnicas aquelas que apresentam algum método seguro e oculto de comunicação entre os nós intermediários.

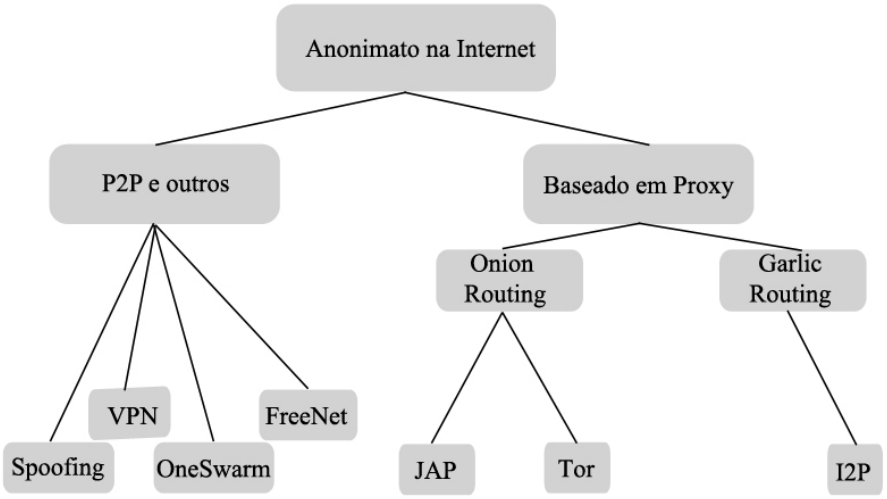

Figura 1. Taxonomia de anonimato na Internet.

Esta seção buscou introduzir a classificação das técnicas a serem apresentadas na Seção 3.

3. Técnicas de Anomimato

São diversas as técnicas de anonimato que podem ser utilizadas na Internet, sendo que algumas visam apenas uma ocultação simples da origem com a intenção de forjar uma fonte. Porém, há também aquelas onde técnicas mais sofisticadas, com a utilização de uma rede de nós confiáveis (rede de amigos), é utilizada. Nesta seção serão apresentadas as mais diversas técnicas, suas funcionalidades e principais características, conforme foi exposto pela Figura 1.

3.1. Baseado em P2P e outros

Primeiramente serão abordadas técnicas mais simples do ponto de vista de sua possível funcionalidade. Apresentar-se-á neste momento as técnicas de *spoofing* de IP e MAC (*Media Access Control*), VPN (*Virtual Private Network*) e os protocolos e ferramentas P2P mais comentadas na atualidade.

IP/MAC Spoofing

Pacotes enviados utilizando o protocolo IP [11] incluem o endereço da origem em seu cabeçalho (no campo *IP Source*), porém este endereço não é verificado pelo protocolo, o que torna simples forjar este endereço, como demonstrado por diversos trabalhos [2, 14, 6]. Este tipo de ataque pode servir tanto para esconder a origem de um ataque de negação de serviço quanto para interceptar um tráfego destinado a outro nó.

A Figura 2 demonstra uma tentativa de ataque de negação de serviço (*Denial of Service* - DoS) cujo atacante tem como fonte seu 168.12.25.5, mas forja o cabeçalho IP para o endereço 156.12.25.4. Desta forma, o destino (companhia XYZ) ao receber o pacote, envia uma resposta ao endereço forjado e não ao original.

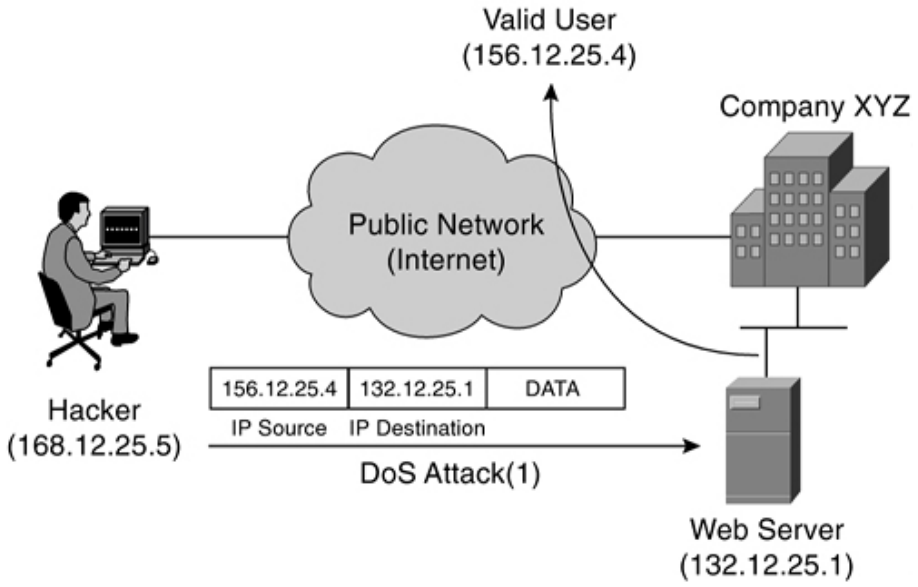

Figura 2. Exemplo de *spoofing* do endereço IP.

Deve-se lembrar que este ataque é muito simples e se utilizado sob conexões persistentes, como é o caso do TCP (*Transmission Control Protocol*), este não funcionaria (a não ser que houvesse a participação e colaboração do nó cujo IP real foi ”spoofado”). Sendo assim, qualquer tipo de validação da origem compromete este ataque de anonimato.

VPN

A utilização de VPN (*Virtual Private Network*) como ocultação de fonte também se demonstra equivocada, porém ainda é muito citada por iniciantes nesta área. Como

se sabe, a VPN realiza apenas a ocultação dos dados, cuja idéia é encaminhar dados de forma segura sob um canal inseguro (como a Internet). Porém, deve-se destacar que esta técnica, quando em conjunto com outra, (*e.g. proxy*) pode ser bem interessante para o ocultamento da fonte e proteção dos dados trafegados.

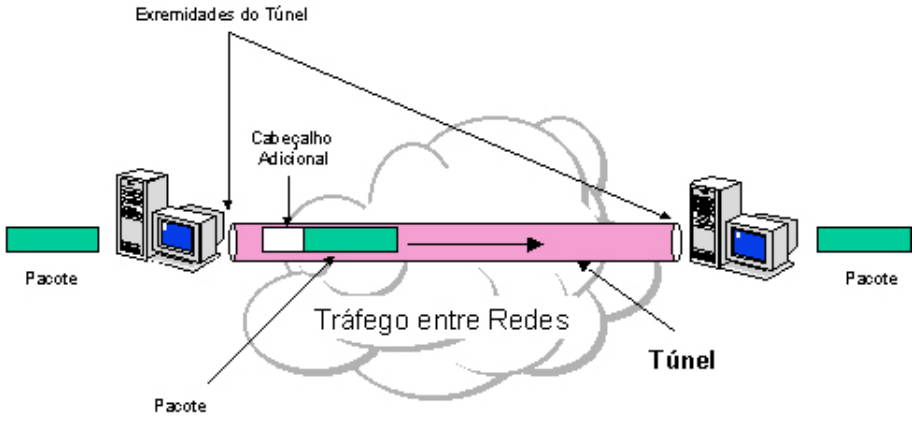

Figura 3. Exemplo de uma VPN.

A Figura 3 mostra a criação do túnel por um protocolo seguro (*i.e.* IPSec), realizada pela VPN e a adição de seu cabeçalho para encaminhamento dos dados sob a rede insegura [12]. Como o cabeçalho deve conter as informações corretas da origem, o IP não pode ser alterado ou oculto, pois, obviamente, impossibilitaria a comunicação de resposta entre as pontas.

OneSwarm

OneSwarm surge como uma proposta P2P onde arquivos podem ser enviados de forma a ocultar a origem, mas mantendo a capacidade de distribuição sem ocultação da fonte (como na proposta Bittorrent), e além destas funcionalidades, é capaz de compartilhar arquivos somente com fontes nas quais o usuário realmente confie [13].

Sua principal motivação parte da visão de que a privacidade na Internet tem se tornado cada vez mais escassa e que protocolos como Bittorrent podem ser facilmente monitorados por terceiros. Sendo assim, é proposto o OneSwarm como uma forma de auxiliar a privacidade e manter um desempenho razoável com relação aos protocolos P2P que hoje não utilizam nenhuma técnica de ocultamento da fonte.

Como nesta rede o consumidor (aquele que recebe um arquivo ou o pedaço de um) é também um produtor (fonte de envio), é possível que ao receber um arquivo o nó altere as configurações de privacidade ligadas àquele arquivo. Sendo assim, apresenta-se a seguir três modos de compartilhamento:

- **Distribuição pública:**
  - da mesma forma como Bittorrent, porém pode ser redistribuído por qualquer outra política.
- **Com permissões:**

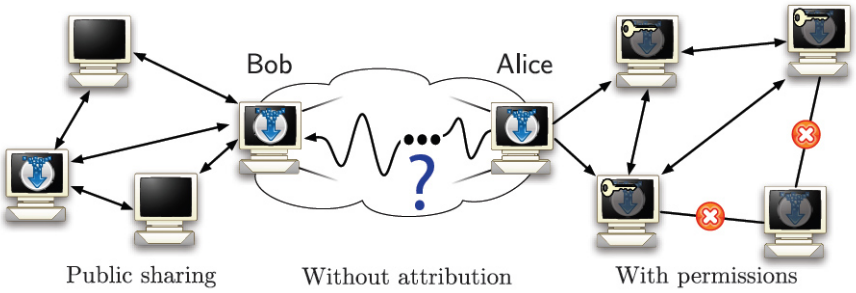

Figura 4. Exemplo das opções compartilhamento para OneSwarm.

  - acesso restrito somente àqueles que se define. e.g. distribuição de fotos de casamento.
- **Sem atribuição:**
  - oculta a fonte ou destino. Encaminha por diversos intermediários para realizar este ocultamento.

Concentrando na parte de ocultação de fonte ou destino, percebe-se que esta técnica é muito parecida com as que serão apresentadas na Seção 3.2.

No OneSwarm a confiança entre os nós que realizam o encaminhamento de forma que a fonte ou o destino sejam ocultados, é realizada pela verificação de chaves RSA 1024 compartilhadas entre os nós. Neste caso a identificação dos nós é somente sua chave compartilhada e não mais o endereço IP. Conforme a Figura 4, após obter o arquivo, Bob pode replicá-lo utilizando a ocultação de fonte ou destino (*without attribution*), e isto é possível por meio de um encaminhamento adicional sobre uma rede *overlay* de nós intermediários. Esta rede *overlay* funciona como a proposta Mix [4], sobrescrevendo o endereço de origem pelos endereços intermediários dos nós que encaminham a mensagem, desta forma, os nós sabem apenas qual o próximo salto, mas não a origem real nem o destino final da mensagem.

FreeNet

Outra solução P2P que fornece possibilidade de anonimato na Internet é a FreeNet [5]. Sua motivação é a manutenção da liberdade de expressão na Internet. Tanto a comunicação quanto a informação armazenada em cada nó é criptografada a fim de manter seu princípio de privacidade. Outra característica relevante desta rede é que cada usuário deve dispor de um espaço para armazenamento de dados em seu computador, com o intuito de replicar e acelerar o acesso a um dado qualquer. Como os dados são armazenados de forma criptografada, o usuário que armazena aquele dado não sabe exatamente qual conteúdo tem em sua máquina neste espaço compartilhado. É interessante também comentar sobre a rede oculta da FreeNet, onde é possível até mesmo utilizar dos serviços HTTP (*Hypertext Transfer Protocol*) que existem nos nós desta rede. Neste caso, a FreeNet não funciona como um *proxy*, permitindo acesso somente aos nós desta rede oculta.

Para manter a resiliência desta proposta, os arquivos são distribuídos e armazenados em diversos nós de forma distribuída, a idéia é que ninguém, além daqueles que

realmente podem ter acesso a um arquivo o tenham.

Após um arquivo ser inserido por um usuário na rede FreeNet, este pode se desconectar, pois esse arquivo já estará distribuído e replicado pelos nós da rede, provendo resiliência à proposta e anonimato àquele que lançou o arquivo em questão.

O roteamento nesta rede é realizado pelo *Key-based Routing* (KBR), que é um método de busca usado em conjunto com as *Distributed Hash Tables* (DHTs). Enquanto as DHTs proveem um método de encontrar um nó que tenha algum pedaço (*chunck*) de dado de um determinado arquivo, KBR provê um método de encontrar o nó mais próximo de acordo com uma métrica (*e.g.* número de saltos, latência).

FreeNet possui dois métodos funcionais, são eles:

- **Opennet:** onde todos os nós da rede FreeNet se comunicam sem restrição de confiança.
- **Darknet:** onde somente nós que compartilham chaves podem se comunicar, incrementando a confiança na relação destes e impedindo possíveis atacantes de resgatar algum arquivo nesta rede.

Um ponto a que se deve destacar nesta rede é que, por se tratar de uma rede paralela, é muito comum que usuários a utilizem para troca de arquivos pessoais de cunho ilegal.

3.2. Baseado em Proxy

O *proxy* pode funcionar como uma solução que oculta o endereço IP real de um *host*. Como exemplo a Figura 5 mostra um usuário sobre o endereço IP 1.1.1.1 que utiliza o *proxy* 2.2.2.2 para acessar a um destino. Esse destino por sua vez acredita que o acesso foi realmente requisitado pelo *host* de IP 2.2.2.2, o que oculta a fonte 1.1.1.1.

Figura 5. Funcionamento de um Proxy.

Questões relativas à real ocultação da fonte neste modo podem ser levantadas. Apenas um nó funcionando como *proxy* pode se mostrar como algo muito simples e de fácil controle. Poderia este *proxy* coletar todas as informações de origem e destino, o que quebraria a ocultação da fonte e também permitiria uma análise do tráfego em questão.

Com o intuito de fortalecer o conceito de ocultação da fonte por meio de *proxy*, várias propostas forma introduzidas. Nessas propostas o roteamento entre uma origem e o destino desejado são realizados em cascata por *proxies* intermediários, incrementando

o anonimato. São essas propostas que serão apresentadas nesta seção, categorizadas em *Onion Routing* (Roteamento Cebola) e uma nova proposta baseada neste, chamada de *Garlic Routing* (Roteamento Alho). Ambas as propostas criam camadas adicionais a cada salto, de forma que o pacote original fique oculto enquanto trafega pela rede.

3.2.1. Onion Routing

Na proposta do *Onion Routing* [15] o pacote é encaminhado por nós intermediários que realizam o papel de *proxies* em cascata, no mesmo esquema Mix [4], utilizando criptografia para ocultar os saltos anteriores. Este processo é realizado a partir da sobreposição de camadas (como em uma cebola), exposto pela Figura 6. Os nós intermediários apenas sabem o salto anterior e qual deve ser seu próximo salto.

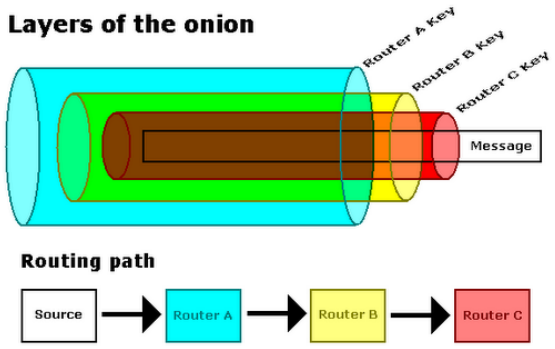

Figura 6. As camadas do Onion Routing.

Porém, uma vulnerabilidade existe no nó final (aquele que entrega o pacote ao destino), neste nó os pacotes passam em sua forma convencional, ou seja, geralmente em texto-puro. Para evitar esta vulnerablidade, é possível realizar algum incremento de segurança (criptografia), como por exemplo, utilizando uma VPN.

Após a conceituação desse roteamento, pode-se introduzir as principais soluções que utilizam esta técnica.

JAP

O JAP (Java Anon Proxy), também chamado de JonDonym ou JonDo, é um clássico do *onion routing*. Seu ambiente é composto pelos chamados Mix (proxies intermediários) e utiliza o Mix Cascade para o efetivo roteamento dos dados. Como é possível observar na Figura 7, caso o usuário não utilize o JAP, ele irá expor seu IP diretamente ao destino, porém utilizando o JAP para sua aplicação, ele tem a possibilidade de ocultar sua fonte.

Perceba que no JAP os Mix são fixos, e ditos como confiáveis, pois não são nós comuns participantes da rede, mas nós eleitos como merecedores de tal confiança. A ideia é manter o roteamento em camadas (*onion routing*) utilizando esses proxies (Mix) e fazer com que muitos usuários saiam com o mesmo endereço IP, impossibilitando o destino de saber qual o real IP da origem e ao mesmo tempo protegendo a análise de tráfego pelo nó (Mix) final uma vez que este é confiável.

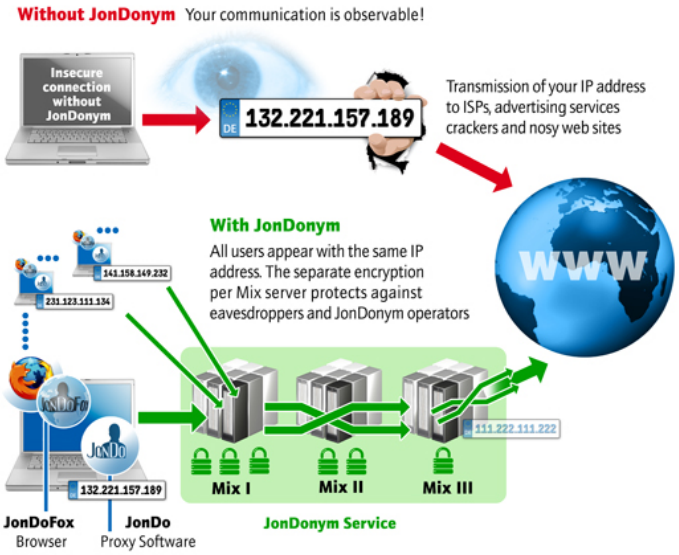

Figura 7. Exemplo de ambiente com e sem JAP.

Um problema desta proposta é exatamente a utilização permanente dos mesmos nós como roteadores intermediários, o que possibilita a detecção do tráfego por uma fonte externa e pode gerar algum tipo de prejuízo no desempenho e segurança em geral.

Tor

Outro sistema extremamente difundido é o Tor [8]. Este utiliza o conceito de Mix Cascade [4] e TCP para o roteamento oculto pelos nós (*proxies*) intermediários. Esses nós intermediários não são fixos e aleatoriamente são escolhidos pelo protocolo do Tor, o que incrementa uma possível falha existente no JAP.

Conforme a Figura 8, Alice deseja se comunicar com Bob e, para tanto, o protocolo do Tor escolhe os nós que serão os saltos intermediários de Alice até seu destino, Bob. Utilizando a técnica de Mix Cascade, os nós intermediários não sabem nem o conteúdo nem a origem e destino real do pacote de Alice. Porém, o roteador final encaminha sem criptografia o pacote de Alice a Bob, o que pode gerar uma falha de segurança considerável. Essa falha pode existir porque a rede é colaborativa, e os nós podem dizer se são apenas clientes, encaminhadores (*relay*) ou ainda um nó de saída (*out-proxy*). Para solucionar este problema, somente com a utilização de uma criptografia em uma camada superior, a fim de que somente, realmente o destino consiga ler o conteúdo do pacote.

Outro problema considerável é que protocolos *Onion Routing* do tipo Tor devem realizar consultas a DNS (*Domain Name System*) antes de enviar seu pacote pela rede caso necessitem resolver um nome de um *host*. Com isso um monitor pode capturar estes pacotes e casar com a informação enviada posteriormente pelo nó Tor para a rede. Porém, este problema pode ser tratado utilizando um *proxy* anterior à resolução de nomes, o que ocultaria todas as possibilidades de detecção da fonte real.

Tor também se utiliza de um nó chamado de diretório central, que mantém a lista de todos os roteadores e se mostra como um ponto confiável para a rede. A identificação

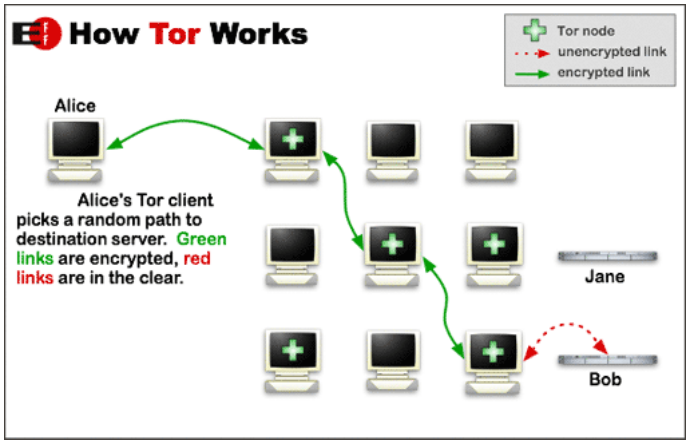

Figura 8. Exemplo de roteamento no Tor.

dos nós é feita por uma chave temporária conforme o *onion routing* e uma outra chave persistente (baseada em certificados TLS - *Transport Layer Security*) de longa duração para identificação geral na rede (armazenados no diretório central), além da associação de uma descrição da capacidade daquele nó.

3.2.2. Garlic Routing

Uma proposta posterior ao *Onion Routing* que visa complementá-lo é o *Garlic Routing*, citado pela primeira vez em [9]. O *Garlic Routing* tem como principal diferença com relação ao *Onion Routing* a possibilidade de agregação de várias mensagens nos roteadores intermediários. Além disso, as mensagens (”dentes” do alho) podem ter opções arbitrárias, tais como a solicitação de inserção de um atraso arbitrário em algum nó, para dificultar o acompanhamento das mensagens trocadas na rede. É possível também incluir tamanhos extras para enganar qualquer observador. Essas modificações dificultam uma possível análise de tráfego.

I2P

A solução baseada em *Garlic Routing* mais difundida atualmente é o I2P (*Invisible Internet Project*) [10, 16]. Sua funcionalidade de agregação de mensagens e distribuição na rede por caminhos alternados confunde um possível atacante. Os roteadores desta rede são identificados por um identificador de persistente, indicado por um ID criptográfico (chaves públicas) [3]. Além desta característica de agregação, o que mais difere o I2P do Tor, por exemplo, é a utilização de múltiplos túneis. Esses túneis são chamados de *inbound* e *outbound*.

- ***Inbound tunnel*:** por onde as mensagens chegam ao usuário/roteador.
- ***Outbound tunnel*:** por onde as mensagens saem do usuário/roteador.

Na Figura 9 os usuários Alice, Bob, Charlie e Dave se comunicam por meio de um roteador I2P. Cada participante tem dois possíveis túneis de entrada (*inbound tunnel* 1, 2, 3, 4, 5 e 6) e dois túneis de saída (*outbound tunnel* - em vermelho). Na imagem os

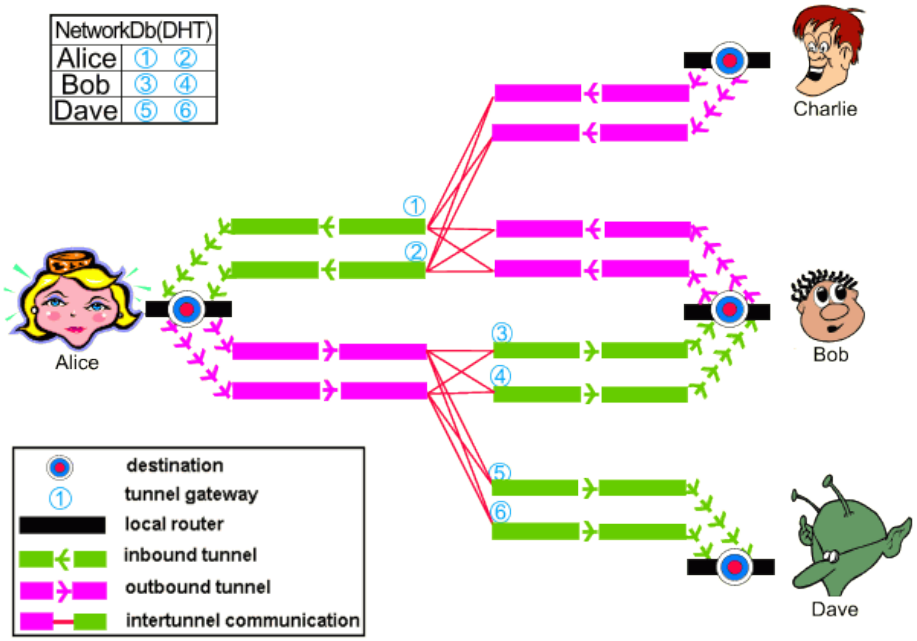

Figura 9. Exemplo de roteamento no I2P.

túneis de entrada de Charlie e os de saída Dave foram suprimidos da imagem por questão de espaço. Sendo assim, quando, por exemplo, Alice deseja enviar uma mensagem a Bob, ela utiliza seus túneis de saída (*outbound*), que são aleatoriamente encaminhados pelos túneis de *gateway* 3 ou 4. Já a resposta de Bob a Alice é realizado de forma contrária utilizando os túneis de saída de Bob (*outbound*) e entrada de Alice (*inbound* - 1 ou 2).

Desta forma concluímos as principais ferramentas e protocolos existentes para anonimato na Internet.

4. Conclusão

Este trabalho apresentou as principais técnicas de anonimato na Internet e as classificou. Todos os sistemas apresentados por este trabalho como forma de anonimato na Internet acrescem a latência e o *overhead* de processamento. Por mais que se espere soluções leves, não há como realizar operações de roteamento intermediário e criptografia sem inserir *overhead*. Como pode ser visto em [7].

Um trabalho interessante com relação a usabilidade de *softwares* que visam prover anonimato na Internet pode ser visto em [1]. Neste trabalho são destacados os principais pontos com relação à instalação, *interface*, documentação e usabilidade das ferramentas pelo usuário.

O questionamento com relação ao intuito de utilização destas ferramentas sempre existirá. Soluções de anonimato na Internet atual têm grande valia, mas continuam a auxiliar àqueles que desejam apenas ocultar coisas contrárias ao bem comum. Por conta disso, diversos trabalhos estudam como encontrar usuários em redes sobrepostas e até

mesmo usuários sob cada uma das técnicas aqui apresentadas. Neste momento o foco deste trabalho não passa por essas características, porém é algo interessante a ser estudado (*e.g.* rastreamento de pacotes, *tracking* etc.).

Referências

[1] ABOU-TAIR, D. E. D. I., PIMENIDIS, L., SCHOMBURG, J., AND WESTERMANN, B. Usability inspection of anonymity networks. In *Proceedings of the 2009 World Congress on Privacy, Security, Trust and the Management of e-Business* (Washington, DC, USA, 2009), CONGRESS '09, IEEE Computer Society, pp. 100–109.

[2] BELLOVIN, S. M. Security problems in the tcp/ip protocol suite. *COMPUTER COMMUNICATIONS REVIEW 19*, 2 (1989), 32–48.

[3] CATALANO, D., DI RAIMONDO, M., FIORE, D., GENNARO, R., AND PUGLISI, O. Fully non-interactive onion routing with forward-secrecy. In *Proceedings of the 9th international conference on Applied cryptography and network security* (Berlin, Heidelberg, 2011), ACNS'11, Springer-Verlag, pp. 255–273.

[4] CHAUM, D. L. Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM 24*, 2 (Feb. 1981), 84–90.

[5] CLARKE, I., SANDBERG, O., WILEY, B., AND HONG, T. W. Freenet: a distributed anonymous information storage and retrieval system. In *International workshop on Designing privacy enhancing technologies: design issues in anonymity and unobservability* (New York, NY, USA, 2001), Springer-Verlag New York, Inc., pp. 46–66.

[6] DE VIVO, M., DE VIVO, G. O., KOENEKE, R., AND ISERN, G. Internet vulnerabilities related to tcp/ip and t/tcp. *SIGCOMM Comput. Commun. Rev. 29*, 1 (Jan. 1999), 81–85.

[7] DHUNGEL, P., STEINER, M., RIMAC, I., HILT, V., AND ROSS, K. W. Waiting for anonymity: Understanding delays in the tor overlay. In *Peer-to-Peer Computing* (2010), IEEE, pp. 1–4.

[8] DINGLEDINE, R., MATHEWSON, N., AND SYVERSON, P. Tor: the second-generation onion router. In *Proceedings of the 13th conference on USENIX Security Symposium - Volume 13* (Berkeley, CA, USA, 2004), SSYM'04, USENIX Association, pp. 21–21.

[9] DINGLEDINE, R. R. The free haven project: Design and deployment of an anonymous secure data haven. In *MASTERS THESIS, MIT* (2000).

[10] I2P. *Invisible Internet Project*, 2012.

[11] INTERNET ENGINEERING TASK FORCE. *RFC 791 Internet Protocol - DARPA Inernet Programm, Protocol Specification*, September 1981.

[12] INTERNET ENGINEERING TASK FORCE. *RFC 2685 Virtual Private Networks Identifier*, September 1999.

[13] ISDAL, T., PIATEK, M., KRISHNAMURTHY, A., AND ANDERSON, T. Privacy-preserving p2p data sharing with oneswarm. *SIGCOMM Comput. Commun. Rev. 41*, 4 (Aug. 2010), –.

[14] NORRIS, E. Analysis of a telnet session hijack via spoofed mac addresses and session resynchronization, 2001.

[15] REED, M. G., SYVERSON, P. F., AND GOLDSCHLAG, D. M. Anonymous connections and onion routing. *IEEE Journal on Selected Areas in Communications 16* (1998), 482–494.

[16] TIMPANARO, J. P., CHRISMENT, I., AND FESTOR, O. I2p's usage characterization. In *Proceedings of the 4th international conference on Traffic Monitoring and Analysis* (Berlin, Heidelberg, 2012), TMA'12, Springer-Verlag, pp. 48–51.