Réseaux de communication anonyme

Sébastien Gambs

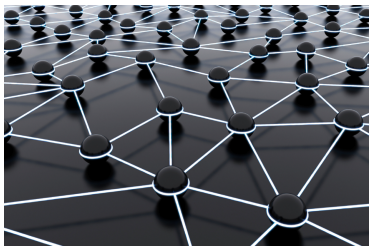
sgambs@irisa.fr

20 novembre 2015

Introduction

Réseau de communication anonyme

- Modèle : chaque noeud dans le réseau est un individu qui peut potentiellement être l'envoyeur ou le receveur d'un message.
- Exemple : chaque ordinateur connecté à Internet est un noeud de ce réseau possédant un identifiant unique (l'adresse IP).
- Réseau de communication anonyme : technologie de protection de la vie privée (PET) permettant de communiquer de manière anonyme dans un réseau, c'est à dire en protégeant l'identité de l'envoyeur et/ou du receveur du message.



- Défi principal : réussir à créer un tel réseau au dessus d'un réseau qui est au départ non-anonyme.

Propriétés d'anonymat

- Anonymat de l'envoyeur et/ou du receveur d'un message : cacher l'identité de l'envoyeur et/ou du receveur d'un message parmi un groupe (ou l'ensemble de la population).
- Exemples de mesure d'anonymat :
  - taille du groupe d'anonymat dans lequel l'individu est caché.
  - incertitude moyenne sur qui est l'auteur d'un message mesuré par une métrique du type entropie.
- Non-chaînabilité : être incapable de relier deux messages à la même identité.
- Remarque : peut concerner deux messages successifs du même individu ou le même message aperçu à deux endroits.
- Non-observabilité : ne pas pouvoir détecter si un message est en train de circuler ou non.
- Remarque: non-observabilité $\Rightarrow$ non-chaînabilité $\Rightarrow$ anonymité

Modèles d'adversaire

- Honnête-mais-curieux (ou *adversaire passif*) :
  - L'adversaire contrôle un ou plusieurs noeuds dans le réseau.
  - Ces noeuds suivent les règles du protocole (pas de triche active) mais enregistrent tous les messages qu'ils voient.
  - Un protocole est sécuritaire dans ce modèle si un noeud (ou une collusion de noeuds) ne peut pas briser une des propriétés d'anonymat malgré les informations enregistrées.
- Malveillant (ou *adversaire actif*) :
  - Les noeuds corrompus par l'adversaire peuvent tricher activement en ne suivant pas exactement les règles du protocole.
  - But possible : apprendre l'entrée des participants honnêtes, biaiser la sortie du protocole ou simplement le faire avorter.
- Espion (local ou global) : ne contrôle aucun noeud du réseau mais à la capacité d'écouter certains liens de communication.

Types de sécurité

- Sécurité calculatoire :
  - Repose sur l'impossibilité pour un adversaire de briser une hypothèse cryptographique en un temps raisonnable pour lui.
  - Exemples : difficulté de factoriser le produit de deux grands nombres premiers, logarithme discret.
  - Si l'adversaire possèdait une puissance de calcul illimitée, il peut briser cette hypothèse par une recherche exhaustive.
- Sécurité au sens de la théorie de l'information :
  - Pas d'hypothèse cryptographique mais plutôt l'impossibilité pour un adversaire de corréler les observations qu'il perçoit avec une entrée spécifique.
  - Plusieurs hypothèses sont possibles et l'adversaire ne dispose d'aucune information pour distinguer laquelle est vraie.
  - Meilleure stratégie pour l'adversaire : deviner une possibilité en choisissant aléatoirement.

Le dîner des cryptographes

Problème du dîner des cryptographes

- Problème introduit par Chaum en 1988.
- Trois cryptographes (Alice, Bob et Charlie) sont assis autour d'une table ronde à l'occasion d'un dîner.
- Le serveur les informe que quelqu'un a payé l'addition du repas.
- Promesse :
  1. soit c'est un (et un seul) des cryptographes qui a payé (mais il ne souhaite pas le révéler publiquement par modestie),
  2. soit c'est la NSA.
- Problème : comment les cryptographes peuvent-ils décider de manière anonyme quelle possibilité est vraie?
- Peut être vu comme une implémentation d'un canal de diffusion (*broadcast*) anonyme.

Modèle

- Chaque cryptographe est un noeud dans un réseau circulaire (en anneau).
- L'entrée de chaque cryptographe est un bit (respectivement a, b et c) qui vaut :
  - 1 si le cryptographe correspondant à ce bit a payé l'addition et
  - 0 sinon.
- De part la promesse, on a :
  1. soit $a \oplus b \oplus c = 1$ (si c'est un des cryptographes qui a réglé l'addition),
  2. soit $a \oplus b \oplus c = 0$ (si c'est la NSA qui a payé).
- But du protocole : calculer le OU-exclusif (XOR) de tous les bits d'entrée de manière anonyme.
- Modèle d'adversaire : honnête-mais-curieux.

Protocole du dîner des cryptographes (DC-net)

Première phase :

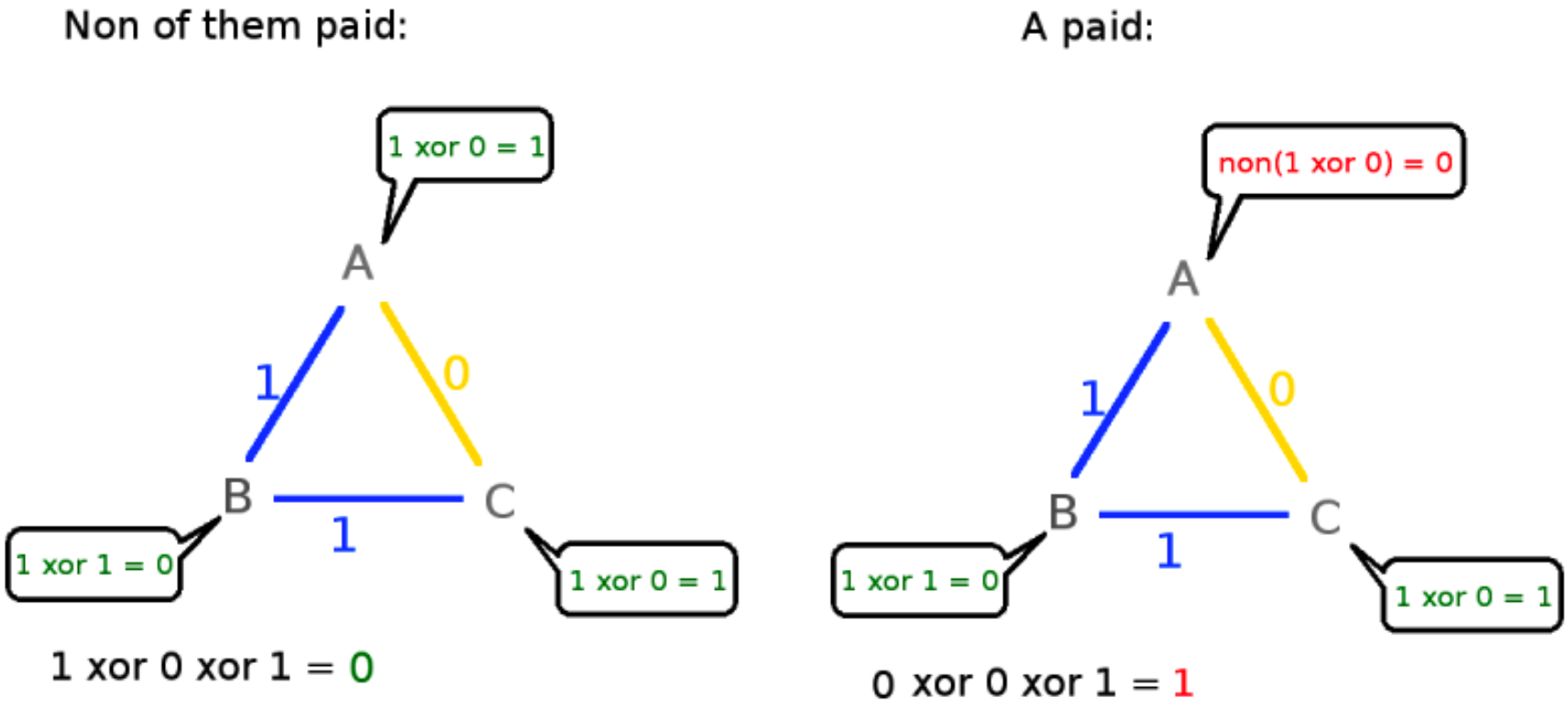
- Chaque participant génère un bit secret aléatoire avec chacun de ses voisins (par exemple en tirant à pile-ou-face).
- Soit r_{AB},r_{AC} et r_{BC}, les bits secrets aléatoires.
- Exemple : r_{AB} est un bit secret connu seulement de Alice et Bob.

Deuxième phase :

- Chaque participant annonce publiquement un bit qui est le ou-exclusif de son bit d'entrée (a, b ou c) et des bits secrets qu'il connait.
- Exemple : Alice annonce publiquement un bit qui est égal à $a \oplus r_{AB} \oplus r_{AC}$.

Sortie du protocole : calculée en faisant le ou-exclusif de tous les bits publiques annoncés.

Illustration du DC-net



Analyse de la sortie

Si on suppose qu'il n'y pas de collision possible alors la sortie du protocole est :

- 1 si un des cryptographes a payé,
- 0 sinon c'est la NSA.

Preuve :

- La sortie du protocole est $(a \oplus r_{AB} \oplus r_{AC}) \oplus (b \oplus r_{AB} \oplus r_{BC}) \oplus (c \oplus r_{AC} \oplus r_{BC})$,
- soit $a \oplus b \oplus c \oplus r_{AB} \oplus r_{AB} \oplus r_{AC} \oplus r_{AC} \oplus r_{AB} \oplus r_{BC} \oplus r_{BC}$
- qui se simplifie en $a \oplus b \oplus c$.
- Or $a \oplus b \oplus c = 1$ si un des cryptographes a payé et 0 sinon.

Anonymat

- Si on suppose que l'adversaire contrôle au maximum un joueur alors celui-ci ne peut rien apprendre sur l'entrée des participants ou sur l'identité de l'éventuel cryptographe qui a réglé la note.
- Preuve : chaque cryptographe anonce publiquement une valeur qui est le ou-exclusif de son entrée plus plusieurs bits aléatoires dont au moins un est inconnu de l'adversaire
  $\Rightarrow$ la vue de l'adversaire est constitué de bits randomisés
  $\Rightarrow$ sécurité au sens de la théorie de l'information
- Remarque : le protocole peut être facilement généralisé à n participants.
- Attaque des voisins : si un noeud est entouré de 2 noeuds en collusion ceux-ci peuvent apprendre son bit d'entrée.

Limitations du DC-net original

- Collision : si deux cryptographes (ou un nombre pair) payent chacun le dîner leur bits vont s'annuler et la sortie du protocole sera 0.
- Nécessité d'avoir un mécanisme pour gérer les collisions.
- Perturbation : si un des participants sait à priori qu'un autre participant a payé il peut néanmoins perturber le protocole en mettant 1 comme bit d'entrée.
- Même effet qu'une collision.
- Solution possible : Dissent est un réseau de communication anonyme actuellement en développement basé sur le DC-net mais qui essaye d'éliminer ses faiblesses.

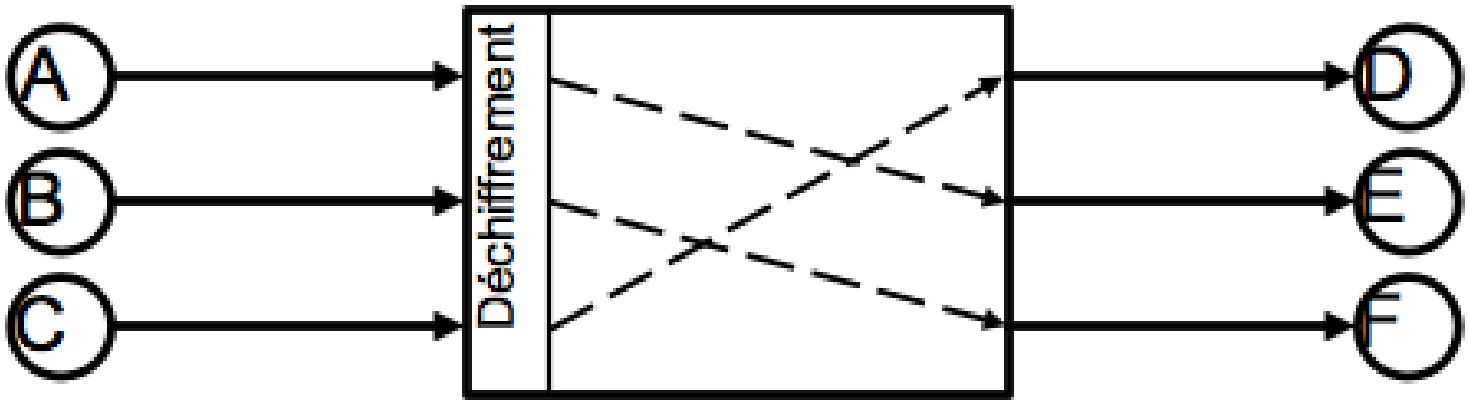
Mixnets

Mix

- Concept introduit par Chaum en 1981 pour empêcher l'analyse de trafic.
- Modèle d'adversaire : espion observant les communications échangés.
- Mix : routeur qui cache le lien entre les messages entrants et sortants par un mécanisme de chiffrement et de permutation des messages.
- Exemple d'application : service de courriel anonyme (Mixmaster).

Fonctionnement d'un Mix simple

1. Reçoit en entrée plusieurs paires du type (*message*, *adresse_du_destinataire*) qui ont été préalablement chiffrées.
2. Déchiffre les messages.
3. Envoie en sortie les messages à leurs destinataires correspondants (possiblement chiffrés).

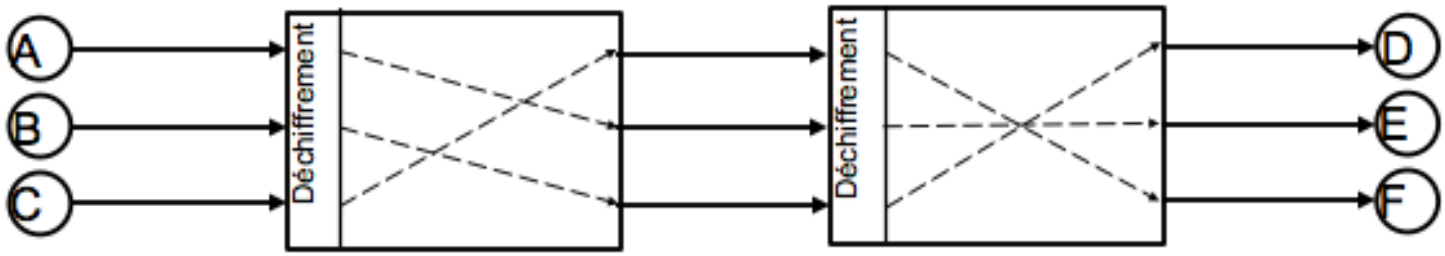


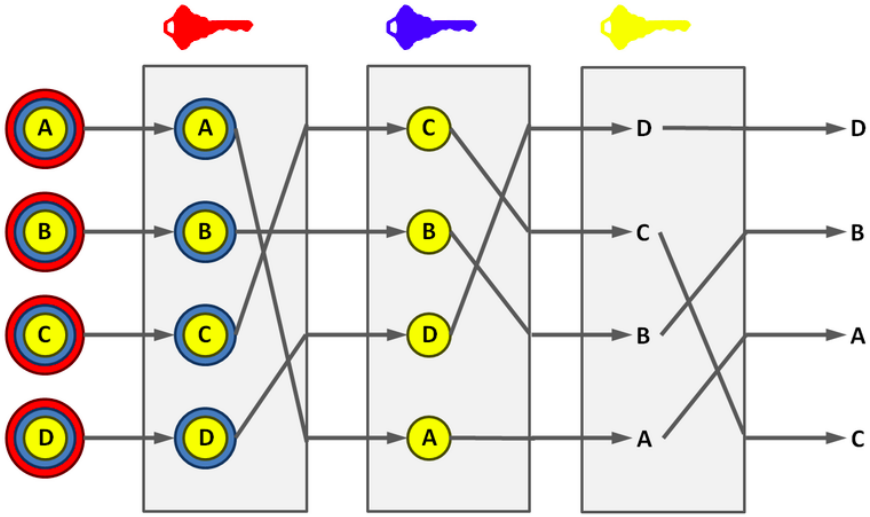
Non-chaînabilité

- Non-chaînabilité : impossibilité pour l'adversaire de faire le lien entre les messages entrants et sortants (sauf s'il contrôle le mix lui-même).
- Possibilité de combiner plusieurs mixs à la suite pour compliquer la tâche de l'adversaire.
- Exemple de deux mixs :
  - Le premier mix peut identifier une communication entre l'envoyeur du message et le second Mix,
  - le second mix peut identifier une communication entre le premier Mix et le receveur,
  - mais aucun des deux mixs ne peut établir de lien entre l'envoyeur et le receveur (sauf collaboration).

Cascade de mixs

- Cascade de mixs : enchaînement successif de plusieurs mixs.





Tor

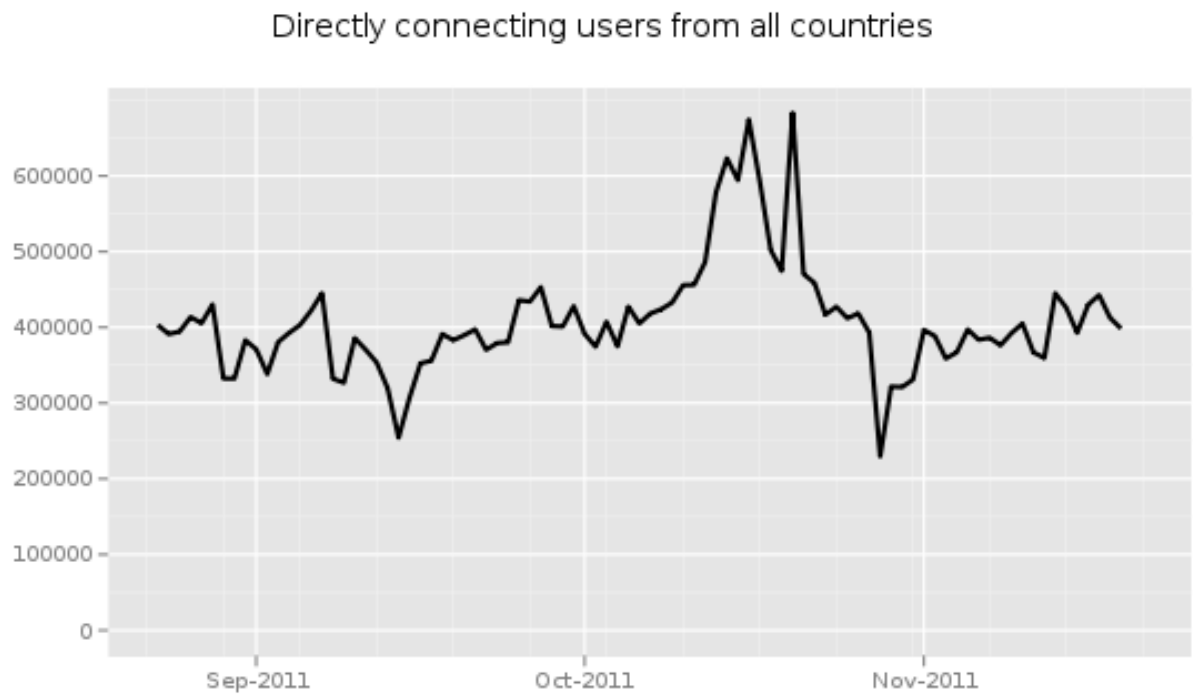
Tor

- Tor (*The Onion Router*) : réseau mondial de communication anonyme organisé en couche autour de routeurs (environ 2500 à l'heure actuelle) qui jouent le rôle de noeud.

- Permet d'anonymiser tout type de communication fait sur Internet.
- Exemples : sites webs visités, messagerie instantanée, courriel.
- Logiciel libre et l'utilisation du réseau est gratuit.
- Implémente un routage de type oignon (seconde génération) pour garantir l'anonymité.

Statistiques d'utilisation de Tor

Daily directly connecting users:



The Tor Project - https://metrics.torproject.org/

Start date (yyyy-mm-dd): 2011-08-23 End date (yyyy-mm-dd): 2011-11-21

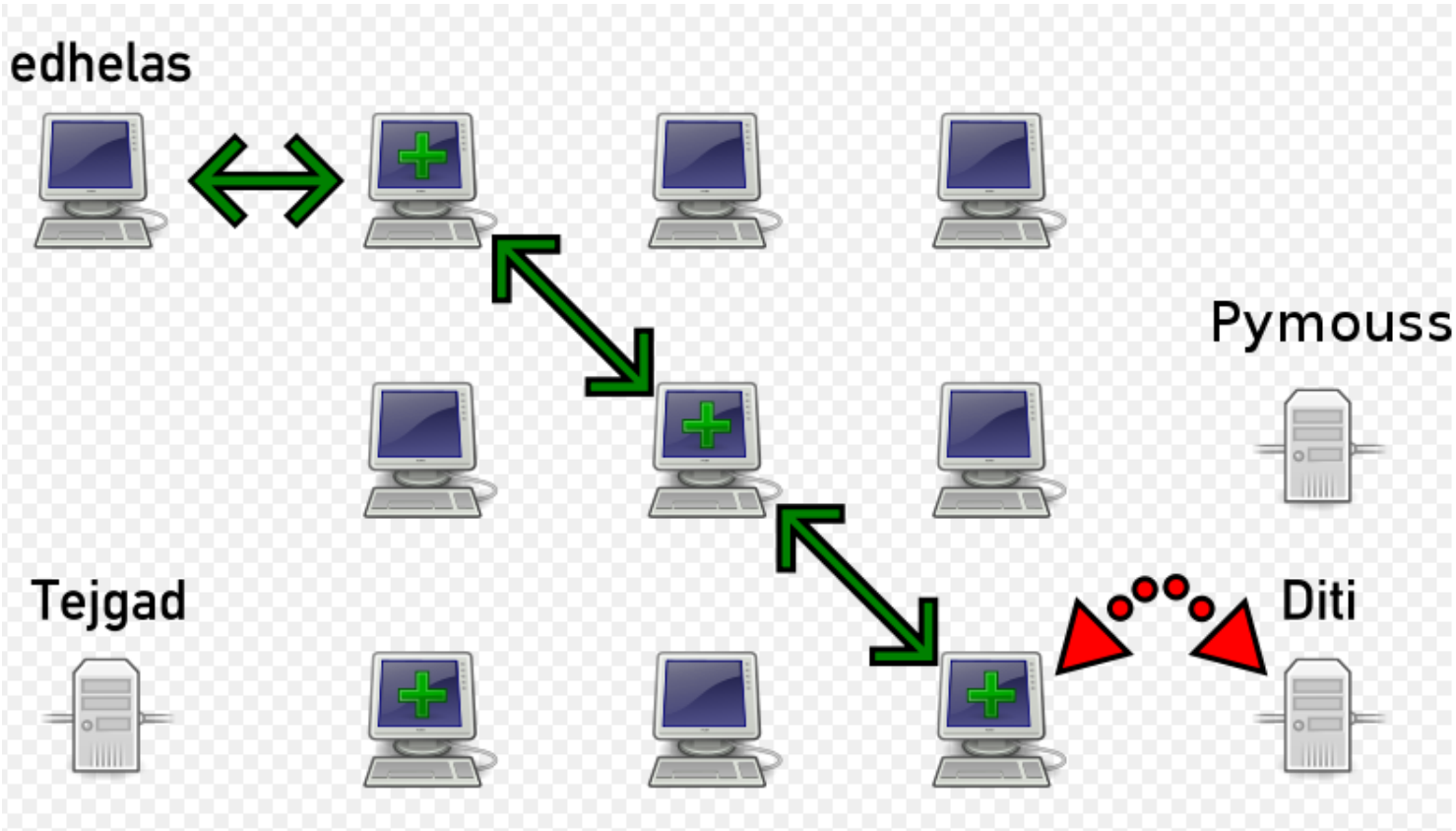
Principe du routage de type oignon

- Le message qui doit être envoyé est chiffré de manière successive par couche.
- Chaque couche est chiffré avec la clé publique d'un routeur oignon présent sur le chemin emprunté par le message (le chemin est décidé a priori par l'envoyeur).
- Lorsqu'un routeur oignon reçoit un message, il "épluche" la couche courante en la déchiffrant avec sa clé secrète.
- Ceci révèle les instructions de routage pour savoir à quel prochain routeur oignon envoyé le message résultant.
- Garantie : protège l'anonymat de l'envoyeur et de receveur du message, ainsi que son contenu.

Construction d'un chemin de routage

- Hypothèse : l'utilisateur connaît la liste des routeurs oignons et de leurs positions dans le réseau.
- L'utilisateur choisit un chemin aléatoire dans le réseau le reliant au noeud avec qui il souhaite communiquer.
- L'utilisateur construit ensuite un "circuit" correspondant à ce chemin tel que dans le circuit chaque noeud connaît seulement l'identité de son successeur et de son prédecesseur.
- Il est possible d'associer un identifiant unique et aléatoire (pseudonyme) au circuit pour être capable de chaîner différents messages envoyés.
- Amélioration de l'efficacité : lors de la construction on peut distribuer des clés secrètes (symmétriques) aux noeuds sur le chemin en les chiffrant avec leurs clés publiques pour sécuriser les liens.

Illustration de la construction du chemin de routage



Chiffrement et déchiffrement d'un message

- Soit un "circuit" composé de n noeuds.
- Chiffrement d'un message :
  1. L'utilisateur chiffre le message une première fois avec la clé publique du noeud n.
  2. Le résultat est chiffré une deuxième fois avec la clé publique du noeud $n-1$.

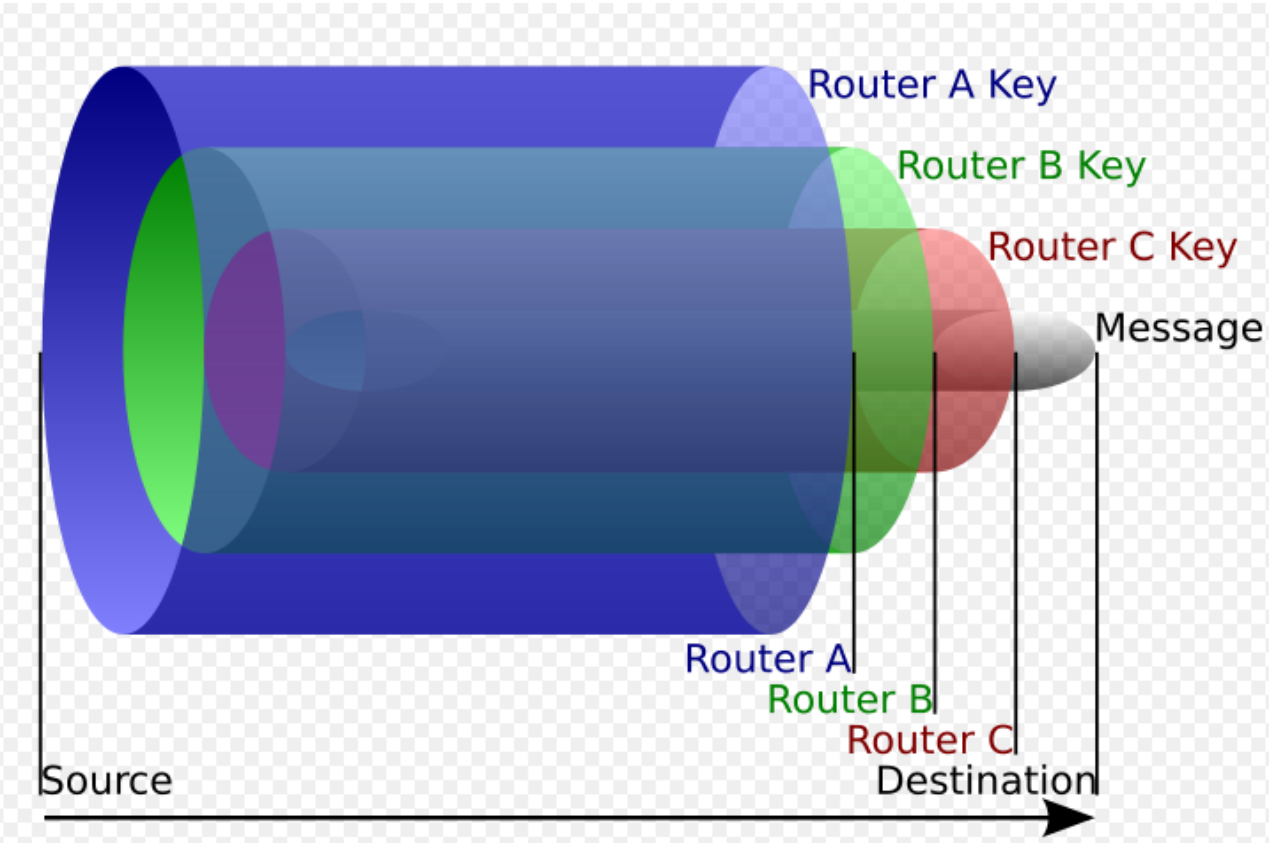
     ...

  N. La dernière fois, le résultat des chiffrements précédents est chiffré avec la clé du noeud 1 sur le chemin.
- Déchiffrement d'un message :
  1. Le noeud 1 déchiffre la première couche et l'envoie au noeud 2.
  2. Le noeud 2 pèle la deuxième couche et l'envoie au noeud 3.

     ...

  N. Le dernier noeud déchiffre le dernier message avec sa clé et récupère le message original.

Exemple “d’oignon”



Anonymat

- Remarque : seul le premier noeud connaît l'adresse IP de l'envoyeur du message.
- Garantie d'anonymat : pour briser l'anonymat de l'envoyeur et du receveur d'un message, l'adversaire doit contrôler tous les noeuds sur le chemin.
- Chemin long $\Rightarrow$ faible probabilité que tous les noeuds du chemin soit corrompus $\Rightarrow$ anonymat plus fort
- Question ouverte : comment implémenteriez vous la propriété de non-observabilité dans un tel réseau?

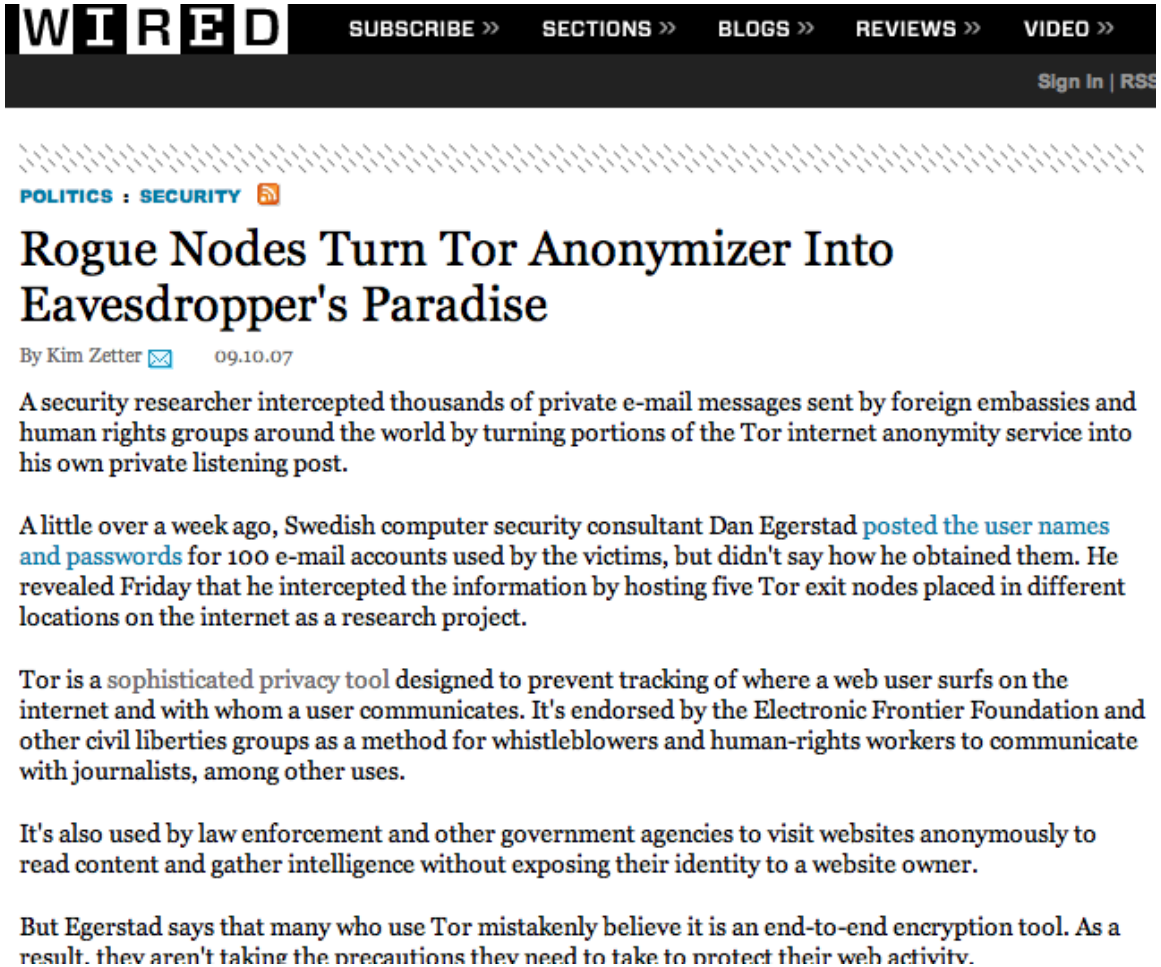
Attaque basée sur le timing

- Attaque basée sur le timing : un espion qui peut écouter les liens de communication sur les points entrants et sortant d'un réseau peut essayer de d'utiliser le timing d'arrivée des paquets échangés pour essayer de chaîner des entrées et sorties.
- Contre-mesure possible : introduction de "faux" paquets.

Empreinte de site web

- Supposons un adversaire capable d'écouter un lien de communication traversé par un chemin sur un réseau de communication anonyme.
- Attaque par empreinte de site web : l'adversaire peut essayer de déduire quel site web est visité en regardant la taille des paquets qui passent et l'intervalle de temps entre les paquets.
- Exemple : un site web contenant de nombreuses images haute résolution aura une signature différente d'un site contenant uniquement du texte.
- Permet d'attaquer non pas directement la propriété d'anonymat mais plutôt de révéler le contenu du message (le site web visité)
  $\Rightarrow$ révèle les intérêts de la personne qui consulte le site web (et donc indirectement son identité)

Tor protège l'anonymat mais pas la confidentialité . . .



WIRED SUBSCRIBE » SECTIONS » BLOGS » REVIEWS » VIDEO »
Sign In | RSS

POLITICS : SECURITY

Rogue Nodes Turn Tor Anonymizer Into Eavesdropper's Paradise

By Kim Zetter 09.10.07

A security researcher intercepted thousands of private e-mail messages sent by foreign embassies and human rights groups around the world by turning portions of the Tor internet anonymity service into his own private listening post.

A little over a week ago, Swedish computer security consultant Dan Egerstad posted the user names and passwords for 100 e-mail accounts used by the victims, but didn't say how he obtained them. He revealed Friday that he intercepted the information by hosting five Tor exit nodes placed in different locations on the internet as a research project.

Tor is a sophisticated privacy tool designed to prevent tracking of where a web user surfs on the internet and with whom a user communicates. It's endorsed by the Electronic Frontier Foundation and other civil liberties groups as a method for whistleblowers and human-rights workers to communicate with journalists, among other uses.

It's also used by law enforcement and other government agencies to visit websites anonymously to read content and gather intelligence without exposing their identity to a website owner.

But Egerstad says that many who use Tor mistakenly believe it is an end-to-end encryption tool. As a result, they aren't taking the precautions they need to take to protect their web activity.

Services cachés

- Tor permet à un utilisateur de mettre sur pied un service caché.
- Exemple : site web dont on cache l'identité du serveur qui l'héberge et qui permet à un utilisateur de s'y connecter de manière anonyme.
- L'utilisateur doit d'abord configurer son serveur localement puis le réseau Tor peut être utilisé pour pointer de manière anonyme sur ce serveur (adresse *.onion*).
- Permet de mettre sur point des services de rendez-vous anonyme (du type canal bidirectionnel anonyme à la RouletteChat).
- Défaut : long délai de connexion.

Crowds

Crowds

- Crowds : protocole de communication anonyme qui protège l'anonymat de l'envoyeur d'un message en le routant de manière aléatoire vers des groupes d'utilisateurs similaires.
- Idée principale : cacher l'origine d'un message en le faisant se fondre dans la foule.

- Système particulièrement adapté pour les réseaux du type pair-à-pair.

Fonctionnement de Crowds

- Initialisation : chaque nouvel utilisateur s'enregistre en tant que membre d'un groupe (appelé "Crowd") en contactant le responsable du groupe.
- Quand un utilisateur rejoint un groupe, tous les membres du groupe en sont notifiés.
- Le responsable du groupe est aussi chargé de la distribution des clés symmétriques assurant la confidentialité entre paires de noeuds.

Algorithme de transmission

1. Lors de la réception d'un message m destiné au noeud P, le noeud courant tire à pile-ou-face avec une pièce avec un biais $p_f > \frac{1}{2}$.
2. Si le résultat est :
   - *face* alors il choisit un noeud uniformément au hasard parmi le groupe et faire suivre le message m (avec le noeud destination P).
   - *pile* alors il envoie directement le message m au noeud P.
3. (Optionnel) enregistrer P dans le cas où un tunnel sera construit.

Innocence probable

- Innocence probable : l'adversaire est incapable de prédire avec plus de 50% de confiance, le noeud qui est l'initiateur d'un message.
- Chaque noeud apparaît comme ayant pu ou non être l'envoyeur d'un message (et donc il est probablement innocent).
- Anonymat : dépend de la taille du groupe ("crowd") et de la probabilité p_f.
- Plus la probabilité p_f est grande, plus l'anonymat de l'envoyeur est protégé mais aussi plus la longueur moyen du chemin généré sera longue.
- Compromis possible entre le niveau d'anonymat désiré et le temps de transmission.

C'est la fin !

Merci pour votre attention.
Questions?