Faiblesses de l'identification dans les espaces numériques ouverts de partage de contenus : le cas des réseaux pair-à-pair

Thibault Cholez, Guillaume Doyen, Isabelle Chrisment, Olivier Festor, Rida Khatoun

Chapitre 3

Faiblesses de l'identification dans les espaces numériques ouverts de partage de contenus : le cas des réseaux pair-à-pair

3.1. Introduction

Au sein du monde numérique de plus en plus axé sur la diffusion de contenus, les réseaux pair-à-pair (dits aussi *Peer to Peer* ou P2P) sont devenus en quelques années un vecteur majeur pour l'échange d'informations générant jusqu'à plus de la moitié du trafic global de l'Internet [IPO 09] permettant à leurs utilisateurs de partager rapidement, et sans coût d'infrastructure, de grandes quantités de données. Parmi les différentes architectures paires à pair complètement distribuées, les tables de hachage distribuées (DHT) ont prouvé aussi bien en théorie qu'en pratique leur capacité à constituer des systèmes d'information performants. Basés sur l'architecture Kademlia, les réseaux P2P tels que KAD ou la partie décentralisée de BitTorrent (Mainline DHT) regroupent ainsi des dizaines de millions d'utilisateurs.

Ces systèmes pair-à-pair présentent cependant des faiblesses en ce qui concerne l'identification des pairs (c'est-à-dire les instances de clients connectés au réseau) ou l'identification des contenus partagés sur le réseau. En effet, si l'absence de composant central apporte au paradigme P2P ses principaux avantages (passage à l'échelle, robustesse, absence de coûts d'infrastructure), elle constitue également une limite en rendant difficile la vérification des identités des différents pairs participant

---

Chapitre rédigé par Thibault CHOLEZ, Guillaume DOYEN, Olivier FESTOR, Isabelle CHRISMENT, Rida KHATOUN.

au réseau et des différents contenus partagés. Ces problèmes d'identification imparfaite sont à l'origine de nombreux dysfonctionnements des réseaux P2P en faisant de ceux-ci un support éventuel pour les activités malveillantes, notamment en facilitant la diffusion de contenus dangereux. Ainsi, l'attaque Sybil [DOU 02] consiste, pour un même attaquant physique, à insérer un grand nombre de « faux » pairs dans le réseau afin d'en altérer le routage des messages ou le stockage des données. En effet, les pairs étant parfaitement autonomes, certains utilisateurs malveillants peuvent créer de multiples identités et détourner le protocole à leurs propres fins, telles que : la surveillance des échanges [STE 07a, MEM 09, CHO 10], la suppression d'information [STE 07a] ou encore le déni de service distribué [NAO 06, STE 07a, WAN 08]. Si ces vulnérabilités pouvant affecter les tables de hachage distribuées sont d'ores et déjà connues et ont été expérimentées ponctuellement, aucune étude à ce jour ne s'est intéressée à recenser leur exploitation en pratique. D'autre part, l'absence d'identification sûre des différents contenus par leur nom induit un second problème majeur affectant ces réseaux : la pollution des contenus [LIA 06, LOC 10, CHO 10]. Dans ce cadre, la pollution consiste à partager un contenu erroné ou dont le nom ne correspond pas au contenu réel, ce qui pose des problèmes de sécurité (diffusion de contenus illégaux, virus) et dégrade significativement les performances du réseau en gaspillant des ressources.

Ce chapitre présente les résultats des travaux de recherche menés principalement durant le projet ACDA-P2P (Approche collaborative pour la détection d'attaques dans les réseaux P2P) qui fut financé par le Groupement d'intérêt scientifique pour la Surveillance, la Sûreté et la Sécurité des Grands Systèmes (GIS 3SGS), ainsi que certains résultats issus du projet MAPE (*Measurement and Analysis of Peer-to-peer Exchanges for pedocriminality fighting and traffic profiling*) financé par l'Agence nationale de la Recherche sous le contrat ANR-07-TLCOM-24. Ceux-ci portent plus précisément sur la détection et la quantification des problèmes de sécurité induits par l'absence d'identification fiable des pairs et des contenus dans les réseaux P2P. Pour cela, nous avons réalisé des explorations d'un réseau P2P largement déployé afin de recenser précisément l'ensemble des pairs et des contenus partagés les plus populaires. Nous avons ensuite analysé les données obtenues afin de détecter les activités suspectes exploitant la faiblesse de l'identification. Ainsi avons-nous constaté pour la première fois la réalité et l'ampleur de certaines attaques publiées et pouvons-nous estimer leur nombre au sein du réseau. Concernant les contenus, nous détectons si ceux-ci sont pollués en étudiant la disparité lexicale des noms proposés par les différents pairs et nous appliquons cette détection afin de quantifier la pollution à l'échelle du réseau.

Ce chapitre est organisé comme suit : la section 3.2 présente le contexte du réseau P2P KAD et les travaux relatifs à la sécurité de ce réseau, plus précisément l'attaque Sybil et la pollution. Nous décrivons ensuite dans la section 3.3 notre explorateur permettant la découverte exhaustive des pairs ainsi que l'analyse des

images du réseau obtenues à travers deux approches de détection des pairs suspects. La section 3.4 s'intéresse à la pollution du réseau et présente notre méthode de détection ainsi que la quantification de la pollution à large échelle. Enfin, la section 3.5 conclut sur les conséquences de l'absence d'identification fiable dans les réseaux P2P actuels.

3.2. Etat de l'art

3.2.1. *Cas d'étude : le réseau P2P KAD*

Afin d'illustrer les problèmes d'identification des pairs et contenus, nous avons étudié le réseau P2P KAD. KAD est un réseau P2P structuré basé sur le protocole de routage Kademlia [MAY 02] et implanté par les clients libres eMule[1] et aMule[2] qui permettent le partage de fichiers entre utilisateurs. Rendu populaire au fil des fermetures des serveurs eDonkey, KAD est principalement utilisé en Europe et en Chine et compte environ 3 millions d'utilisateurs simultanés, ce qui en fait l'un des plus importants réseaux P2P déployés.

Dans KAD, chaque pair ainsi que chaque information indexée dans le réseau possèdent un identifiant « KADID » de 128 bits définissant sa place sur la DHT. Le routage est basé sur la métrique XOR grâce à laquelle on mesure la distance entre deux identifiants. La table de routage de chaque pair est organisée en un arbre dont les feuilles sont constituées de groupes de taille constante de K contacts (K=10), la distance entre les contacts retenus et le pair courant étant divisée par deux (1 bit supplémentaire commun) à chaque niveau de l'arbre. Ainsi le niveau i représente une portion du réseau de taille $n/2^i$, donc d'autant plus petite que celle-ci est proche du pair courant. Cette organisation permet de localiser efficacement les identifiants recherchés en O(log n) messages, n étant la taille du réseau.

En tant que support au partage de fichiers, la fonction principale de la DHT de KAD est d'indexer des mots-clés et des fichiers selon la procédure présentée par la figure 3.1. Lorsqu'un fichier est partagé (dans l'exemple, le fichier vidéo nommé *matrix_revolution.avi*), son contenu ainsi que chaque mot-clé constituant le nom du fichier sont hachés par une fonction MD4 (donnant les identifiants 32, 54, 87 pour chacun des mots-clés et le fichier). Les identifiants ainsi générés sont ensuite publiés sur le réseau. Les pairs chargés de l'indexation d'une information sont les dix pairs dont les identifiants sont les plus proches de celui de l'information. L'assignation des identifiants des pairs n'est donc pas strictement contrainte, bien qu'utilisant

1. www.emule-project.net.
2. www.amule.org.

normalement une fonction aléatoire, alors que les identifiants des mots-clés et fichiers indexés sont obtenus par la fonction de hachage MD4.

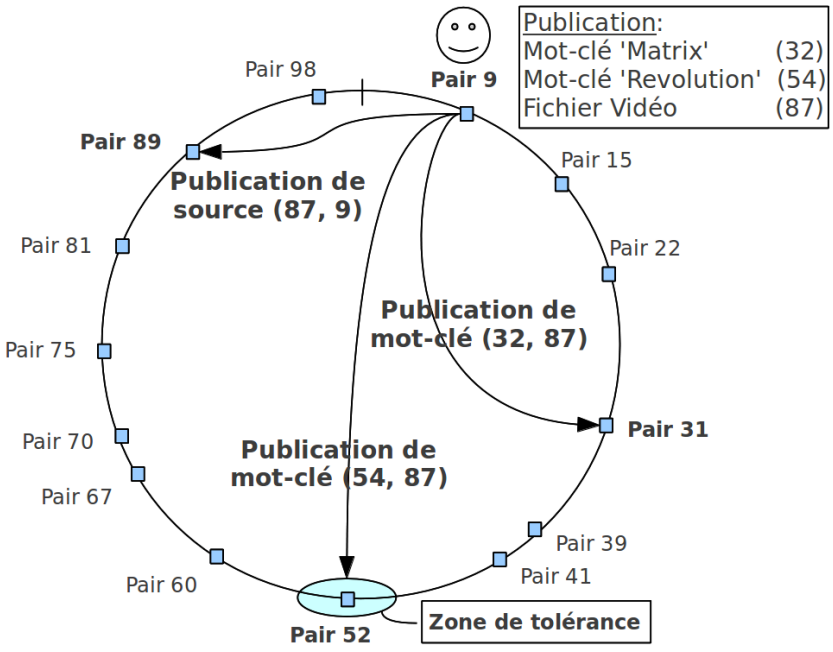


Figure 3.1. *Indexation à deux niveaux sur KAD*

Un mécanisme de double indexation permet de retrouver un fichier correspondant à un ensemble de mots-clés. Pour publier un fichier, deux types de requêtes sont nécessaires :

– les requêtes *KADEMLIA2_PUBLISH_KEY_REQ* sont envoyées vers l'identifiant des mots-clés et associent un mot-clé (32 ou 54) avec un fichier (87) ;

– les requêtes *KADEMLIA2_PUBLISH_SOURCE_REQ* sont envoyées vers l'identifiant du fichier (87) et associent un fichier avec une source (le pair 9 le partageant).

La réalisation de services (publication ou recherche) se fait en deux étapes. Dans un premier temps, le processus de localisation trouve les pairs les plus proches de l'identifiant de l'information visée (en émettant des requêtes *KADEMLIA2_REQ* de

manière itérative), puis les requêtes spécifiques au service demandé sont envoyées à ces pairs. On notera à cette étape que si l'identifiant d'un mot-clé ou d'un fichier est obtenu en calculant l'empreinte de celui-ci (hash MD4), l'identifiant des pairs est lui calculé aléatoirement et n'est donc pas vérifiable. Cette faiblesse est notamment exploitée par l'attaque Sybil.

3.2.2. *Faiblesse de l'identification des utilisateurs : l'attaque Sybil*

3.2.2.1. *Principe*

Les pairs constituant un réseau sont des entités autonomes et par définition non fiables (pannes, comportements malveillants). Les réseaux P2P sont cependant capables d'assurer un service fiable par trois éléments :

– le fait que les pairs soient indépendants les uns des autres ;

– le fait que les pairs soient aléatoirement répartis sur l'espace des identifiants ;

– l'utilisation de mécanismes de réplication.

La réplication est primordiale tant au niveau routage (entrées multiples dans la table de routage) afin de maintenir la connectivité des pairs, qu'au niveau service (indexation d'une donnée répliquée sur plusieurs pairs) afin d'assurer sa continuité. La réplication, associée au fait que les pairs soient indépendants et aléatoirement répartis, assure une distribution des services P2P au niveau géographique (une donnée est indexée par des pairs de différents pays) et limite l'impact des pairs malveillants : tant que le nombre de pairs malveillants coordonnés reste faible à l'échelle du réseau, une donnée est majoritairement indexée sur des pairs sains. Le service devient ainsi robuste aux défaillances ponctuelles des pairs ou de parties du réseau IP. L'attaque Sybil met à mal les deux premières hypothèses et, en leur absence, les mécanismes de réplication ne suffisent plus à assurer la fiabilité des services P2P.

L'attaque Sybil, telle que décrite initialement par Douceur [DOU 02], consiste pour une même entité à insérer un grand nombre de pairs dans le réseau. Les pairs ainsi générés ne sont pas indépendants et l'entité qui les a créés peut représenter une part significative du réseau. Une donnée répliquée en théorie sur plusieurs pairs indépendants peut de ce fait être uniquement indexée par des Sybils. La faiblesse des réseaux P2P permettant cette attaque est le peu de contraintes imposées sur la génération des identités des pairs. Une même entité peut ainsi créer autant d'identités différentes connectées au réseau que ses ressources le lui permettent. Douceur montre qu'en l'absence d'une autorité de certification centralisée des identités, l'attaque Sybil ne peut être évitée ni par une identification réalisée indépendamment par chaque pair, ni même réalisée conjointement. Cependant, des

mécanismes peuvent être conçus pour limiter l'ampleur de l'attaque en la rendant plus difficile.

Telle que définie dans [DOU 02], l'attaque Sybil viole seulement l'indépendance des pairs. En pratique, ses nombreuses mises en œuvre impliquent également un positionnement non aléatoire des pairs insérés, si bien qu'une attaque Sybil comprend par extension un placement stratégique des Sybils sur la DHT. Le fait d'insérer les pairs à des endroits spécifiques du réseau rend l'attaque plus efficace en augmentant localement la présence de l'attaquant dans le réseau. Ceci permet plusieurs applications introduites par [CAS 02] telles que la prise de contrôle d'une donnée indexée sur une DHT (tous les réplicats sont stockés par les pairs insérés) ou l'isolation d'un pair du réseau (tous ses contacts sont des pairs insérés).

3.2.2.2. *Applications*

L'article [STE 07a] présente une attaque Sybil réalisée à grande échelle sur le réseau KAD. L'attaque nécessite deux étapes. Les auteurs utilisent tout d'abord un explorateur afin de découvrir l'ensemble des pairs du réseau, puis ils injectent les attaquants directement dans la table de routage de chaque pair en créant les identifiants des attaquants de manière à occuper une place précise dans la table de routage de la victime. Les auteurs ont ainsi inséré depuis une unique machine 2^{16} attaquants dans une zone de la DHT contenant d'ordinaire 8 000 pairs (c'est-à-dire 1/256^{e} du réseau). Cette attaque leur permit d'intercepter la grande majorité des requêtes de recherche et de publication à destination de cette zone, et ainsi d'en contrôler les services. Cette attaque fut cependant très intrusive, mais les auteurs ont montré que des attaques beaucoup plus localisées sont également possibles.

Celles-ci consistent à prendre le contrôle de données indexées sur la DHT en choisissant précisément l'identifiant des pairs insérés de manière à les rendre responsables de la donnée ciblée. Dans Kademlia, les attaquants choisissent leur identifiant de manière à être proches de la donnée selon la métrique XOR. Les auteurs de [STE 07a] ont ainsi montré qu'il était possible d'éclipser une donnée très populaire du réseau KAD, en l'occurrence le mot-clé « the », avec seulement 32 attaquants. Les attaquants insérés deviennent ainsi responsables de la donnée en recevant toutes les requêtes de publication relatives. Il leur suffit de dénier les demandes de recherche pour l'identifiant ciblé pour que les pairs du réseau ne puissent plus accéder à l'information éclipsée. Toute recherche de contenus utilisant le mot-clé éclipsé aboutit à l'absence de résultats, bien que les fichiers soient toujours partagés par les pairs. L'article [KOH 09] perfectionne encore davantage l'attaque éclipse dans KAD en créant une chaîne d'attaquants dont la proximité est toujours croissante avec l'identifiant ciblé. De cette manière, la recherche d'un pair

progresse indéfiniment vers l'identifiant sans jamais l'atteindre et l'information n'est pas trouvée quand le délai alloué à la recherche expire.

Dès lors, certains mécanismes de protection ont été implantés pour protéger la table de routage de telles attaques [CHO 09]. De nouvelles contraintes empêchent dorénavant deux pairs affichant une même adresse IP d'être insérés dans une même table de routage. De même, deux pairs appartenant au même sous-réseau ne peuvent pas être trop proches dans une même table de routage. Si cette tentative d'améliorer l'identification des pairs en prenant en compte leur adresse IP est louable, elle est cependant insuffisante. Nous avons montré dans nos précédents travaux [CHO 10] que les attaques ciblées peuvent utiliser des nœuds distribués du point de vue du réseau IP et continuer d'être efficaces avec peu de ressources. Le schéma 3.2 montre les échanges de messages nécessaires à la réalisation d'un service sur KAD lorsqu'une référence est attaquée. Les pairs malveillants sont ainsi insérés plus proches que n'importe quels autres de la ressource visée (96 bits en commun) et coopèrent pour attirer les requêtes de service.

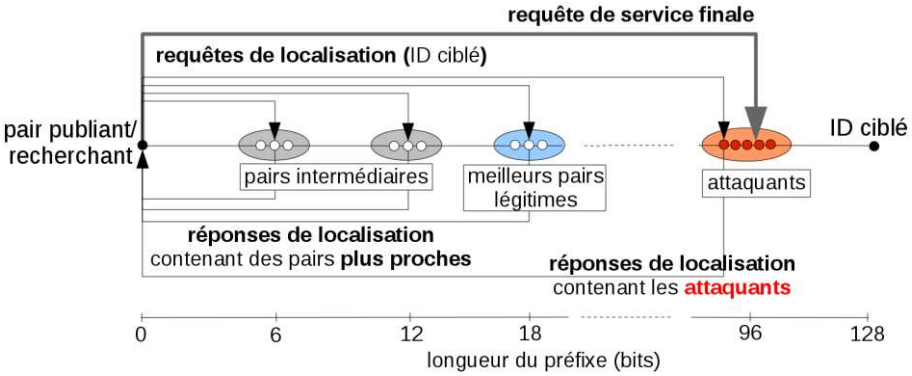


Figure 3.2. *Prise de contrôle d'une référence sur la DHT de KAD*

D'autres applications malveillantes de l'attaque Sybil sont également possibles : les auteurs de [STE 07a] ont également expérimenté, tout comme [NAO 06], un déni de service distribué en injectant systématiquement l'adresse IP d'une machine victime dans les réponses émises par les Sybils et générant ainsi plus de 100 Mbit/sec de trafic. Les auteurs de [LIA 06] ont montré que la DHT d'Overnet pouvait être polluée efficacement par l'insertion de nœuds autour de certains mots-clés. Ce problème affecte également KAD [LOC 10]. Nous avons montré dans [CHO 10] que l'attaque locale permettait en outre de polluer efficacement le réseau en générant à faible coût de faux fichiers très attractifs ce qui peut amener les utilisateurs à télécharger des contenus indésirables et illégaux (virus, contenu pédophile, etc.) à leur insu.

Les auteurs de [CAS 02] et [SIN 06] présentent une autre forme d'attaque éclipse ne visant pas à faire disparaître une donnée de la DHT, mais à déconnecter un ou plusieurs pairs du reste du réseau ou à partitionner celui-ci. Pour réaliser cette attaque, étant donné un ensemble de pairs victimes à déconnecter, les Sybils doivent être insérés de manière à remplir intégralement la table de routage de ces pairs. Tout accès des victimes au réseau P2P est alors contrôlé par les attaquants. Une telle attaque a été mise en œuvre avec succès contre le réseau KAD [WAN 08] et a montré qu'il était ainsi possible de partitionner le réseau avec peu de ressources. Au-delà des DHT, l'attaque Sybil a également été prouvée comme étant efficace pour perturber le transfert d'un fichier au sein d'un essaim du réseau BitTorrent [KON 07]. Les attaquants insérés dans le *swarm* peuvent alors mentir sur la disponibilité des parties du fichier et en ralentir ainsi la diffusion. Cette attaque est d'autant plus efficace que le nombre d'attaquants est élevé et que ceux-ci sont insérés tôt dans la vie de du *swarm*.

Enfin, et contrairement aux précédentes applications, l'attaque Sybil n'est pas toujours utilisée à des fins malveillantes. En effet, certaines méthodes de supervision des réseaux P2P sont basées sur l'injection de sondes. Les nœuds ainsi insérés en des points spécifiques constituent autant de sondes capables de surveiller les messages échangés au sein du réseau P2P KAD. [STE 07a] surveille ainsi une portion complète de la DHT, [CHO 10] s'intéresse à des mots-clés spécifiques et annonce des pots de miel alors que [MEM 09] place des sondes de manière à recevoir une copie du trafic émis vers chaque pair du réseau. Or, l'injection de multiples sondes dans le réseau P2P par une entité souhaitant le superviser est équivalente à une attaque Sybil. De telles méthodes ont été appliquées avec succès sur d'autres réseaux P2P dont [KLE 04, ACO 07], BitTorrent [FAL 07, WOL 10]. Les pairs insérés dans ces expériences ont adopté un comportement normal (mis à part la supervision des messages) et n'ont, par conséquent, pas dégradé le réseau. Ces pratiques posent néanmoins des problèmes relatifs à la vie privée des utilisateurs du réseau.

3.2.3. *Faiblesse de l'identification des contenus : la pollution*

3.2.3.1. *Différentes formes de pollution*

Le second problème majeur affectant les réseaux P2P de partage de fichiers est la pollution en leur sein. Un fichier est dit pollué si le contenu fourni par le service P2P ne correspond pas à la description présentée à l'utilisateur. Plusieurs formes de pollution existent. Le contenu peut ainsi être valide, mais sans rapport avec la description, partiellement dégradé, inexploitable une fois téléchargé ou encore, complètement fictif. Dans tous les cas, la pollution est liée à l'incapacité du système

P2P à proposer une identification fiable des contenus pour ses utilisateurs, aboutissant à l'impossibilité de lier la source initiale à son contenu.

La pollution dégrade notablement la qualité de services des réseaux P2P : d'une part, car l'utilisateur doit vérifier le contenu du fichier et rechercher un autre fichier sain en cas de pollution, et d'autre part, car la diffusion de la pollution consomme inutilement les ressources limitées du réseau P2P. La pollution pose également des problèmes de sécurité. Ainsi, un contenu illégal ou malveillant (virus) peut être diffusé avec le nom de fichiers exécutables légitimes. Nous présentons dans cette section les différentes recherches menées sur la pollution des réseaux P2P, en nous intéressant à la compréhension du phénomène et à son impact sur la sécurité.

La pollution des réseaux P2P étant un sujet vaste et complexe, nous proposons tout d'abord de recenser les différentes formes de pollution ainsi que leurs moyens possibles de mise en œuvre afin de bien positionner notre contribution dans ce domaine (voir section 3.4). Nous distinguons notamment deux grands types de pollution selon qu'ils impliquent directement la corruption des données (*content pollution*) ou celle des métadonnées (*metadata pollution*), ce qui est illustré par la partie gauche de la figure 3.3.

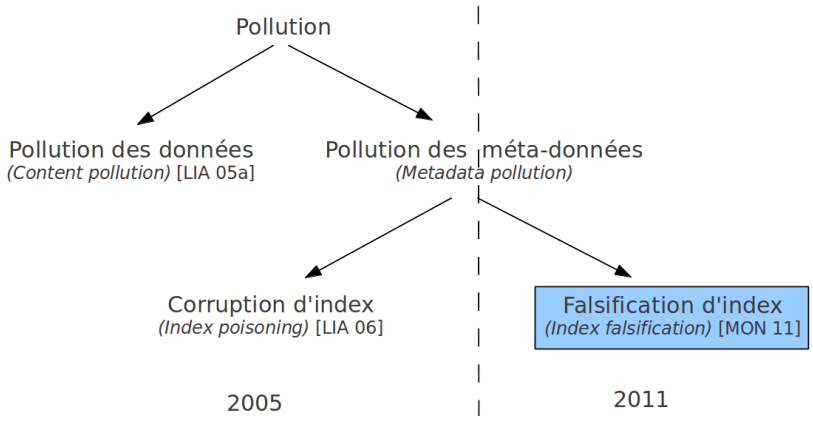


Figure 3.3. *Taxonomie des différentes formes de pollution*

3.2.3.2. *Pollution des données*

La première forme de pollution des données consiste à partager des fichiers dont le contenu correspond à la description, mais est fortement dégradé. Cette forme de pollution a été largement constatée sur le réseau Kazaa [LIA 05a]. La dégradation peut avoir plusieurs formes telles que l'ajout de bruit, la fin prématurée du contenu, l'annonce d'un message, etc. Une autre forme de pollution des contenus vise à

dégrader les performances de téléchargement d'un fichier existant non pollué en partageant de fausses pièces d'information. Cette forme de pollution est notamment utilisée dans le cadre de BitTorrent dont l'indexation des fichiers partagés est contrôlée par des sites Web qui limitent l'indexation de fichiers indésirables.

L'article [CHR 05] différencie la pollution normale qui consiste pour un pair à partager par inadvertance un fichier inutilisable, de l'empoisonnement (*poisoning*), qui consiste à injecter volontairement des leurres (fichiers corrompus) pour un contenu donné afin de réduire son intérêt. Les auteurs évaluent alors la disponibilité réelle de quinze contenus dans quatre réseaux P2P (Gnutella, FastTrack, eDonkey et Overnet) victimes de différentes stratégies de pollution. L'injection de nombreux faux fichiers uniques, comparable à une pollution involontaire, n'affecte presque pas la disponibilité des contenus réels. Pour nuire à la disponibilité des fichiers réels, il faut que le nombre de faux fichiers injectés soit extrêmement élevé (99 faux fichiers pour 1 réel), de manière à remplir les réponses de recherche sans que des fichiers légitimes ne puissent être sélectionnés, ce qui nécessite l'injection de dizaines de milliers de fichiers. Cette forme de pollution est donc négligeable. La seconde stratégie de pollution consiste à annoncer peu de faux fichiers différents, mais par de nombreux pairs, en les renouvelant périodiquement. Les faux fichiers sont alors mieux classés dans les résultats de recherches, difficilement détectables, et se diffusent facilement. Quelques centaines d'annonces permettent ainsi de réduire la disponibilité des fichiers réels au profit des faux.

L'article [TUV 10] présente une autre forme de pollution, beaucoup plus élaborée qui est basée sur la collision de la fonction de hachage MD4 utilisée dans le réseau P2P eDonkey (mais la fonction MD5 présente également cette vulnérabilité). Les auteurs montrent qu'il est possible de générer instantanément un exécutable dont l'identifiant entre en collision avec un autre fichier partagé dans le réseau. Cette attaque est donc un moyen de diffuser des virus à l'insu des utilisateurs, même si ces derniers ont localisé la ressource *via* un URI, c'est-à-dire, en identifiant directement le fichier par son empreinte sans consultation préalable d'un mécanisme d'indexation pouvant être corrompu. Une contre-mesure possible est la validation des fichiers par l'empreinte AICH également utilisée dans le réseau eD2k pour corriger les erreurs, ou, à long terme, un changement de la fonction de hachage, mais ceci briserait la rétrocompatibilité entre clients. Cette attaque est cependant difficile à réaliser, la pollution du mécanisme d'indexation est beaucoup plus abordable et répandue.

Du point de vue du pollueur, ces formes de pollution ont le désavantage de requérir, au moins dans un premier temps, des ressources importantes pour diffuser rapidement les données constituant les fichiers corrompus. Dans le cadre des réseaux P2P basés sur des DHTs, l'indexation des contenus partagés est réalisée par les pairs eux-mêmes, et elle est peu contrôlée. Cette faiblesse associée à l'impossibilité

d'identifier la liaison entre un fichier et son nom permet d'autres formes de pollution plus efficaces exploitant les métadonnées utilisées par les mécanismes d'indexation.

3.2.3.3. *Pollution des métadonnées*

La corruption du mécanisme d'indexation (*index poisoning*), consiste à diffuser des informations erronées dans l'index (centralisé ou distribué) des réseaux P2P au sujet des fichiers ciblés. Cette forme de pollution est également appelée « metadata pollution » dans l'article [LIA 05a]. Le pollueur crée, pour des mots-clés donnés, de fausses références de fichiers dans l'index ou de faux contacts (IP, port). Afin de sembler populaire, chaque faux fichier peut être annoncé de nombreuses fois. La ressource ainsi polluée apparaît comme inaccessible, n'étant en réalité partagée par aucune source. Contrairement à l'injection de fichiers pollués, cette méthode nécessite très peu de bande passante, seuls des messages de signalisation sont nécessaires. Les auteurs de [LIA 06] étudient cette pollution dans deux réseaux P2P : FastTrack et Overnet qui opéraient une DHT similaire à celle de KAD. Ils ont recensé ainsi toutes les versions et copies de fichiers publiés pour dix contenus afin d'estimer leur niveau de pollution et les stratégies employées. Une simple analyse du nombre de publications, pour un même contenu, émises par chaque pair permet d'identifier les pollueurs et indirectement les fausses entrées. Cette estimation est confirmée par une analyse automatisée des fichiers téléchargés. Les pollueurs ainsi identifiés représentent 7 % des pairs, mais 77 % des fichiers publiés. Le coût d'une telle pollution est négligeable, en particulier pour Overnet dont les messages d'annonce sont envoyés par UDP. L'article [LOC 10] a montré que cette forme de pollution peut également affecter KAD. Si le nombre de faux fichiers annoncés pour un mot-clé donné est suffisamment élevé, ceux-ci peuvent de plus saturer la table d'indexation des pairs chargés dudit mot-clé et empêcher ainsi le référencement de fichiers légitimes.

Les auteurs [LIA 06] décrivent également une autre stratégie de pollution plus efficace pouvant être réalisée en insérant des nœuds pollueurs dans le réseau, de manière à les rendre responsables de l'indexation des contenus ciblés et ainsi générer directement des réponses corrompues. Concernant le réseau Overnet, l'injection de pairs malveillants, dans la DHT partageant l'identifiant du mot-clé ou fichier à polluer, permettrait de réaliser cette pollution. Cette étude omet cependant une autre forme de pollution importante des métadonnées pouvant affecter le mécanisme d'indexation. Le mécanisme d'indexation peut en effet être corrompu en mélangeant l'indexation de fichiers existants, plutôt qu'en insérant des références de fichiers inexistants. Cette pollution peut être simplement réalisée en changeant le nom d'un fichier, puis en le partageant sur le réseau. Un objet pollué par cette méthode pourra être téléchargé, mais son contenu est sans relation avec le nom du fichier sélectionné par l'utilisateur. Cette pollution est plus intrusive, car un fichier non souhaité est téléchargé par l'utilisateur, et engendre d'importants problèmes de

sécurité (contenu offensant, virus). Par ailleurs, cette pollution n'est pas détectable par les métriques présentées dans [LIA 06], car elle ne nécessite pas la création de nombreuses entrées erronées.

Il est à noter que l'indexation des contenus dans les espaces numériques de partage de fichiers peut également se faire de manière centralisée, notamment grâce à des sites Web. C'est le cas pour le réseau P2P BitTorrent (*via* l'indexation de *tracker*), mais également de eDonkey/KAD dont les contenus peuvent être identifiés directement par des URI. Cette forme d'indexation est moins sujette à la pollution, car elle ne dépend pas des annonces ou de l'indexation de multiples pairs, mais de tels sites Web (The Pirate Bay, eMule-Paradise, etc.) sont régulièrement l'objet de poursuites par les ayants droit des contenus indexés.

3.2.3.4. *Diffusion de la pollution*

Afin de comprendre les facteurs influençant la diffusion de la pollution, l'article [DUM 05] modélise et simule deux formes de pollution, l'une ciblée sur un fichier particulier, l'autre affectant toutes les requêtes du réseau indifféremment, et ce, pour différents types d'architecture P2P. Les simulations montrent que la diffusion d'une pollution ciblée de faux fichiers est étroitement liée à la vigilance des utilisateurs et à leur propension à rendre disponibles les fichiers téléchargés, et ce, indifféremment des architectures. La pollution globale simulée dans l'article est cependant assez peu réaliste. Elle consiste à insérer aléatoirement des pairs corrompant les recherches de sources (annonçant des pairs lents ou les attaquants eux-mêmes). Les architectures P2P sans hiérarchie sont plus résistantes à l'insertion aléatoire de pairs malveillants, car moins de nœuds sont contactés lors de la recherche. Cette étude est cependant trop générale dans les stratégies de pollution simulées et dans la représentation des réseaux P2P pour bien appréhender la pollution, d'où la nécessité de mesures réelles.

L'article [LIA 05a] établit un bilan précis de la pollution affectant le réseau P2P Kazaa. Pour cela, les auteurs utilisent un explorateur émettant des requêtes sur tous les super-nœuds du réseau et découvrant tous les fichiers partagés étant donné un ensemble de mots-clés, puis ils appliquent un algorithme de détection de pollution sur les quelques contenus étudiés (analyse du format d'encodage, de la durée, etc.). L'analyse nécessite le téléchargement des fichiers, mais est peu sujette aux erreurs d'appréciation. Il apparaît que les fichiers populaires sont extrêmement pollués (62 % des versions et 73 % des copies), malgré le mécanisme de notation des fichiers fourni par les clients, alors que d'autres fichiers ne sont pas du tout affectés.

Les auteurs de [LIA 05a] identifient les sources originelles de la pollution dans Kazaa comme étant des sociétés polluant intentionnellement le réseau pour en réduire l'attrait. La pollution est ensuite diffusée par les pairs eux-mêmes. L'article [LEE 06] propose ainsi un modèle de propagation de la pollution au sein du réseau

P2P Kazaa, basé sur une étude du comportement des utilisateurs. Les 30 utilisateurs participant à l'étude à travers des questionnaires et l'utilisation d'un client instrumenté ont montré une faible réactivité vis-à-vis de la pollution, malgré une bonne connaissance des réseaux P2P et du problème. Ce manque de vigilance se traduit par une vérification tardive (plusieurs heures après obtention) ou une vérification trop superficielle (écoute incomplète) des fichiers téléchargés et il participe largement à diffuser la pollution. Le modèle de propagation montre qu'une différence de vigilance de 20 % peut décupler la pollution initiale et que, globalement, la pollution peut quadrupler la consommation de ressources dans le réseau.

3.2.4. *Bilan : identification et sécurité des réseaux P2P*

Au terme de cet état de l'art, il apparaît que l'attaque Sybil, résultant de l'incapacité des réseaux P2P à vérifier l'identité des pairs, est un des problèmes de sécurité majeurs des réseaux P2P. Cette menace pèse indifféremment sur toutes les architectures P2P, cependant, elle s'est avérée particulièrement néfaste dans le cadre des réseaux P2P structurés où des applications ont montré qu'il était possible de faire disparaître des contenus indexés sur le réseau ou encore de partitionner celui-ci. Bien que ceci n'ait pas été réellement mesuré, le contrôle du mécanisme d'indexation par une attaque Sybil peut également permettre la diffusion massive de pollution [LIA 06]. Par ailleurs, la stratégie de pollution mélangeant l'indexation des fichiers du fait de l'impossibilité d'associer de manière fiable un contenu à son nom soulève des problèmes importants de sécurité et a été très peu étudiée dans la littérature.

Ces deux grands problèmes affectant la sécurité des réseaux P2P que sont l'attaque Sybil et la pollution ont été largement décrits, mais n'ont jamais été mesurés sur un réseau P2P réel, ce qui constitue un profond manque à l'état de l'art actuel. Les sections suivantes présentent nos travaux [CHO 11, MON 11] visant à détecter et quantifier précisément l'exploitation qui est faite de ces faiblesses au sein d'un réseau P2P largement déployé, à savoir KAD.

3.3. Détection des positionnements suspects de pairs

3.3.1. *Introduction*

Nous avons vu dans la section précédente que la principale vulnérabilité pouvant être exploitée pour attaquer les réseaux P2P est une attaque Sybil localisée consistant en l'insertion de nœuds en des positions spécifiques sur la DHT. Nous proposons dans cette section de détecter et de recenser les pairs suspects dans le

réseau P2P KAD pouvant traduire ce comportement. Pour cela, nous réalisons une cartographie du réseau grâce à un explorateur spécifiquement conçu pour obtenir une image très précise de la DHT. Nous analysons ensuite les résultats afin de détecter deux types de positionnements suspects selon qu'ils impliquent localement un groupe de pairs malveillant ou uniquement un seul pair. Nous constatons ainsi pour la première fois la réalité de certaines attaques publiées et pouvons estimer leur nombre au sein du réseau.

3.3.1.1. *Travaux relatifs à l'exploration des DHT*

Un explorateur ou *crawler* est un outil capable de découvrir l'ensemble des pairs d'un réseau et de stocker les différentes informations les concernant.

Plusieurs explorations du réseau KAD ont déjà été réalisées à diverses fins. Les auteurs de [WAN 08] et [STE 07a] découvrent ainsi les pairs du réseau à des fins d'attaque. Pour chaque pair découvert, ils interrogent ce dernier en émettant de nombreuses requêtes de localisation (*Kademlia Request*) vers des identifiants précalculés de manière à obtenir tous les contacts de la table de routage du pair interrogé. Ces informations servent ensuite à insérer des Sybils [WAN 08] ou à corrompre les références de contacts existants [STE 07a]. Utilisant *Blizzard*, le même explorateur que [WAN 08], les auteurs de [STE 07b] réalisent des explorations périodiques de la DHT de manière à étudier certaines caractéristiques des pairs dans le temps.

Les auteurs de [YU 09] utilisent une autre approche basée sur l'interrogation de contacts par des requêtes d'amorçage (*Bootstrap Request*). Cette approche est censée être plus performante (20 contacts retournés par requête d'amorçage contre 11 pour celle de localisation). Cependant les contacts obtenus sont choisis aléatoirement dans la table alors que les requêtes de localisation spécifient une adresse cible permettant de contrôler le parcours des tables. Les résultats de cette exploration ont mis en évidence un nombre important de pairs (20 %) partageant leur identifiant dont les auteurs étudient les causes possibles.

Si de nombreuses observations du réseau KAD ont été réalisées, aucune jusqu'à présent ne s'est intéressée à recenser les attaques pouvant affecter la DHT. De même, aucune étude n'estime l'efficacité de l'explorateur mis en œuvre dont les algorithmes sont peu détaillés quand ils sont mentionnés.

3.3.2. *Exploration du réseau KAD*

3.3.2.1. *Méthode d'exploration*

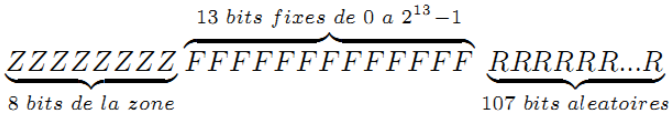
La conception de notre explorateur vise deux objectifs : d'une part, obtenir une vision précise du réseau et d'autre part, limiter l'empreinte de l'exploration sur le réseau en limitant le nombre de requêtes envoyées à chaque pair. Ceci permet en outre d'obtenir une exploration compatible avec les limitations implantées dans les derniers clients, contrairement aux précédentes stratégies d'exploration désormais limitées par la protection contre l'inondation qui empêche un pair de recevoir rapidement des messages d'une même source. Notre méthode d'exploration se divise en trois phases décrites ci-après.

3.3.2.1.1. Amorçage

La phase d'amorçage sert à obtenir une première image imprécise de l'ensemble de la DHT. Pour cela, des requêtes d'amorçage (*Bootstrap*) sont émises. Les requêtes d'amorçage permettent d'obtenir 20 contacts tirés aléatoirement dans la table de routage du pair sollicité et sont donc parfaitement adaptées à une première découverte globale de la DHT. De nouveaux contacts sont ainsi progressivement interrogés au fur et à mesure des réponses jusqu'à ce que 500 000 contacts aient été découverts dont au moins 500 par zone[3]. Au-delà de cette valeur, les contacts retournés étant sélectionnés au hasard, il est de plus en plus difficile d'apprendre de nouveaux contacts par cette méthode.

3.3.2.1.2. Exploration complète

Ensuite, chaque zone est explorée avec précision grâce aux requêtes de localisation (*Kademlia Request*). Un pair ainsi interrogé retourne les 4 contacts connus étant les plus proches de l'identifiant spécifié en paramètre. Afin de découvrir l'ensemble des pairs, nous générons 2^{21} soit environ 2 millions de « KADIDs cibles » uniformément répartis, et nous envoyons pour chacun d'eux une requête de localisation au pair le plus proche déjà découvert. Ainsi, 2^{13} ($2^{21}/2^{8}$) KADIDs cibles sont générés dans chaque zone selon le format :



3. Une zone est une subdivision artificielle de l'espace d'adressage basée sur le premier octet de poids fort des identifiants (de *0x00* à *0xFF*).

où *Z*, *F* et *R* désignent respectivement des bits de zone, les bits fixés et ceux tirés aléatoirement une fois.

3.3.2.1.3. Seconde passe

Dès qu'une zone a été explorée, c'est-à-dire quand tous les KADIDs cibles de cette zone ont été envoyés, une seconde exploration de celle-ci a lieu pour en améliorer la cartographie. Pour chaque contact précédemment découvert, on calcule alors son voisin le plus proche dont on extrait ensuite le préfixe commun de longueur *x* bits entre les deux KADIDs. On construit ensuite un nouveau « KADID cible » partageant ce préfixe et où les (128-*x*) bits restants sont aléatoires. Une requête de localisation pour ce KADID cible est finalement envoyée au contact. Cette phase permet de découvrir quelques contacts manqués lors de l'exploration complète. L'exploration se termine lorsque tous les contacts ont ainsi été interrogés sur leur voisinage immédiat.

Les performances de l'algorithme d'exploration décrit ci-dessus ont été validées par plusieurs expériences réalisées sur le réseau KAD [CHO 12].

3.3.2.2. *Cartographie obtenue*

3.3.2.2.1. Informations enregistrées

Pour chaque pair découvert, nous enregistrons les informations suivantes : <KADID, adresse IP [4], port TCP, port UDP, version de KAD, état du pair>.

La version de KAD fait référence à la version du protocole implantée par le client, l'état du pair est P(*possible*), T(*tried*) ou R(*responded*) selon respectivement que le contact a juste été découvert, a été contacté ou a répondu :

[...]
<32FFF76959F6A7095347FB338B304330, #.#.#.#, 38060, 16905, 0, T>
<32FFFC5C4D5AE9A082871FF68B1F0D9C, #.#.#.#, 5149, 1025, 4, R>
<32FFFC5C4D5AE9A082871FF68B1F0D9C, #.#.#.#, 5149, 5159, 4, P>
Zone 33: 15196 contacts
<3300048A90460A8AAC3DD2FF542ADF98, #.#.#.#, 12399, 39949, 9, R>
<3300083A0480CFA91B8C142401DD26F2, #.#.#.#, 5611, 5621, 8, T>
<330018506569424D7CBA7133F437EDC8, #.#.#.#, 6647, 6657, 8, P>

4. Certaines adresses IP sont anonymisées dans le cadre de cet ouvrage.

<33002596F7AAAA4348FB4349F0A14FA4, #.#.#.#, 46318, 61632, 9, R>
<33002EF905E27753B1900BC602D29C20, #.#.#.#, 19774, 19774, 8, T>
<33004546934FABE9685674DE1598548F, #.#.#.#, 51478, 52073, 9, R>
[...]

3.3.2.2.2. Résultats généraux

L'exploration d'une zone compte entre 13 000 et 17 000 contacts, le nombre total de pairs recensé à un instant donné allant de 3,3 millions à 4,3 millions selon le jour et l'heure de l'exploration. D'un point de vue macroscopique, la répartition des pairs sur l'ensemble de l'espace d'adressage de la DHT est bien uniforme (figure 3.4), conformément au comportement attendu de la majorité des pairs légitimes générant aléatoirement leur identifiant à la première connexion.

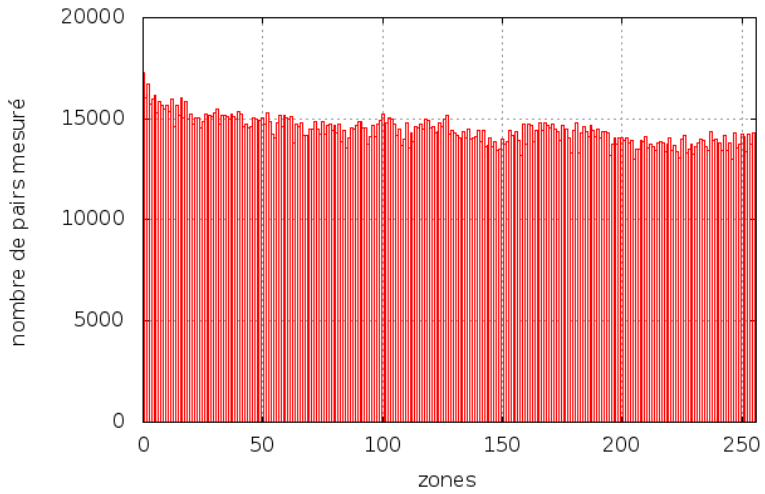


Figure 3.4. *Répartition des pairs sur la DHT*

Nous analysons plus précisément les résultats d'exploration dans la sous-section suivante, avec pour objectif de détecter les placements traduisant des comportements déviants. Les résultats obtenus pour les différentes explorations réalisées étant similaires, nous illustrons notre analyse avec les données d'une exploration réalisée le 8 juillet 2010 et comptant 3 688 932 pairs.

3.3.3. *Détection des pairs suspects*

Comme expliqué précédemment, une attaque sur la DHT implique l'insertion d'un ou plusieurs pairs à proximité de l'identifiant ciblé, afin d'attirer tout ou partie des requêtes à son attention. Pour une meilleure efficacité, plusieurs pairs peuvent être insérés conjointement afin d'attirer davantage de requêtes.

3.3.3.1. *Détection par densité des pairs*

Notre première analyse s'intéresse à localiser de tels groupes de pairs sur la DHT. Nous cherchons ainsi à détecter les couples de pairs dont la distance trop proche traduit un placement intentionnel à proximité d'un tiers identifiant plutôt qu'un choix aléatoire de leurs identifiants :

$$F(x) = N \div 2^x \qquad [3.1]$$

Soit F la fonction donnant le nombre moyen de pairs partageant x bits avec un pair courant étant donné un nombre total N de pairs dans le réseau. Nous considérons un nombre de 4 millions de pairs connectés simultanément. Le tableau 3.1 en présente certaines valeurs pour $N = 4*10^6$ et $x \in [1;128]$. De plus, le préfixe moyen partagé entre deux pairs consécutifs est de $d_{moy} = \log2(N) = 21.93$ bits.

Nombre de bits en commun	**Nombre moyen de pairs**
1	2 000 000
8	15 625
12	976.5
16	61
18	15.25
20	3.8
24	0.24
28	0.015
32	$9.32*10^{-4}$
64	$2.17*10^{-13}$
96	$5.05*10^{-23}$
128	$1.17*10^{-32}$

Tableau 3.1. *Nombre moyen de pairs partageant un préfixe avec un identifiant donné pour une DHT de 4 millions*

Etant donné notre exploration de la DHT, nous avons calculé le préfixe commun entre chaque pair et son plus proche voisin, les résultats sont présentés dans la figure 3.5. Si les préfixes jusqu'à 35 bits sont communément partagés entre voisins et ne permettent pas de détecter les attaques, les contacts partageant davantage de bits traduisent un placement intentionnel. Le premier graphe de la figure 3.6 illustre cette déviation de la norme théorique (équation 3.1) pour les contacts partageant entre 22 et 45 bits. Plus le préfixe commun est élevé, plus l'espérance de trouver de tels voisins est faible et traduit un placement intentionnel, ce qui est illustré par le second graphe de la figure 3.6. Nous avons ainsi relevé 426 groupes de contacts anormalement proches (partageant un préfixe entre 35 et 127 bits) et traduisant autant d'attaques groupées potentielles.

Nous avons par ailleurs réalisé cette analyse sur une seconde exploration de KAD effectuée en avril 2011 et durant laquelle 2 074 groupes de pairs suspects ont pu être mis en évidence. Les groupes d'attaquants montrent en outre des motifs d'attaque évidents en partageant un préfixe identique (40 bits), en utilisant des adresses IP appartenant au même sous-réseau ou encore des ports spécifiques.

L'exemple ci-dessous illustre deux groupes de ces pairs. Ceci tend à prouver que les attaques affectant le réseau évoluent dans le temps, les résultats d'explorations éloignées dans le temps étant différents :

Prefix « 4A9D8C877700000000000000000000 », length 40, shared by 6 contacts:
<4A9D8C87774AF8C551FE78BDDC3F5A37, 123.144.174.128, 10875, 10875, 8, T>
<4A9D8C877780DFB9985E75EE92AD1C68, 123.144.160.21, 10875, 10875, 8, T>
<4A9D8C877780DFB9985E75EE92AD1C68, 123.145.184.122, 10875, 10875, 8, T>
<4A9D8C877797D58D4C21B5BD5224F067, 123.144.160.98, 10875, 10875, 8, T>
<4A9D8C877797D58D4C21B5BD5224F067, 123.144.167.199, 10875, 10875, 8, T>
<4A9D8C8777F0F03BD1FE123548E269D2, 123.144.163.209, 10839, 10839, 0, R>

Prefix « 4A9D8C8777800000000000000000000 », length 41, shared by 4 contacts:
<4A9D8C877780DFB9985E75EE92AD1C68, 123.145.184.122, 10875, 10875, 8, T>
<4A9D8C877797D58D4C21B5BD5224F067, 123.144.160.98, 10875, 10875, 8, T>
<4A9D8C877797D58D4C21B5BD5224F067, 123.144.167.199, 10875, 10875, 8, T>
<4A9D8C8777F0F03BD1FE123548E269D2, 123.144.163.209, 10839, 10839, 0, R>

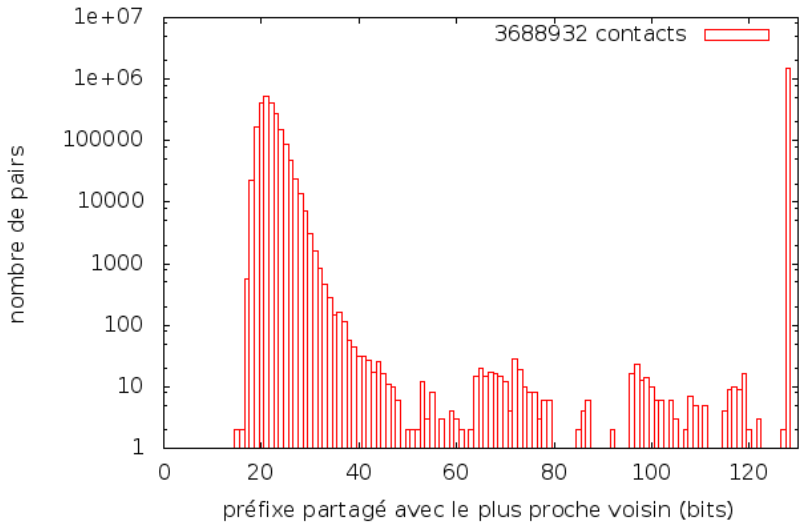


Figure 3.5. *Répartition des préfixes entre voisins sur la DHT*

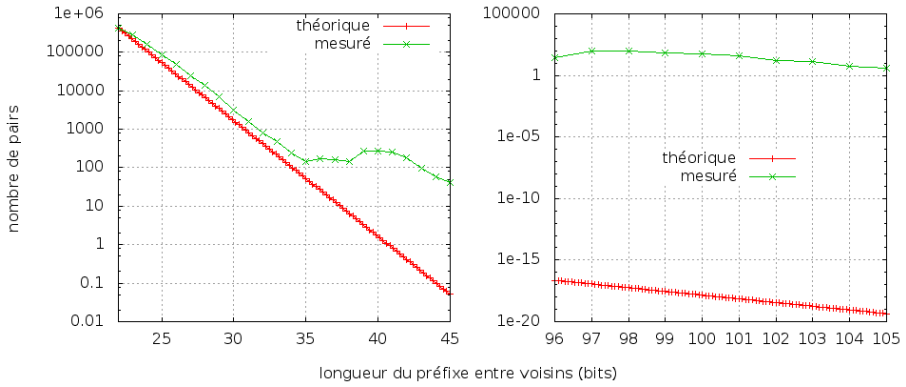


Figure 3.6. *Nombre moyen théorique et mesuré de pairs en fonction du préfixe partagé*

Identifiants partagés :

Quelle que soit l'exploration, l'écart le plus important concerne le préfixe de 128 bits (1 million de pairs) qui correspond aux pairs partageant exactement le même identifiant et mérite une analyse à part. En effet, sur les 3 688 932 pairs trouvés lors de l'exploration, on ne dénombre après analyse que 2 613 963 KADIDs différents.

Tout comme [YU 09], nous constatons donc l'existence de KADIDs partagés par plusieurs pairs. Plus précisément, parmi les KADIDs relevés :

– 82,36 % (2 152 900) des KADIDs sont utilisés par un pair unique ;

– 17,64 % (461 063) des KADIDs sont partagés par plusieurs pairs dont :
  - 10,42 % des KADIDs sont commun à 2 pairs ;
  - 2,85 % des KADIDs sont commun à 3 pairs ;
  - les pourcentages décroissant jusqu'à 1 KADID partagé par 259 pairs.

Le partage de préfixe peut traduire une attaque, mais peut également traduire un changement bénin de configuration d'un pair. En effet, un pair changeant d'adresse IP (allocation dynamique d'adresse, mobilité), ou de port de communication durant sa connexion au réseau apparaîtra deux fois avec le même identifiant le temps que la DHT mette à jour ses références. Afin d'éviter de compter ces cas, nous supprimons de la liste des identifiants suspects les cas pour lesquels deux pairs partagent un identifiant et où seule l'adresse IP ou seul le port changent entre les deux pairs. Ainsi parmi les KADIDs partagés entre deux pairs (272 149) :

– 49,73 % ne diffèrent que par l'adresse IP (ports UDP et TCP identiques) ;

– 26,91 % ne diffèrent que par le port UDP ;

– 1,44 % ne diffèrent que par le port TCP ;

– 21,92 % sont suspects.

Par cette méthode, 248 569 identifiants différents peuvent être suspectés. Malgré les précautions prises, ce chiffre peut être soumis à de faux positifs. Nous proposons donc une seconde estimation des attaques affectant KAD qui est plus fiable, car prenant en compte les identifiants des contenus, et non uniquement ceux des pairs.

3.3.3.2. Détection par proximité aux ressources

Les analyses précédentes ont une limite importante : elles permettent d'identifier des attributions d'identifiants suspects sans pour autant pouvoir les corréler à un contenu précis. Par ailleurs, les analyses précédentes étant basées sur des proximités entre pairs, au moins deux pairs doivent être insérés pour être détectés, les attaques n'impliquant qu'un pair passent inaperçues.

Une manière fiable de détecter les attaques est donc de pouvoir mettre en évidence la proximité anormale des pairs malveillants par rapport à une ressource plutôt que la proximité des pairs entre eux. La difficulté de cette approche est que les identifiants des ressources ne sont pas connus *a priori*. Pour appliquer cette méthode, nous avons extrait des mots-clés de contenus pouvant être partagés sur KAD depuis plusieurs sources d'information (meilleures ventes Amazon, iTunes, fichiers populaires sur ThePirateBay). Nous avons ensuite calculé l'identifiant de chacun des mots-clés composant les différents titres par la fonction MD4 utilisée par

KAD. Nous avons finalement recherché les contacts étant anormalement proches de ces identifiants (partageant un préfixe supérieur à 30 bits) dans les données des explorations. Un extrait des résultats est donné ci-après :

[...]
twilight 4D62D26BB2A686195DA7078D3720F60A
<4D62D26BB2A686195DA7078D3720F632, X.Y.#.#, 7290, 7294, 8, R> [prefix = 122]
soundtrack AC213377BB53F608390BD94A6AE6DD35
<AC213377BB53F608390BD94A82582F42, #.#.#.#, 5003, 5002, 8, R> [prefix = 96]
harry 770CF5279AB34348C8FECF9672747B94
<770CF5279AB34348C8FECF96524D8CDE, #.#.#.#, 5003, 5002, 8, P> [prefix = 98]
robin B9DF47E5BFAD75F8EE5E3F50EA217983
<B9DF47E5BFAD75F8EE5E3F5051F34AA8, #.#.#.#, 5003, 5002, 8, R> [prefix = 96]
<B9DF47E5BFAD75F8EE5E3F50EA21799F, X.Y.#.#, 7290, 7294, 8, R> [prefix = 123]
[...]

216/888 of the proposed keywords are targeted with at least 96 bits by:
44 IP addresses (showing 2119 unique KADIDs in the whole crawler's data)
41 subnets /24 (showing 2155 unique KADIDs in the whole crawler's data)

Sur les 888 mots-clés utilisés pour cette analyse, un quart d'entre eux avaient un pair proche partageant au moins 96 bits ce qui, étant donné l'espérance de trouver un pair légitime avec un tel préfixe (voir tableau 3.1) traduit sans équivoque un placement intentionnel et un comportement suspect. Un échantillon de ces mots-clés est donné dans le tableau 3.2, certains faisant référence à un contenu explicite, d'autres étant plus génériques.

Pour les pairs suspects ainsi détectés, nous avons recherché leur présence sur l'ensemble de la DHT afin de découvrir d'autres identifiants ciblés éventuels et absents de la liste initiale de nos mots-clés. Nous avons ainsi relevé que les seuls mots-clés recherchés ne représentent que 10 % de la présence de ces clients (adresse IP + port) sur la DHT. En comptant les 216 identifiants de mots-clés initiaux, ces clients sont au total présents sur 2 119 KADIDs.

Ce résultat montre clairement que de nombreux contenus de la DHT sont attaqués, parmi les plus populaires. De plus, des configurations d'attaques émergent rapidement des données. Par exemple, parmi les 216 identifiants, 205 sont ciblés par des pairs ayant exactement les ports suivants : UDP=5003, TCP=5002, un préfixe de 96 bits, mais des adresses IP distribuées sur plusieurs réseaux. Un autre attaquant

probable cible 16 identifiants parmi les 216 en utilisant des pairs ayant exactement les ports : UDP=7290, TCP=7294, un préfixe de 122 bits et une adresse IP venant d'un sous réseau spécifique (les adresses réseau considérées sont codées sur 16 bits, de la forme X.Y.#.#).

Mot-clé	**Meilleur préfixe**
avatar	126
invictus	123
sherlock	122
princess	122
Frog	98
ncis	96
nero	96
nine	122
love	122
american	97
russian	97
the	96
black	96
pirate	96
…	…

Tableau 3.2. *Exemples de mots-clés probablement attaqués avec le préfixe du plus proche contact trouvé*

Bien que cette méthode de détection soit fiable, elle a des limites, notamment quant au jeu de caractères utilisé par les mots-clés. Ceux considérés pour notre expérience utilisent en effet l'alphabet latin, or, KAD est pour moitié utilisé en Asie. Les pairs ciblant spécifiquement des contenus décrits avec des caractères asiatiques peuvent échapper à cette analyse. Par ailleurs, d'autres attaques peuvent cibler exclusivement les fichiers et non les mots-clés.

3.3.4. *Bilan*

Alors que plusieurs attaques pouvant affecter le réseau KAD ont été décrites dans de précédents travaux et que de nombreuses observations de ce réseau ont déjà

été réalisées, aucune d'entre elles ne s'était intéressée jusqu'alors aux questions de sécurité affectant la DHT. Afin d'estimer les positionnements anormaux des pairs pouvant traduire des attaques, nous avons tout d'abord développé et évalué un explorateur capable de découvrir précisément la DHT de KAD, malgré les limitations récemment incluses dans les clients.

Une première analyse considérant la proximité entre les identifiants des pairs a mis en évidence des regroupements de pairs anormaux, quelques pairs étant trop proches les uns des autres (426 en juillet 2010, 2 074 en avril 2011), mais la grande majorité d'entre eux partageant un même identifiant (248 569). Une seconde analyse basée sur l'étude de mots-clés populaires a mis en évidence qu'une grande proportion de ceux-ci est attaquée. Les pairs impliqués sont d'ailleurs présents sur de nombreux identifiants de la DHT (2119) et des configurations d'attaques peuvent être clairement mises en évidence. Concernant les mots-clés ciblés, les attaquants insèrent un seul pair extrêmement proche du contenu (96 bits ou 122 bits communs), mais ne semblent en revanche pas réaliser d'attaques impliquant plusieurs pairs.

Les deux approches de détection utilisées sont complémentaires. La première, basée sur l'analyse des distances inter-pairs, permet une détection des attaques sans nécessiter la connaissance des contenus ciblés, mais ne détecte que des attaques massives (c'est-à-dire où plusieurs pairs sont insérés). La seconde, basée sur l'analyse des distances pairs-contenus, permet de détecter des attaquants isolés, mais nécessite la connaissance *a priori* du contenu ciblé.

3.4. Détection des noms de fichiers suspects

3.4.1. *Introduction*

Nous avons décrit et détecté, dans la précédente section, des attaques reposant sur l'insertion ciblée de nœuds au sein de KAD malgré les plus récentes protections. Les nœuds ainsi placés peuvent réaliser plusieurs actions malveillantes (surveillance, attaque éclipse, déni de service, etc.) et nous avons détecté de nombreux placements suspects durant certaines périodes de mesure indiquant que le réseau est ponctuellement la cible d'attaques.

Cependant, les attaques internes nécessitant l'insertion de nœuds ne sont pas les seuls problèmes de sécurité affectant les réseaux P2P. De précédentes études [LIA 05a, LIA 06] ont en effet révélé que les réseaux P2P sont largement victimes de pollution. Cependant, ces études datent de 2005 et ont été réalisées sur des réseaux cibles aujourd'hui obsolètes tels qu'Overnet ou Kazaa. Dès lors, on peut légitimement se demander si la quantification de la pollution et les différentes formes relevées alors correspondent aux réseaux actuellement opérationnels tels que

KAD. Il est important d'évaluer quels sont les types de pollution actuels ainsi que leur niveau de propagation afin de savoir comment, et dans quelle mesure, les dizaines de millions d'utilisateurs utilisant les tables de hachage distribuées publiques sont affectés par cette menace.

Nous montrons ici qu'une nouvelle forme de pollution, non décrite dans l'état de l'art, sévit actuellement et s'avère être particulièrement dangereuse. Celle-ci provient directement d'un problème d'identification, non pas des pairs, mais des contenus. Nous l'appelons falsification d'indexation (ou mélange d'indexation). Cette pollution consiste à indexer un même fichier sous différents noms et, par conséquent, différents mots-clés, qui ne sont pas liés au contenu réel du fichier. La falsification d'indexation est un cas particulier de pollution des métadonnées (*metadata pollution*), ce qui est illustré par la figure 3.3. Pour un titre donné, le fichier pollué est publié de nombreuses fois par de fausses sources afin de le rendre populaire. Cette forme de pollution est plus néfaste que les précédentes pour deux raisons : d'abord, car elle aboutit au téléchargement complet d'un contenu indésirable et gaspille par conséquent inutilement des ressources, mais surtout, parce que le contenu téléchargé peut être dangereux pour l'utilisateur. Le contenu réel peut ainsi être un virus ou une vidéo pouvant gravement heurter la sensibilité de l'utilisateur (contenu pornographique, pédophile, etc.). Par ailleurs, cette forme de pollution génère de faux positifs lors de la supervision des fichiers illégaux dont l'indexation est ainsi falsifiée. Or, si les utilisateurs peuvent expérimenter régulièrement cette forme de pollution sur un réseau P2P tel que KAD, aucune étude à ce jour n'a décrit ni quantifié cette forme de pollution.

3.4.2. *Détection de la pollution*

3.4.2.1. *Visibilité de la falsification d'indexation*

Le principe de cette pollution repose sur le fait d'associer de nombreux noms différents à un même fichier. Cependant, le schéma d'indexation à deux niveaux de KAD rend très difficile la détection de cette pollution. En effet, si l'on se place du point de vue des pairs responsables d'un mot-clé, ces derniers ne reçoivent que les publications, pour un fichier pollué donné, incluant le mot-clé dont ils ont la charge dans leur nom. Cette contrainte forte sur la présence d'un mot-clé empêche de détecter la falsification d'indexation au niveau des mots-clés. Dans le cas où deux noms complètement différents sont associés à un fichier, ce dernier sera indexé au travers des mots-clés sur des ensembles de pairs complètement disjoints. L'interrogation des pairs responsables des mots-clés ne permet donc pas de détecter la falsification d'indexation. Nous avons confirmé cette hypothèse en plaçant une sonde à proximité du mot-clé « avatar ». Tous les noms de fichiers recueillis par la sonde contiennent effectivement le mot-clé « avatar » et semblent donc tous

proposer le bon contenu alors que ce mot-clé est en réalité fortement affecté par la pollution.

Si l'on se place du point de vue des pairs responsables d'indexer les sources d'un fichier, ceux-ci ne sont pas non plus capables d'appréhender la falsification de l'indexation. En effet, bien que l'ensemble des sources publie le fichier pollué sur les quelques pairs proches de son identifiant, le nom du fichier partagé ne fait pas partie des informations publiées lors de l'envoi d'une requête de type « *KADEMLIA2_PUBLISH_SOURCE_REQ* » associant une source (adresse IP, port) à un fichier (identifiant MD4). Le nom du fichier partagé est une information uniquement associée aux requêtes de type « *KADEMLIA2_PUBLISH_KEY_REQ* » afin de guider le choix de l'utilisateur lors de la présentation des différents fichiers disponibles pour un mot-clé. L'interrogation de la DHT de KAD ne permet donc pas de détecter la falsification d'indexation. Les pairs indexant les mots-clés ne voient pas les noms de fichiers contradictoires et les pairs responsables des sources n'ont pas connaissance des noms de fichiers associés.

Il est cependant possible de détecter la falsification d'indexation grâce à une fonctionnalité des clients KAD qui est externe à la DHT et disponible lors du téléchargement d'un fichier. Lorsqu'un fichier doit être téléchargé, les sources potentielles sont découvertes par interrogation de la DHT (par l'envoi de messages *KADEMLIA2_SEARCH_SOURCE_REQ*) puis, une connexion TCP vers chacune des sources potentielles est initiée afin de commencer le téléchargement. Les pairs acceptant la connexion constituent les sources réelles (disponibles à un moment donné) et, selon leur charge, partagent une partie du fichier ou placent la demande dans une file d'attente. Une requête spécifique peut alors être envoyée aux sources réelles par l'intermédiaire de la connexion TCP ouverte afin de connaître le nom par lequel ces sources partagent le fichier demandé. Cette information est accessible dans la fenêtre « Détails du fichier » de l'interface graphique. Grâce à ces informations, des noms contradictoires annoncés par les différentes sources d'un même fichier peuvent apparaître et la falsification d'indexation peut ainsi être constatée.

La capture d'écran 3.7 montre ainsi les noms annoncés pour un fichier sain alors que la capture 3.8 montre un fichier dont l'indexation est corrompue. Concernant le fichier sain, nous constatons que la majorité des sources (22) annoncent exactement le nom de fichier requis par l'utilisateur et les autres sources annoncent une variante minime du même nom en gardant de nombreux mots-clés en commun.

En revanche, concernant le fichier pollué, aucun des pairs n'annonce le nom de fichier désiré par l'utilisateur, à savoir « *Indiana Jones et les Aventuriers de l'Arche perdue* », ni même ne s'accorde sur un autre nom, tous les noms de fichier annoncés étant différents les uns des autres et ne partageant aucun mot-clé commun.

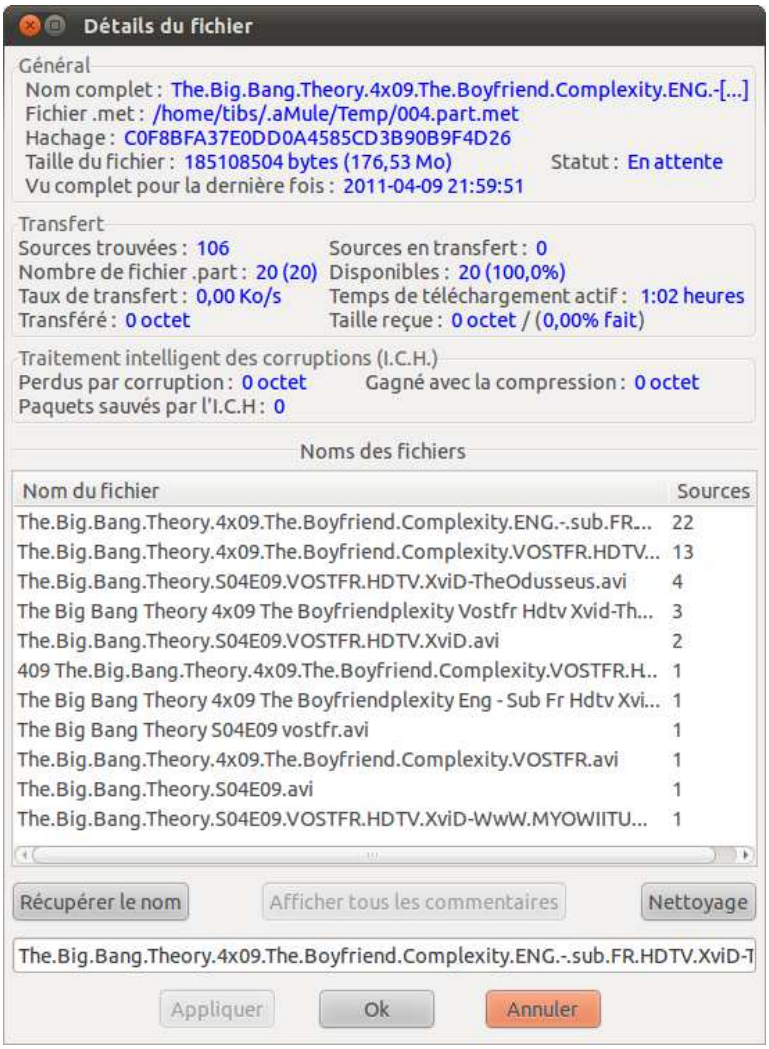


Figure 3.7. *Noms de fichier annoncés par les sources d'un fichier sain*

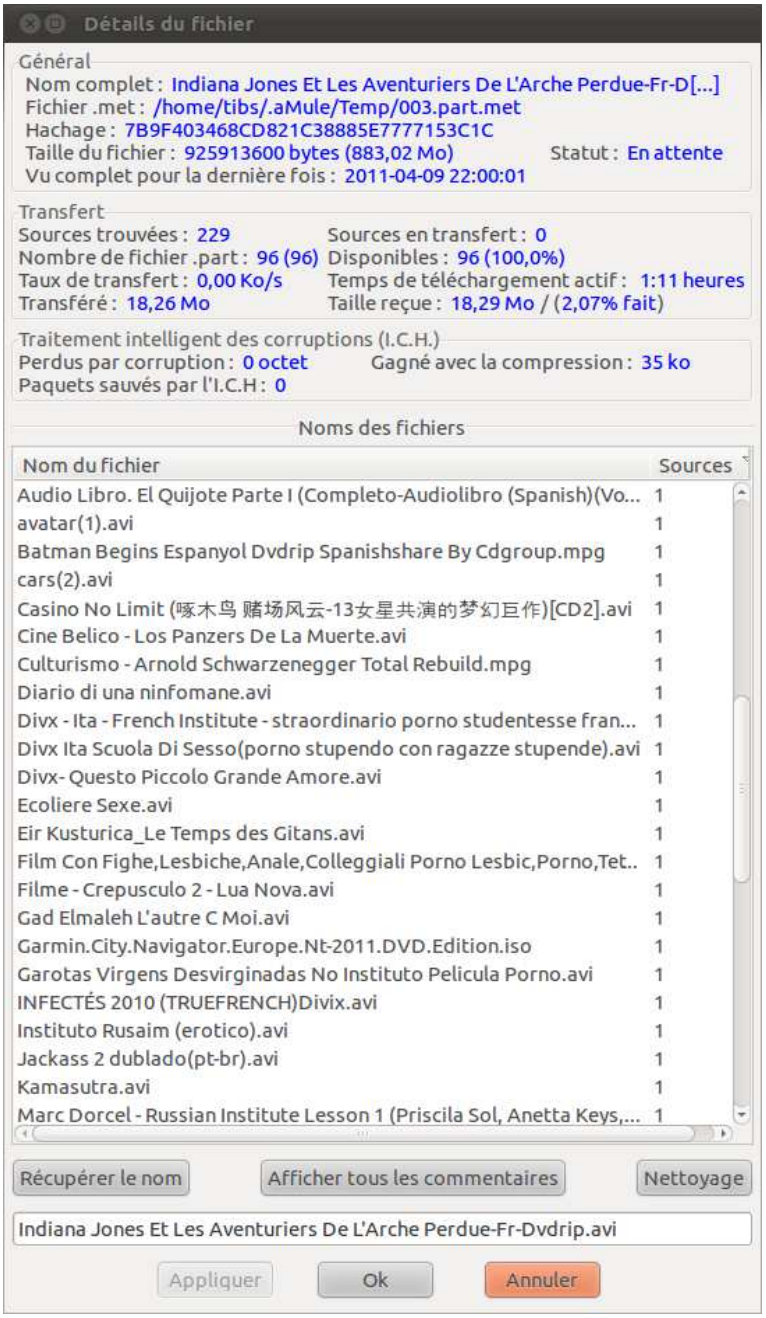


Figure 3.8. *Noms de fichier annoncés par les sources d'un fichier pollué*

Notre solution pour détecter la pollution consiste en l'analyse de la disparité lexicale des noms proposés par les différents pairs. Une approche similaire a été proposée par les auteurs de [LAT 08] et [GUI 09] pour lutter contre la corruption de l'indexation basée sur l'analyse des noms de fichiers en considérant les mots-clés communs aux différents noms de fichiers associés au contenu. Cependant, les métriques proposées n'étaient pas assez discriminantes pour détecter la pollution de manière fiable. En effet, il suffisait notamment qu'une seule source partage le fichier avec un nom différent pour que ce fichier soit considéré comme pollué, ce qui semble être difficilement applicable étant donné le nombre de sources des fichiers populaires. Plusieurs points doivent donc être améliorés pour en faire une solution viable contre la pollution, notamment la prise en compte du nombre de pairs affichant chaque nom de fichier trouvé ou encore l'obtention des différents noms sans exploration ni supervision du réseau.

3.4.2.2. *Métrique de détection*

Etant donné un fichier dont le contenu est identifié par son empreinte MD4, nous souhaitons savoir si celui-ci est fiable ou fait l'objet d'une pollution par falsification d'indexation en analysant les différents noms de fichier donnés par les sources contactées. Nous utilisons pour cela une métrique capable d'apprécier la similarité entre deux ensembles de mots comme la distance de Jaccard [MAN 99]. Soit X et Y deux ensembles de mots-clés, X étant l'ensemble des mots composant le nom de fichier choisi par l'utilisateur et Y étant l'ensemble des mots composant un des noms de fichier donné par les sources, l'indice de similarité de Tversky [TVE 77] *S(X,Y)* est une valeur entre 0 et 1 définie par la formule suivante :

$$S(X,Y) = \frac{|X \cap Y|}{|X \cap Y| + \alpha * |X - Y| + \beta * |Y - X|} \quad [3.2]$$

Pour détecter la pollution, nous utilisons un cas particulier de cette formule en fixant $\alpha = \beta = 0.5$, car aucun des deux noms ne peut être privilégié et considéré comme une référence, ce qui produit le coefficient de similarité de Dice [MAN 99] défini par :

$$S(X,Y) = \frac{|X \cap Y|}{|X \cap Y| + 0.5 * |X - Y| + 0.5 * |Y - X|} = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad [3.3]$$

Si les deux noms de fichiers sont identiques, le coefficient résultant est de 1, si les deux noms sont complètement disjoints, le coefficient résultant vaut 0. Pour attribuer un indice de pollution *P* à un fichier *X*, nous calculons la différence entre 1

et la moyenne de l'ensemble des indices de similarité obtenus par les différents noms de fichiers trouvés, soit :

$$P(X) = 1 - \frac{\sum_{i=1}^{n} S(X, Y_i)}{n} \quad [3.4]$$

3.4.3. *Quantification et caractérisation de la pollution*

3.4.3.1. *Exploration des fichiers*

Afin de quantifier la pollution affectant le réseau P2P KAD, nous avons collecté les informations présentées ci-avant pour de nombreux fichiers. L'accès aux noms de fichiers contradictoires est très coûteux puisqu'il nécessite :

- 1) une recherche de fichiers sur la DHT ;
- 2) une recherche de sources potentielles sur la DHT ;
- 3) l'établissement d'une connexion TCP pour chaque source réelle.

Nous limitons pour cela notre collecte d'information à l'étude des 100 contenus les plus téléchargés de 2010 selon l'un des principaux sites d'indexation de Torrents[5] recevant plus de 100 millions de recherches par an[6]. Pour chacun des contenus, nous collectons les différents noms associés aux 20 fichiers les plus populaires trouvés dans KAD, c'est-à-dire dont le nombre de sources estimé à l'issue de la recherche par mots-clés est le plus important. Nous estimons donc la pollution du réseau à partir d'une base de 2 000 fichiers parmi les plus populaires.

Une fois le téléchargement d'un fichier lancé, la découverte des sources réelles est progressive et prend un certain temps que nous souhaitons estimer avant de lancer la collecte des différents noms pour les 2 000 fichiers. Sur un échantillon de 150 fichiers, nous avons ainsi mesuré l'évolution du nombre de sources réelles obtenues pendant une heure et dont le graphique 3.9 illustre les dix premières minutes. Il apparaît qu'en moyenne plus de 97 % des sources trouvées à l'issue d'une heure sont obtenues dès 300 secondes. Pour collecter les données nécessaires à la quantification de la pollution, nous instrumentons un client KAD recherchant séquentiellement les mots-clés de notre jeu de données, lançant pour chacun d'eux le téléchargement des 20 fichiers les plus populaires et enregistrant, après 300 secondes d'attente, les différents noms de fichiers obtenus par interrogation des sources réelles trouvées.

5. www.kickasstorrents.com.
6. http://torrentfreak.com.

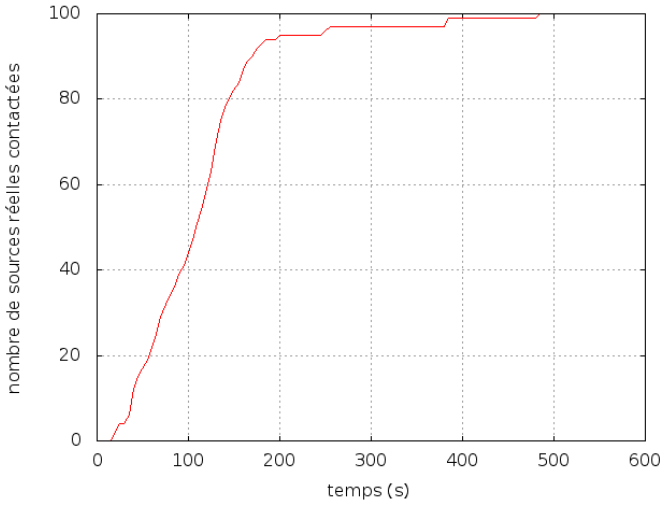


Figure 3.9. *Nombre moyen de sources réelles trouvées en fonction du temps*

3.4.3.2. *Résultats*

Après avoir collecté les différents noms possibles des 2 000 fichiers, nous avons appliqué notre métrique à chacun d'eux. Le graphique 3.10 montre la distribution du nombre de fichiers en fonction des indices de pollution obtenus. Nous remarquons que la métrique est bien discriminante en établissant la majorité des scores aux extrémités de l'échelle ce qui facilite l'établissement de seuils de détection. Nous définissons ainsi trois ensembles à partir de deux seuils. Les fichiers dont le score de pollution est : supérieur à 0.7 sont considérés comme pollués ; entre 0.7 et 0.3 ils sont estimés comme peut-être pollués ; inférieur à 0.3, ils sont perçus comme sains.

L'application de ces seuils à l'ensemble des 2 000 fichiers populaires étudiés donne la répartition présentée figure 3.11. Plus de 41 % des contenus populaires sont ainsi clairement pollués par la falsification d'indexation. De plus, pour 21 % des fichiers, aucune source réelle n'a pu être trouvée malgré le fait que les fichiers apparaissent avec un nombre élevé de sources estimées lors de la recherche par mot-clé, ceci correspond donc à corruption du mécanisme d'indexation par de fausses références [LIA 06]. En considérant les deux formes de pollution, plus de 62 % des fichiers sont pollués, bien que la falsification d'indexation soit la plus néfaste des deux formes de pollutions constatées. Seuls 29 % des fichiers peuvent être considérés comme parfaitement sains ; le doute subsistant par rapport à la métrique pour moins de 10 % des fichiers.

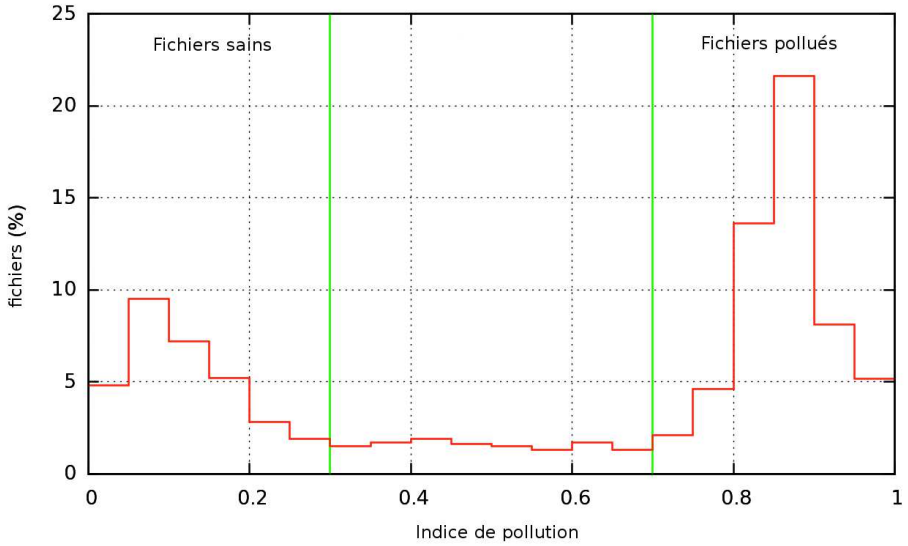


Figure 3.10. *Distribution du nombre de fichiers en fonction de l'indice de pollution obtenu*

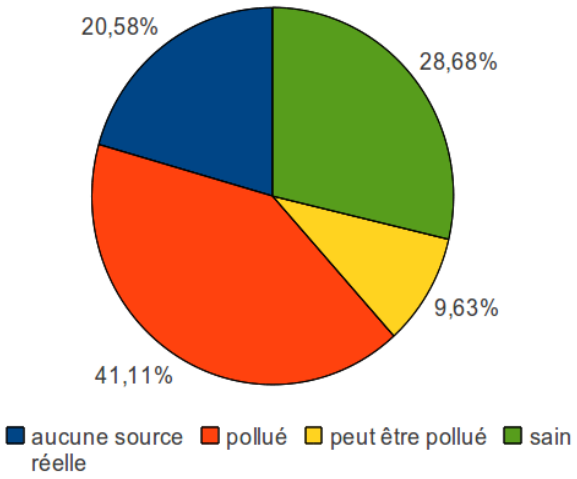


Figure 3.11. *Quantification de la pollution des contenus dans le réseau P2P KAD*

Afin d'évaluer la précision de notre métrique, nous avons procédé à l'analyse des noms de fichiers obtenus par dix experts. Chacun d'eux a évalué 100 fichiers dont des sources réelles ont été trouvées, soit 10 % de l'échantillon global, ils ont classé, tout comme la métrique, chaque fichier parmi les trois catégories : pollué, peut-être pollué ou sain. Les experts ont eu cependant très peu recours à la catégorie « peut-être pollué » qui représente moins de 1 % des fichiers analysés. Parmi les fichiers classés dans cette catégorie par la métrique, 78 % ont été jugés sains par les experts et 12 % pollués. L'avis des experts nous permet également de calculer le taux d'erreur de notre métrique. Ainsi, les faux positifs sont les fichiers considérés comme pollués par la métrique et sains par les experts. A l'inverse, les faux négatifs sont les fichiers considérés comme sains par la métrique et pollués par les experts. Le tableau 3.3 donne les valeurs des taux d'erreur. Au regard du taux d'erreur très faible constaté, la métrique que nous proposons détecte très bien la pollution par falsification d'indexation. La proportion de fichiers sains est en revanche sous-évaluée, car une partie d'entre eux est considérée comme « peut être pollué » et un faible taux de faux positifs subsiste.

% de faux positifs	% de faux négatifs
3,78	0,88

Tableau 3.3. *Taux d'erreur de la métrique de détection de la falsification d'indexation*

Nous sommes cependant capables d'identifier les rares cas de faux positifs. Il s'agit de versions localisées de films dont le titre a été traduit. Ainsi, il peut arriver qu'une partie des utilisateurs nomme le fichier par le titre original du film alors qu'une autre partie utilise la version localisée du titre. Le tableau 3.4 illustre ce problème pour un fichier ayant mal été identifié par notre méthode. Du point de vue de la métrique, les deux noms ne semblent pas liés, car les mots-clés sont complètement différents, seule une connaissance sémantique des mots traduits ou des différents noms existant pour un même film permet de considérer que le contenu est sain, ce qu'ont bien identifié les experts.

Nom de fichier choisi (X)	Nom de fichier majoritaire des sources (Y)
Il Cigno Nero Sub Ita.avi	Black.Swan.2010.DVDSCR.XviD-TiMKY.avi

Tableau 3.4. *Exemple de faux positif lors de la détection de la pollution*

Nous avons également regardé comment la pollution affectait chacun des 100 contenus étudiés au travers des 20 fichiers considérés pour chacun d'eux. Il apparaît que tous les contenus sont affectés avec entre 5/20 et 20/20 fichiers pollués par la falsification d'indexation. Ainsi, le contenu le moins pollué est « *the big bang theory*

» avec 5 fichiers considérés pollués, alors que les 20 fichiers sélectionnés pour « *avatar* » étaient pollués. La pollution n'affecte pas uniquement des contenus soumis aux droits d'auteur puisque le mot-clé « *ubuntu* » désignant une distribution Linux sous licence GNU GPL est également largement contaminé avec 15 fichiers pollués sur les 20.

Pour terminer, nous avons évalué la proportion de fichiers affectés par le mélange d'indexation dont au moins un des titres proposés indique un contenu pédophile ou pornographique. Nous avons recherché pour cela les mots-clés pédophiles explicites, et des mots-clés à caractère pornographiques utilisés par le logiciel ProCon[7] pour réaliser un contrôle parental. Les résultats sont présentés dans le tableau 3.5. Il apparaît que la majorité des fichiers pollués par la falsification d'indexation sont potentiellement des fichiers pornographiques. Plus grave, 8.8 % des fichiers pollués peuvent contenir au final un contenu pédophile ce qui, étant donné la popularité des fichiers étudiés, implique une diffusion importante de ces contenus illégaux à l'insu des utilisateurs et peut générer de nombreux faux positifs en cas de supervision inadaptée.

	% de fichiers pédophiles	**% de fichiers pornographiques**
pour tous fichiers	3,6	21,1
pour les fichiers pollués	8,8	55,7

Tableau 3.5. *Taux d'erreur de la métrique de détection de la falsification d'indexation*

3.5. Conclusion

3.5.1. *Bilan*

Nous nous sommes intéressés tout au long de ce chapitre aux conséquences de l'identification non fiable des pairs et des contenus proposés par les systèmes P2P de partage de fichiers. Nous avons tout d'abord relevé deux problèmes majeurs décrits dans l'état de l'art que sont l'attaque Sybil et la pollution des contenus. La première attaque résulte de l'absence d'identification forte des pairs autorisant des pairs malveillants à créer de multiples identités, à des emplacements spécifiques, et leur permettant de détourner le mécanisme d'indexation du système P2P. La seconde attaque provient principalement de l'absence de liens entre le contenu et le nom d'un fichier permettant à un pair malveillant de falsifier l'indexation d'un contenu en le liant à de multiples noms. Ces deux attaques sont critiques pour la sécurité des

7. https://addons.mozilla.org/fr/firefox/addon/procon-latte.

utilisateurs du réseau qui peuvent être amenés à télécharger des contenus illégaux (par exemple à caractère pédophile) à leur insu. Ces attaques portent également atteinte à la vie privée des utilisateurs et aux performances du réseau.

Pour la première fois, nos travaux ont permis de constater la réalité de ces attaques au sein d'un réseau P2P largement déployé, à savoir KAD. Nous avons ainsi été en mesure de détecter de nombreux pairs suspects dont les positionnements anormaux traduisent probablement des attaques. Pour cela, nous avons utilisé deux analyses, considérant d'abord la proximité entre les identifiants des pairs, puis entre les identifiants de pairs et ceux de contenus précalculés. Il s'avère que de nombreux pairs suspects sont présents dans le réseau et de nombreux contenus populaires sont ciblés. Ces contenus populaires sont également largement affectés par la nouvelle forme de pollution que nous avons mise en évidence qui est la falsification d'indexation. En effet, nous avons élaboré une métrique fiable, basée sur le coefficient de similarité de Dice, permettant de détecter cette pollution et nous avons mesuré que plus de 41 % des fichiers, parmi 2 000 des plus populaires, sont victimes de cette forme de pollution.

Les diverses solutions existant à ce jour pour pallier ces problèmes sont malheureusement encore insuffisantes.

3.5.2. *Solutions possibles*

En ce qui concerne l'attaque Sybil, les effets de celle-ci peuvent être limités par l'adjonction de mécanismes de protection visant à limiter le nombre de Sybils et donc leur capacité de nuisance. De nombreuses solutions ont été imaginées pour limiter cette attaque [URD 11] par l'ajout de contraintes à l'insertion d'un pair dans le réseau. Ces protections peuvent être classées en trois catégories : les contraintes structurelles [CAS 02, SIN 04], les vérifications de ressources [ROW 07, YU 06] et la certification des identifiants [AIE 08, LES 08]. Cependant, les systèmes développés montrent de nombreuses limites quand on souhaite protéger un réseau déjà déployé : la plupart imposent des contraintes devant exister dès la création de la DHT (réseaux sociaux [YU 06], certification distribuée [LES 08]), d'autres font intervenir une entité centrale de certification des identifiants [CAS 02] qui est incompatible avec le paradigme pair-à-pair, ou introduisent un surcoût important (puzzle cryptographique [ROW 07]). La solution que nous avons proposée [CHO 11] en plus d'être extrêmement efficace, est la seule répondant aux contraintes de KAD – qui doit rester entièrement distribué et garder une rétrocompatibilité avec les anciennes versions des clients. Cependant, même si les effets de l'attaque Sybil peuvent être contenus, la cause première ne peut être résolue sans le recours à un système de certification authentifiant les pairs du réseau comme l'affirment [DOU 02, CAS 02] ou encore [STE 07a].

De même, plusieurs solutions ont été envisagées pour lutter contre la pollution. L'article [LIA 05a] classe les solutions potentielles en deux catégories : celles nécessitant le téléchargement du fichier (vérification manuelle ou semi-automatique) et celles pouvant être appliquées préalablement (systèmes de confiance ou réputation) et permettent de protéger les utilisateurs en détectant la pollution *a priori*. Les articles [LIA 05b] et [LIA 06] détectent les sous-réseaux diffusant la pollution sur Kazaa après une exploration complète du réseau. Ceci étant impossible à réaliser à l'échelle d'un client, les adresses IP des pairs suspects sont alors diffusées sous forme de liste noire. Les mécanismes de réputation [COS 07], appliqués aux pairs ou aux contenus, sont inefficaces, car ils peuvent être facilement corrompus. Enfin, certaines solutions reposent sur l'implication des pairs indexant pour détecter la pollution et sont malheureusement pleinement vulnérables à l'attaque Sybil. Ces solutions ne peuvent donc être envisagées que conjointement à d'autres protections contre l'attaque Sybil telles que [CHO 11]. Ainsi, l'article [SHI 09] présente une approche basée sur le filtrage collaboratif des fichiers pollués par les pairs qui sont alors chargés de filtrer les annonces erronées pour les contenus qu'ils indexent. De même, [DEN 09] s'attaque à la source de la falsification d'indexation en proposant d'authentifier les noms donnés aux contenus. Chaque donnée indexée par un pair est signée par la source originale qui crée une preuve de l'association, mais cette méthode n'est pas en revanche rétrocompatible.

Nous sommes donc forcés de constater que le problème de l'identification des contenus distribués et de leur source peut être difficilement résolu au sein de l'Internet actuel étant donné le difficile recours aux autorités de certification centralisées et leur contradiction par nature avec le paradigme pair-à-pair. Les systèmes P2P ont, malgré certaines faiblesses, néanmoins montré leur grande efficacité pour répondre aux besoins de diffusion des contenus numériques. Les limites actuelles motivent à plus long terme le développement de nouvelles approches de diffusion de l'information telles que l'approche de *Content Centric Networking* [JAC 07]. Dans ce paradigme, tout contenu peut être répliqué n'importe où dans le réseau et diffusé de manière pair-à-pair par des nœuds non fiables, l'élément adressable au cœur du réseau devenant le contenu lui-même et non son hôte dont il n'est plus nécessaire de connaître l'adresse IP. Chaque contenu est ainsi directement adressable sur le réseau, sécurisé et autosuffisant : son intégrité peut être vérifiée et il peut être authentifié par sa source originelle tout comme le nom qui lui a été associé [SME 09]. Tirant parti des enseignements sur les limites des réseaux actuels (IP et P2P), l'approche proposée par CCN introduit la notion de confiance dans le contenu lui-même et non plus dans le système qui le délivre, ce qui semble être la solution la plus prometteuse à ce jour pour résoudre les différents problèmes d'identification affectant la distribution des contenus numériques par des systèmes P2P.

3.6. Bibliographie

[ACO 07] Acosta W., Chandra S., « Trace driven analysis of the long term evolution of gnutella peer-to-peer traffic », *Proceedings of the Passive and Active Measurment Conference* (PAM'07), Louvain-la-neuve, Belgique, 2007.

[AIE 08] Aiello L.M., Milanesio M., Ruffo G., Schifanella R., « Tempering Kademlia with a Robust Identity Based System », *Proceedings of the 8t^h International Conference on Peer-to-Peer Computing* (P2P'08), p. 30–39, IEEE Computer Society, Washington, Etats-Unis, 2008.

[CAS 02] Castro M., Druschel P., Ganesh A., Rowstron A., Wallach D.S., « Secure routing for structured peer-to-peer overlay networks », *SIGOPS Oper. Syst. Rev.*, vol. 36, n°SI, p. 299-314, ACM, 2002.

[CHO 09] Cholez T., Chrisment I., FEstor O., « Evaluation of Sybil Attacks Protection Schemes in KAD », *3rd International Conference on Autonomous Infrastructure, Management and Security - AIMS 2009 Scalability of Networks and Services*, Enschede, Pays-Bas, « Lecture Notes », *Computer Science*, vol. 5637, p. 70-82, Université of Twente, Springer, 2009.

[CHO 10] Cholez T., Chrisment I., Festor O., « Monitoring and Controlling Content Access in KAD », *International Conference on Communications - ICC 2010*, IEEE, Capetown Afrique du Sud, mai 2010.

[CHO 11] Cholez T., Hénard C., Chrisment I., Festor O., Doyen G., Khatou N.R., « Détection de pairs suspects dans le réseau pair à pair KAD », *6^e Conf. sur la Sécurité des architectures réseaux et systèmes d'information* (SAR-SSI 2011), Financement GIS - 3SGS - Projet ACDAP2P, IEEE, La Rochelle, France, mai 2011.

[CHO 12] Cholez T., Chrisment I., Festor O., Doyen G., « Detection and Mitigation of Localized Attacks in a Widely Deployed P2P Network », *Journal of Peer-to-Peer Networking and Applications, Special Issue on Experimental Evaluation of Peer-to-Peer Applications*, Springer, New York, mai 2012.

[CHR 05] Christin N., Weigend A.S., Chuang J., « Content availability, pollution and poisoning in file sharing peer-to-peer networks », *Proceedings of the 6th ACM conference on Electronic commerce* (EC'05), New York, Etats-Unis, ACM, p. 68-77, 2005.

[COS 07] Costa C., Almeida J., « Reputation Systems for Fighting Pollution in Peer-to-Peer File Sharing Systems », *Proceedings of the 7th IEEE International Conference on Peer-to-Peer Computing* (P2P'07), Washington, Etats-Unis, IEEE Computer Society, p. 53-60, 2007.

[DEN 09] Deng L., He Y., Xu Z., « Combating Index Poisoning in P2P File Sharing », *Proceedings of the 3rd International Conference and Workshops on Advances in Information Security and Assurance* (ISA'09), Berlin, p. 358-367, Springer-Verlag, Heidelberg, 2009.

[DOU 02] Douceur J.R., « The Sybil Attack », *Revised Papers from the First International Workshop on Peer-to-Peer Systems* (IPTPS '01), Londres, Royaume Uni, Springer-Verlag, p. 251-260, 2002.

[DUM 05] Dumitriu D., Knightly E., Kuzmanovic A., Stoica I., Zwaenepoel W., « Denial-of-service resilience in peer-to-peer file sharing systems », SIGMETRICS *Perform. Eval. Rev.*, vol. 33, n°1, p. 38-49, ACM, 2005.

[FAL 07] Falkner J., Piatek M., John J.P., Krishnamurthy A., Anderson T., « Profiling a million user dht », *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement* (IMC 07), New York, Etats-Unis, ACM, p. 129-134, 2007.

[GUI 09] Loup Guillaume J., Latapy M., Magnien C., Valadon G., *Content Rating and Fake Detection System, Measurement and Analysis of P2P Activity Against Paedophile Content Project*, Rapport, Laboratoire d'Informatique de Paris 6 - CNRS Université Pierre et Marie Curie, 2009.

[IPO 09] IPOQUE, « Internet Study 2008/2009 », www.ipoque.com/resources/internet- studies/ Internet- study- 2008_2009, 2009.

[JAC 07] Jacobson V., Mosko M., Smetters D., Garcia-Luna-Aceves J.J., « Content-Centric Networking : Whitepaper Describing Future Assurable Global Networks », *Response to* DARPA RFI SN07-12, 2007.

[KLE 04] Klemm A., Lindemann C., Vernon M.K., Waldhorst O.P., « Characterizing the query behavior in peer-to-peer file sharing systems », *Proceedings of the 4th ACM SIGCOMM Conference on Internet measurement* (IMC'04), New York, Etats-Unis, ACM, p. 55-67, 2004.

[KOH 09] Kohnen M., Leske M., Rathgeb E.P., « Conducting and Optimizing Eclipse Attacks in the KAD Peer-to-Peer Network », *Proceedings of the 8th International IFIP-TC 6 Networking Conference* (Networking'09), Berlin, p. 104-116, Springer-Verlag, Heidelberg, 2009.

[KON 07] Konrath M.A., Barcellos M.P., Mansilha R.B., « Attacking a Swarm with a Band of Liars : evaluating the impact of attacks on BitTorrent », *Proceedings of the Seventh IEEE International Conference on Peer-to-Peer Computing* (P2P'07), Washington, Etats-Unis, p. 37-44, IEEE Computer Society, 2007.

[LAT 08] Latapy M., Magnien C., Valadon G., *First report on database specification and access including content rating and fake detection system*, Rapport, *Measurement and Analysis of P2P Activity Against Paedophile Content Project*, Laboratoire d'Informatique de Paris 6 - CNRS Université Pierre et Marie Curie, 2008.

[LEE 06] Lee U., Choi M., Cho J., Sanadidi M.Y., Gerla M., « Understanding Pollution Dynamics in P2P File Sharing », *Proceedings of the 5th International Workshop on Peer-toPeer Systems* (IPTPS'06), Santa Barbara, Etats-Unis, 2006.

[LES 08] Lesueur F., Mé L., Tong V.V.T., « A Sybil-Resistant Admission Control Coupling SybilGuard with Distributed Certification », *Proceedings of the 4th International Workshop on Collaborative Peer-to-Peer Systems* (COPS), Rome, Italie, IEEE Computer Society, juin 2008.

[LIA 05a] Liang J., Kumar R., Xi Y., Ross K.W., « Pollution in P2P File Sharing Systems », p. 1174-1185, *IEEE Infocom*, 2005.

[LIA 05b] Liang J., Naoumov N., Ross K.W., « Efficient Blacklisting and Pollution-Level Estimation in P2P File-Sharing Systems », Cho K., Jacquet P. (dir.), de *Lecture Notes in Computer Science*, AINTEC, vol. 3837, p. 1-21, Springer, New York, 2005.

[LIA 06] Liang J., Naoumov N., Ross K.W., « The Index Poisoning Attack in P2P File Sharing Systems », *IEEE Infocom*, IEEE Computer Society, 2006.

[LOC 10] Locher T., Mysicka D., Schmid S., Wattenhofer R., « Poisoning the Kad Network », *11th International Conference on Distributed Computing and Networking* (ICDCN), Calcutta, Inde, janvier 2010.

[MAN 99] Manning C.D., Schütze H., *Foundations of statistical natural language processing*, MIT Press, Cambridge (MA), 1999.

[MAY 02] Maymounkov P., Mazières D., « Kademlia : A Peer-to-Peer Information System Based on the XOR Metric », *Revised Papers from the1st International Workshop on Peer-to-Peer Systems* (IPTPS'01), p. 53-65, Londres, Royaume Uni, Springer-Verlag, 2002.

[MEM 09] Memon G., Rejaie R., Guo Y., Stutzbach D., « Large-Scale Monitoring of DHT Traffic », *International Workshop on Peer-to-Peer Systems* (IPTPS), Boston, Etats-Unis, avril 2009.

[MON 11] Montassier G., Cholez T., Doyen G., Khatoun R., Chrisment I., Festor O., « Content Pollution Quantification in Large P2P networks : a Measurement Study on KAD », *Proceedings of the 11th IEEE International Conference on Peer-to-Peer Computing* (P2P'11), Kyoto, Japon, IEEE, août 2011.

[NAO 06] Naoumov N., Ross K., « Exploiting P2P systems for DDoS attacks », *Proceedings of the 1st international Conference on Scalable information systems* (InfoScale'06), p. 47, New York, Etats-Unis, ACM, 2006.

[ROW 07] Rowaihy H., Enck W., Mcdaniel P., Porta T.L., « Limiting Sybil Attacks in Structured P2P Networks », *IEEE Infocom*, p. 2596-2600, IEEE Computer Society, 2007.

[SHI 09] Shin K., Reeves D.S., Rhee I., Song Y., *Winnowing : Protecting P2P Systems Against Pollution By Cooperative Index Filtering*, Tech report n°TR-2009-2, Department of Computer Science, North Carolina State University, 2009.

[SIN 04] Singh A., Castro M., Druschel P., Rowstron A., « Defending against eclipse attacks on overlay networks », *Proceedings of the 11th workshop on ACM SIGOPS European workshop* (EW 11), p. 21, New York, Etats-Unis, ACM, 2004.

[SIN 06] Singh A., Ngan T.W., Druschel P., Wallach D.S., « Eclipse Attacks on Overlay Networks : Threats and Defenses », *Proceedings, 25th IEEE International Conference on Computer Communications* (INFOCOM 2006), p. 1-12, Barcelone, Espagne, 2006.

[SME 09] Smetters D., Jacobson V., *Securing Network Content*, Rapport, PARC, octobre 2009.

[STE 07a] Steiner M., En-Najjary T., Biersack E.W., « Exploiting KAD : possible uses and misuses », *SIGCOMM Comput. Commun. Rev.*, vol. 37, n°5, p. 65-70, ACM, 2007.

[STE 07b] Steiner M., En-Najjary T., Biersack E.W., « A global view of KAD », *ACM SIGCOMM Internet Measurement Conference* (IMC 2007), San Diego, Etats-Unis, 23-26 octobre, 2007.

[TUV 10] Tuvian U., Porat L., *Hash collision Attack Vectors on the eD2k P2P Network*, Rapport, Interdisciplinary Center, Herzliya, Israël, 2010.

[TVE 77] Tversky A., « Features of Similarity », *Psychological Review*, vol. 84, p. 327-352, 1977.

[URD 11] Urdaneta G., Pierre G., Van Steen M., « A Survey of DHT Security Techniques », *ACM Computing Surveys*, vol. 43, n°2, juin 2011, www.globule.org/publi/ SDST_acmcs 2009.html.

[WAN 08] Wang P., Tyra J., Chan-Tin E., Malchow T., Kune D.F., Hopper N., Kim Y., « Attacking the Kad network », *Proceedings of the 4th international conference on Security and privacy in communication networks* (SecureComm 08), p. 1-10, New York, Etats-Unis, ACM, 2008.

[WOL 10] Wolchok S., Halderman J.A., « Crawling BitTorrent DHTs for Fun and Profit », *Proceeding 4th USENIX Workshop on Offensive Technologies* (WOOT), Washington, Etats-Unis, août 2010.

[YU 06] Yu H., Kaminsky M., Gibbons P.B., Flaxman A., « SybilGuard : defending against sybil attacks *via* social networks », *Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications* (SIGCOMM'06), p. 267-278, New York, Etats-Unis, ACM, 2006.

[YU 09] Yu J., Fang C., Xu J., Chang E.-C., Li Z., « ID Repetition in KAD », dans Schulzrinne H., Aberer K., Datta A. (dir.), *Peer-to-Peer Computing*, p. 111-120, IEEE, 2009.