

# Détection optimale à base de distribution laplacienne quantifiée

Cathel ZITZMANN, Rémi COGRANNE, Lionel FILLATRE, Igor NIKIFOROV, Florent RETRAINT et Philippe CORNU

ICD - LM2S - Université de Technologie de Troyes (UTT) - UMR STMR - CNRS  
12, rue Marie Curie - B.P. 2060 - 10010 Troyes Cedex - France  
prenom.nom@utt.fr

**Résumé** – Cet article s’intéresse à la détection d’informations cachées dans les bits de poids faible des coefficients DCT d’une image JPEG. Ce problème est difficile pour deux raisons principales : d’abord, le nombre de coefficients DCT utilisables pour dissimuler les bits cachés est aléatoire ; ensuite, la compression des images JPEG induit une forte quantification des coefficients DCT. Le test statistique optimal proposé dans ce papier est basé sur la modélisation des coefficients DCT avec une distribution laplacienne quantifiée. Il est adapté aux variations du pas de quantification entre les coefficients DCT et il prend explicitement en compte le caractère aléatoire du nombre de coefficients DCT utilisables. Ce test maximise la probabilité de détecter des informations cachées en garantissant un faible nombre de fausses alarmes.

**Abstract** – This paper focuses on the detection of hidden information embedded in bits of the DCT coefficients of a JPEG image. This problem is difficult for two main reasons: first, the number of DCT coefficients used to conceal the hidden bits is random, then the JPEG image compression induces a strong quantization of DCT coefficients. The optimal statistical test proposed in this paper is based on the modeling of DCT coefficients with a quantified Laplacian distribution. It is adapted to variations between the quantization DCT coefficients and it explicitly takes into account the randomness of the number of DCT coefficients used. This test maximizes the probability of detecting hidden information by ensuring a low number of false alarms.

## 1 Introduction et contribution

Détecter la présence d’une information cachée dans un médium numérique est un défi important pour les forces de l’ordre. Dans ce contexte, il est nécessaire que les probabilités de fausse alarme et de non détection soient maîtrisées. Cet article se propose d’étudier la détection de la méthode de remplacement des bits de poids faible (LSB) dans des images naturelles compressées au format JPEG. Lors de la compression, l’image est découpée en blocs de  $8 \times 8$  pixels et la transformée en cosinus discrète (DCT) est appliquée à chaque bloc. Dans chaque bloc figurent 64 coefficients correspondant à différentes fréquences. Les coefficients DCT sont ensuite quantifiés à l’aide d’une matrice de quantification qui figure dans l’en-tête du fichier. Lors de l’insertion d’un message, l’algorithme de dissimulation remplace les LSB des coefficients DCT quantifiés avec les bits du message. Afin de construire le détecteur, il faut disposer d’un modèle de la distribution des coefficients DCT quantifiés ainsi que d’un modèle statistique de la dissimulation. Ainsi, un détecteur basé sur la modélisation laplacienne de la distribution des coefficients DCT est proposé. Dans la littérature, plusieurs approches existent pour détecter ce type de dissimulation [6], et certaines utilisent tous les coefficients DCT utilisables de l’image sans tenir compte des fréquences auxquelles correspondent ces coefficients [4]. Dans cet article, nous proposons des bases théoriques pour un détecteur prenant en compte les spécificités de l’algorithme de dissimulation et du format de compression JPEG.

Les contributions de cet article sont :

- Une fonction de décision adaptée à chaque fréquence en tenant compte du pas de quantification, du paramètre de la distribution laplacienne et du nombre de coefficients utiles.
- La synthèse d’un détecteur adapté à un nombre de mesures aléatoires pour un modèle laplacien des vecteurs de coefficients DCT,

## 2 Positionnement du problème

Soit une image découpée en  $n$  blocs de 64 coefficients. Soit  $V_k = \{v_{k,1}, \dots, v_{k,n}\}$  un vecteur de longueur  $n$  composé de la valeur du  $k^{\text{ième}}$  coefficient de chaque bloc (voir Fig. 1), avec  $v_{k,i} \in \{\lfloor \frac{-1024}{\Delta_k} \rfloor, \dots, \lceil \frac{1023}{\Delta_k} \rceil\}$ ,  $\forall k = 2, \dots, 64$  et  $\Delta_k$  le pas de quantification associé au  $k^{\text{ième}}$  coefficient.

L’algorithme Jsteg [5] insère un message dans les bits de poids faible des coefficients DCT d’une image au format JPEG. Le principe de l’algorithme (Fig. 2) est de remplacer les bits de poids faible par les bits du message.

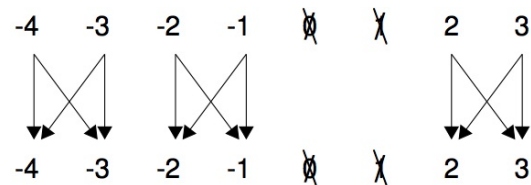


FIGURE 2 – Insertion dans les coefficients DCT  $v_{k,i}$ .

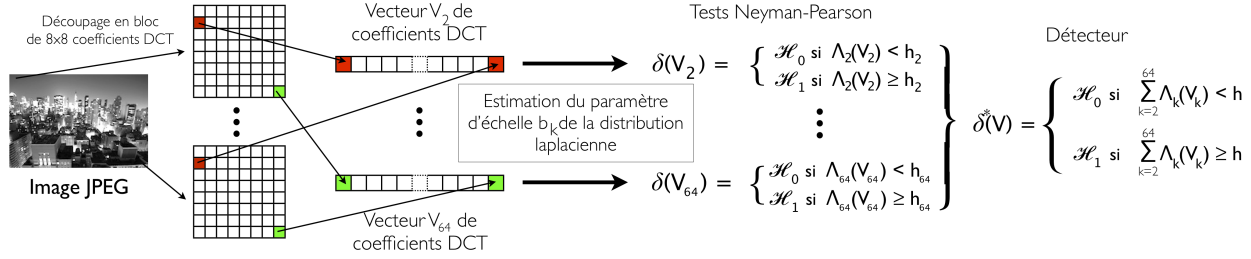


FIGURE 1 – Vers un détecteur optimal

La compression JPEG a tendance à créer de nombreux 0 dans les hautes fréquences. Si un 1 était inséré dans un 0, il prendrait la valeur 1 et la présence d'un message pourrait être décelée facilement. De ce fait, le message ne sera pas inséré dans les 0. Comme l'insertion se fait dans des paires de valeurs (voir Fig. 2), les 1 ne sont donc pas utilisés non plus. De plus, le premier coefficient n'est pas utilisé pour l'insertion, car sa modification impacterait visuellement l'image. Ainsi, \$V\_1\$ n'est pas utilisé pour la détection et les 0 et les 1 sont des valeurs interdites. Par conséquent, le nombre de coefficients utiles de chaque vecteur est représenté par une variable aléatoire \$n\_k \leq n\$, qui dépend de l'image analysée (principalement de son contenu) et de la matrice de quantification. On peut montrer que \$n\_k\$ est une variable aléatoire positive à valeurs entières qui suit une loi binômiale \$\mathcal{B}(n, p\_k^\*)\$, où \$p\_k^\*\$ est la probabilité qu'un coefficient \$v\_{k,i}\$ du vecteur \$V\_k\$ soit utilisable, i.e. \$v\_{k,i} \notin \{0, 1\}\$. Par convention, les \$n\_k\$ premiers éléments du vecteur sont des valeurs utilisables et les \$n - n\_k\$ derniers sont les valeurs interdites.

Afin de construire un détecteur, nous avons besoin de connaître la distribution des coefficients DCT d'une image naturelle. Il s'agit d'un problème intéressant et ouvert dans de nombreux domaines. En effet, Reininger et Gibson [2], Lam et Goodman [1] ainsi que Smoot et Rowe [3] proposent une distribution laplacienne, Eggerton et Srinath une distribution de Cauchy et Müller une distribution gaussienne généralisée. Dans ce papier, comme dans d'autres articles [7] traitant de la détection d'informations cachées, il a été choisi de modéliser les coefficients DCT d'une image naturelle avec une distribution laplacienne. La distribution laplacienne continue est donnée par : \$f\_{b,a}(x) = \frac{1}{2b} \exp\left(\frac{-|x-a|}{b}\right)\$, \$\forall x \in \mathbb{R}\$. Il a été montré [1] que la distribution des coefficients DCT est symétrique de moyenne \$a = 0\$. Ainsi, seul le paramètre d'échelle \$b\$ de la distribution laplacienne est à estimer.

La densité de probabilité \$\text{Lap}(b\_k, \Delta\_k)\$ du \$k^{\text{ième}}\$ coefficient DCT modélisé par une distribution laplacienne "discrète" quantifiée avec le pas de quantification \$\Delta\_k\$ est obtenue par :

$$\begin{aligned} P(y = v_{k,i} | y \notin \{0, 1\}) &= q_{v_{k,i}}(b_k, \Delta_k) \\ &= q_{-v_{k,i}}(b_k, \Delta_k) \\ &= \exp\left(\frac{-\Delta_k |v_{k,i}|}{b_k}\right) \sinh\left(\frac{\Delta_k}{2b_k}\right). \end{aligned}$$

Le but du problème de détection est de construire un test statistique qui détecte les informations cachées à partir de l'en-

semble des vecteurs de coefficients DCT en supposant qu'ils sont tous indépendants. L'impact sur le détecteur des phénomènes de corrélation (entre les vecteurs et dans les vecteurs) sera étudié ultérieurement. À présent, tous les calculs seront conditionnés par le fait que les \$v\_{k,i}\$ doivent être des valeurs utilisables.

Enfin, la densité de probabilité notée \$\text{Lap}(b\_k, \Delta\_k, R)\$ d'un coefficient DCT après insertion par la méthode d'écrasement des LSB avec un taux d'insertion \$R\$ vaut :

$$\begin{aligned} P^R(y = v_{k,i} | y \notin \{0, 1\}) &= q_{v_{k,i}}^R(b_k, \Delta_k) \\ &= \left(1 - \frac{R}{2}\right) q_{v_{k,i}}(b_k, \Delta_k) + \frac{R}{2} q_{v_{k,i} - \eta(v_{k,i})}(b_k, \Delta_k), \end{aligned}$$

$$\text{avec } \eta(v_{k,i}) = \begin{cases} 1 & \text{si } v_{k,i} \text{ est impair,} \\ -1 & \text{si } v_{k,i} \text{ est pair.} \end{cases}$$

Le taux d'insertion correspond à la probabilité qu'un LSB soit remplacé par un bit du message.

Lors de l'analyse d'une image, deux hypothèses sont envisageables : \$\mathcal{H}\_0 = \{\text{l'image est saine}\}\$ et \$\mathcal{H}\_1 = \{\text{l'image est stéganographiée avec un taux d'insertion } R\}\$. Le test d'hypothèses considéré s'écrit sous la forme suivante :

$$\begin{aligned} \mathcal{H}_0 &: \{v_{k,i} \sim \text{Lap}(b_k, \Delta_k), \forall k=2, \dots, 64, \forall i=1, \dots, n_k\} \\ \mathcal{H}_1 &: \{v_{k,i} \sim \text{Lap}(b_k, \Delta_k, R), \forall k=2, \dots, 64, \forall i=1, \dots, n_k\}. \end{aligned} \quad (1)$$

### 3 Détecteur optimal

On cherche le test le plus puissant dans la classe \$K\_{\alpha\_0} = \{\delta^\* : P\_{\mathcal{H}\_0}(\delta^\*(V) = \mathcal{H}\_1) < \alpha\_0\}\$, où \$V = \{V\_2, \dots, V\_{64}\}\$ permettant de résoudre le problème (1). Le paramètre \$b\_k\$ est estimé au préalable, et \$\Delta\_k\$ figure dans l'en-tête du fichier, c'est pourquoi nous considérons ces paramètres connus. Ce test est donné par le lemme de Neyman-Pearson et consiste en la règle de décision suivante :

$$\delta^*(V) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda^*(V) = \sum_{k=2}^{64} \Lambda_k(V_k) < h \\ \mathcal{H}_1 & \text{si } \Lambda^*(V) = \sum_{k=2}^{64} \Lambda_k(V_k) \geq h \end{cases} \quad (2)$$

où

$$\begin{aligned} \Lambda_k(V_k) &= \Lambda_k(v_{k,1}, \dots, v_{k,n_k}) \\ &= \sum_{i=1}^{n_k} \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(\frac{\Delta_k}{b_k} \underbrace{(|v_{k,i}| - |v_{k,i} - \eta(v_{k,i})|)}_{=\zeta_{k,i}}\right)\right) \end{aligned}$$

$$\text{avec } \zeta_{k,i} = \begin{cases} 1 & \text{si } \begin{cases} v_{k,i} > 0 \text{ et } v_{k,i} \text{ impair,} \\ v_{k,i} < 0 \text{ et } v_{k,i} \text{ pair,} \end{cases} \\ -1 & \text{si } \begin{cases} v_{k,i} < 0 \text{ et } v_{k,i} \text{ impair,} \\ v_{k,i} > 0 \text{ et } v_{k,i} \text{ pair,} \end{cases} \end{cases} \quad (3)$$

et le seuil  $h$  est la solution de l'équation  $P_0(\Lambda^*(V) \geq h) = \alpha_0$ . Le théorème suivant établit les probabilités statistiques du test de Neyman-Pearson.

**Théorème 1** *Le test (2) atteint la puissance  $\beta_{\delta^*}$  lorsque  $n$  tend vers l'infini, pour un seuil de décision  $h$  fixé par  $\alpha_0$  :*

$$\beta_{\delta^*} \simeq 1 - \Phi\left(\frac{\sigma_0^*}{\sigma_1^*} \Phi^{-1}(1 - \alpha_0) + \frac{\mu_0^* - \mu_1^*}{\sigma_1^*}\right)$$

et  $h \simeq \sigma_0^* \Phi^{-1}(1 - \alpha_0) + \mu_0^*$ , où  $\mu_j^*$  et  $\sigma_j^{*2}$  sont définis dans la section suivante.

Ce théorème est asymptotique car  $n$  est proportionnel au nombre de pixels de l'image, et pour une grande image,  $n$  tend vers l'infini.

## 4 Schéma de démonstration du théorème

Afin de prendre en compte la présence de valeurs interdites dans les vecteurs de coefficients DCT, il est nécessaire de considérer la variable aléatoire  $n_k$  dans le calcul des performances du détecteur.

L'expression du log-rapport de vraisemblance pour une fréquence est :

$$\begin{aligned} \Lambda_{k,j}(v_{k,1}, \dots, v_{k,n_k}) &= \sum_{i=1}^{n_k} X_{k,i} \\ &= \sum_{i=1}^{n_k} \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(\frac{\Delta_k}{b_k} (\zeta_{k,i})\right)\right) \end{aligned}$$

avec  $\zeta_{k,i}$  la variable aléatoire qui vaut 1 avec la probabilité  $p_{k,j}$  et qui vaut  $-1$  avec la probabilité  $1 - p_{k,j}$  sous l'hypothèse  $\mathcal{H}_j$ ;  $p_{k,j}$  étant la probabilité que  $v_{k,i}$  soit positif et impair ou négatif et pair sous l'hypothèse  $\mathcal{H}_j$  (voir eq. 3).

Ainsi,

$$\begin{aligned} \Lambda_{k,j}(v_{k,1}, \dots, v_{k,n_k}) &= n_k a_k + (c_k - a_k) \sum_{i=1}^{n_k} Y_{k,i} \\ &= n_k a_k + (c_k - a_k) S_{n_k}, \end{aligned}$$

avec

$$a_k = \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(-\frac{\Delta_k}{b_k}\right)\right),$$

et

$$c_k = \log\left(1 - \frac{R}{2} + \frac{R}{2} \exp\left(\frac{\Delta_k}{b_k}\right)\right).$$

Les variables aléatoires  $Y_{k,i} = \frac{X_{k,i} - a_k}{c_k - a_k}$  sont indépendantes, identiquement distribuées et suivent une loi de Bernoulli d'espérance  $p_{k,j}$  et de variance  $p_{k,j}(1 - p_{k,j})$ .

Posons  $S_{n_k} = \sum_{i=1}^{n_k} Y_{k,i}$ . Comme les  $Y_{k,i}$  suivent la loi de Bernoulli  $\mathcal{B}(1, p_{k,j})$ , et que la variable aléatoire  $n_k$  suit la loi binômiale  $\mathcal{B}(n, p_{k,j}^*)$ , il peut être montré que  $S_{n_k}$  suit la loi binômiale  $\mathcal{B}(n, p_k^* p_{k,j})$ .

Comme  $n$  est grand et que  $p_k^* p_{k,j}$  et  $1 - p_k^* p_{k,j}$  sont de même ordre de grandeur, la loi binômiale est très bien approchée par une loi normale. Ainsi, sous l'hypothèse  $\mathcal{H}_j$ , pour chaque fréquence, nous obtenons :

$$\frac{\frac{\Lambda_{k,j} - n_k a_k}{c_k - a_k} - n p_k^* p_{k,j}}{\sqrt{n p_k^* p_{k,j} (1 - p_k^* p_{k,j})}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

et

$$\frac{\Lambda_{k,j} - \mu_{k,j}}{\sigma_{k,j}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1),$$

avec

$$\mu_{k,j} = n p_k^* [a_k (1 - p_{k,j}) + c_k p_{k,j}]$$

et

$$\sigma_{k,j}^2 = n p_k^* [a_k^2 (1 - p_k^*) + (c_k - a_k)^2 p_k^* p_{k,j} (1 - p_k^* p_{k,j})],$$

où  $\xrightarrow[n \rightarrow \infty]{\mathcal{L}}$  désigne la convergence en loi lorsque  $n$  tend vers l'infini.

La somme de plusieurs lois normales indépendantes étant une loi normale, pour le détecteur optimal, sous l'hypothèse  $\mathcal{H}_j$ , on obtient :

$$\frac{\Lambda^*(V) - \mu_j^*}{\sigma_j^*} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1),$$

où  $\mu_k^* = \sum_{k=2}^{64} \mu_{k,j}$  et  $\sigma_j^{*2} = \sum_{k=2}^{64} \sigma_{k,j}^2$ .

## 5 Expérimentations numériques

Afin d'illustrer les résultats des sections précédentes, nous utilisons l'image `mandrill` (Fig. 3) qui fait partie d'une base de données de référence<sup>1</sup>. Il s'agit d'une image couleur de taille  $512 \times 512$  au format `TIFF`. Elle a été compressée au format `JPEG` en utilisant `imagemagick` avec une qualité de compression de 70 et mise en dégradé de gris.

Dans la pratique, on souhaite analyser une image, c'est à dire, déterminer si un message a été inséré dans l'image avec un taux d'insertion  $R = 0.2$ . La méthode repose sur différentes étapes.

Tout d'abord, les vecteurs  $V_k = \{v_{k,1}, \dots, v_{k,n}\}$  contenant les coefficients DCT de chaque fréquence  $k$  sont extraits ainsi que les pas de quantification  $\Delta_k$  correspondants. Ensuite, les paramètres  $b_k$  de la distribution laplacienne sont estimés par la méthode du maximum de vraisemblance (Fig. 4) :

$$\hat{b}_k = \frac{1}{n} \sum_{i=1}^n |v_{k,i}|.$$

Cet estimateur a de bonnes propriétés : son biais est nul et sa variance faible.

1. <http://sipi.usc.edu/database/>

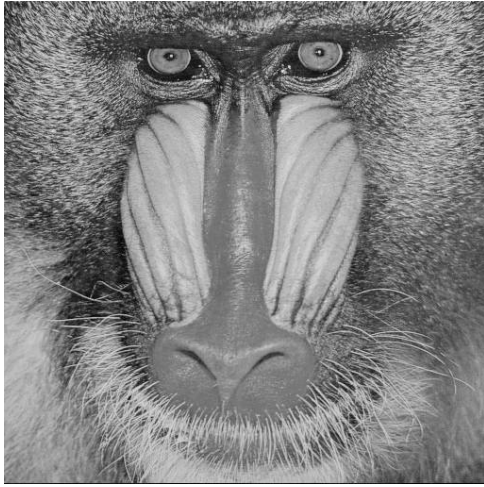


FIGURE 3 – Mandrill.jpg en dégradé de gris.

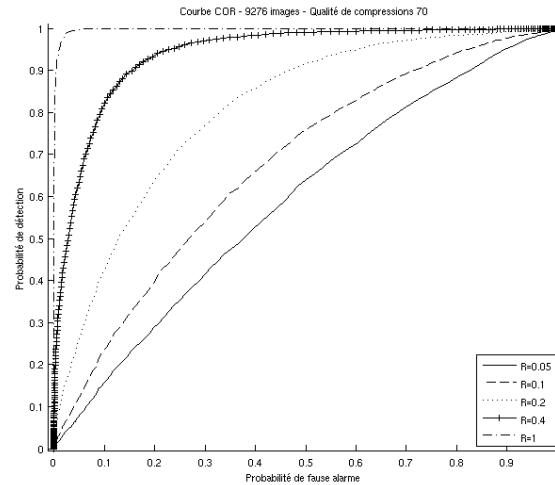


FIGURE 5 – Courbe COR

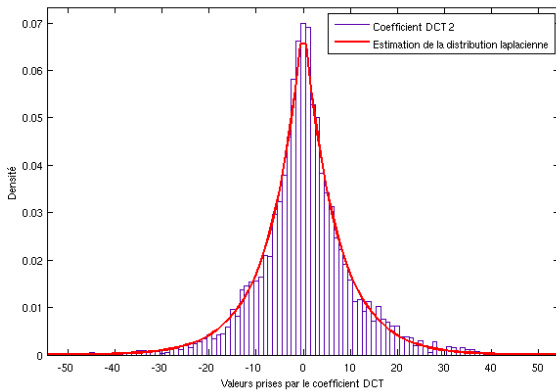


FIGURE 4 – Distribution empirique du 2<sup>ème</sup> coefficient DCT et estimation de la distribution laplacienne correspondante.

Dans cet exemple, les fréquences utilisées par le détecteur optimal sont celles dont la probabilité  $p_k^*$  est supérieure à 0.6. Cette sélection permet de ne pas prendre en compte des vecteurs dont le paramètre de la distribution laplacienne aurait pu être mal estimé.

Le détecteur a ensuite été testé sur la base de données Boss<sup>2</sup> contenant 10000 images de taille  $512 \times 512$  au format PGM. Ces images ont été converties avec `imagemagick` au format JPEG avec une qualité de compression de 70. La figure 5 représente la courbe COR obtenue pour les images de la base Boss pour différents taux d'insertion  $R$ . Le cas  $R = 0.05$  correspond au pire cas au sens où il s'agit du taux d'insertion minimum que l'on cherche à détecter en pratique et le cas  $R = 1$  quant à lui est le cas le plus favorable. Les 10000 images de la base ont été testées, mais seulement 9276 comportent au moins un vecteur  $V_k$  qui respectait la condition  $p_k^* > 0.6$ .

## 6 Conclusion et perspectives

Cet article permet de poser les bases théoriques d'un détecteur d'informations cachées dans les bits de poids faible des coefficients DCT d'une image JPEG. Ce travail pourra être poursuivi par l'étude de l'impact des corrélations entre les vecteurs de coefficients DCT sur ce détecteur, et par la modélisation statistique d'algorithmes de dissimulation plus complexes.

## Références

- [1] E.Y. Lam et al. A Mathematical Analysis of the DCT Coefficient Distributions for Images. *IEEE Trans. on IP*, 9(10), 2000.
- [2] R. Reininger et al. Distributions of the Two-Dimensional DCT Coefficients for Images. *IEEE Trans. on Comm.*, 31(6), 1983.
- [3] S.Smoot et al. Study of DCT Coefficient Distributions. In *In Proc. of the SPIE Symp. on EI*, 1996.
- [4] T. Zhang et al. A Fast and Effective Steganalytic Technique Against JSteg-like Algorithms. In *SAC*, pages 307–311, 2003.
- [5] D. Upham. Jsteg (1993). <http://www.nic.funet.fi/pub/crypt/steganography/jpeg-jsteg-v4.diff.gz>.
- [6] Andreas Westfeld. Generic adoption of spatial steganalysis to transformed domain. In *Information Hiding*. 2008.
- [7] M. Wu, Z. Zhu, and S. Jin. A New Steganalytic Algorithm for Detecting Jsteg. In *ICCNMC*, pages 1073–1082, 2005.