Codes correcteurs et sécurité en stéganographie pour les images et les vidéos

THESE : oct. 2010 – oct. 2013
financement ministériel

Marc Chaumont & William Puech
LIRMM (Laboratoire d'Informatique, de Robotique et Microélectronique de Montpellier)
Equipe ICAR, 161 rue Ada, 34392 Montpellier cedex 5 - France
Tel : +33 4.67.41.85.14, Fax : +33 4.67.41.85.00
Marc.Chaumont@lirmm.fr, William.Puech@lirmm.fr

Mots clefs : codes détecteurs et correcteurs d'erreurs, codes BCH et Reed-Solomon..., sécurité, image, vidéo.

L'idée du *codage matriciel* (en anglais : « Matrix encoding ») a été introduite en stéganographie par Cranddall [Crandall 1998] en 1998. La première implémentation a ensuite été proposée par Westfeld avec l'algorithme de stéganographie F5 [Westfeld 2001]. L'objectif est de transmettre un message au sein d'une image via la modification de l'image, mais avec la contrainte de minimiser le nombre de coefficients de l'image modifiés. Plus précisément, le *codage matriciel* consiste à détourner l'utilisation classique des codes détecteurs et correcteurs d'erreur en bloc. L'idée consiste du côté décodeur (c'est-à-dire à la réception de l'image) à calculer les syndromes de chaque bloc de coefficients à partir de la matrice de contrôle du code correcteur. Le syndrome correspond au message qui est contenu dans l'image. Toute l'astuce consiste donc, du côté codeur (c'est-à-dire à l'émission de l'image), à modifier l'image de sorte que les syndromes calculés au décodeur représentent le message et également de sorte que l'image soit le moins modifiée.

Depuis 2001, des codes plus performants ont été utilisés pour réaliser le *codage matriciel* : algorithme Modified Matrix Encoding : MME [Kim et al. 2007], algorithme FastBCH [Zhang et al. 2009], [Sachnev et al. 2009], algorithme basé Reed-Solomon (RS) [Fontaine and Galand 2009]... L'insertion est toujours basée sur le calcul de syndrome, mais les codes utilisés possèdent une efficacité d'insertion e = $\frac{\textit{nombre de bits du message}}{\textit{nombre de coefficients modifiés}}$ meilleure que celle du code de Hamming utilisé dans F5, pour un même nombre de bits insérés. Une autre évolution des *codes matriciels* a été également proposée à travers les « codes à papier mouillé » [Fridrich et al. 2005] et consiste à sélectionner les sites d'insertion du côté codeur, mais avec un décodeur ignorant les sites sélectionnés.

Depuis 2008, l'étude de la sécurité des schémas stéganographiques se clarifie (la sécurité pour la stéganographie consiste à être indétectable). La sécurité est évaluée :

- soit de manière aveugle (on ne connait pas l'algorithme de stéganographie) en calculant la distance de Fisher [Ker 2009] ou bien évaluée de manière empirique en comparant plusieurs schémas (F5, MME, PQe+PQt [Fridrich et al. 2007], FastBCH, RS…) avec le « classifieur » de l'état de l'art de Pevny et Fridrich [Pevny and Fridrich 2007] ou la méthode proposée dans [Pevny and Fridrich 2008].
- soit de manière ciblée (on connait l'algorithme de stéganographie) en déterminant les « traces » statistiques dues à l'insertion [Pevny et al. 2009].

Après un état de l'art des approches par *codage matriciel*, et par *code à papier mouillé* [Fridrich 2009], le candidat reprendra et poursuivra le travail autour des schémas basés syndrome : FastBCH [Zhang et al. 2009], [Sachnev et al. 2009], et Reed-Solomon (RS) [Fontaine and Galand 2009]... L'utilisation d'autres codes correcteurs pourra également être envisagée. La sécurité des schémas ainsi proposés sera ensuite évaluée de manière aveugle et/ou ciblée.

La thèse doit donc aboutir à la proposition d'améliorations de schémas existants et/ou la proposition de nouveaux schémas et elle doit également apporter des réponses aux questions suivantes : dans quelle mesure le codage matriciel et le codage par codes à papier mouillés est sûr, comment évaluer proprement la sécurité, comment faire plus sûr et est-ce possible ?

Une collaboration avec le laboratoire Coréen « Multmedia Security Lab », dirigé par Hyoung Joong Kim, est envisagée.

[Crandall 1998] R. Crandall: Some notes on steganography, Posted on Steganography Mailing List (1998), http://os.inf.tu-dresden.de/~westfeld/crandall.pdf.

[Westfeld 2001] A. Westfeld: "High capacity despite better steganalysis (F5 – a steganographic algorithm)". In: Moskowitz, I.S. (ed.) IH 2001. LNCS, vol. 2137, pp. 289–302. Springer, Heidelberg (2001).

[Kim et al. 2007] Y. Kim, Z. Duric, D. Richards: "Modified matrix encoding technique for minimal distortion steganography". In: Camenisch, J.L., Collberg, C.S., Johnson, N.F., Sallee, P. (eds.) IH 2006. LNCS, vol. 4437, pp. 314–327 (2007).

[Zhang et al. 2009] R. Zhang, V. Sanchev, H. J. Kim: "Fast BCH Syndrome Coding for Steganography". In: Katzenbeisser, S. and Sadeghi, A.-R (Ed.) Information Hiding 2009, IH'2009, LNCS 5806, pp. 48-58, 2009, Springer-Verlag Berlin Heidelberg 2009.

[Sachnev et al. 2009] V. Sachnev, H.J. Kim and R. Zhang: "Security Less Detectable JPEG Steganography Method Based on Heuristic Optimization and BCH Syndrome Coding", The 11th ACM Workshop on Multimedia and Security, *MM&Sec'09,* September 7–8, 2009, Princeton, New Jersey, USA.

[Zhang et al. 2008] W. Zhang, X. Zhang, and S. Wang. "Maximizing steganographic embedding efficiency by combining Hamming codes and wet paper codes". In K. Solanki, K. Sullivan, and U. Madhow, editors, *Information Hiding, 10th International Workshop*, Lecture Notes in Computer Science, pages 60–71, Santa Barbara, CA, June 19–21, 2008. Springer-Verlag, New York.

[Fridrich et al. 2007] Fridrich, Pevny, T., Kodovsky, J.: Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities. In: Dittmann, J., Fridrich, J. (eds.) Proceedings of the 9th ACM Multimedia & Security Workshop, Dallas, TX, September 20–21, pp. 3–14 (2007).

[Pevny and Fridrich 2007] Pevny, T., Fridrich, J.: Merging Markov and DCT features for multi-class JPEG steganalysis. In: Delp, E.J., Wong, P.W. (eds.) Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, January 29 – February 1, vol. 6505, pp. 3-1–3-14 (2007).

[Pevny and Fridrich 2008] Pevny, T., Fridrich, J.: Benchmarking for Steganography In : K. Solanki, K. Sullivan, and U. Madhow (Eds.): Proceeding International Hiding IH'2008, LNCS 5284, pp. 251–267, Springer-Verlag Berlin Heidelberg 2008.

[Fontaine and Galand 2009] C. Fontaine and F. Galand: "How Reed-Solomon Codes Can Improve Steganographic Schemes", Hindawi Publishing Corporation EURASIP Journal on Information Security Volume 2009, Article ID 274845, 10 pages doi:10.1155/2009/274845.

[Fridrich et al. 2005] J. Fridrich, M. Goljan, and D. Soukal. "Efficient wet paper codes". In M. Barni, editor, *Proceedings, Information Hiding, 7th International Workshop, IH 2005, Barcelona, Spain, June 6-8, 2005*, LNCS. Springer, Berlin, 2006.

[Pevny et al. 2009] T. Pevny, P. Bas and J. Fridrich, Steganalysis by Subtractive Pixel Adjacency Matrix, Proc. ACM Multimedia and Security Workshop, Princeton, NJ, September 7–8, pp. 75–84, 2009.

[Fridrich 2009] J Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, 464 pages, Cambridge University Press, 1st edition, December 31, 2009.

[Ker 2009] Ker, A.: Estimating Steganographic Fisher Information in real images. In: Proc. 11th Information Hiding Workshop 2009.