Подход за защита на електронни документи срещу несанкциониран достъп

Георги Върбанов, Гео Кунев

Approach to protection of electronic documents against unauthorized access: *In this article we present a modified algorithm for protection of electronic documents using watermarks. The idea is to assign to each document a unique identification number of the server, specifying users' rights for downloading. The proposed method, using watermarking allows certain corrections of errors in the watermark detection as well as enhanced opportunities for building a message with a greater length and low rates of modulation*

Key words: *Watermarks, Error Correction Code.*

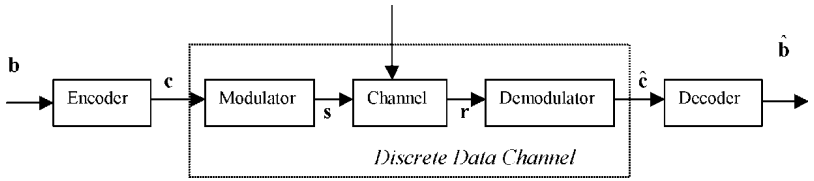
ВЪВЕДЕНИЕ

В статията предлагаме модифициран алгоритъм за защита на електронни документи използвайки воден знак. Основната идея е присвояването на уникален номер на всеки документ, които се изтегля от сървър делегираш правата на отделните потребители за изтегляне на съответния документ. Предложения метод използва за влагането на воден знак един метод познат ни от комуникациите с възможност на корекция на грешки при извличането и възстановяването на водния знак в противовес на вграждане на по-дълго съобщение в документа.Използуват се възможностите на кодовете с работещи с разредени матрици от данни, които постигат много добри резултати(LDPC).

Тези документи са налични само в pdf формат. Всички известни документи са представени или преобразувани в този формат(doc,txt, jpg, png, bmp и други). Няма ограничения в използването на криптиране за по-добра сигурност. Вграждането на водния знак се използува за менажмент на изтеглянето на електроните документи от сървъра.

ИЗЛОЖЕНИЕ

На фигура.1 е показана блоковата схема на примерна комуникационна система работеща с корекция на грешките при предаване. Подобна схема се използува за влагане на воден знак. В сравнение със стандартната схема за влагане на воден знак има 2 нови блока-кодер и декодер. Съответно в предавателната(маркиращата) и декодиращата(приемната страна) Предавателя добавя битове за корекция на грешки, а декодера проверява и редуцира такива ако са възникнали при извличането на водния знак.



Фиг. 1. Комуникационна система с корекция на грешки

Двоичната поредица е маркиращото съобщение. То преминава през целия предавателен и приемен тракт, след като се кодира в двоична последователност C. Вектора на C е записан в изображението използващ някои от алгоритмите за поставяне на воден знак чрез модулиране на сигнали, които симулират преминаване на сигнал през комуникационен канал.

Симулацията на канално разпространение се полага от това, че изображението може да бъде подложено на най-различни атаки за разрушаване на вградения воден

знак. Полученият шум в декодера се дължи именно на тези атаки. В резултат се получава възтановена последователност R. Има 2 възможни решения на демодулатора твърдо и меко(софт) .При първото се получава 0 и 1 на изхода, а при второто се получават по няколко нива за 0 и 1 съответно. В статията се използва по гъвкавото решение [3][4][7].

Това се постига чрез кавантоване на демодулирания сигнал на 8 нива. За алгоритъмът за кодиране и декодиране за решение на задачата е използван намиращ-широко приложение в JSM технологиите за разпространение на слаби сигнали. Блоковата схема е показана на фигура 4.

Алгоритъм за влагане на воден знак

Вмъкване на воден знак

Стъпка1 генериране на двоична последователност b с дължина T_k.

Стъпка 2 това е кодирането с добавяне на битове за проверка b с дължина върху k-брои блкове T_k (n,k) чрез $LDPC$. Кодираната последователност е C с дължина T_n.

Стъпка 3 повторение на операцията всеки кодиран бит се повтаря ch пъти C_C дължина генерирана $T_n * ch = T_C$.

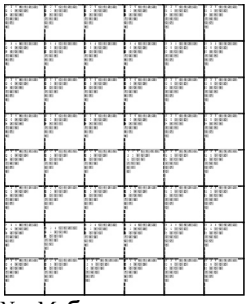
Стъпка 4 модулация на получената последователност в спектъра на генерира се T_k ± 1 псевдо случайна последователност рс с дължина T_C използваща ключ K_i умножен по B_C с T_k Рс за получаване на водния знак W_C (това представляват персоналния ключ IDs и данни).

Стъпка 5 реализация на DWT до 2 ниво и избор на LL потдомеин за маркирането фигура 3.

Стъпка 6 генериране на маска за влагане на водния знак W_C чрез използване на ключ $K2$ в резултат на което се получава маска W_m.

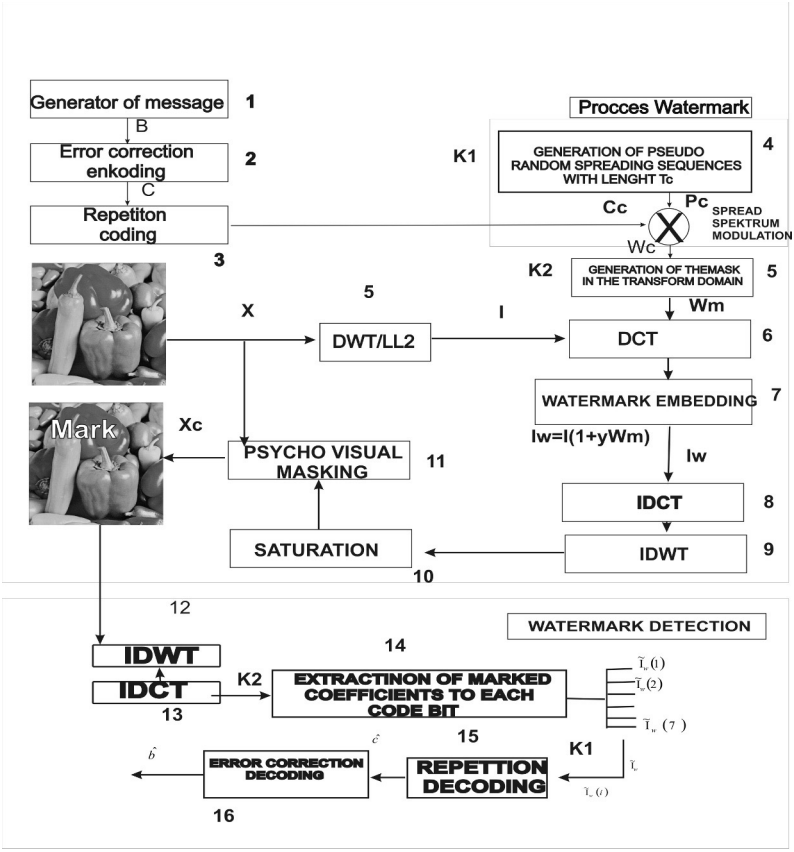
1	3	4	10	11	21	23	
2	5	9	12	20			
6	8	13	19				
7	14	18					
15	17						
16							

Фиг. 2. Маркирани коефициенти в блок от 8x8 в регион от DCT



Фиг. 3. $N \times M$ **блока с размер 8x8 в** DCT **област**

Маркират се по 16 коефицента(в тях са включени и битовете за проверка) от всеки блок избрани псевдослучайно намиращи се в нискочестотния потдиапазон в зиг заг. Всеки от посочените коефицентите се формира, като разлика формирана, както при Koch and Zhao[9]. Във всеки блок от 8x8 се селектират 16 честотни коефициенти и се разполагат върху целия домеин LL от wavelet разпределението. На фигура 3 е показано разпределението по блокове върху целия LL потдомеин по псевдо-случаен принцип.



Фиг. 4. Блокова схема на алгоритъма на влагане на воден знак

Стъпка 7 вграждане на водния знак W_m в трансформираната област I. За вграждането на водния знак W_m използваме добавящ алгоритъм:

$$I_m = I(1+\gamma W_M), |\gamma| < 1 \quad (1)$$

Стойноста на се избира с компромисен вариант между стабилност на водния знак и допустими изкривявания на изображението. По-високи стойности подобряват здравината на знака, но увеличават внесения шум.

Стъпка 8 е инверсо преобазуване в честотната област $IDCT$.

Стъпка 9 е обратно вавалет преобразуване $IDWT$.

Стъпка 10 е елиминиране на насищането за стойности по-големи от 255 се присвоява максималната 255 и за по-малки от 0 се присвоява 0.

Стъпка 11 е отчитане на влиянието на психовизуалните особености на човешкото зрение чрез измерване и оценка на качеството маскирането в пространственото разпространение на сигнала, т.е. можем да си представим, че маскираното изображение X_w преминава през психовизуална система "Human Visual System" (HVS).

Механизъм за откриване на водния знак

Кодовете работещи с разредени матрици "Low Density Parity Check" ($LDPC$)[4][5] са линеини кодове по оношение на проверката за четност на вгражданите матрици от данни H, които имат много малко значещи елементи равни на 1, като повечето са 0. $A(d, p, y)$ матрица с проверка на четност и има ρ елемнта еквивалентни на 1 във всяка колона и във всеки един ред, където y е поредния номер на реда и d е неговата дължина реципрочно на колоните е (dp/y) и стойноста на кода е $(y-p)/y$.

Извличане на маркираните коефицценти

Има различни алгоритми за извличане на водния знак. В случая се придържаме към схемата разработена от Coch and Zhao.[2][9]

Алгоритъма работи с псевдослучайно избрани коефиценти в 8x8 DCT блокове, в които се селектират 2 случайни коефицента в средночестотния диапазон.

С fb е означен избрания случаен блок 8x8 DCT, а fb(m1,n1), fb(m2,n2c) са означени селектираните коефиценти:

$$\mathrm{A}_b = \left|f_b\left(m_{1,}, n_{1,}\right)\right| - \left|f_b\left(m_{2,}, n_{2,}\right)\right| \tag{2}$$

За вграждането на един бит от поредицата на водния знак двойката на селектираните битоове се модифицира, като дистанция:

$$\Delta_H = \left\{\begin{matrix} \geq q \\ \leq -q \end{matrix} ako \begin{matrix} w=1 \\ w=0 \end{matrix}\right\} \tag{3}$$

Правото 8x8 точково DCT преобразуване е дефинирано в [3],[8], където q е параметър контролиращ силата на влагане съответно извличане на съответния бит и обратно косиносово DCT(IDCT) 8x8 точково преобразование е определено в [3][9]:

$$\mathrm{I}(u,v) = \frac{\zeta(u)}{2} \cdot \frac{\zeta(v)}{2} \sum_{k=0}^{7} \sum_{l=0}^{7} \mathrm{X}(k,l) \cdot \cos\left(\frac{(2\kappa+1)u\pi}{16}\right) \cdot \cos\left(\frac{(2l+1)v\pi}{16}\right) \tag{4}$$

А инверсното 8x8 DCT(IDCT) е определено като:

$$\mathrm{X}(k,l) = \sum_{u=0}^{7} \sum_{v=0}^{7} \frac{\zeta(u)}{2} \cdot \frac{\zeta(v)}{2} \mathrm{I}(u,v) \cdot \cos\left(\frac{(2\kappa+1)u\pi}{16}\right) \cdot \cos\left(\frac{(2l+1)v\pi}{16}\right) \tag{5}$$

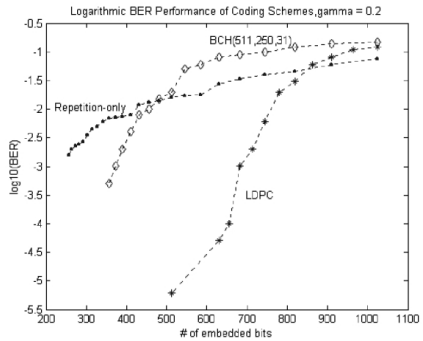
$$\textit{където} \quad k,l,u,v \quad \varepsilon = \{0,1,2,3,4,5,6,7\}; u \quad \zeta(u) = \begin{pmatrix} \frac{1}{\sqrt{2}} for & u=0 \\ 1 for & u>0 \end{pmatrix}, \zeta(v) = \begin{pmatrix} \frac{1}{\sqrt{2}} for & v=0 \\ 1 for & v>0 \end{pmatrix} \tag{6}$$

Това установява дължината на вектора ch ,-като $\tilde{\mathrm{I}}^1{}_w(i)$, където I е между 1 и T_n.

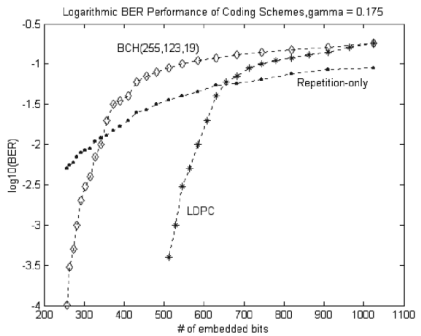
$$\tilde{\mathrm{I}}_w = \left[\tilde{\mathrm{I}}_w(1), \tilde{\mathrm{I}}_w(2) .. \tilde{\mathrm{I}}_w(\mathrm{T}_n)\right] where \quad \tilde{\mathrm{I}}_w(i) = \left[\tilde{\mathrm{I}}^1{}_w(i), \tilde{\mathrm{I}}^2{}_w(i) .. \tilde{\mathrm{I}}^{ch}{}_w(i)\right] \tag{7}$$

ЗАКЛЮЧЕНИЕ

Експерименталните резултати са показани на фигури 5 и 6 и те са валидни само за RAW изображение без да вазимодейиства с друг алгоритъм за маркиране. За ралиазация е използвана програмата Matlab 7.



Фиг. 5.



Фиг. 6.

Експериментите са проведени върху познатите ни като стандартни за тестване изображения на Lena, Barbon, Peppers и други

Не се симулира правенето на анализ на пакетите на принт сървъра. Целта е да се постигнат добри резултати при извличането на водния знак. Вижда се, че се получават добри резултати около 500 бита дължина на вградената последователност и се постига корекция до 2 % на грешката. Съответно BER е ~0.2,.което е близо до доказаното от Shannon ограничение [1][6].

В таблица 1 са показани реузлтатите от внесеното съотношение сигнал шум в изображението от вложения знак и неговото съотношение на шум "Peak signal to noise ratio" ($PSNR$) and "Watermark to document ratio" (WDR)

Таблица 1. $PSNR$ and WDR Results

Y	***PSNR (dB)***	***WDR (dB)***
0.2	*41.65*	*-36.74*
0.225	*40.65*	*-35.74*
0.25	*39.75*	*-34.84*

Теоретичната грешка BER е еквивалентна на:

$$Q\left(\sqrt{SNR}\right) = 1/2 erfc\left(\sqrt{SNR/2}\right) \quad (8)$$

ЛИТЕРАТУРА

[1] Balado F., J. R. Hernandez and F. Perez, Approaching the capacity limit in image watermarking, Signal Processing vol.81 number 6 June (2001), pp. 1215-1238

[2] Cox, J., M. L. Miller and J. A. Bloom, Digital Watermarking, Morgan Kauffman Publishers, ISBN 1-55860-714-5, 2002

[3] Gallager, R. G., "Low Density Parity Check Codes", IRE Trans. Information Theory, vol. IT-8, pp. 21-28, Jan 1962

[4] Darbon, J., B. Sankur and H. Maitre, " Error Correcting Code Performance For Watermark Protection, Security and Watermarking of Multimedia Contents III, SPIE,vol. 4314, San Jose (CA,USA), 20-25 Jan 2001

[5] MacKay, D J.C, "Good Error-Correcting Codes Based on Very Sparse Matrices", IEEE Trans. Information Theory, vol. 45, pp. 399-431, March 1999.

[6] Swanson, M.D,B.Zhu and A.H.Tewfik, "Transparent robust image watermarking", SPIE Conf. on Visual Communications and Image Proc., vol.3, pages 211-214, ISBN: 0-7803-3259-8, 16-19 Sep 1996.

[7] Zinger, S., Z. Jin, H. Maitre and B. Sankur, "Optimization of Watermarking Performances Using Error Correcting Codes and Repetition", CMS'2001: Communications and Multimedia Security, May 2001, 229-240, Darmsadt, Germany.

[8] Wilson, S. G., Digital Modulation and Coding,Prentice Hall 1996, ISBN 0-13-210071-1.

[9] Zhao, J. and E. Koch,"Embedding Robust Labels into Images for Copyright Protection", Proc. Of Intellectual Congress on Intellectual Property Rights for Specialized Information Knowledge and New Technologies, , pp.242-251, Aug. 21-25, 1995.,Vienna, Austria

За контакти:

гл. ас. Георги Б. Върбанов, Катедра “Компютърна наука и технологии”, Технически университет “-Варна, тел.: 052-383614, e-mail: gvarbanov@gmail.com,

гл. ас. д-р Гео В. Кунев Катедра “Компютърна наука и технологии” тел:052 383638 geo_qnew@hotmail.com

Докладът е рецензиран.