DOI УДК 004.056.2

С.Б. Вепрев, С.А. Нестерович

ИСПОЛЬЗОВАНИЕ РИСУНКОВ ФОРМАТА *BMP* ДЛЯ СКРЫТОЙ ПЕРЕДАЧИ ТЕКСТОВОЙ ИНФОРМАЦИИ

В статье описан метод скрытой записи текстовой информации с помощью изображения. Приведено описание основной идеи и метода современной стеганографии для скрытой записи текстовой информации в теле рисунка формата bmp.

Ключевые слова: *информационная безопасность, стеганография, защита данных.*

S.B. Veprev, S.A. Nesterovich

USING BMP FORMAT FIGURES FOR HIDDEN TRANSMISSION OF TEXT INFORMATION

The article describes the method of hidden recording of textual information using an image. The description of the main idea and the method of modern steganography for the hidden recording of textual information in the body of the figure in bmp format is given.

Keywords: *information security, steganography, data protection.*

Стеганография, латино-греческое сочетание слов, означающих в совокупности «тайнопись», является одним из самых интересных способов защиты информации [1]. Стеганография предполагает, что передаваемый текст «растворяется» в сообщении большего размера с совершенно «посторонним» смыслом. Но если взять и извлечь из него некоторые символы по определенному закону, например каждый второй или третий, и т.д., получим вполне конкретное тайное сообщение.

Существует множество стеганографических идей. Так, например, можно составить текст [2], в котором следует читать только заглавные буквы, которые и несут секретную информацию. Можно текст написать молоком. Тогда его можно прочитать, подогрев лист письма и т.д.

Следует отметить, что использование стеганографических методов защиты данных в совокупности с шифрованием может дать значительный эффект. В данной статье рассматривается возможность использования графических объектов в формате bmp для передачи данных в документе с использованием рисунка. Выберем 24-разрядный формат bmp. Формат файлов bmp представлен на рис. 1.

Изображение в теле файла сохраняется построчно ***снизу вверх***. Для хранения каждой строки выделяется кратное 4 количество байт. В незначащих байтах хранится «мусор». При хранении изображения в формате True Color каждому пикселу соответствуют три последовательных байта, хранящих составляющие цвета B, G и R (blue, green и red).

Для примера передачи данных в теле файла с рисунком формата bmp выберем следующую традиционную структуру документа, представленного на бланке организации, который отражает:

1) в виде рисунка – неизменный текст с атрибутикой, названием и реквизитами предприятия, включая: логотип, полное наименование организации, информацию о способах связи, юридический адрес, контактный номер телефона/факса, адрес электронной почты;

Смещение	Длина поля	Описание поля (что здесь находится)
Заголовок файла		
0	2	Код 4D42h - Буквы 'BM'
2	4	Размер файла в байтах
6	2	0 (Резервное поле)
8	2	0 (Резервное поле)
10	4	Смещение, с которого начинается само изображение (растр)
Заголовок BITMAP (Информация об изображении)		
14	4	Размер заголовка BITMAP (в байтах) равно 40
18	4	Ширина изображения в пикселях
22	4	Высота изображения в пикселях
26	2	Число плоскостей, должно быть 1
28	2	Бит/пиксел: 1, 4, 8 или 24
30	4	Тип сжатия
34	4	0 или размер сжатого изображения в байтах
38	4	Горизонтальное разрешение, пиксел/м
42	4	Вертикальное разрешение, пиксел/м
46	4	Количество используемых цветов
50	4	Количество "важных" цветов
Палитра (Карта цветов для N цветов), если используется		
54	4*N	Палитра

Рис. 1. Формат файлов bmp

2) в виде текста:
• реквизиты адресата;
• причина отправки документа;
• непосредственно текст документа;
• исполнитель документа и его телефон;
3) подпись должностного лица.

В данном документе нас будут интересовать изображение логотипа с атрибутикой и текст самого письма. Пример такого документа приведен на рис. 2.

Рисунок (выделенный логотип Минкомсвязи с соответствующими атрибутами) также представим в виде последовательности числовых данных. Напомним, что каждому пикселю соответствуют 3 байта. В среднем, такой рисунок в формате **bmp** имеет размер от 100 до 300 КБ. Нам потребуется (будем ориентироваться) объем рисунка в 256 КБ, что соответствует 87 381 (512 × 512/3) пикселю. Если рисунок будет несколько меньше, то его всегда можно дополнить до требуемого объема.

В приведенном документе (рис. 2) логотип занимает 272 КБ, что нас вполне устраивает (для интереса, проверьте его размер, сохранив рисунок логотипа в формате (24-разрядный рисунок *.bmp;*.dib).

Каждому пикселю мы будем в соответствии ставить некоторый символ. Размер в 87 381 символ соответствует примерно 48 стандартным страницам обычного текста шрифта Times New Roman с размером 14 и полуторным интервалом (1 800 символов на страницу). То есть, потенциально можно совместно с рисунком в одностраничном документе передавать документ, содержащий до 87 381 символ.

Если рассматривать как пример вариант, когда надо будет зашифровать только одну страницу текста, то тогда будет кодироваться только эта одна страница текста.

МИНИСТЕРСТВО СВЯЗИ
И МАССОВЫХ КОММУНИКАЦИЙ
РОССИЙСКОЙ ФЕДЕРАЦИИ
(Минкомсвязь России)

СТАТС-СЕКРЕТАРЬ –
ЗАМЕСТИТЕЛЬ МИНИСТРА

Тверская ул., 7, Москва, 125375
Справочная: 8-(495)-771-8100

01.02.2011 № 007-007

Наименование организации, которой отправляется данный документ

на № исх-111-1-11 от 01.01.2011

О результатах проведения оценки откорректированного проекта плана информатизации

В соответствии с обращением Вашей организации (письмо от 01.01.2011 г. № исх-111-1-11) Минкомсвязь России провело оценку откорректированного в части финансирования мероприятий по информатизации в соответствии с объемами бюджетных ассигнований проекта плана информатизации на 2002-2005 годы и направляет свое заключение.

Приложение: Заключение на 11 л. В 1 экз.

Петров А.А.

Исп. Иванов И.И.
+7(495)111-11-11

Рис. 2. Традиционная структура делового документа

Для того чтобы изменение чисел в описании пикселя не было особо заметным, выделим из каждой последовательности из трех байт по три последних бита первых двух байтов и два бита третьего байта. Это повлияет на изменение цвета не более чем на 6%. В этом случае, с точки зрения возможностей обычного человеческого зрения, сравнение цветов пикселей останется незамеченным. Данный вывод можно обосновать еще и тем, что для сравнения цветов потребуется оригинал, который предоставлен не будет, так как он нам не нужен в принципе. Следует также отметить, что для данного метода чем «грязнее» рисунок, тем сложнее будет определить наличие внедрения в него некоторых изменений.

Например, для последовательностей «110111001 00101011 1010111» (исходная последовательность, определяющая цвет пикселя) и «11011111 00101101 1010100» (модифицированная последовательность, содержащая символ) изменения будут незначительными. Для приведенного примера, вычитая побайтно по модулю 2 из первой последовательности вторую, получим значение: 110 – для первого байта, 110 – для второго байта и 11 для третьего байта. Итого 11011011, что равно числу 219 и соответствует заглавной букве **Ы**. Сравнение цветов показано на рис. 3.

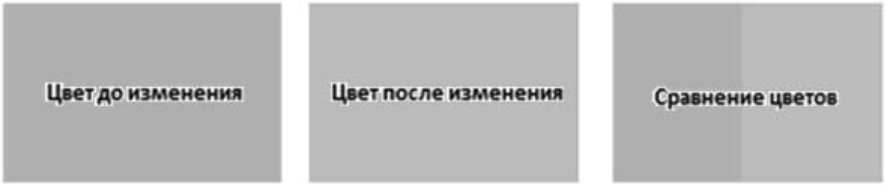

Рис. 3. Сравнение цветов в рисунке до и после модификации

После преобразования выделяем в каждом байте 3, 3 и 2 бита, суммируем в соответствии с адресом пикселя в рисунке и получаем модификации пикселей, соответствующих вводимому тексту. Таким образом, вначале мы строим массив чисел из «остатков» по 3, 3 и 2 бита для каждого цвета каждого пикселя. Получаем массив длиной в 87 381 элемент. Затем мы выделяем некоторую последовательность байтов в этом массиве длиной в наш текст.

Отметим, что желательно, чтобы адреса байтов в данной последовательности более или менее равномерно покрывали весь массив. Для построения такой последовательности можно, например, использовать обычное аффинное преобразование: $\mathbf{t = mod_{87381}\,(ax + b)}$. Поскольку число 87 381 простым не является, то задавать множитель «а» можно в пределах от 2 до 87 380, но исключив числа, являющиеся его делителями (это 3, 7, 9, 21, 57, 63, 73, 133, 171, 219, 399, 511, 657, 1 197, 1 387, 1 533, 4 161, 4 599, 9 709 и 12 483). Константу "**b**" можно задавать произвольно в пределах от 0 до 87 380. Элемент по адресу (номеру) **x** будет переставлен на место по адресу (номеру) $\mathbf{y = mod_{87381}\,(ax + b)}$. Элемент по адресу **y** будет переставлен на место по адресу $\mathbf{z = mod_{87381}\,(ay + b)}$ и т.д. Таким образом, можно без наложений полностью заполнить поле длиной в 87 381 символ. Для такой последовательности нам требуется определить:

– адрес первого байта, с которого мы начнем создание нашей последовательности (то есть, осуществим циклический сдвиг последовательности на заданное количество байтов);

– множитель;

– суммируемую константу.

Уточняем, что разряженную последовательность можно вообще не строить, но тогда возникает опасение, что искаженность рисунка может быть видна. На рис. 4

показан аналог результата ввода такой разряженной последовательности, преобразованной по правилу У = 6Х + 5, для таблицы 16 × 16 (в примере воспользуемся тем, что число 257 является простым).

В совокупности, эти три параметра (сдвиг, множитель и константа) можно определить и как секретный ключ, но в контексте данной статьи это не принципиально. Более того, каждый вводимый текст можно специально начинать именно с перечисления указанных параметров с последующим вводом необходимых данных. Это позволит на приемном конце без труда восстановить массив исходных данных (текстовое сообщение). Построение последовательности имеет назначение не запутать противника, а просто скрытно внедрить в рисунок текст сообщения.

Отметим, что в случае необходимости засекречивания передаваемых данных возможно перед введением этих данных в тело рисунка предварительно осуществить их шифрование. Но это уже другая задача.

Исходный текст

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
2	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
3	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
4	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
5	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
6	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
7	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
8	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144
9	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
10	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176
11	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192
12	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208
13	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224
14	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
15	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256

Текст, преобразованный по правилу Y = 6X+5, со сдвигом на 32 позиции

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
1	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256
2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
4	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
5	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
6	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
7	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
8	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
9	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
10	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144
11	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
12	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176
13	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192
14	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208
15	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224

Рис. 4. Результат «размытия» вводимого текста

Выводы

Описанный метод является достаточно эффективным для скрытой передачи данных.

Описанный метод не следует считать алгоритмом шифрования, так как каких-либо специальных методов трансформации и преобразования данных в нем не используется.

При использовании описанного способа скрытой передачи дополнительных данных, внедренных в типовой деловой документ, можно смело утверждать, что:

– в том случае, если посторонний человек будет просматривать такой одностраничный документ и не будет заранее знать о том, что в нем содержатся дополнительные сведения, у него не возникнет мысли о том, что данный документ дополнительно может содержать до 48 страниц текстовой информации;

– достаточно затруднительно, даже зная о том, что в документе передаются дополнительные сведения, извлечь их из документа. Это связано с необходимостью знать, как декодируются файлы формата **bmp** и как строится массив рабочих данных для их декодирования.

Литература

1. *Аграновский А.В*. Стеганография, цифровые водяные знаки и стеганоанализ. – М. : Вузовская книга, 2009. – 220 с.

2. *Грибунин В.Г., Оков И.Н., Туринцев И.В*. Цифровая стеганография. – М. : Солон-Пресс, 2009. – 265 с.

3. *Федосеев В.А*. Цифровые водяные знаки и стеганография : учебное пособие с заданиями для практических и лабораторных работ. – Самара : СГАУ, 2015. – 128 с.

4. *Конахович Г.Ф., Пузыренко А.Ю*. Компьютерная стеганография. Теория и практика. – М. : МК-Пресс, 2006. – 288 с.

References

1. *Agranovskiy, A.V.* Steganografiya, tsifrovye vodyanye znaki i steganoanaliz. – M. : Vuzovskaya kniga, 2009. – 220 s.

2. *Gribunin, V.G., Okov, I.N., Turintsev, I.V.* Tsifrovaya steganografiya. – M. : Solon-Press, 2009. – 265 s.

3. *Fedoseev, V.A*. Tsifrovye vodyanye znaki i steganografiya : uchebnoe posobie s zadaniyami dlya prakticheskikh i laboratornykh rabot. – Samara : SGAU, 2015. – 128 s.

4. *Konakhovich, G.F., Puzyrenko, A.Yu.* Komp'yuternaya steganografiya. Teoriya i praktika. – M. : MK-Press, 2006. – 288 s.