Секція 2. ІНФОРМАЦІЙНА БЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ: СОЦІАЛЬНИЙ, ПРАВОВИЙ І ТЕХНІЧНИЙ АСПЕКТ

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕФЕКТИВНОСТІ СТЕГАНОАНАЛІТИЧНИХ АЛГОРИТМІВ ДЛЯ ЦИФРОВИХ ЗОБРАЖЕНЬ

Ахмаметьєва Г.В.

кандидат технічних наук, доцент кафедри кібербезпеки Національного університету «Одеська юридична академія»

Розвиток інформаційних технологій та впровадження цифрового документообігу в усі сфери людської діяльності сприяє накопиченню значної кількості конфіденційної інформації на серверах компаній, підприємств, організацій, що потребує її додаткового захисту. За останні роки був зроблений великий крок вперед в області стеганографії, яка досліджує засоби прихованого захисту інформації шляхом вбудови цифрових водяних знаків, організацію каналів прихованого зв'язку. Саме прихована передача повідомлень за допомогою стеганографічних методів становить найбільшу небезпеку, оскільки виток конфіденційних даних може призвести до катастрофічних наслідків – зловмисник може отримати доступ до персональних даних працівників і клієнтів установ та використати їх для скоєння злочинів. Особливу небезпеку становить використання стеганографічних методів в умовах воєнного стану, коли чутлива критична інформація може опинитися в руках ворога шляхом її прихованої передачі через відкриті месенджери або інтернет-ресурси у вигляді невинного зображення або відео. Для запобігання витоку інформації через приховані канали зв'язку слід застосовувати стеганоаналітичні алгоритми, спрямовані на виявлення факту наявності вбудованого повідомлення в будь-які цифрові контейнери.

Одним з найбільш використовуваних контейнерів в стеганографії є цифрове зображення завдяки наявності в ньому надмірної інформації та можливості приховування значного обсягу даних. Однак з метою ускладнення виявлення наявності додаткової інформації в контейнері стеганоаналітичними алгоритмами, часто вкладення повідомлення відбувається з малою пропускною спроможністю прихованого каналу зв'язку (ППС). Це значно зменшує обсяг даних, що слід передати, однак забезпечує більш надійний захист повідомлення, оскільки його наявність в контейнері складніше виявити.

Однією з ефективних розробок для виявлення вкладення в цифрових зображеннях є стеганоаналітичний метод Color Triads, опис якого наведений в [1], що показав високу ефективність для визначення вбудованих повідомлень методом LSB при малих значеннях ППС (менше 0,5 біт/піксель). В даній роботі проведений експеримент виявлення факту наявності повідомлень, вбудованих стеганографічними алгоритмами WOW [2], S-UNIWARD [3], MiPOD [4], за допомогою алгоритмічної реалізації методу Color Triads, та порівняння результатів з

сучасними аналогами. Вибір алгоритмів WOW, S-UNIWARD та MiPOD обумовлений їх доступністю в мережі та широким використанням в наукових публікаціях при тестуванні стеганоаналітичних алгоритмів.

При проведенні обчислювального експерименту вбудовування повідомлення в цифрові зображення здійснюється в одну довільно обрану складову кольору з ППС 0.4, 0.2 і 0.1 біт/піксель алгоритмами S-UNIWARD, MiPOD і WOW. За результатами стеганоаналізу сформованих стеганоповідомлень були отримані помилки першого роду (False Negative FN), наведені в таблиці 1, які вказують на пропуск стеганоповідомлень. Помилки другого роду (False Positive FP) – визначення незаповненого контейнера як стеганоповідомлення, склали 0,49%.

Таблиця 1. Помилки першого роду (FN) виявлення наявності/відсутності повідомлення в цифровому зображенні алгоритмічною реалізацією методу Color Triads, %

Стеганографічний алгоритм	ППС, біт/піксель			
	0,4	0,2	0,1	0,05
S-UNIWARD	0,16	0,66	4,11	14,45
MiPOD	0,33	1,32	8,70	14,12
WOW	0	0,82	3,78	11,33

З таблиці 1 видно, що стеганоаналітичний алгоритм Color Triads забезпечує високу ефективність виявлення вбудованих повідомлень при значеннях ППС 0,1 біт/піксель та вище методами S-UNIWARD, WOW та MiPOD.

Порівняння ефективності виявлення наявності повідомлень будемо визначати за точності детектування наявності/відсутності повідомлень в зображенні серед всіх зображень, які включають як стеганоповідомлення, так і незаповнені контейнери. Результати порівняння ефективності для стеганоаналітичного алгоритму Color Triads і сучасних аналогів наведені в таблиці 2.

Таблиця 2. Порівняння точності детектування цифрових зображень стеганоаналітичним алгоритмом Color Triads та сучасними аналогами, %

	ППС, біт / піксель	SRM + EC [5]	CNN [5]	CNN + SRM + EC [5]	SCA-TLU-CNN [6]	FLD [7]	LSMR [8]	LASSO [9]	Color Triads
S-UNIWARD	0.4	77,95	90,95	84,13	87,19	–	–	–	99,67
	0.2	–	–	–	77,76	66,58	66,58	63,29	99,43
	0.1	59,00	76,64	61,12	67,80	–	–	–	97,70
	0.05	–	–	–	60,00	–	–	–	92,53
MiPOD	0.4	76,11	90,74	84,35	–	–	–	–	99,59
	0.2	–	–	–	–	66,79	66,93	63,31	99,09
	0.1	57,87	78,12	59,76	–	–	–	–	95,41
	0.05	–	–	–	–	–	–	–	92,69
WOW	0.4	–	–	–	90,41	–	–	–	99,51
	0.2	–	–	–	83,09	67,11	67,33	63,06	99,35
	0.1	–	–	–	75,58	–	–	–	97,87
	0.05	–	–	–	65,5	–	–	–	94,09

З таблиці 2 видно, що алгоритм Color Triads набагато перевищує результати аналогів: точність виявлення наявності/відсутності вбудованого повідомлення перевищує 92% навіть для ППС 0,05 біт/піксель.

В роботі наведений порівняльний аналіз розробленого раніше автором стеганоаналітичного алгоритму з сучасними аналогами. Показано, що алгоритм Color Triads є ефективнішими за сучасних аналогів, особливо при значеннях ППС 0,1 біт/піксель та вище.

Список використаних джерел:

1. Akhmametieva, A.: Steganalysis of digital contents, based on the analysis of unique color triplets. Annales Mathematicae et Informaticae, no. 47, 3-18 (2017).
2. Holub, V., Fridrich, J.: Designing Steganographic Distortion Using Directional Filters. In: 2012 IEEE International Workshop on Information Forensics and Security (WIFS 2012), pp. 1-6. IEEE, Tenerife, Canary Islands (2012).
3. Holub, V., Fridrich, J., Denemark, T.: Universal Distortion Function for Steganography in an Arbitrary Domain. EURASIP Journal on Information Security (Section: SI: Revised Se-lected Papers of ACM IH and MMS 2013), 1-13 (2014).
4. Sedighi, V., Cogranne, R., Fridrich, J.: Content-Adaptive Steganography by Minimizing Statistical Detectability. IEEE Transactions on Information Forensics and Security, vol.11, iss.2, 221-234 (2016).
5. Couchot, J.-F., Couturier, R., Salomon, M.: Improving Blind Steganalysis in Spatial Domain Using a Criterion to Choose the Appropriate Steganalyzer Between CNN and SRM+EC. In IFIP International Conference on ICT Systems Security and Privacy Protection, pp. 327-340. Springer, Rome, Italy (2017).
6. Jian Ye, Jiangqun Ni, Yang Yi: Deep Learning Hierarchical Representations for Image Steganalysis. IEEE Transactions on Information Forensics and Security, vol.12, iss.11, 2545-2557 (2017).
7. Kodovsky, J., Fridrich, J., Holub, V.: Ensemble classifiers for steganalysis of digital media. Information Forensics and Security, IEEE Transactions, vol.7, no.2, 432-444 (2012).
8. Fong, D.C.-L., Saunders, M.: LSMR: An iterative algorithm for sparse least-squares prob-lems. SIAM Journal on Scientific Computing, vol. 33, no. 5, 2950-2971 (2011).
9. Cogranne, R., Sedighi, V., Fridrich, J., Pevný, T.: Is Ensemble Classifier Needed for Steganalysis in High-Dimensional Feature Spaces? IEEE International Workshop on Information Forensics and Security (WIFS 2015), pp.1-6. IEEE, Rome, Italy (2015).