

рівномірний розподіл; 2) біноміальний розподіл; 3) геометричний розподіл. Результати показали високу точність прогнозування на рівні визначення максимального значення відносної похибки моделювання.

### **Порівняння різних моделей характеристичних векторів для JPEG-стеганоаналізу**

УДК 004.056;  
004.415.24

Наталія Кошкіна

*Інститут кібернетики імені В.М. Глушкова НАН України,  
nata.koshkina@gmail.com*

При практичному стеганоаналізі основними критеріями вибору методу детектування прихованих повідомлень є його точність та швидкодія. Проте здійснити порівняння методів навіть одного класу, базуючись на даних з наукових публікацій, досить складно через відмінності в умовах чисельних експериментів. Тому, дослідження, що дозволяють порівняти ефективність сучасних методів стеганоаналізу в однакових умовах, є актуальними, а їх результати можуть бути використані для практичної організації систем виявлення стеганографічних приховань. Також вони можуть слугувати базою для подальшого вдосконалення стеганоаналітичних методів і систем.

*Метою даної роботи є здійснення порівняльного аналізу ефективності різних статистичних моделей характеристичних векторів для стеганоаналітичних систем на базі машинного навчання, на основі якого можна буде вибрати оптимальну модель для заданих практичних умов.*

Для проведення чисельних експериментів було обрано три програми, що здатні здійснювати приховування у файли формату JPEG: Jsteg, Jphide та Steganos Privacy Suite 2012 (модуль Срут&Hide). Вихідний тестовий набір містив 1330 кольорових зображень розмірами 512×384 пікселі, в кожне з них вищезгаданими програмами вкраплювалося 1 Кб випадкового тексту.

Обравши класичний класифікатор – метод опорних векторів (англ. SVM, support vector machine) та зафіксувавши однакові умови експериментів ми прорахували точність та швидкодію стеганоаналізу на базі наступних відомих статистичних моделей для JPEG-контейнерів: CHEN, CC-CHEN, LIU, CC-PEV, CC-C300, GFR, DCTR.

В таблиці 1 наведена кількість елементів характеристичного вектора для кожної з вищезазначених моделей, а також швидкість створення характеристичного вектора, отримана на ПК з процесором Intel Core i5-661 3.33 ГГц і 8 Гб ОЗУ та усереднена за 100 обчисленнями.

Таблиця 1

№ п / п	Параметри Модель	Кількість елементів вектора	Швидкість обчислення вектора, сек
1	CHEN	486	0.2
2	CC-CHEN	972	0.9
3	LIU	216	44.2
4	CC-PEV	548	1.5
5	CC-C300	48600	1.1
6	GFR	17000	6.2
7	DCTR	8000	2.0

Швидкість обчислення характеристичних векторів впливає на швидкість роботи стеганоаналітичної системи як на етапі її навчання, так і на етапі детектування стегановкладок. Особливо критичною ця характеристика є для систем реального часу. Як бачимо з даних таблиці 1 різниця в швидкості обчислення характеристичних векторів для різних моделей досить помітна. Найшвидше обчислюються характеристичні вектори для моделі CHEN, найповільніше – для LIU. Тобто коли в моделі LIU буде опрацьований тільки один файл, в моделі CHEN таких файлів буде більше ніж 200.

Точність виявлення стеганоконтейнерів, яка була усереднена за 10 повторами стеганоаналізу з випадковим поділом контейнерів на навчальну та контрольну вибірки, представлена у таблиці 2. Також в дужках після точності наведено усереднену кількість хибно позитивних та хибнонегативних тривог для 557 пустих та 557 заповнених контейнерів контрольної вибірки.

Таблиця 2

Атака на Модель	Jsteg	Jphide	Steganos
1 CHEN	97.9% (0; 24)	83.1% (89; 99)	83.1% (67; 50)
2 CC-CHEN	98.0% (0; 23)	88.2% (61; 71)	94.9% (32; 25)
3 LIU	99.8% (0; 2)	88.8% (57; 68)	98.0% (5; 17)
4 CC-PEV	84.0% (75; 104)	76.9% (79; 179)	76.9% (81; 177)
5 CC-C300	99.1% (4; 6)	91.6% (51; 43)	95.7% (25; 24)

6	GFR	96.3% (19; 22)	91.6% (66; 27)	92.0% (40; 50)
7	DCTR	98.6% (13; 3)	93.1% (45; 32)	97.1% (14; 18)

За результатами цих експериментів в трійку лідерів потрапили моделі LIU, CC-C300 та DCTR. Найкращу точність при виявленні стеганоконтейнерів, створених програмами Jsteg та Steganos Privacy Suite 2012 забезпечила модель LIU, при виявленні Jphide стегановкладок – модель DCTR. Найгірша точність класифікації для всіх трьох варіантів стеганопрограм отримана для моделі CC-PEV. Також незалежно від стеганопрограми модель CC-CHEN має перевагу над вихідною моделлю CHEN.

Зауважимо, що у випадках, коли допустиме зниження швидкодії, є резерви покращення точності. Так, можна не обирати якусь одну модель формування характеристичних векторів, а використовувати декілька ефективних для даного типу контейнерів та стеганоперетворення. При цьому можна як комбінувати чи усереднювати результати на базі різних моделей (наприклад, байєсівське усереднення), так і навчити окрему модель тому, яку саме з наявних моделей використати для передбачення (наприклад, дерево прийняття рішень).

Крім того в загальному випадку стегоаналітична система повинна мати можливість працювати в режимі мультикласифікації, тому було виконане також порівняння ефективності і для подібних сценаріїв.

В подальшому планується перевірити точність розглянутих та інших статистичних моделей при виявленні вдосконалених стеганоперетворень, таких як nsF5, YASS, HUGO, WOW, UNIWARD.