MODERN STGANOGRAPHIC APPROACHES IN SOCIAL NETWORKS

HRISTO I. PARASKEVOV, ALEKSANDAR Y. STEFANOV

ABSTRACT: *Online social networks (OSN) have gained great popularity among Internet users because of the opportunities they provide for easy sharing of news, opinions, multimedia content and various kinds of activities. In this regard, the OSN has become an important factor in the formation of positions and attitudes in modern society. Logical is the reaction of some governments and organizations both to using the abovementioned qualities of social networks and to make efforts to censor them. Two interrelated issues become more and more relevant: the development of steganographic methods using the OSN's specificities to overcome imposed censorship, and the creation of a set of measures to perform adequate stealth analysis to reveal built-in secret, malicious communication channels.*

KEYWORDS: *steganography, social networks, hiding data*

СЪВРЕМЕННИ СТЕГАНОГРАФСКИ ПОДХОДИ В СОЦИАЛНИТЕ МРЕЖИ*

ХРИСТО И. ПАРАСКЕВОВ, АЛЕКСАНДЪР Й. СТЕФАНОВ

1 Въведение

Социалните мрежи предоставят множество и разнообразни възможности за изграждане на стеганографски тайни канали. Важни моменти са огромният брой потребители и съответстващия обем на обменяната информация. Подходите при реализирането на стеганографски канали могат да се разгледат в две основни направления:

✓ методи експлоатиращи спецификите на OSN при публикуване и обмен на съдържание;

✓ методи използващи спецификите на OSN при отразяване на поведението на потребителите.

Следва да се отбележи, че е налице и трета възможност – разработване на хибридни методи, комбиниращи наличните възможности както по съдържание, така и по поведение. В настоящата статия е направен обзор на съществуващи към момента подходи и възможности за използване на стеганография в онлайн социални мрежи.

Едни от най-масово използваните стеганографски контейнери са графичните изображения. Известни са множество методи, приложими към различни графични формати, позволяващи получаване на добри резултати [1]. Графичните изображения са основна част от публикуваната в онлайн социалните мрежи потребителска информация. В тази връзка е логично да се проучи приложимостта на гореспоменатите методи в тази среда. Проведените изследвания обаче показват, че част от онлайн социалните мрежи са въвели ограничения относно поддържаните от тях графични формати. Нещо повече честа

* Докладът е частично финансиран със средства от проект РД-08-159/09.02.2018

практика е извършването на обработка на публикуваните от потребителя изображения. Съгласно резултатите публикувани в [2], прилаганите политики се различават значително за различните платформи: например Google+ не обработва изображенията ако резолюцията им не надвишава 2048x2048, в тази връзка използването на всеки графичен стеганографски инструмент е допустимо. При Twitter ситуацията е по-различна - заличават се някои от полетата, съдържащи метаданни, което прави неприложими методите, които ги експлоатират. Същевременно интегритета на изображения по-малки от 1024x768 се запазва. Подобно на Twitter Facebook премахва метаданните от изображението, освен това обаче го подлага на допълнителни обработки - модификация на размера, стойностите на част пикселите, компресията. Всичко това води до сериозни предизвикателства по отношение използването на графична стеганография в тази социална мрежа. Публикуваните в [3] резултати кореспондират с направените по-горе наблюдения. Въпреки ограниченията налагани от Facebook, в резултат на проведени изследвания, публикувани в [4] е установено, че при избор на подходящ контейнер, графичната стеганография е приложима и в тази социална мрежа. Най-добри резултати в това отношение са постигнати със специализирания софтуер JP Hide & Seek (JPHS) - до 60 % успеваемост, като положителни резултати са постигнати и с Steg и Steghide.

2 Стеганографски подходи в социалните мрежи

Перспективен подход за приложение на стеганографията в социалните мрежи е описан в [5]. Реализираният метод е наречен StegHash и използва наличния в множество социални мрежи елемент наречен hashtag. Под hashtag се разбира обикновено етикет състоящ се от дума, предхождана от символа "#". Обикновено се "прикача" към публикация (снимка, видео, текст) като показва нейната относимост към дадена тематика. Особено полезен е в социални мрежи, в които е налице ограничение в дължината на публикуваните текстови съобщения.

Описаният метод, разширява възможностите на използването на графични изображения, видеофайлове и др. като контейнери на тайното съобщение. Добавянето на хаштагове към такива обекти позволява изграждане на връзки между тях, увеличавайки многократно размера на скритото съобщение. На всеки набор от n на брой хаштага, съответстват n! на брой пермутации, които са индивидуални индекси на всяко съобщение. Притежателите на секретния ключ (паролата), които разполагат със секретния генератор (функцията) могат да изградят връзка между тези индекси и в последствие да проследят веригата между отделните контейнери, съдържащи отделните части на тайното съобщение [5]. Методът е реализиран в две от популярните онлайн социални мрежи Twitter и Instagram, като са отчетени особеностите им по отношение на прилаганата към публикуваните изображения обработка. Установено е, че различните платформи имат различен подход по отношение възможностите за извършване на търсене по прикрепените към обекта (контейнера) хаштагове - докато Twitter няма ограничения по отношение на броя им, в Instagram се допуска търсене само по един. При експериментите е използван малък набор от хаштагове (3,4,5,6) с оглед избягването на нарушаване на работата на социалната мрежа посредством генериране на солиден допълнителен трафик, което би провокирало активирането на съответните политики на сигурност. Регистрирани са обнадеждаващи резултати за малки по обем съобщение (100% успеваемост) като за по-големи е необходимо да се извършат допълнителни оптимизации.

Едно от приложенията на стеганографията в социалните мрежи е насочено към осигуряване на защита на персоналните данни, предоставени от потребителя в неговия

профил. В [6] е описана система за такава защита означена като NOYB (none of your bussiness). При регистрирането и използването на онлайн социална мрежа потребителят в общия случай първоначално въвежда и периодично допълва и коригира, голямо количество лична информация - име, адрес, месторабота, семейни връзки, снимков материал. Тази информация става достъпна за определен от потребителя кръг от хора, в съответствие с възприетата и прилагана от социалната мрежа политика за сигурност. Масово използваната социална мрежа Facebook например, изисква потребителят да използва същото име, което използва и в ежедневието си, както и да посочва точна информация за себе [7]. В тази връзка при спазване на тези изисквания потребителят посочва достоверна информация, която обаче би могла да стане достояние на не съвсем добронамерени лица. Информацията може да бъде използвана както за извършване на измами така и за осъществяване на определени политически и бизнес цели. Широко известен е скандалът със събирането и обработването на лични данни на Facebook потребители от Cambridge Analytica. Посредством приложението "thisisyourdigitiallife" [8] организацията автоматизирано е събирала и обработвала персонална информация за изграждане на "психографичен" профил на повече от 87 милиона потребители [9]. Според твърденията на представители на компанията събраната информация може с висока точност да предскаже високочувствителна лична информация като: сексуална ориентация, етническа принадлежност, религиозни и политически възгледи, ниво на интелигентност, използване на психоактивни вещества, емоционален статус, възраст и пол [8]. В тази връзка е обяснимо защо част от потребителите предпочитат да се регистрират под фалшива самоличност и да посочват неверни данни в контактната си информация. Разбира се при установяването на този факт от страна на Facebook, потребителите са заплашени от закриване на профилите им. В предложената в [6] система NOYB персоналната информация се разделя на "атоми". Всеки от тях може да се криптира, като по този начин информацията става достъпна само за запознатите с ключа и алгоритъма на криптиране. Разбира се повечето от социалните мрежи не биха толерирали директно въвеждане на криптирани данни. Тук на помощ идва стеганографията. Посредством предварително генериран речник се извършва замяна на отделните "атоми" на реалните потребителските данни със съответните им стойности в речника, генериран на псевдослучаен принцип. Например потребителските данни (име, пол, възраст) - (Алис, Ж, 25) се разделят на два "атома" - (Алис, Ж) и (25). Първият "атом" се заменя с (Боб, М), а вторият с (28). Така за масовия потребител профилът принадлежи на Боб, 28 годишен мъж, а за "просветените" е ясно, че се касае за Алис, 25 годишна жена. Посредством този подход се реализират няколко цели:

✓ защита на личните данни - за масовия потребител остава в тайна истинската самоличност на ползвателя на защитения профил;

✓ защитата може да се реализира без да е необходимо съдействие от страна на онлайн социалната платформа. Същевременно ако политиката на администраторите на социалната мрежа е да оказват съдействие, защитата може да се приложи към всички потребители;

✓ системата е трудно разкриваема дори за враждебно настроена платформа, особено в случаите на социална мрежа с голям брой потребители. За разкриването ѝ е необходимо анализиране на немалък брой "заподозрени" профили с цел откриване на данни уличаващи потребителя в използване на неистинска идентичност.

Разбира се гореописания подход има и своите недостатъци. Част от потребителските данни не могат да бъдат заменени с фиктивни, поради факта, че трябва да бъдат уникални

и се използват за обратна връзка при регистрацията. Това в особена сила важи за телефонния номер и адреса на електронната поща.

Друг възможен подход за реализиране на стеганографски канал е използването на популярната в онлайн социалните мрежи възможност за реакция на потребител по отношение на достъпни за него публикации [10]. В този случай се реализира таен канал въз основа на поведението на потребителя. Носител на тайната информация могат да бъдат поставени "лайкове" (харесвания) на определени публикации. Предложеният подход може да бъде разширен като в реализирането на канала участват: време (както астрономическо така и относително), на поставяне на реакциите, поредност на реакциите (за множество статии), тип ("харесване", "любов", "ха-ха", "Еха", "тъга", "гняв"). Допълнително разширяване на канала може да се постигне посредством менажиране на множество профили от страна на изпращача на тайното съобщение, при което се манипулират поредността, времената и типовете на реакциите на контролираните от него профили. Реакциите могат да бъдат както по отношение на публикации на получателя на тайното съобщение, така и по отношение на публикации на трета, неутрална към комуникацията страна.

По отношение на споделената текстова информация под формата на публикации и разменената такава в чат комуникация например би могла да се използва текстова стеганография. Основна трудност при нея е липсата на излишък от информация, какъвто е налице при графичните и звуковите файлове. [11] Един обикновен текстов документ (например .txt формат) се записва, обработва и възпроизвежда от компютърната система във вид идентичен или много близък до това, което се изобразява пред погледа на потребителя. По-перспективни в това отношение са файлови формати поддържащи форматиране на текста - шрифт, размер, отместване, разстояние между редове и символи.

Методите приложими в текстовата стеганография могат да бъдат разделени на две основни групи:

- ✓ манипулация на форматирането на текста;
- ✓ манипулиране на значението на текста.

Методите от втората група имат своите ограничения [11]:

- синтактичен - поставяне или пропускане на синтактични знаци (, . : ; и др.) на съответните места в текста. Количеството информация, което може да бъде скрито по този начин е в пряка връзка с броя на знаците и в тази връзка като цяло е силно ограничено). Съответно неправилното поставяне на голям брой излишни знаци би събудило основателни подозрения за наличието на целенасочено манипулиране на текста;

- семантичен - използване на синоними на определени думи, от предварително договорен речник, с цел предаване на скрита информация:

Логическа "0"	Логическа "1"
Голям	Значителен
Студен	Хладен
Хубав	Красив
Мотор	Двигател

Предвид наличието на нюанси в значението на отделните думи от синонимната двойка, са налице ситуации, при които директната замяна на едната дума с другата е не винаги удачна - например в словосъчетанието "големи усилия" замяната със значителен е възможна - "значителни усилия". Не така стои въпросът при "голяма сграда", в случая по-

подходяща била алтернативата "огромна сграда", а не "значителна сграда". В тази връзка при създаването на речници би следвало да се отчитат гореспоменатите особености, което от своя страна би се изразило в усложняване на структурата на речника (например добавяне на алтернативни думи в двете колони).

Логическа "0"	Логическа "1"
Голям, важен	Огромен, значителен

- използване на абревиатури или съкращения. В ежедневието си често използваме абревиатури като алтернатива на пълното изписване на наименования (Шуменски университет - ШУ; мобилен телефон - GSM и др.). Подобна е ситуацията с използването на съкращения при провеждане на чат комуникация "не знам" се замяна често с "нзн" или "nz", "какво правиш" с "к пр". Аналогично при е и при англоезичните чатове - "be right back" с "brb", "face to face" с "F2F" и т.н. Подобно на семантичния метод използването на пълното наименование или съкращението (абревиатурата) може се експлоатира за предаване на скрито съобщение. Количеството информация и в този случай е ограничено по обем. Тук отново е налице необходимостта от предварително договаряне на речници между двете страни.

- Начин на изписване на думите (спелуване) - при този метод се използва обстоятелството, че едни и същи думи в английския език се изписват по различен начин в Британския и Американския варианти на езика.

AmericanSpelling	BritishSpelling
Favorite	Favourite
Criticize	Criticise
Fulfill	Fulfil
Center	centre

И в този случай количеството информация, което може да бъде скрито е малко поради малкия брой думи, отговарящи на горното условие.

- Генериране на правописни грешки в текста - предвид публикуването в социалните мрежи на разнообразна информация от широк кръг от потребители с различна степен на грамотност е налице възможност да се реализира таен канал посредством преднамерено и контролирано допускане на правописни грешки в публикацията или кореспонденцията. Откриването на подобен канал би изисквало изучаване и анализ на проявената от потребителя езикова грамотност с цел преценка дали се касае за неволно или целенасочено допускане на съответните езикови грешки.

Въпреки ограниченията по отношения на реализирания посредством гореописаните методи таен канал, те са директно приложими както при публикуване на текстова информация в потребителския профил така и при осъществяване на чат комуникация между двама и повече потребители.

По-перспективни по отношение на капацитета на тайния канал са методите на текстова стеганография използващи манипулиране на форматирането на текста:

- ✓ промяна на разстоянието между редовете;
- ✓ промяна на разстоянието между буквите;
- ✓ добавяне на празни интервали между буквите;

✓ използване на емотикони - на различните емотикони съответстват различни стойности ":-)" - "0", ":-(" - "1", ":-*" - "01", ":-P" - "10" и т.н. [12];

Следва да се отбележи, че за разлика от последните два от гореописаните методи, първия и втория не са директно приложими при публикуване на текст в профила или чат комуникация. За да се използват е необходимо обработения текст да бъде записан във формат, поддържащ осъществените манипулации (например .doc или .docx) и изпратен като файл до друг потребител.

Друга възможност за използване на текстова стеганография е генерирането и последващо разпращане или публикуване на спам съобщения. В самото съобщение могат да бъдат приложени методи, манипулиращи както значението така и форматирането на текста. Предвид огромния брой спам съобщения циркулиращи в Интернет в това число и социалните мрежи (например "Ако не разпратиш това съобщение на n на брой потребители в следващия m на брой часа, ще се случи ..." предхождано от някаква история/информация) е налице възможност за реализиране на таен канал с немалък капацитет, особено при комбиниране на няколко метода.

Прилагането на криптографски алгоритми повишава устойчивостта на стеганографските методи[13, 14].

3 Заключение

Описаните подходи и методи дават основание да се счита, че при използването на социални мрежи могат да се реализират перспективни стеганографски канали. Усилията могат да бъдат насочени към създаване на хибриден метод, при който скритото съобщение се вгражда в контейнер с голям капацитет (графично изображение, видео или звуков файл), намиращ се на подходящо място в Интернет пространството (сайтове за видео споделяне, облачни структури за съхранение на информация и др.). Същевременно другият важен компонент – стегоключа се предава между заинтересуваните страни при използване на набор (матрица) от профили, посредством контролирано и синхронизирано въздействие върху поведението им – лайкване на публикация (време, поредност, тип), споделяне на публикация (време, поредност, местоположение, достъпност), активиране и деактивиране на комуникационно приложение, добавяне и премахване на потребители в списъка, манипулиране на публикувани или изпращани текстови съобщения и др. За реализирането на такъв метод е необходимо:

✓ извършване на проучване на достатъчен брой социални мрежи, комуникационни програми, сайтове за видео споделяне;

✓ избор на подходящи социална мрежа, комуникатор(и), сайт(ове) за видео споделяне;

✓ генериране на необходимия брой потребителски профили;

✓ изпълване с подходящо съдържание на профилите;

✓ провеждане на съответните експерименти с цел определяне оптимален контейнер за съобщението и метод за предаване на стегоключа.

Важен момент при експлоатирането на стеганографски методи в социални мрежи е своевременното реагиране и пренастройване на метода съобразно динамиката във функционалността на използваната среда.

ЛИТЕРАТУРА:

[1] Станев, Ст. Стеганологична защита на информацията. Шумен: Университетско издателство „Еп. К. Преславски“, 2013. ISBN: 978-954-577-825-4.

[2] N. ‚Jianxia, I. Singh, Harsha V. Madhyastha, Srikanth V. Krishnamurthy, Guohong Caox, and Prasant Mohapatra. Secret Message Sharing Using Online Social Media

[3] Еминов Д., С. Хасанова, Д. Тончев - Стеганография в онлайн социални мрежи – MATTEX, 2014, том I, стр.173-179

[4] D. Tejas, J. Hiney, Using Facebook for Image Steganography

[5] Szczypiorski, K, StegHash: New Method for Information Hiding in Open Social Networks

[6] Guha S., K. Tang, P. Francis - NOYB: Privacy in Online Social Networks

[7] Facebook Inc. Условия за ползване – "https://web.facebook.com/terms.php?_rdc=1&_rdr", 20 юли 2018 г.

[8] Sanders, J. and D. Patterson, Facebook data privacy scandal: A cheat sheet. https://www.techrepublic.com/article/facebook-data-privacy-scandal-a-cheat-sheet/, 21 юли 2018 г.

[9] Whittaker, Z. Facebook: Cambridge Analytica took a lot more data than first thought. https://www.zdnet.com/article/facebook-confirms-cambridge-analytica-took-more-data-than-first-thought/, 20 юли 2018 г.

[10] Zhang X., - Behavior steganography in social network

[11] Roy S., M. Manasmita - A Novel Approach to Format Based Text Steganography

[12] Chandragiri A., P. Cooper, Y. Liu and Q. Liu - Implementing Secure Communication on Short Text Messaging

[13] Kordov, K., Stoyanov, B., (2017), Least Signi cant Bit Steganography using Hitzl-Zele Chaotic Map, International Journal of Electronics and Telecommunications, 63(4), 417-422. doi: https://doi.org/10.1515/eletel-2017-0061

[14] Stoyanov, B., Zhelezov, S, Kordov, K., (2016), Least signi cant bit image steganography algorithm based on chaotic rotation equations, Comptes rendus de l'Acad emie bulgare des Sciences, 69(7), 845-850.

Христо Иванов Параскевов
Шуменски университет „Епископ Константин Преславски“,
E-mail: h.paraskevov@shu.bg

Александър Йонков Стефанов
Шуменски университет „Епископ Константин Преславски“,
E-mail: a.stefanov@shu.bg - докторант