Безпека життєдіяльності та ліквідація наслідків надзвичайних ситуацій

УДК 519.2:004.9

Д.В. Рудченко

Харківський національний університет радіоелектроніки, Харків

АНАЛІЗ ДАНИХ СОЦІАЛЬНИХ МЕРЕЖ ДЛЯ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ

Стаття присвячена вирішенню маркетингових завдань та питань кібербезпеки щодо аналізу інформації отриманої у соціальних мережах. Описано основні загрози соціальних мереж. Представлено різноманіття даних та методи їх збору для проведення аналізу. Розроблено програмне забезпечення для збору та аналізу даних із соціальної мережі Вконтакті, яке дає швидкість обробки навіть для великих обсягів і точність кінцевого результату. Проведено збір та аналіз даних, визначення користувачів, які є центром поширення інформації.

Ключові слова: *аналіз соціальних мереж, загрози, виявлення експертів, кібербезпека.*

Вступ

Соціальні мережі – це феномен сьогоднішнього дня. Збільшення числа користувачів і особистий характер даних призводить до проблем з безпекою та конфіденційністю. Переваги використання соціальних мереж полягають у вільному і швидкому зв'язку з друзями зазвичай у вигляді об'єктів, таких як пости, картинки, відео і тексти. Ще одна особливість – створення мереж: друзів, колег, членів сім'ї. Проблема пов'язана з можливістю відстеження дій, відносин і з обмеженою можливістю управляти критичними діями, такими як видалення даних. Інформація, керована різними видами соціальних мереж, дуже цікава. Головним чином за її потенціал для формування загального особистого профілю. Агрегування інформації із загальнодоступних профілів дуже корисно для конкретних цілей, таких як побудова стратегії маркетингу і виявлення груп осіб, пов'язаних із забороненими організаціями.

Загрози в соціальних мережах

В соціальних мережах існує цілий ряд різних загроз. В [1] описані загрози для соціальної мережі ділових кіл, в [2–5] описані загрози для соціальних мереж і деякі профілактичні заходи.

1. Друзі. Довіра до тих, хто входить в список «друзів», завжди вище, ніж випадковим людям. З одного боку, це добре, оскільки формується лояльна аудиторія навколо компанії, бренду або людини. З іншого боку, це можливість для зловмисників.

2. Можливість заміни людини або маскарад: напевно, не зовсім ясно, хто приховує свої дії за ім'ям друзів або ховається за фотографіями друзів в профілі соціальної мережі. Можливо по IP-адресою відправника зібрати про нього принаймні деяку інформацію в кореспонденції по електронній пошті, яка не працює в соціальній мережі.

3. Використання сервісів скорочення URL-адрес. В останні роки особливо популярні послуги скорочення URL-адрес, що дозволяють маскувати небажану адресу сайту під короткою посиланням. Фактично, домен перенаправляє відвідувача. Сьогодні йде активна боротьба з цими ризиками – служба скорочення URL-адрес почала використовувати вдосконалені механізми для виявлення спаму та інших загроз. Однак для користувачів соціальних мереж ця загроза зберігається: спокусливі повідомлення і пропозиції від уже відомих контактів, які були зламані, часто призводять до завантаження шкідливого ПО або відображення небажаних веб-сторінок.

4. Використання тих самих імен користувачів і паролів в корпоративній мережі і зовнішніх соціальних ресурсах – ця атака також відома як «Daisy Chain». В результаті злом профілів соціальної мережі користувачів значно підвищує ризик проникнення на корпоративні ресурси від імені одного зі співробітників компанії.

5. Веб-атака. Соціальні мережі можуть використовуватися хакерами для організації атак через вразливості в браузерах, а також XSS / CSRF-атак. Інструментами для таких атак можуть бути троянські програми, підроблені антивіруси, соціальні черви, шкідливі JavaScript і HTML-код, які використовуються для поширення власних списків друзів і інших. Їх головна мета – увійти в інформаційну систему відвідувача соціальної мережі, його робочої станції або пристрою і закріпитися в ній. Для захисту використовуються такі традиційні інструменти, як антивірусне програмне забезпечення, здатне працювати в реальному часі і блокувати завантаження шкідливого коду.

6. Витік інформації і компрометація поведінки співробітників компанії. Соціальні мережі можуть

використовуватися для організації витоків важливої інформації компанії, а також для підриву її репутації. Така атака може вестись від внутрішніх співробітників, які незадоволені керівництвом або спеціально вбудованими інсайдерами. У соціальних мережах люди часто поводяться зовсім інакше, ніж в корпоративному середовищі спілкування, і, можливо, шокуюче публікування і грубі репліки можуть завдати певної шкоди репутації їх роботодавців. DLP-системи і продукти для аналізу публікацій в Інтернеті, призначені для захисту від цих загроз.

7. Advanced Persistent Threat (APT)-атаки. Виходячи з вищесказаного, соціальні мережі можуть використовуватися як шлюз або джерело загрози організації, службі або іншому підрозділу для кваліфікованих хакерів з найсучаснішим і просунутим шкідливим кодом, методами атак і хакерською методологією взагалі.

8. Зростання трафіку, особливо при перегляді відеоджерел.

9. Зміст з ознаками підбурювання до расової, етнічної або релігійної ненависті, пропаганда тоталітарних сект.

10. Пропаганда і публічне виправдання тероризму [6].

11. Кібер-приниження і кіберзалякування.

13. Популяризація та поширення наркотиків.

Збір даних

Розрізняють пасивний (прямий пошук) і активний (спілкування) способи отримання інформації [7].

Прямий пошук – пошук, наприклад, за ключовими словами з використанням пошукових сервісів самої мережі або зовнішніх пошукових систем.

Дані про людину – все, що може про себе повідомити сама досліджувана людина або сказати про неї люди, які її знають:

– установчі дані людини – ПІБ, дата і місце народження, фото. Причому, вважаючи, що таким чином захищає свої дані, залишає на одному форумі ім'я, на іншому форумі дату народження, на третьому форумі icq і при цьому всюди реєструється під одним ніком;

– компетенції – освіту, досвід роботи, досягнення (з подробицями і деталізацією, що трудова книжка не потрібна). І знову в одному місці один комплект, в іншому трохи змінений, в третьому ще з якоюсь зміною. Потрібно інформацію зібрати, проаналізувати та виявити нестиковки, в яких міститься найцікавіше;

– зв'язки: родинні, дружні, робочі. Особливості особистості – переваги, хобі, погляди, переконання, висловлювання на форумах, в блогах і мікроблогах з потрібних проблем: фактично можна скласти повне уявлення про досліджуваний об’єкт;

– дані про компанії – як про компанію, так і про співробітників;

– контактні і установчі дані самої компанії;

– хто співробітники, хто менеджери, хто виконавці, як з ким зв'язатися, в тому числі і скривджені співробітники, і незадоволені клієнти;

– внутрішня обстановка офісу, взаємини всередині колективу. розміри офісу, його наповнення майном і співробітниками, активність співробітників і активність телефонних переговорів, присутність клієнтів і робота з ними, корпоративний стиль, місця проведення корпоративних заходів і стиль їх проведення – це є непрямим свідченням розміру і прибутковості компанії;

– дані про продукт (послуги) – виробник, споживач, зацікавлені особи;

– особливості продукту (найрізноманітніші), його властивості, якість і т.д;

– аналоги продукту і його замінники, їх особливості;

– дані про подію, ситуацію, проблеми;

– що, де, коли відбулося, хто учасники і суть події, причому, з різних точок зору і ставлення до події;

– «підводні» течії і неявні сили, залучені в події можуть бути ідентифіковані як самими популярними подіями, так і на основі аналізу медіа-активності різних груп;

– наслідки очевидні і не явні можуть бути виявлені прямим спілкуванням з учасниками і свідками і на основі даних прояву інтересу до події різних сил;

– з фотографій: люди розміщують те, що вони вважають цікавим, важливим, гідним. Таким чином, користувачі неявно хваляться майном, причетністю до події, зв'язком з важливою людиною. по деталях з фотографій можна дізнатися: місце роботи об'єкта; основні робочі проекти – на задньому плані може бути шафа з вартими в ряд справами, на корінцях яких назви проектів; персональні дані по фото авто (номер потрібно зафарбовувати), адреси перебування по фото на тлі будинків, фото біля свого будинку (дачі, фазенди) з обов'язковим попаданням адресної таблички.

– за метаданих gps координати місця, дата-час;

– з особистих повідомлень (на «стінах»);

– дата народження, якщо хоч хтось привітав користувача «в прямому ефірі»;

– місця відпочинку об'єкта при обміні думками про них зі знайомими;

– стиль поведінки і спілкування, загальний тренд висловлювань з тих чи інших питань;

– факти біографії (а також участь у подіях) стають об'єктом обговорення;

– зв'язки з іншими людьми або організаціями виявляються через їх згадки в спілкуванні;

– спілкування: користувачі соціальних мереж мають величезну кількість інформації, яку вони не опублікували, яку можна запитати безпосередньо або спровокувати обговорення даного питання.

Як шукати об’єкт.

Прямий пошук – це пошук особистої сторінки за допомогою вбудованої пошукової системи за випадковим збігом кількох ознак (ПІБ + дата народження; або ПІБ + місце проживання). Такий пошук по різному

організовано в різних мережах, але він є скрізь, оскільки цей пошук є фундаментом розвитку даної мережі.

Функція для пошуку – у соціальних мережах є різні оператори роботи з пошуком (у Фейсбук це "|", (вертикальний слеш або логічне "АБО"), який допомагає вивести результати, які містять хоча б одну частину пошукового запиту.

Крім того, існує багато спеціалізованих сервісів для пошуку в соціальних мережах (Poiski, 123people – показує результат пошуку по блогам і фотохостингу, по посиланнях і мікроблогу, відео, документів, доменів, телефонними номерами, електронними адресами та іншої інформації, Yoname – пошук за профілями і повідомленнями користувачів певних соціальних мереж, Wink – індексує тільки «соціальні» сайти, де інформація забезпечується тематичними тегами L. Пошук здійснюється по вже відфільтрованій і класифікованій інформації, Bing-social – що корректно шукає через Інтернет].

Через групи – це пошук об'єкта в групах, утворених користувачами (робота, навчання, місця відпочинку, місця служби, групи за інтересами і за захопленнями). Провокування самого об'єкта – можна підштовхнути об'єкт до відповідної реакції на ваші дії (висловлювання) і тим самим ідентифікувати.

Аналіз соціальних мереж

Виявлення експертів в мережах.

Соціальна мережа може бути інструментом для пошуку експертів в конкретній галузі. Виявлення експертів пов'язано з проблемами визначення довіри і розподілу впливу, а також з проблемою поширення інформації в мережі. З цієї точки зору поширення експертного впливу транзитивно, тобто вплив передається від одного вузла до іншого, зменшуючись з кожним залученим вузлом експертів [8–9]. Більш детальний огляд методів виявлення експертів можна знайти, наприклад, в [10].

Результати.

Отримання великих наборів даних соціальних мереж є необхідністю для проведення їх аналізу. В ході роботи була розроблена програма збору даних про друзів, групи і записи зі стіни користувача з соціальній мережі Вконтакті. Розроблена програма підтримує багатопотокове скачування для збільшення швидкості роботи. Збір даних здійснювався з соціальної мережі ВКонтакті. Для збору використовувалися методи VK API для розробників додатків (https://vk.com/dev/methods).

Для аналізу даних соціальних мереж було зроблено чотири етапи роботи (рис. 1–4):

1. Виявлення 20-ти найбільш популярних користувачів (знаходження більшого числа зв'язків користувача серед дерева соціальної групи).

2. Побудова графа на основі результатів першого етапу роботи по матриці суміжності

Вершинами графа є акаунти користувачів з вибірки, а ребрами – зв’язки між ними.

3. Виявлення 20-ти користувачів з найбільшою кількістю друзів.

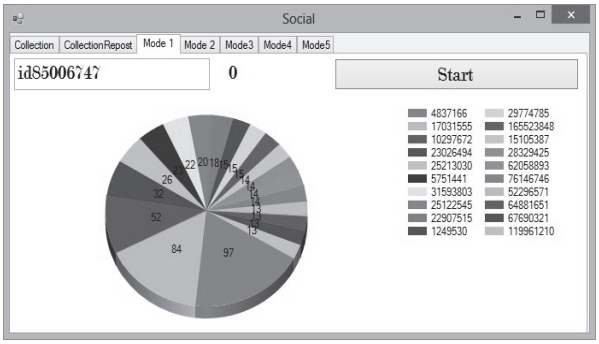

Рис. 1. Результати роботи першого етапу роботи

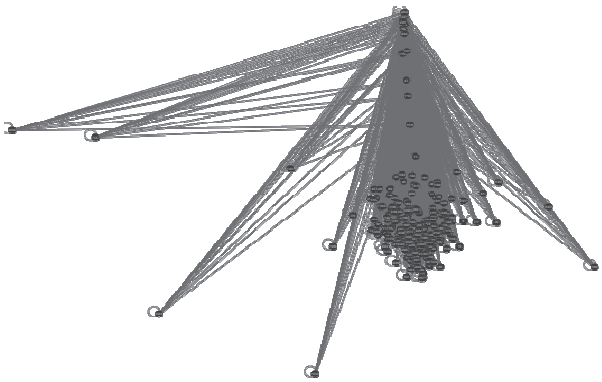

Рис. 2. Результат побудови графа

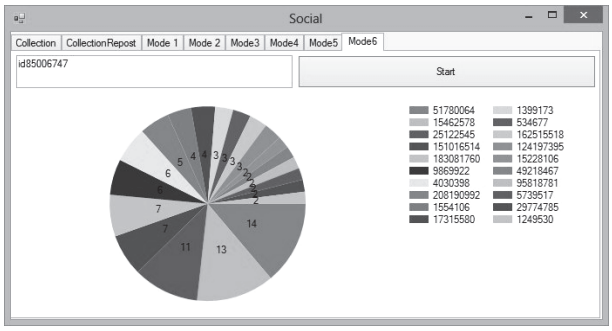

Рис. 3. Результати роботи другого етапу роботи (у відсотках)

4. Виявлення 20-ти користувачів з найбільшою кількістю репостів за 100 актуальними записами на стіні кожного користувача.

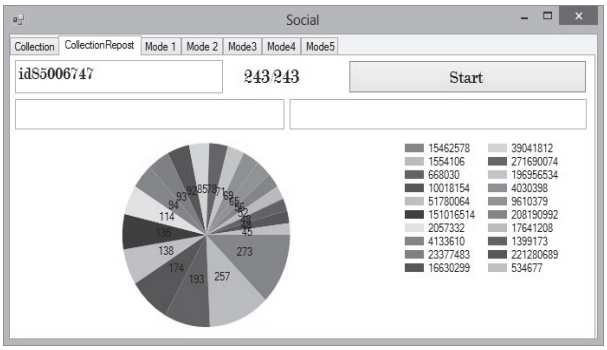

Рис. 4. Результати роботи четвертого етапу роботи

Висновки

В роботі деякі питання пов'язані з аналізом даних соціальної мережі Вконтакті. Для цього було написано програмне забезпечення для збору та аналізу даних цієї соціальної мережі. Після цього було проведено збір даних тестової вибірки користувачів соціальної мережі, який включає в себе: айді, їх друзів та кількість репостів для кожного з них. Основуючись на зібраних даних було знайдено акаунт користувача з найбільшою кількістю друзів, пошук лідера шляхом вибору найбільш популярного акаунта серед його друзів та друзів його друзів. Також був проведений пошук експерта шляхом виявлення акаунта користувача з найбільшою кількістю репостів. Також в роботі було побудовано граф зв'язків користувача соціальної мережі, його друзів та друзів його друзів. Підсумовуючи, можна сказати, що результати даної роботи можна застосовувати для визначення користувачів, які є центром поширення інформації.

Список літератури

1. Palo Alto Networks. Top 10 social networking threats. [Electronic resourse] / Palo Alto Networks // Electronic data. – [Network, 2017]. Mode of access: World Wide Web: http://www.networkworld.com/article/2213704/collaboration-social/top-10-social-networking-threats.html. Accessed 07 March 2017.

2. Carley K. Destabilizing networks / K. Carley, J. Lee, D. Krackhardt // Connections. – Vol. 24, № 3. – 2002. – P. 79-92.

3. Дзюндзюк В.Б. Віртуальні співтовариства: потенційна загроза для національної безпеки / В.Б. Дзюндзюк // Державне будівництво [Електронне видання]. – 2011. – № 1. – Режим доступу до журн.: http:// www.kbuapa.kharkov.ua. – Назва з екрану.

4. Матвиенко Ю.А. Деструктивные сетевые социальные структуры как средство информационной войны и угроза безопасности России / Ю.А. Матвиенко // Информационно-аналитический портал «Геополитика» [Электронный ресурс]. – 2011. – Режим доступу до журн.: http://old.geopolitica.ru/Articles/1218. – Загл. с экрана.

5. Shantanu Ghosh. Top seven social media threats. [Electronic resourse] / Shantanu Ghosh // Electronic data. – [Computerweekly, 2017]. Mode of access: World Wide Web: http://www.computerweekly.com/tip/Top-seven-social-media-threats. Accessed 07 March 2017.

6. Stohl C. Networks of Terror: Theoretical Assumptions and Pragmatic Consequences / C. Stohl, M. Stohl // Communication Theory. – Vol. 17, 93. – 2007. – P. 124.

7. Ігор Нежданов. Особливості соц. мереж. Що таке соціальна мережа [Електронний ресурс] / Ігор Нежданов // Електронні дані. [Гугл, 2017] Доступ: World Wide Web: http://goo.gl/M9Rfuv.

8. Bonchi F., Castillo C., Jaimes A. Social network analysis and mining for business applications. / F.Bonchi, C.Castillo, A.Jaimes // ACM Trans Intell. Syst. Technol. – T. 2, № 3. – 2011. – P. 1-37.

9. Charu C. Social network data analytics / C. Charu // Springer Science & Business Media, 2012. – 486 p.

10. Укустов С.С. Подход к решению проблемы идентификации влиятельных разработчиков в социальной сети gitkhab / С.С. Укустов, А.Г. Кравец // Труды Волгоградского государственного технического университета. – 2012. – Вып. 15 (102). – С. 61-66.

Надійшла до редколегії 12.01.2017

Рецензент: д-р техн. наук проф. Л.О. Кіріченко, Харківський національний університет радіоелектроніки Харків.

АНАЛИЗ ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ

Д.В. Рудченко

Статья посвящена решению маркетинговых задач и вопросов кибербезопасности по анализу информации полученной в социальных сетях. Описаны основные угрозы социальных сетей. Представлены многообразие данных и методы их сбора для проведения анализа. Разработано программное обеспечение для сбора и анализа данных из социальной сети Вконтакте, которое дает скорость обработки даже для больших объемов и точность конечного результата. Проведен сбор и анализ данных, определение пользователей, которые являются центром распространения информации.

Ключевые слова: *анализ социальных сетей, угрозы, выявленные экспертов, кибербезопасность.*

SOCIAL NETWORK DATA ANALYSIS FOR SECURITY

D. Rudchenko

The article is devoted to the solution of marketing problems and cyber security issues in the analysis of information received in social networks. The main threats of social networks are described. A variety of data and methods for their collection for analysis are presented. A software was developed for the collection and analysis of data from the social network Vkontakte, which gives the processing speed even for large volumes and the accuracy of the final result. The collection and analysis of data, identification of users, which are the center for dissemination of information.

Keywords: *analysis of social networks, threats, identified experts, cybersecurity.*