УДК 007.51:004.5

Войтович О.П.
Вінницький національний технічний університет

Дудатьєв А.В.
Вінницький національний технічний університет

Головенько В.О.
Вінницький національний технічний університет

МОДЕЛІ ТА ЗАСІБ ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВИХ ОБЛІКОВИХ ЗАПИСІВ У СОЦІАЛЬНИХ МЕРЕЖАХ

У статті запропоновано метрики ознак фейкових облікових записів у соціальних мережах. Розроблені структурні моделі, що дозволяють виявити фейкові облікові записи у соціальній мережі. На основі запропонованих моделей та використанні рейтингових оцінок розроблено систему підтримки прийняття рішень при виявленні фейкових облікових записів. Результати експериментальних досліджень показали достовірність прийняття рішення близько 0,8.

Ключові слова: *фейкові облікові записи, соціальні мережі, кібербезпека, інформаційні війни, рейтингові оцінки.*

Постановка проблеми. Соціальні мережі є специфічною ареною проведення спеціальних інформаційних операцій, зокрема інформаційно-психологічних операцій, що спрямовуються на суспільство [1; 2]. Сотні мільйонів людей по всьому світу вже активно користуються соціальними мережами для спілкування, перегляду новин тощо, проте велика частка користувачів використовують соціальні мережі як інструмент маніпуляції індивідуальною та суспільною свідомістю за допомогою інформаційних вкидів [1]. Для цього вони використовують фейкові облікові записи, на яких відсутня або міститься неправдива інформація про користувача. Незалежно від конкретних цілей тих, хто створює фейкові облікові записи, їх використання спрямоване, як правило, на зміну суспільної думки в тій чи іншій формі.

Аналіз досліджень. Фейковий обліковий запис (фейк) – це обліковий запис у соціальній мережі з неправдивою інформацією про користувача, власника даної сторінки.

Використовувати фейковий профіль можна як із легальних причин (наприклад для продажу товарів у соціальних мережах тощо) [3], так і зі зловмисною метою (шахрайства, маніпуляції, заборонений контент тощо) [3].

Метою підроблених (фейкових) облікових записів в інформаційних війнах є введення в оману інших користувачів, маніпулювання їхньою поведінкою, причому Інтернет дає можливість здійснювати таке маніпулювання суспільством в цілому [4], і фейкові облікові записи, стали використовуватись як джерело інформаційно-психологічних операцій в інформаційній війні [5].

У задачі аналізу контенту в рамках протидії інформаційній війні важливо те, що певний обліковий запис розповсюджує певний контент (генерує репости, лайки, коментарі) за бажанням або за винагороду. Також існують спеціальні біржі облікових записів, де предметом продажу є готовий обліковий запис з потрібними замовнику параметрами. Виконавці створюють фейковий обліковий запис, наповнюють його контентом, заповнюють його, створюючи ілюзію великого числа друзів, високої активності, входження в потрібні групи тощо. Після цього налаштовують обліковий запис під вимоги замовника (соціальні особливості, переваги, стиль поведінки) і продають замовнику, який вже використовує такі фейкові облікові записи на свій розсуд [6].

Але формалізованих моделей, на базі яких можна розробити інструментарій для аналізу облікових записів, не існує, тому дослідження в цьому напрямку є актуальними.

Постановка завдання. Відповідно метою статті є розробка моделей та засобу для дослідження метрик облікових записів користувачів у соціальній мережі Facebook та створення системи

підтримки прийняття рішень при визначенні фейкових облікових записів у соціальній мережі. Для досягнення мети необхідно вирішити такі задачі: проаналізувати існуючих підходів аналізу облікових записів, розробити структурні моделі ознак фейкових облікових записів, на основі яких розроблено програму автоматизації перевірки облікових записів.

Метрики облікових записів у соціальній мережі. Дослідження показали [5; 7–11], що можна виділити такі основні категорії ознак фейкових облікових записів: лайки, персональні дані, статуси та посилання, друзі, фото, дата народження.

Лайки (*LIKES*) за ознаками можна поділити на їх кількість (*QUANTITY*) та хто їх залишив на сторінці користувача (*FROM*). У свою чергу, залишити лайки можуть як друзі (*Friends*), так і незнайомці (*Strangers*). Для визначення фейковості профілю також має значення кількість лайків (*NumberOfLikes*). Якщо в користувача під певним постом кількість лайків більша за кількість його друзів (*NumberOfFriends*), це може свідчити про те, що користувач отримав ці лайки незвичним шляхом. Відсутність лайків (*NumberOfLikes = 0*) на сторінці вказує на «ізоляцію» користувача, що також може свідчити про його фейковість. Структурна модель метрик у категорії «Лайки» показана на рис. 1.

Параметри моделі ознак фейковості у категорії «Лайки» можна записати у вигляді кортежів:

LIKES = {FROM; QUANTITY}
FROM = {friends; strangers}
QUANTITY = {NumberOfLikes = 0;
NumberOfLikes<NumberOfFriends;
NumberOfLikes>NumberOfFriends}

Персональна інформація на сторінці користувача (*PERSONAL INFORMATION ABOUT USER*) [8] може сказати чимало про фейковість або справжність профілю. Для подальшого аналізу персональну інформацію можна поділити на ім'я користувача (*USER`S NAME*), кількість інформації (QUANTITY OF INFORMATION), суперечливу інформацію (*CONTRADICTORY INFORMATION*) та приватну інформацію (*PRIVATE INFORMATION*).

Ім'я користувача дослідити важко, оскільки існує чимало людей з таким самим ім'ям та прізвищем. Проте варто перевірити ім'я на предмет його співпадіння з ім'ям видатної людини (*IsCelebrityName*). Також слід звернути увагу на те, чи належить ім'я користувача (*UserNameCountry*) до типових імен країни цього користувача (*UserCounrty*).

Відсутність персональної інформації у профілі (*InfoAboutUser*), незначна інформація про інтереси (*NumberOfIntersts*) та групи користувача (*NumberOfGroups*) свідчить про те, що користувач не хоче, щоб його могли ідентифікувати інші користувачі, а отже це теж є ознакою фейковості.

Суперечливість інформації на сторінці є одним з найбільш достовірних показників фейковості, проте і потребує складного аналізу. Наприклад, інформація в постах користувача (*PostsInfo*) не відповідає інформації, зазначеній у профілі (*InfoAboutProfile*), або користувач знаходиться у групах (*InfoAboutGroups*), які не відповідають його зазначеним інтересам (*InfoAboutIntersts*).

До персональної інформації відносять електронну пошту (*IsMailExist*) та номер мобільного телефону (*IsPhoneExist*). Користувачі рідко виставляють таку інформацію у відкритий доступ, на відміну від фейкових облікових записів та спеціально створених рекламних профілів.

Структурна модель метрик у категорії «Персональна інформація про користувача» показана на рис. 2.

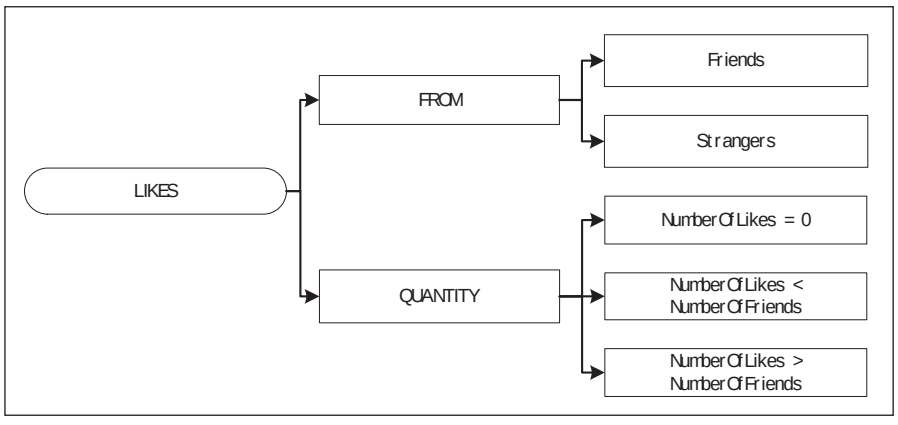

Рис. 1. Структурна модель ознак фейковості у категорії «Лайки»

Параметри моделі ознак фейковості у категорії «Персональна інформація про користувача» можна записати у вигляді кортежів:

PERSONAL INFORMATION ABOUT USER = {USER NAME; NUMBER OF INFORMATION; CONTRADICTORY INFORMATION; PRIVATE INFORMATION}

USER NAME = {IsCelebrityName; UserNameCountry/UserCountry}

NUMBER OF INFORMATION = {InfoAboutUser; NumberOfInterests; NumberOfGroups}

CONTRADICTORY INFORMATION = {PlaceOfStudy/PlaceOfBorn; PlaceOfWork/PlaceOfBorn; ProfileInfo/PostsInfoAboutUser; InfoAboutGroups/InfoAboutInterests; PostsInfo/InfoAboutInterests}

PRIVATE INFORMATION = {isPhoneExists; isEmailExists}

Статуси та пости (*STATUSES AND POSTS ON PAGE*) на сторінці користувача аналізуються як одне ціле, оскільки вони відрізняються лише розміщенням у профілі. Їх можна аналізувати за такими ознаками: за частотою редагування/додавання (*UPDATES*) та за коментарями (*COMMENTS*). Статуси та пости іноді використовуються у якості реклами (*Advertising*) [9].

Частота редагування/додавання постів та статусів (*UpdateFrequency*) вказує на активність користувача. Якщо пости/статуси додаються рідко або дуже часто – це є однією з ознак фейковості. Якщо користувач давно додав пост/статус і протягом тривалого часу не оновлює, існує імовірність того, що цей обліковий запис є фейковим.

Кількість коментарів (*QUANTITY*) також вказує на активність самого профілю. Їх відсутність або надмірна кількість найчастіше буває саме у фейків. Коментарі можуть залишити (*FROM*) як друзі користувача (*Friends*), так і незнайомці (*Strangers*).

Структурна модель метрик у категорії «Статуси та пости» показана на рис. 3.

Параметри моделі ознак фейковості у категорії «Статуси та пости» можна записати у вигляді кортежів:

STATUSANDPOSTS ON PAGE = {POSTS; STATUSES}

POSTS = {Advertising; UPDATES; COMMENTS}

STATUSES = {Advertising; UPDATES; COMMENTS}

UPDATES = {UpdateFrequency; NumberOfPosts}

COMMENTS = {QUANTITY; FROM}

UANTITY = {NumberOfComments;

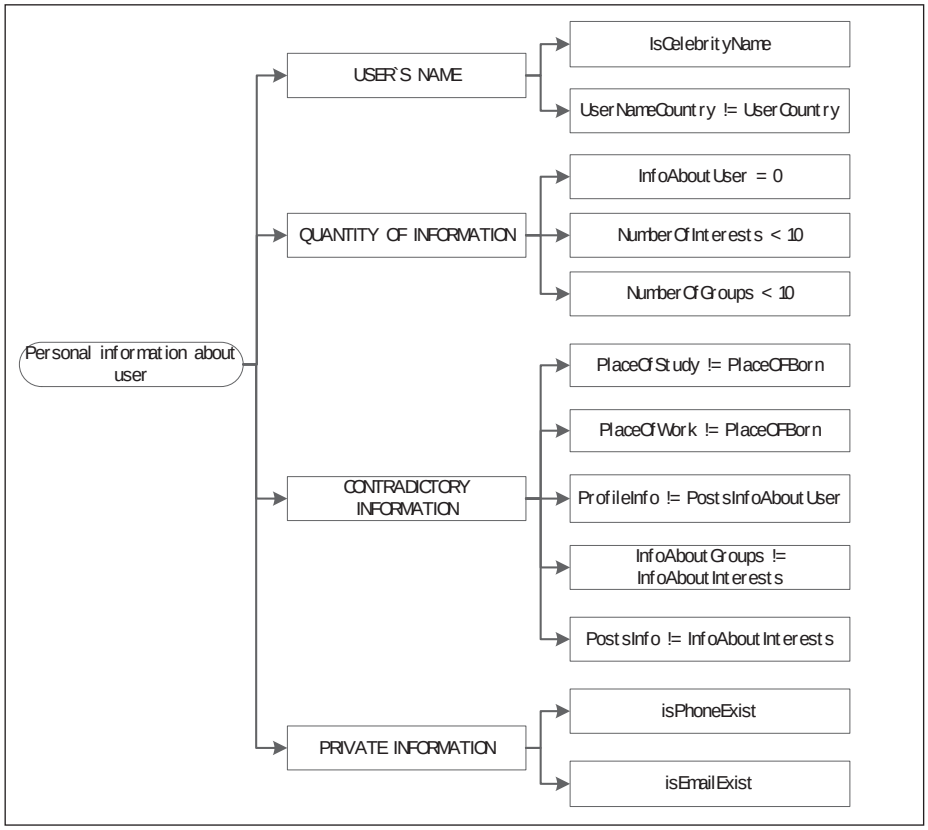

Рис. 2. Структурна модель ознак фейковості у категорії «Персональна інформація про користувача»

NumberOfComments / NumberOfFriends}
FROM = {Friends; Strangers}

Друзі користувача (*FRIENDS*) грають доволі значну роль у визначенні фейка, оскільки вони вказують як на активність профілю в соціальній мережі, так і на коло інтересів користувача [10].

Залежність фейковості від кількості друзів (*QUANTITY*) користувача проаналізувати важко, тому що для того, щоб зробити висновок про фейковість, необхідно аналізувати самих друзів. Наприклад, якщо в списку друзів користувача є фейки (*IsFriendFake*), є імовірність, що і сам користувач – фейк. Якщо користувач не має друзів (*NumberOfFriends = 0*), існує велика імовірність, що його профіль використовується не для спілкування, а для інших цілей. Велика кількість друзів (*Max(NumberOfFriends)*) за короткий проміжок часу (*Min(timeline)*) після створення профілю також викликає підозри, тому, скоріше за все, такий профіль є фейковим. Структурна модель метрик у категорії «Друзі» показана на рис.4.

Параметри моделі ознак фейковості у категорії «Друзі» можна записати у вигляді кортежів:

FRIENDS = {QUANTITY;
INFORMATIONABOUTFRIENDS}
QUANTITY = {NumberOfLikes;
NumberOfFriends; NumberOfLikes/
NumberOfFriends; Max(NumberOfFriends);

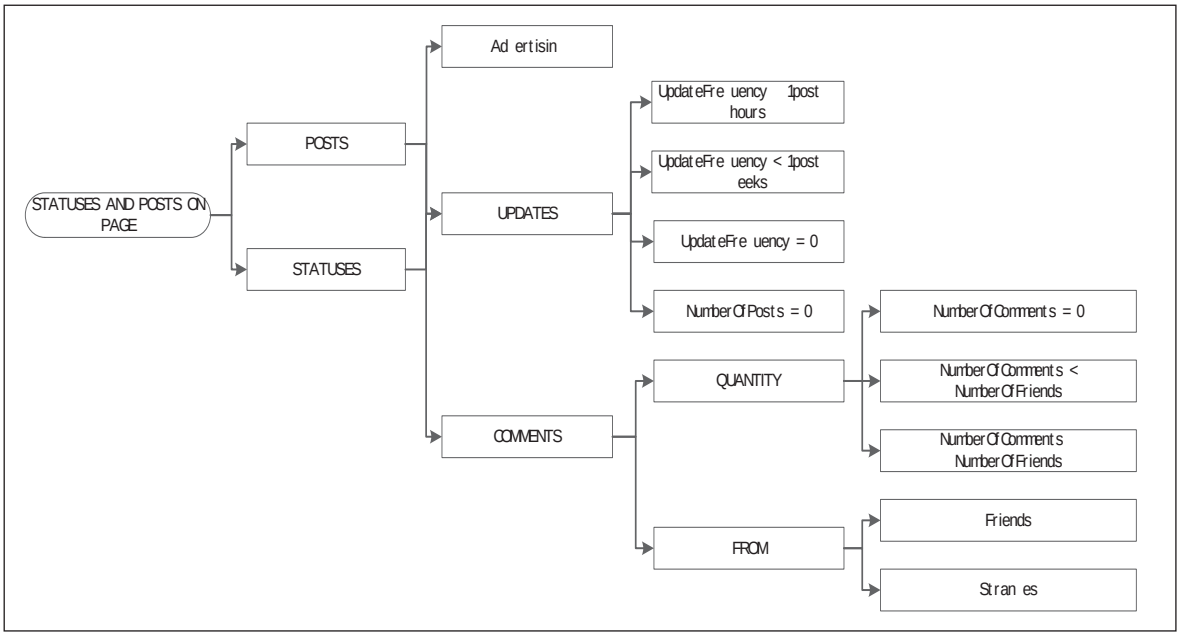

Рис. 3. Структурна модель ознак фейковості у категорії «Статуси та пости»

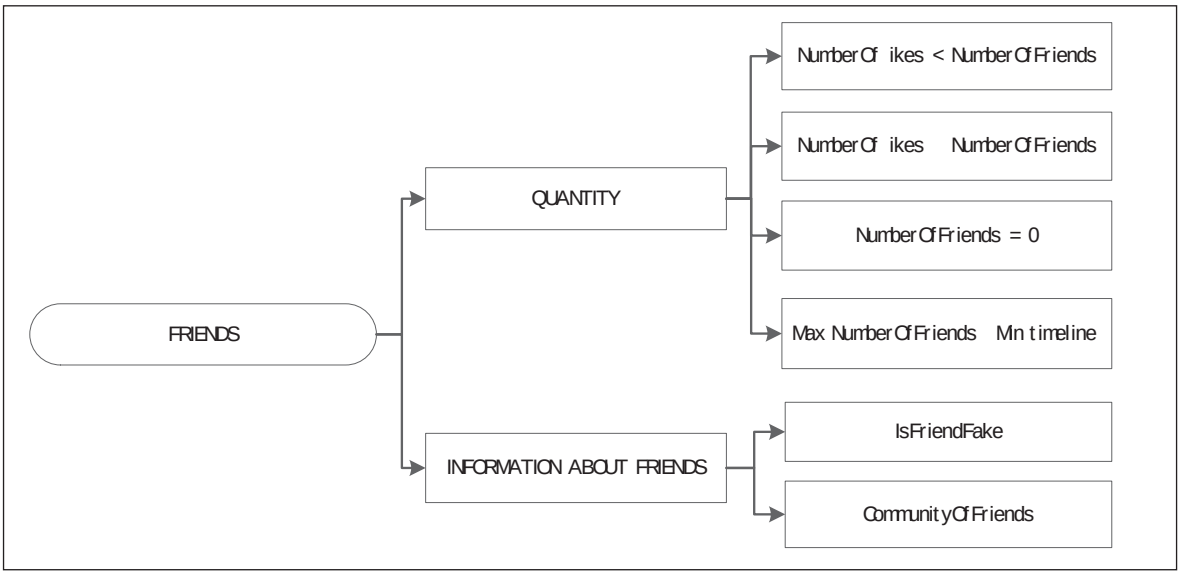

Рис. 4. Структурна модель ознак фейковості у категорії «Друзі»

Min(timeline)}
INFO ABOUT FRIENDS = {IsFriendFake;
CommunityOfFriends}

Аналіз *фотографій користувача* (*PHOTO*) відіграє найважливішу і, водночас, найважчу частину дослідження фейковості облікового запису. По-перше, відсутність фотографій як на аватарі (*AVATAR*), так і в альбомах (*PROFILE*) вже свідчать про те, що даний обліковий запис є фейковим. По-друге, за наявності фотографій на сторінці їх потрібно аналізувати на предмет співпадіння з іншими зображеннями в Інтернеті (*PICTURES ON THE INTERNET*) або з фотографіями інших профілів (*StrangeProfiles*), оскільки користувач міг завантажити замість своїх фотографії знаменитостей (*Celebrities*), тварин (*Animals*) або інших об'єктів (*OtherPictures*). Кількість фотографій (*QuantityOfPhotos*) також є важливим показником, оскільки надмірна або замала кількість фотографій вказує на неправдивість фотографій або неактивність користувача відповідно.

Структурна модель метрик у категорії «Фото» показана на рис. 5.

Параметри моделі ознак фейковості у категорії «Фото» можна записати у вигляді кортежів:

PHOTO = {PROFILE; AVATAR}
PROFILE = {QUANTITY; COINCIDENCE}
AVATAR = {QUANTITY; COINCIDENCE}
QUANTITY = {NumberOfPhotos}
COINCIDENCE = {TheUserOtherProfilePhoto;
StrangeProfiles; PICTURE ON THE INTERNET}
PICTURE ON THE INTERNET = {Celebrities;
Animals; OtherPictures}

Дата народження (*DATE OF BIRTH*) має ознаки, які можуть вказувати на фейковість сторінки. Часто користувачі фейкових облікових записів не приділяють уваги детальному заповненню сторінки та залишають дату народження за замовчуванням (зазвичай 1 січня). Також можлива ситуація, коли вік користувача (*DateOfBirth*) підлягає сумніву або не співпадає з іншими датами на сторінці (*DatesOnProfile*). Наприклад, користувачу 15 років, проте інша інформація на сторінці свідчить, що він закінчив ВУЗ 10 років тому.

Структурна модель метрик у категорії «Дата народження» показана на рис.5.

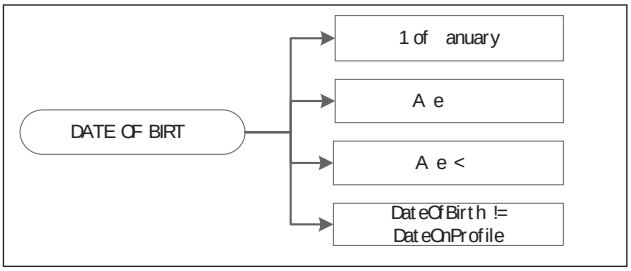

Рис. 6. Структурна модель ознак фейковості у категорії «Дата народження»

Параметри моделі ознак фейковості у категорії «Дата народження» можна записати у вигляді кортежів:

DATEOFBIRTH = {DateOfBirth; DateOfBirth/
DateOnProfile}

Звичайно, окремо ці критерії не можуть однозначно вказувати на «фейковість» облікового запису, оскільки лише аналіз їх об'єднання може поставити під сумнів справжність облікового запису. Для більш точного визначення статусу облікового запису необхідно використовувати аналіз з використанням якомога більшої кількості критеріїв.

У цій статті не враховані інші важливі параметри облікових записів, наприклад час створення сторінки, швидкість формування кола друзів тощо, ці та інші параметри будуть враховані в подальших дослідженнях.

Система підтримки прийняття рішення на основі рейтингових оцінок. Для прийняття

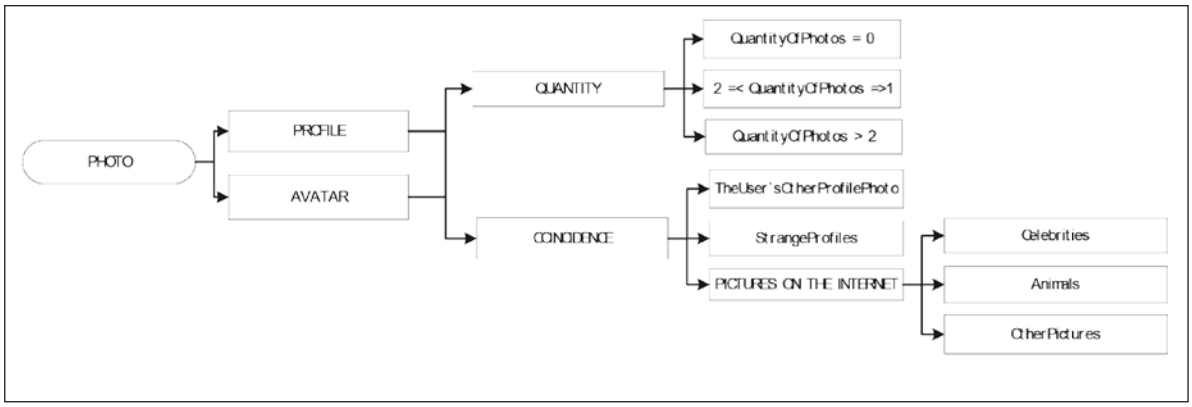

Рис. 5. Структурна модель ознак фейковості у категорії «Фото»

рішення щодо фейковості облікового запису запропоновано використовувати метод рейтингових оцінок, що дозволяє оцінювати інформацією, яка поділяється за категоріями, та враховує коефіцієнти значимості параметрів.

Нехай система оцінюється за n показниками, x_i – значення i-го показника. Представленням рейтингової оцінки є лінійна згортка, математична модель якої записується у вигляді [12]:

$$F = \sum_{i=1}^{n} \lambda_i x_i ,$$

де, λ вага i-го показника, що визначається експертом.

На основі рейтингового підходу і багатофакторного аналізу розробляється група рейтингових оцінок показників і встановлюється зв'язок між ними. При цьому застосування рейтингового підходу передбачає, що рейтингові оцінки присвоюються усім групам факторів.

Нехай система ознак, що оцінюється, описується на основі заданого набору показників, таких як $X = (x_1, ..., x_i, ..., x_n)$. Показники можуть бути різнорідними: числовими, логічними, лексичними, векторними тощо. Для оперування різнорідними показниками для кожного з них введено нормовану функцію, яка будь-яке значення показників x_i переводить у множину дійсних значень на відрізку [0; 1], тоді $0 \le x_i \le 1$. Нормування може призвести до завищення або заниження фактичного показника, проте цей негативний ефект нормування нейтралізується за допомогою введення для кожного показника вагового коефіцієнта або індикатора значимості, що визначається емпіричним способом (наприклад, методом експертних оцінок). При існуванні залежності показників між собою необхідно врахувати змішані зведення, де коефіцієнтами є коефіцієнти кореляції відповідної пари.

Якщо вагові коефіцієнти підбирати за умови нормування, то цільова функція виконуватиме роль рейтингу на відповідному рівні ієрархії системи. Для того, щоб процес рейтингового підходу мав максимальний ефект, до нього мають бути залучені показники всіх факторів. Ця умова автоматично виконується, коли рейтингова оцінка на кожному рівні ієрархії співпадає з цільовою функцією. Інтегральна рейтингова оцінка відображає пріоритети показників. Формування цих показників, а отже і формування рейтингової оцінки, проводиться експертами.

Експериментальні дослідження. Для роботи з даними соціальної мережі Facebook було обрано мову програмування Python та бібліотеку Facebook-SDK [13, 14]. Для того, щоб отримати доступ до інформації про користувача у соціальній мережі

Таблиця 1

Результати аналізу облікових записів користувачів

Користувач	Статус облікового запису	Результат програми, балів	Висновок програми
Vitalii Holovenko	Справжній	37,8	Справжній
Татьяна Головенько	Справжній	45,9	Не визначено
OleksandrTopchii	Фейк	67,5	Фейк
IvanVorobyov	Справжній	19,8	Справжній
Alex Rudyk	Фейк	90,0	Фейк
Ольга Гнатюк	Справжній	21,6	Справжній
Петро Петрович	Фейк	96,3	Фейк
Andrii Beatle	Фейк	62,1	Фейк
Жека Олейник	Справжній	34,2	Справжній
Владислав Круговой	Справжній	54,0	Не визначено
Олеся Войтович	Справжній	48,6	Не визначено
Talii Santie	Фейк	77,4	Фейк
Jenny Rahl	Фейк	70.2	Фейк
Sergey Hubchakevych	Справжній	28,8	Справжній
Георгий Выфв	Фейк	87,3	Фейк
Alice Black	Фейк	66,6	Фейк
Liliana Vess	Фейк	69,3	Фейк
Konrad Von H.	Фейк	67,5	Фейк
Сергей Тарахта	Справжній	24,3	Справжній
Иван Петров	Фейк	44,8	Справжній
Fin Age	Фейк	65,7	Фейк

Facebook, необхідно отримати токен автентифікації, у якому вказані права розробника на доступ до даних у соціальній мережі [14].

Як показник обрано систему балів, що свідчить про фейковість облікового запису користувача. Кожен з параметрів під час аналізу залежно від умови отримує певну кількість балів а від 1 до 5. Так, 100 балів показує, що обліковий запис є фейковим, а 0 балів – справжнім. Для різних категорій було обрано різні вагові коефіцієнти, опираючись на експертні знання. Якщо результат дослідження становить від 10 до 45 балів, то система приймає рішення , що обліковий запис є справжнім, від 55 до 100 балів – обліковий запис є фейковим. Проте якщо результат було отримано у межах від 45 до 55 балів – необхідно провести додаткові дослідження. У результаті роботи програмного засобу на екран виводиться інформація про фейковість чи справжність облікового запису користувача.

Для тестування перевірено різні облікові записи користувачів у соціальній мережі Facebook, серед яких були як фейкові, так і справжні облікові записи. Наприклад, обліковий запис користувача VitaliiHolovenko є справжнім, але деяка інформація про користувача відсутня. У результаті перевірки програмний запис видає результат 37,8 балів, отже визначає його як справжній. Результати аналізу 21 облікового запису наведено у табл. 1.

Достовірність прийняття рішення становить близько 81%.

Висновки. Розглянуто та проаналізовано основні метрики соціальної мережі Facebook, на основі яких можна визначити фейковий обліковий запис. Проаналізовано кожну з метрик за їх можливими параметрами та впливом на статус облікового запису. Кожну з метрик віднесено до відповідних категорій для подальшої зручності їх аналізу.

Запропоновано структурну модель ознак для виявлення фейкових облікових записів за означеними метриками, яка включає в себе такі категорії як: лайки, персональні дані, статуси та посилання, друзі, фото, дата народження.

Розроблена система підтримки прийняття рішень, яка реалізує виявлення фейкових облікових записів у соціальній мережі Facebook на основі рейтингових оцінок. Експериментальні дослідження показали достовірність прийняття рішення системою 0,8.

Список літератури:

1. Voitovych O., Holovenko V. Research of social networks as a source of information in warfare. Inżynier XXI wieku projectujemy przyszlosc: monografia / pod red: Jacek Rysiński. Bielsko-Biała, 2016. С. 111–119.

2. Дудатьєв А. В., Войтович О. П. Інформаційна безпека соціотехнічних систем: Модель інформаційного впливу. Інформаційні технології та комп'ютерна інженерія. 2017. № 38. С. 16–21.

3. Коршунов А., Белобородов И., Бузун Н. Анализ социальных сетей: методы и приложения. Труды Института системного программирования РАН. 2014. Т. 26. № 1. С. 439-456.

4. Дудатьєв А.В. Комплексна інформаційна безпека СТС: моделі впливу та захисту : монографія. Вінниця: ВНТУ, 2017. 128 с.

5. Нежданов И.Ю. Технологии информационных войн в Интернете URL: http://bash.rosmu.ru/activity/attach/events/1283/01.pdf (дата звернення: 10.01.2017).

6. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. М.: Физматлит, 2010. 228 с.

7. Michal Kosinski, Sandra C. Matz, Samuel D. Gosling, Vesselin Popov, David Stillwe. Facebook as a Research Tool for the Social Sciences. Opportunities, Challenges, Ethical Considerations, and Practical Guidelines. American Psychologist. 2015. Vol. 70. No. 6. 543-556 pp.

8. Aaron Aguis. 10 Metrics to Track for Social Media Success. URL: https://www.socialmediaexaminer.com/10-metrics-to-track-for-social-media-success/ (дата звернення: 10.01.2017).

9. Батура Т.В., Копылова Н.С., Мурзин Ф.А., Проскуряков А.В. Методы анализа данных из социальных сетей. Вестник НГУ. Серия: Информационные технологии. 2013. Т. 11. Вып. 3. С. 5–21.

10. Берни Хоган. Анализ социальных сетей в интернете, 2013. URL: https://postnauka.ru/longreads/20259 (дата звернення: 10.01.2017).

11. Горчинская О., Ривкин А. Анализ данных социальных сетей. Открытые системы. СУБД. 2015. № 3. С. 22.

12. Худяков Ю.Г., Николайкин Н.И., Андрусов В.Э. Управление опасностями производственной среды: Монография. М.: ООО «Проспект», 2017. 122 с.

13. API Reference URL: https://facebook-sdk.readthedocs.io/en/latest/api.html (дата звернення: 10.01.2017).

14. Find your Facebook ID URL: https://findmyfbid.com/ (дата звернення: 10.01.2017).

МОДЕЛИ И СРЕДСТВО ОПРЕДЕЛЕНИЯ ФЕЙКОВЫХ УЧЕТНЫХ ЗАПИСЕЙ В СОЦИАЛЬНЫХ СЕТЯХ

В статье предложены метрики признаков фейковых учетных записей в социальных сетях. Разработанные структурные модели, позволяющие определить фейковые учетные записи в социальной сети. На основе предложенных моделей и рейтинговых оценок разработана система поддержки принятия решений при определении фейковых учетных записей. Результаты экспериментальных исследований показали достоверность принятия решения около 0,8.

Ключевые слова: *фейковые учетные записи, социальные сети, кибербезопасность, информационные войны, рейтинговая оценка.*

MODELS AND INSTRUMENT FOR DETECTION OF FAKE ACCOUNTS IN SOCIAL NETWORKS

The metrics for social network fake account detection are proposed in the article. Structure model that allow detection of fake account in the social network is designed. Based on the proposed models and the ranking score, a decision making support system for detection of fake accounts in social network is developed. Experimental research shows accuracy of system decision making about 0.8.

Key words: *fake accounts, social networks, cyber security, information wars, ranking score.*