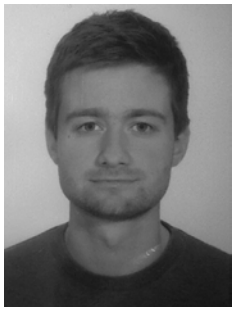
© Nechayev O. Methods of detecting and analysis criminal networks based on billing information of cellular mobile operator // Ukrainian Scientific Journal of Information Security, 2015, vol. 21, issue 3, p. 236-244.

МЕТОДИ ВИЯВЛЕННЯ ТА АНАЛІЗУ КРИМІНАЛЬНИХ МЕРЕЖ СФОРМОВАНИХ НА ОСНОВІ БІЛІНГОВОЇ ІНФОРМАЦІЇ ОПЕРАТОРІВ МОБІЛЬНОГО ЗВ'ЯЗКУ

Олександр Нечаєв

Інститут спеціального зв'язку та захисту інформації НТУ України «КПІ», Україна



НЕЧАЄВ Олександр Олексійович

Рік та місце народження: 1988 рік, м. Київ, Україна.
Освіта: Інститут спеціального зв'язку та захисту інформації НТУУ «КПІ», 2010 рік.
Посада: аспірант.
Наукові інтереси: теорія складних мереж, аналіз соціальних графів.
E-mail: a4455395@gmail.com

Анотація. *Сучасній людині все складніше уявити себе без використання таких засобів комунікації як Інтернет та мобільний зв'язок. Майже всі сучасні злочини зі стадії готування до вчинення здійснюються з використанням електронних засобів зв'язку та залишають гетерогенні сліди в інформаційному просторі. Окремі з таких слідів у встановленому законом порядку збираються та акумулюються правоохоронними органами. Великі об'єми цих даних не дають можливості їх неавтоматичної обробки. Не так давно на стику соціології та теорії складних мереж виник окремий науковий напрям – аналіз соціального графу, підвидом якого є аналіз кримінальних мереж. В даній роботі запропонована архітектура експертної системи з виявлення та аналізу організованих злочинних угрупувань на основі автоматизованої обробки білінгової інформації операторів мобільного зв'язку. До уваги пропонується метод, завдяки якому можливо ідентифікувати кримінальні угрупування із сукупності звичайних соціальних контактів в мережах телефонного зв'язку. Також, приводяться результати застосування розробленого методу на реальних соціальних та кримінальних мережах. Описуються методи планування ефективного деструктивного впливу та проведення активних заходів проти організованої злочинності. Окрім того, автором пропонується метод викриття внутрішньої структури кримінальних мереж, заснований на модифікації відомого алгоритму пошуку релевантних веб-сторінок.*

Ключові слова: *виявлення ознак кримінальних мереж, злочинні угрупування, мережі телефонних контактів, ранжирування вузлів, планування деструктивного впливу.*

Вступ

На даний час важко уявити сучасний світ без використання таких засобів комунікацій як Інтернет та мобільний зв'язок. Окрім того, не менш ніж в побуті, електронні засоби комунікації використовуються для координації та вчинення злочинних дій. Відомо, що всі вони залишають гетерогенні сліди в інформаційному просторі, частина яких у встановленому законом порядку, акумулюються правоохоронними та іншими спеціальними органами для виконання задач із забезпечення громадської та державної безпеки. Поміж тим, занадто великі об'єми такого роду інформації виключають можливість її якісної обробки людськими ресурсами.

Не так давно на стику соціології та теорії складних мереж виник новий науковий напрям – аналіз соціальних мереж, та його підвид – аналіз кримінальних мереж, методи яких починають активно впроваджуватися в роботу правоохоронних органів та спецслужб розвинутих країн світу. Найбільш явним представником кримінальної мережі у вказаному контексті є мережі, побудовані на основі статистичної білінгової інформації операторів мобільного зв'язку (час виклику, тривалість, ідентифікатори співрозмовників, тощо).

В даній роботі представлена теоретична основа для створення автоматизованої системи з виявлення кримінальних мереж та їх аналізу з метою викриття їх структури та планування ефективних деструктивних дій для їх подальшого знищення або припинення протиправної діяльності.

Аналіз існуючих досліджень

Початкові дослідження в області аналізу соціальних мереж (підвидом яких є кримінальні мережі) проведені в 60-х роках Мільграмом і Траверсом [1], де аналізуються характеристики реальних складних мереж, проводяться соціальні експерименти в реальному світі, та представлена так звана «модель малого світу». Згідно цієї моделі доводяться докази, що, не зважаючи на велику розмірність, складні мережі показують ряд загальних властивостей. Наприклад, ступеневий розподіл

соціальної мережі підкоряється степеневому закону, хоча і асимптотично, тобто відношення вузлів графу, що мають k зв'язків між собою, до числа всіх вершин для великих значень визначається як: $P(k) \sim k^{-\gamma}$, де γ - це константа, значення якої знаходиться зазвичай у межах $2 < \gamma < 3$, однак інколи значення γ може бути поза цими межами. Також в таких мережах існує відносно короткий шлях, що з'єднує довільну пару вузлів.

Барабаши А. [2] довів модель зростання мережі, яка була декілька разів успішно використана в World Wide Web, телекомунікаційних мереж, тощо. Автор довів, що реальні складні мережі мають одну і ту ж динаміку зростання, так звану пільгову прихильність: нові вузли частіше з'єднуються з вузлами з великим ступенем (кількістю контактів), ніж з не великим. Ці досліди спричинили виникнення поняття «безмасштабності» мережі (тобто специфічного розподілу вузлів та їх відповідного розвитку), дозволяючи наявність вузлів-концентраторів (вузлів-зірок) та посередників (тих, через які проходить максимальна кількість мінімальних шляхів між довільними парами вузлів). Таким чином «ефект малого світу», привілейована модель зростання, та степеневий розподіл ступенів характеризують структуру соціальної мережі [5].

Перші роботи з дослідження злочинних організацій методами теорії складних мереж належать Малколму Спарроу [3]. Він визначив чотири риси, які відрізняють кримінальні мережі від інших соціальних структур:

1. Обмежений розмір. Кримінальні мережі переважно складаються не більше ніж з 1-2 тисяч вузлів.

2. Неповнота інформації. Кримінальні мережі неминуче містять порожні фрагменти інформації та помилкову інформацію.

3. Невизначені кордони. Важко визначити всі відносини окремих вузлів.

4. Динаміка. Нові стійкі з'єднання здійснюються лише з вузлами схожих мережевих структур.

Завдяки М. Спарроу інші дослідники взялись за дослідження кримінальних мереж як підвиду соціальних. На приклад, Бейкер і Фаулкнер [5] вивчали злочинні мережі, порівнюючи їх розвиток із законами, притаманним електричному обладнанню. Арквілла і Ронфельд [4] узагальнили попереднє дослідження шляхом введення концепції мережевого протистояння тероризму. Останні явно демонструють різницю між звичайними соціальними мережами та кримінальними.

Перше практичне застосування методів теорії складних мереж в аналізі злочинних організацій здійснено Валдісом Кребсом, який розкрив структуру Аль-Каїди шляхом аналізу мережі телефонних контактів осіб, які мали місце в подіях 11 вересня 2001 року в США [8]. В. Кребсом також було вперше застосовано візуалізацію кримінальних мереж, що довело її важливість в експертному аналізі. Завдяки цьому випадку стартувало широке застосування теорії складних в діяльності правоохоронних та розвідувальних органів розвинутих країн світу. Результати таких дій періодично з'являються в пресі, але специфічні механізми аналізу саме кримінальних мереж природно підпадають під гриф державної або комерційної таємниці.

У той час, аналіз складних мереж різної природи є об'єктом досліджень багатьох науковців, результати яких відкриті. Окремі з цих досліджень можуть застосовуватись у протидії злочинності та тероризму. З урахуванням викладеного, є сенс сформувати архітектуру експертної системи з виявлення та аналізу кримінальних мереж, в якій об'єднати ті новітні розробки теорії складних мереж, які зможуть ефективно застосовуватись в даному напрямку.

Майже всі сучасні злочини залишають гетерогенні сліди в інформаційному просторі, які можливо використовувати для формування складних мереж, та ефективного їх застосування у викритті злочинних угрупувань та їх нейтралізації. Визначимо види первинної інформації, на основі якою можуть бути викриті кримінальні мережі:

- Інтернет-листування електронною поштою;
- білінгова інформація операторів мобільного зв'язку;
- засоби Інтернет-спілкування (Skype, Viber, Whatsapp, Telegram тощо);
- спеціальні бази даних спецслужб та правоохоронних органів.

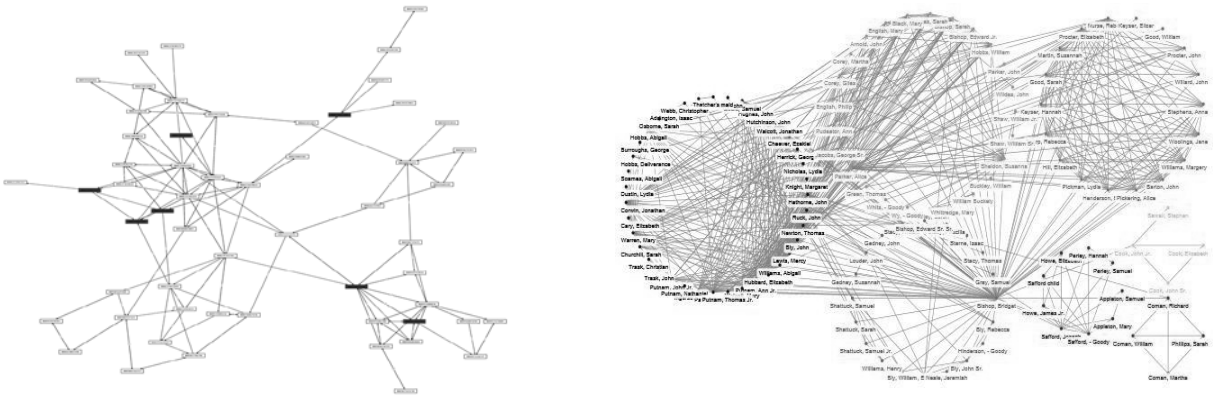
Всю вище перелічену інформацію отримують та зберігають у встановленому законом порядку спеціальні та правоохоронні органи. В даній роботі основною для побудови складної мережі буде вважатися білінгова інформація операторів мобільного зв'язку (яка містить статистичну інформацію телефонних контактів, таку як тривалість виклику, номери учасників бесіди, дата та час з'єднання, місцезнаходження абонентів на час сеансу, тощо).

Згідно законодавства більшості країн білінгову інформацію правоохоронні органи можуть використовувати в рамках оперативно-розшукової або іншої справи стосовно абонентів та їх зв'язків, в діях яких вбачаються ознаки готування або скоєння злочину; або інформацію щодо телефонних з'єднань абонентів, які знаходились в межах певної базової станції мобільного зв'язку в зоні дії якої здійснено готування до скоєння або скоєно злочин.

Таким чином в загальному випадку є два види первинної інформації для виявлення кримінальної мережі:

1. Мережа телефонних контактів відомих функціонерів злочинного угрупування, на основі аналізу контактів яких необхідно виявити повну структуру кримінальної мережі (рис. 1.а).

2. Мережа телефонних контактів, частиною якої є злочинне угрупування, яке необхідно ідентифікувати (рис. 1.б).



а) *б)*

Рис. 1. Мережі: а) злочинців, де виділені відомі правоохоронним органам, не виділені – ті, що слід виявити; б) з виділеними групами, одна з яких є злочинним угрупуванням

Для виявлення злочинної мережі з випадку зображеного на рис. 1а необхідно виділити ряд правил, згідно яких проводити ітераційне виявлення спільників встановлених функціонерів злочинного угрупування до тих пір, доки не буди виявлено жодного нового члена. Ці правила можливо визначити лише на основі збору статистичної інформації щодо характеру телефонних контактів серед членів реальних кримінальних мереж. Оскільки такі дані не містяться у відкритому доступі та на даний час не існує алгоритму моделювання кримінальних мереж, властивості яких були б наближені до реальних, вирішення даного випадку в роботі розглядатися не буде.

У випадку на рис 1б виявлення кримінальної проходить в два етапи:

1) Виявлення соціальних груп. Існує багато таких алгоритмів, кожен з яких має свої переваги та недоліки [20].

2) Ідентифікація серед виявлених груп підмережі, яка за певними ознаками є злочинним угрупуванням. На даний час у відкритих джерелах відсутні методи математичного вирішення цього питання.

В. Каширін в роботі «Аналіз та моделювання кримінальних мереж» демонструє різницю в степеневому розподілі злочинних мереж різного типу та розмірності [21]. Порівнюючи його дані з результатами досліджень Ксіомі Ванга, який проводив дослідження даного коефіцієнта на реальних складних мережах різної природи [22] можна безперечно вказати, що кримінальні мережі відносяться до так званих «безмасштабних» мереж, але не мають специфічного розподілення ступенів вузлів, на основі якого їх можна ідентифікувати серед інших мереж.

Sarvari H. та Abozinadah E. в експериментах з реальними кримінальними мережами, сформованими на основі контактів відомих членів злочинних угрупувань засобами соціальних мереж та електронної пошти приділяють увагу показникам центральності злочинних мереж [23]. Проте, ці коефіцієнти також не дозволяють однозначно ідентифікувати мережу як кримінальну.

Нехай якимось чином вдалось ідентифікувати кримінальну мережу. Визначимося з рядом питань, відповіді на які є найбільш цікавими для правоохоронних органів, та які необхідно дати в результаті безпосереднього аналізу цієї мережі:

1) викриття структури організованого злочинного угрупування, визначення її лідерів та певних груп, які виконують конкретні злочинні функції;

2) визначення найбільш оптимального та ефективного шляху деструктивного впливу на кримінальну мережу з метою припинення її протиправної діяльності або її існування;

3) прогнозування розвитку злочинної мережі та планування проведення стосовно неї активних оперативних заходів (впровадження в ОЗУ інформаційного джерела, придбання джерела із злочинного середовища тощо).

Питання викриття шляху ефективного деструктивного впливу на злочинну організацію є безперечно найбільш цікавим, а з'ясування структури самої мережі та прогнозування її розвитку є лише допоміжною інформацію яка допоможе у плануванні заходів саме зі знищення мережі.

У теорії графів під ефективним деструктивним впливом на мережу розуміється втрата зв'язності, коли мережа ділиться на два або більше не з'єднаних між собою частин [24]. У випадку з кримінальними мережами під атакою на вузол розуміється нейтралізація особи причетної до протиправної діяльності шляхом його захоплення або іншої дії, яка призведе до втрати контакту особи з іншими членами угрупування.

Природно, всі реальні мережі після вдалої атаки намагаються відновити свою зв'язність шляхом створення нових з'єднань [24]. В криміналістиці не часто трапляється шанс схопити все злочинне угрупування і частіше на свободі залишаються окремі функціонери, доказову базу на яких здобути не вдалось. Таким чином, планування атак та прогнозування подальшого розвитку мережі є дуже складним питанням та потребує значної уваги з боку дослідників.

Експериментальні дослідження автора [17] та інших дослідників [1-4] довели, що кримінальні мережі, як і інші безмасштабні мережі, нечутливі до випадкових атак, а також до атак на вузли з найбільшим ступенем (кількістю зав'язків).

Метою даної роботи є розробка архітектури експертної системи для вирішення наступних задач:

1) Виявлення кримінальних мереж на основі аналізу білінгової інформації операторів мобільного зв'язку.

2) Аналіз кримінальних мереж з метою:

- викриття внутрішньої структури ОЗУ;
- визначення лідерів;
- прогнозування розвитку злочинної мережі;
- визначення прихованих інформаційних джерел ОЗУ;
- планування активних оперативних заходів (на приклад впровадження інформаційного джерела в злочинну мережу);
- планування ефективних деструктивних дій та атак з метою припинення протиправної діяльності мережі або припинення її існування.

Основна частина дослідження

Виявлення кримінальних мереж.

Під соціальною групою або субграфом будемо розуміти сукупність вузлів (абонентів мобільного зв'язку), де внутрішньо групові зв'язки більш щільні ніж зовнішні (рис. 2). Також, буде використовуватись припущення, згідно з яким кожний вузол може входити тільки в одну групу.

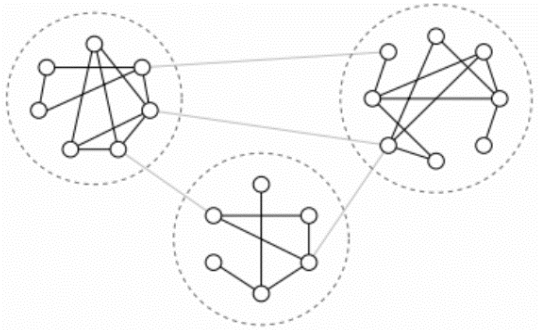


Рис. 2. Приклад соціальної мережі, в яку входить 3 групи

Таким чином, ми маємо мережу, з виділеними соціальними групами, однією з яких є кримінальна мережа, яку необхідно виявити. Як було зазначено у попередньому розділі, кримінальні мережі не відрізняються від звичайних соціальних мереж певними значеннями таких базових показників теорії складних мереж як розподілення ступенів вузлів та різноманітних коефіцієнтів центральності [25]. Поміж тим, багатьма дослідниками, включаючи автора даної роботи проводились дослідження з розробки методів виявлення неявних або прихованих зав'язків в мережах злочинців [14,18].

Спарроу М. виділяв наявність помилкової інформації, як одну з основних властивостей кримінальних угрупувань [4]. Наявність цієї помилкової інформації спричинена наміром приховати певний контакт, у зв'язку з чим в мережі виникають так звані аномалії. Аномалії притаманні мережам різної природи, та їх виявлення і аналіз застосовується для розуміння топології мереж, викриття їх вразливості в вірусології, медицині, дослідженні комп'ютерних мереж, тощо [26].

В даному випадку ми будемо розглядати саме соціальну аномалію, яка виникає з високої ймовірності існування зв'язку між двома особами при його фактичній відсутності.

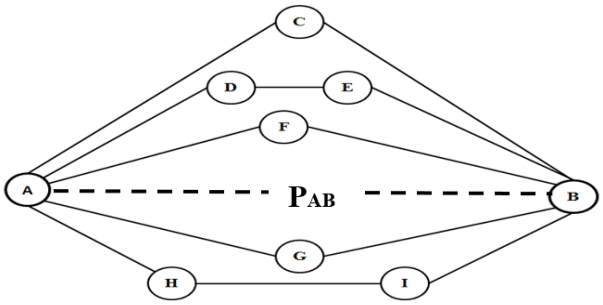


Рис. 3. Схема відображення неявного зв'язку між вузлами **A** та **B**

На рис. 3 Зображена схема контактів соціальних акторів (вузлів) **A** та **B**. Дані вузли мають багато спільних контактів та, відповідно, опосередкованих зав'язків. У той же час, фактичний зв'язок між ними відсутній. Саме таку ситуацію ми будемо називати соціальною аномалією.

Такі феномени, трапляються і в звичайних соціальних мережах. Наприклад, люди (або родичі) посварились між собою та навіть при великій кількості спільних знайомих не підтримують прямого між собою контакту.

Для математичної реалізації цього методу уявимо мережу, в якій всі існуючі зв'язки представлені ймовірностями їх існування та можуть бути знайдені за певними змістовними ознаками, або отримані експертним шляхом. В цьому випадку мережа може бути представлена у вигляді так званої «матриці неявних зв'язків», елементами якої є ймовірності зв'язків між вузлами.

Основа запропонованого методу – обчислення ймовірностей зав'язків між всіма парами вузлів мережі з урахуванням ваги прямих (при їх наявності) та сумарної ваги опосередкованих зв'язків між парою через один, два, або більше вузлів. Розглянемо для пояснення методу найбільш простий випадок - мережу, що складається з трьох вузлів V_1, V_2 та V_3, з'єднаних між собою зв'язками з ймовірностями $p_{1,2}$, $p_{1,3}$ та $p_{2,3}$. Ймовірність зв'язку між вузлами V_1 та V_2 буде вираховуватись за формулою:

$$p_{1,2} = 1 - (1 - p_{1,2})(1 - p_{1,3}p_{2,3}). \quad (1)$$

Якщо шляхи між вузлами i та j можливі через декілька вузлів, то загальна формула ймовірності зв'язку між ними буде мати такий вигляд:

$$p_{i,j} = 1 - (1 - p_{i,j}) \prod_{k \neq i,j} (1 - p_{i,k}p_{k,j}) \prod_{k \neq i \neq l \neq j} (1 - p_{i,k}p_{k,l}p_{l,j}). \quad (2)$$

Для випробування методу було проведено експерименти на наступних реальних соціальних мережах:

- мережа контактів профілів Facebook з вільної бази Стенфордського університету (4039 вузлів, 88234 зв'язки);
- мережа контактів профілів Twitter з вільної бази Стенфордського університету (вузлів 81306, 1768149 зв'язки);

- Ернонівська мережа контактів електронною поштою з вільної бази Стенфордського університету (вузлів 36692, 183831 зв'язок);

- мережа контактів «Карате клуб» (вузлів 34, зв'язків 78), яка за своїми властивостями наближена до терористичних для класично використовується для відповідних дослідницьких експериментів [19];

- мережа контактів терористичного формування Аль-Каїда (вузлів 19, зв'язків 54), які мали місце в трагічних подіях 11.09.2001 року [8].

Етапи експерименту:

1) Виявлення субграфів в мережі за алгоритмом Louvain [27] (окрім терористичних мереж, оскільки вони вже є сформованими групами).

2) Визначення середньої кількісті соціальних аномалій (відсутніх зв'язків між парами вузлів, вірогідність існування яких перевищує середню вірогідність існування реально існуючого зв'язка).

Для виявлення груп та обчислення коефіцієнтів використовувалась мова програмування Python. В результаті кластеризації не злочиних мереж алгоритмом Louvain середній розмір субграфа мережі склав 53,2 вузли.

Результати експерименту приведені в табл. 1.

Таблиця 1

Кількість соціальних аномалій в терористичних та звичайних мережах

№	**Мережа**	**Середня кількість соціальних аномалій**
1	Facebook	1,3
2	Twitter	1,1
3	Ernon emails	0,8
4	Karate club	5
5	Al-qaeda	3

Видно як сильно відрізняється кількість соціальних аномалій в реальних та злочинних мережах. На жаль у відкритому доступі не має достатньої кількості даних для відтворення кримінальних мереж щоб виконати більш масштабний та об'єктивний експеримент. Проте за наявними даними запропонований метод себе виправдовує.

Аналіз кримінальних мереж.

Планування активних оперативних заходів.

Всі природні мережі знаходяться в постійному русі та намагаються досягти цільності [1,2]. Не виключенням є і терористичні мережі. Дж. Барроса довів в своїх дослідженнях, що реальні складні мережі мають більший шанс до з'єднання з новими вузлами в місцях зі слабким інформаційним потоком [30]. При аналізі кримінальних мереж з метою визначення частини мережі для сприятливого впровадження інформаційного джерела дане дослідження набуває значної ваги.

Соціологом Р. Бйортом була запропонована концепція структурних ям як альтернатива теорії соціального капіталу [16]. В області аналізу кримінальних мереж, структурна яма розуміється під місцем, де частини мережі не зв'язані між собою, або їх зв'язок занадто слабкий (рис. 4).

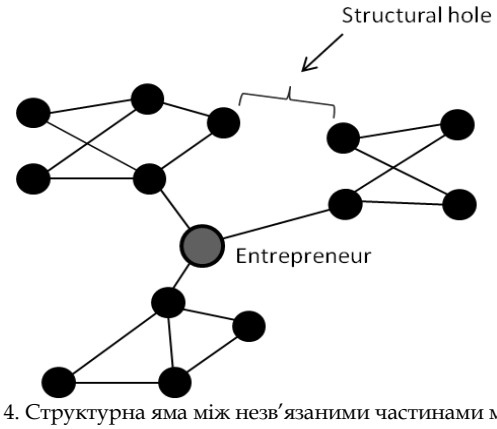


Рис. 4. Структурна яма між незв'язаними частинами мережі

«Ентерпренером» в нашому випадку виступає інформаційне джерело, для якого необхідно обрахувати сприятливе місце впровадження в мережу з максимальними показниками вірогідності його прийняття членами ОЗУ. При цьому, успіх операції буде залежати не тільки від самого «намагання» мережі посилити інформаційний потік в певній своїй частині, а й від розміру соціальних груп між якими планується впровадити джерело.

Планування атак на злочинну мережу.

Як приводилось вище, показником ефективної атаки на мережу є втрата зв'язності – коли після видалення одного або декількох вузлів або зв'язків між вузлами мережа ділиться на дві не зв'язані частини [24]. В більшості дослідах проводяться моделювання атак на вузли мережі. Хоча існує й інший вид атак, який необхідно враховувати в аналізі злочинних мереж – атака на зв'язок. Нижче розглянуті обидва випадки.

Часто може здаватись, що найбільш ефективною атакою є видалення вузлів з найбільшим числом зв'язків (ступінь вузла). Поміж тим, експериментально автором було доведено, що це є у більшості випадків найефективнішим [18]. Найшвидше мережі втрачають зв'язність при видаленні вузлів за коефіцієнтом посередництва.

Коефіцієнт посередництва (Betweenness) – показник теорії складних мереж, який вказує скільки найкоротших шляхів між всіма можливими парами вузлів проходить через конкретний вузол (рис. 5).

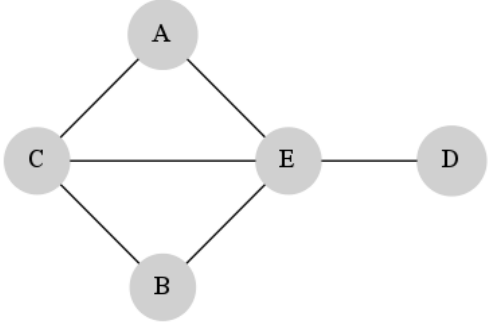


Рис. 5. Мережа, в якій вузол **E** має найбільше значення показника посередництва

Посередництво для вузла n розраховується за формулою:

$$B(n) = \sum_{i \neq j \neq n \in V} \frac{\sigma_{ij}(n)}{\sigma_{ij}}, \quad (3)$$

де σ_{ij} – кількість найкоротших шляхів між вузлами i та j, $\sigma_{ij}(n)$ – кількість найкоротших шляхів між вузлами i та j, які проходять через вузол n, V - множина всіх вузлів.

З рис. 7 видно, що вузол **E** має найбільший показник посередництва, та його видалення призведе до ділення мережі на дві частини. У даному прикладі один і той же вузол має найбільший коефіцієнт посередництва та найбільшу ступінь. Однак це притаманно тільки невеликим мережам, які складаються з не більше ніж 100 вузлів. У великих природних мережах найбільші значення цих двох показників належать, як правило, різним вузлам.

Моделювання атак на вузол часто застосовується через його ефективність та відносну простоту реалізації. Атака на зв'язок є більш тонкою. Для прикладу можливо схопити одного або декількох з членів ОЗУ, та мережа на деякий час може втратити зв'язність (видалення вузла). У той же час, під видаленням зв'язку між вузлами кримінальної мережі розуміється поширення певної достовірної або хибної інформації, в результаті чого дві особи перестають підтримувати між собою зв'язок. Дане рішення використовується, коли не має вагомих підстав для затримання члена ОЗУ, та у той же час не має іншого шляху для припинення протиправної діяльності злочинного угрупування.

У деяких випадках, атака на один зв'язок є достатньою для втрати зв'язності мережі. В теорії складних мереж існує спеціальний термін, який відображає такий випадок. Він має назву «міст» (рис. 6).

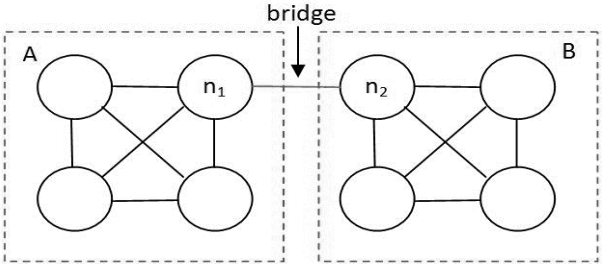


Рис. 6. Міст – зв'язок, який є з'єднувальною ланкою між двома частинами мережі або групами

З урахуванням викладеного, при плануванні атак на мережу необхідно враховувати можливість як атак на вузли, так і на зв'язки, що їх з'єднують.

Викриття внутрішньої структури кримінальної мережі.

Відомо, що злочинні угрупування, розмір яких перевищує десяток членів є квазіієрархічними структурами, тобто мають лідера, та розподілення конкретних функціональних ролей [28]. Для викриття внутрішньої структури кримінальної мережі та визначення її лідерів автором пропонується метод ранжирування вузлів, ітеративний принцип якого можливо сформулювати наступними словами: «лідером у мережі є той, хто віддає найбільшу кількість доручень тим, хто, у свою чергу, теж віддають багато (але менше за перших) доручень, і т.д. …».

Математична реалізація даного принципу наближена до алгоритму HITS (Hyperlink-Induced Topic Search), запропонованого Дж. Клейнбергом для пошуку найбільш вагомих веб-сторінок відповідно запиту користувача на основі інформації, яка визначається множиною гіперпосилань [29]. Згідно зазначеного алгоритму для кожного вузла мережі (в даному випадку веб-сторінки) обраховуються взаємопов'язані показники авторства (auth) та портальності (hub) за формулами:

$$hub(A_i) = \sum_{A_i \to A_j} auth(A_j), \quad auth(A_i) = \sum_{A_j \to A_i} hub(A_j). \quad (4)$$

Неможливість застосування класичного алгоритму Дж. Клейнберга полягає в тому, що він не враховує вагу зв'язків між вузлами, що в мережах телефонних контактів грає важливу роль. На приклад, очевидна різниця в стійкості зв'язку між парою абонентів які спілкувались 100 разів за місяць і тими, які спілкувалися лише один раз. Таким чином, автором пропонується використання модифікованого алгоритму HITS:

$$hub(A_i) = \sum_{A_i \to A_j} auth(A_j) \cdot \log_2 E_{ij}, \quad auth(A_i) = \sum_{A_j \to A_i} hub(A_j) \cdot \log_2 E_{ji}. \quad (5)$$

Поміж тим, експериментальним шляхом автором було з'ясовано, що навіть модифікований алгоритм допускає помилки через велику кількість надлишкових зв'язків в мережі. Виходячи з цього для коректного результату необхідно попередньо відкинути частину зв'язків, які для даного дослідження не мають сенсу.

Під квазіієрархічною мережею розуміється певна організація, в якій соціальні актори знаходяться у певних відносинах між собою. А саме:

- керівник – підлеглий;
- нейтральні відносини (на приклад працівники сусідніх відділів).

У направленому графі, де кожний зв'язок має напрямок (абонент А дзвонить абоненту В), на основі вищевикладеного можна виділити три види зв'язків:

- керівник → підлеглий;
- підлеглий → керівник;
- підлеглий ↔ підлеглий.

Для реалізації алгоритму важливою інформацією буде зв'язок керівник → підлеглий, або іншими словами відпрацювання доручення керівником підлеглому. Інша інформація буде заважати та негативно впливати на результат.

Для попередньої фільтрації мережі від небажаної «шумової» інформації автором запропонована модель, що заснована на наступних загальних гіпотезах [17]:

- керівник частіше зв'язується із підлеглим, тобто вага вихідного зв'язку мережі від керівника має перевищувати вагу вхідного зв'язку, що йде від підлеглого;

- керівник зв'язується із підлеглим хаотично, тоді коли у нього виникає потреба. Тобто нормована дисперсія розподілу інтервалів часу зв'язку у напрямку «керівник → підлеглий» приймає відносно великі значення;

- підлеглий звітує перед керівником щодо виконання відпрацьованого доручення з певною закономірною періодичністю. Розподіл інтервалів часу зв'язку контактів «підлеглий → керівник» (звітування), які починаються з попереднього зв'язку «керівник → підлеглий» (відпрацювання доручення) має рівномірний розподіл.

На рис. 7 зображений приклад часового розподілу контактів між керівним вузлом (**A**) та підлеглим (**B**). Розподіл часових інтервалів між дорученнями керівника до підлеглого не є сталим. Поміж тим, розподіл інтервалів часу між дорученнями керівника та звітуваннями підлеглого є близьким до сталого. Дисперсія σ^2 для інтервалів часу, що застосовується для формальної оцінки у подібних випадках розраховується за формулою:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(T_i - \langle T \rangle\right)^2, \tag{6}$$

де n – кількість інтервалів часу T_i, $\langle T_i \rangle$ – середнє значення інтервалу часу.

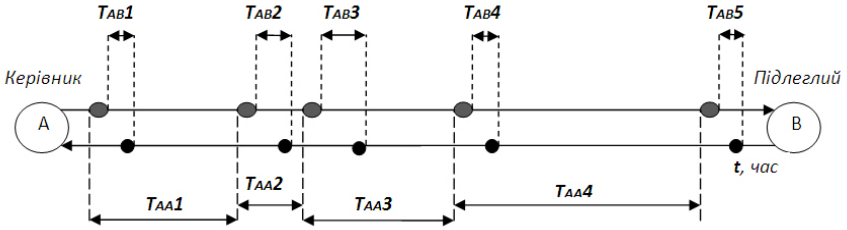
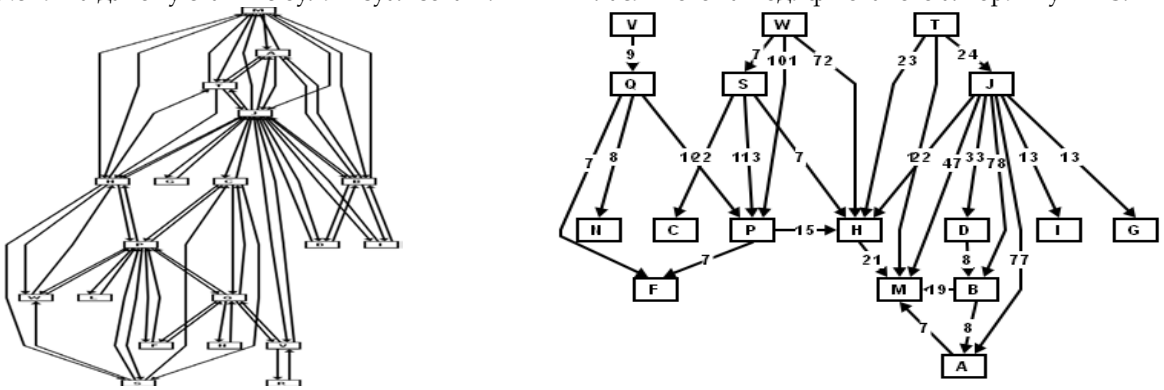


Рис. 7. Періодичність вхідних та вихідних контактів між керівником (**A**) та підлеглим (**B**)

Для демонстрації різниці у використанні класичного та модифікованого алгоритмів взято мережу з 19 вузлів, між якими завідомо відомі їх соціальні ролі (рис. 8а). Вузли в мережі мають між собою двосторонні контакти з певною вагою, які для збереження наочності схеми на даному етапі не були візуалізовані.

На рис. 8б зображена однонаправлена мережа з відкинутими зв'язками підлеглий→керівник. Також, візуалізовано ваги зв'язків. Далі, для відфільтрованої мережі розраховані показники авторства для кожного соціального актора із застосуванням класичного та модифікованого алгоритму HITS.



а) *б)*

Рис. 8. Мережі: а) з двосторонніми зв'язками; б) з відфільтрованими зв'язками

Таблиця 2

Результати ранжування вузлів в мережі за класичним та модифікованим алгоритмами для вузлів з високим значенням показника авторства

	№	Вузол	Auth		№	Вузол	Auth
Класичний HITS	1	J	1	Модифікований HITS	1	T	1
	2	T	0,51606		2	J	0,49843
	3	B	0,50811		3	W	0,48652
	4	H	0,48723		4	B	0,48513
	5	A	0,31251		5	P	0,38136
	6	S	0,29636		6	S	0,3695
	7	M	0,27271		7	A	0,30829
	8	W	0,26119		8	C	0,28531
	9	P	0,23921		9	M	0,26766

З результатів експерименту, приведених в Табл. 2 видно значну різницю розрахунків класичним та модифікованим алгоритмами. З чого випливає підтвердження важливості врахування ваги зв'язків між вузлами, як одного з ключових факторів структурного аналізу соціальних мереж.

Висновки

В останні роки по всьому світу спостерігаються регулярні терористичні атаки проти мирного населення. Добре розуміння структури терористичних та злочинних організацій може значно допомогти попередженню таких випадків. Розроблена архітектура експертної системи з аналізу кримінальних мереж, сформованих на основі

білінгової інформації операторів мобільного зв'язку експериментально доводить свою необхідність у використанні правоохоронними органами для підвищення ефективності боротьби зі злочинністю.

В роботі демонструються конкретні особливості, які відрізняють кримінальні мережі від звичайних соціальних мереж. Розроблений метод дозволяє ідентифікувати ознаки кримінальної мережі, що підтверджується рядом експериментів з реальними соціальними та злочинними мережами.

Також, пропонується алгоритм викриття внутрішньої структури злочинного угрупування з використанням модифікованого для цих цілей метода Дж. Клейнберга, який класично використовується для ранжирування релевантних веб-сторінок. Дана модифікація робить можливим математично виявити лідерів ОЗУ та певні групи, зосереджені на виконанні тих чи інших функцій.

Окрім того, автором пропонується використання соціологічного поняття «структурних ям» для планування впровадження інформаційного джерела в злочинну мережу. Показано як структурні ями відрізняються від виявлення секторів складної мережі зі слабким інформаційним потоком, показано переваги їх використання.

Об'єднання приведених в роботі методів складає архітектуру експертної системи виявлення та аналізу кримінальних мереж, сформованих на основі статистичної інформації телефонних з'єднань абонентів мобільного зв'язку, впровадження якої в правоохоронних органах надає значні переваги у боротьбі зі злочинністю та тероризмом.

Література

[1] Travers, Jeffrey & Stanley Milgram. An Experimental Study of the Small World Problem, Sociometry, Vol. 32, No. 4, 1969, p. 425-443.

[2] Barabasi, Albert-Laszly and Reka Albert, Emergence of scaling in random networks, Science, 286, October 15, 1999, p. 509-512.

[3] Sparrow M.K. The application of network analysis to criminal intelligence: An assessment of the prospects. Social Networks, 13(3), 1991, p. 251–274.

[4] Arquilla J., Ronfeldt D. Networks and netwars: The future of terror, crime, and militancy. Survival, 44(2), 2001, p. 251–274.

[5] Baker W., Faulkner R. The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am. Social. Rev., 58, 1993.

[6] Ferrara E., De Meo P., Catanese S., Fiumara G. Detecting criminal organizations in mobile phone networks. Expert Systems with Applications, 41(13), 2014, p. 5733–5750.

[7] Klerks P., Smeets E. The network paradigm applied to criminal organizations: Theoretical nitpicking or a relevant doctrine for investigators. Connections, 24, 2001, p 53–65.

[8] Krebs V. Mapping networks of terrorist cells. Connections, 24(3), 2002, p. 43–52.

[9] Morselli C. Assessing vulnerable and strategic positions in a criminal network. Journal of Contemporary Criminal Justice, 26(4), 2010, p. 382–392.

[10] Schneider F., Feldmann A., Krishnamurthy B., Willinger W. Understanding online social network usage from a network perspective. In Proc. 9th SIGCOMM conference on Internet measurement conference, 2009, p. 35-48.

[11] Zang H., Baccelli F., Bolot J. Bayesian inference for localization in cellular networks. In 2010 Proceedings IEEE INFOCOM, IEEE, 2010, p. 1–9.

[12] Xu J., Chen H. Criminal network analysis and visualization. Comm. ACM, 48(6), 2005, p. 100–107.

[13] Yang C., Chen H., Hong K. Visualization of large category map for internet browsing. Decis. Support Syst., 35(1), Apr. 2003, pp. 89–102.

[14] Yang C., Liu N., Sageman M. Analyzing the terrorist social networks with visualization tools. In Intelligence & security informatics, 2006, pp. 189–198.

[15] Freeman L. Set of measures of centrality based on betweenness. Sociometry, 1977, p. 35–41.

[16] Burt Ronald S. Structural Holes and Good Ideas. American Journal of Sociology, 110 (2), 2004, p. 349–399.

[17] Ланде Д.В., Нечаєв О.О. Алгоритм ранжирування вузлів квазіієрархічних мереж соціального характеру // Проблеми інформатизації та управління, 49(1), 2015, С. 46-50.

[18] Ланде Д.В., Нечаєв О.О. Відновлюваність зв'язків у безмасштабних мережах // Реєстрація, зберігання і обробка даних. – 2012. – Т. 14, № 3. – С. 92-98.

[19] Zachary W. An information flow model for conflict and fission in small groups, Journal of Anthropological Research 33, 452-473 (1977).

[20] Mehjabin Khatoon W., Aisha Banu A Survey on Community Detection Methods in Social Networks, I.J. Education and Management Engineering, 2015, 1, p. 8-18.

[21] Каширин В. Анализ и моделирование криминальных сетей. – СПб НИУ ИТМО, «Итнернет и современное общество». – 2012. – 24 с.

[22] Xiaomin Wang. Deciding on the type of the degree distribution of a graph (network) from traceroute-like measurements, 2011, International Journal of Computer Networks & Communications, vol. 4 (3), pp. 151-167.

[23] Hamed Sarvari, Ehab Abozinadah, Alex Mbaziira. Constructing and Analyzing Criminal Networks,2014 , Security and Privacy Workshops, 84-91.

[24] Nisha Chaurasia, Akhilesh Tiwari, Efficient Algorithm for Destabilization of Terrorist Networks, 2013, International Journal of Information Technology & Computer Scien;Nov2013, Vol. 5 Issue 12, p. 21.

[25] Hyoungshick Kim, Ross Anderson, Temporal node centrality in complex networks, 2012, Computer Laboratory, University of Cambridge, 15, p. 85.

[26] Carlos Garc´ıa C., Andreas Vost, Jochen Kogel, Distributed Anomaly Detection with Network Flow Data, 07.2015, Detecting Network-wide Anomalies, ISAR conference.

[27] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment, Vol. 2008, No. 10.

[28] U.S. Army Training and Doctrine Command, A Military Guide to Terrorism in the Twenty-First Century, 2008, Chapter 3, p. 1-14.

[29] Kleinberg J. Authoritative sources in a hyperlinked environment // Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, Philadelphia, PA, 1998. – p. 668-677.

[30] Barros J. Information Flows in Complex Networks, 2009, Information Theory and Statistical Learning, p. 267-287.

УДК 004.724.4 (045)

Нечаев А.А. Методы выявления и анализа сетей преступников, построенных на основе биллинговой информации операторов мобильной связи

Аннотация. *Современному человеку все сложней представить себя без использования таких средств коммуникации как Интернет и мобильная связь. Почти все преступления, включая стадию приготовления, совершаются с использованием электронных средств связи, которые оставляют гетерогенные следы в информационном пространстве. Отдельные из таких следов в установленном законом порядке собираются и аккумулируются правоохранительными органами. Большие объемы данных не позволяют обрабатывать их вручную. Не так давно на стыке социологии и теории сложных сетей возникло отдельное научное направление – анализ социального графа, и как его подвид - анализ криминальных сетей. В данной работе предложена архитектура экспертной системы по выявлению и анализу организованных преступных группировок на основе автоматической обработки биллинговой информации операторов мобильной связи. Предлагается метод, благодаря которому возникает возможность выявления преступных группировок из совокупности простых социальных контактов в сетях телефонной связи. Также, приводятся результаты использования разработанного метода на реальных социальных и криминальных сетях. Описываются методы планирования эффективного деструктивного воздействия и проведения активных мероприятий против организованной преступности. Кроме того, автором предлагается метод анализа внутренней структуры криминальных сетей, основанный на модификации известного алгоритма поиска релевантных веб-страниц.*

Ключевые слова: *выявление признаков криминальных сетей, преступные группировки, сеть телефонных контактов, ранжирование узлов, планирование деструктивного воздействия.*

Nechayev O. Methods of detecting and analysis criminal networks based on billing information of cellular mobile operator

Abstract*. Nowadays it is difficult to imagine a modern person without such means of communication as the Internet and mobile communications. Almost all modern crimes starting from preparation and to commitment are carried out by using electronic means of communication and leave heterogeneous traces in cyberspace. In accordance with the law some of these tracks are collected and accumulated by law enforcement agencies. Because of large volumes of data they can't be processed manually. Not long ago a separate scientific direction – analysis of social networks, with analysis of criminal networks as a subdivision, appeared at the crossing of sociology and the theory of complex networks. This paper proposes the structure of expert system aimed at detection and analysis of organized criminal groups on the basis of automatic data processing of billing information of mobile operators. A method is proposed which allows identify criminal groups based on a pool of regular social contacts in telephone communication networks. The proposed method was tested in real social and criminal networks and results are given in this paper. Methods of effective destructive actions planning and implementation of active operational measures in relation to organized crime are described. In addition, the author proposes the method of disclosing internal structure of criminal networks, based on the modification of the famous search algorithm of relevant web pages.*

Key words*: detecting criminal networks, offender organizations, mobile phone networks, ranking nodes, destructive influence planning.*

Отримано *21 жовтня 2015 року*, затверджено редколегією *11 листопада 2015 року*