

КОНКУРЕНТНА РОЗВІДКА ТА УПРАВЛІННЯ ЗНАННЯМИ / BUSINESS INTELLIGENCE & KNOWLEDGE MANAGEMENT

МЕТОД ВИЯВЛЕННЯ PR-ВПЛИВУ ЧЕРЕЗ ІНТЕРНЕТ ЗМІ

Мирослав Рябий¹, Олег Хатян², Сергій Багацький¹

¹ТОВ «Миронівський хлібопродукт», Україна

²Компанія WisdomWell, Україна



РЯБИЙ Мирослав Олександрович, к.т.н.

Рік та місце народження: 1987 рік, с. Салиха, Київська область, Україна.

Освіта: Національний авіаційний університет, 2009 рік.

Посада: фахівець відділу інформаційної безпеки з 2015 року.

Наукові інтереси: криптографія, квантова криптографія, національна безпека, прогнозування загроз та викликів, дослідження інформаційного простору.

Публікації: більше 50 наукових публікацій, серед яких наукові статті, матеріали та тези доповідей на конференціях, навчальні посібники, патенти.

E-mail: m.ryabyy@ukr.net



ХАТЯН Олег Анатолійович

Рік та місце народження: 1968 рік, м. Київ, Україна.

Освіта: НТУУ «КПІ», 1996 рік.

Посада: директор центру впровадження інформаційних технологій з 2014 року.

Наукові інтереси: національна безпека, інформаційна безпека, прогнозування загроз та викликів, дослідження інформаційного простору, експертні системи.

Публікації: більше 20 наукових публікацій, серед яких наукові статті, матеріали і тези доповідей на конференціях та навчальний посібник.

E-mail: ohatyan@gmail.com



БАГАЦЬКИЙ Сергій Петрович

Рік та місце народження: 1979 рік, м. Городище, Черкаська область, Україна.

Освіта: Харківський військовий університет, 2001 рік.

Посада: начальник відділу інформаційної безпеки з 2014 року.

Наукові інтереси: інформаційна безпека, прогнозування загроз та викликів, дослідження інформаційного простору, експертні системи.

Публікації: більше 5 наукових публікацій.

E-mail: s.bagatskiy@mhp.com.ua

Анотація. Розглядається інформаційний простір, утворений інформаційним потоком повідомлень Інтернет ЗМІ. В основі роботи лежить гіпотеза про відмінність трендів тематичних інформаційних потоків типу неупередженого інформування та «PR-вплив» у характеристичному просторі лінгвістичних ознак. Наведено метод розподілу загального інформаційного потоку на складові тематичні потоки, та спосіб відокремлення тематичних потоків з лінгвістичними ознаками «PR-впливу». Наведені дані попередніх емпіричних досліджень дозволяють оцінити граничні умови для прийняття рішення щодо показника наявності «PR-впливу» тематичного потоку. Також аналіз наведених результатів показує наявність зв'язку окремих тематичних трендів, отриманих у результаті застосування нашого метода. Цей факт відкриває можливості теорії складних мереж для реконструкції топології реального інформаційного простору та проведення якісного аналізу інформаційних потоків.

Ключові слова: інформаційний простір, інформаційний потік, інформаційний вплив, інформаційна операція, метод виявлення PR-впливу, тематичний тренд, тематичний вектор, складні мережі.

Постановка проблеми у загальному вигляді

Поточний аналіз матеріалів Інтернет ЗМІ дозволяє дійти висновку про наявність розподілу інформаційного потоку (ІП) за певними типами, які мають принципову відмінність процесу формування інформаційного простору з точки зору завдань. Серед них:

- неупереджене інформування – метою є відображення інформаційного сліду подій, процесів, явищ реального світу, реакції суспільства та держави на ці події (робота державних інститутів);

- спрямований інформаційний вплив – завжди має в основі чітко сформульовану цільову функцію;

- аналітичні матеріали – посідають особливе місце серед усіх матеріалів Інтернет ЗМІ, мають на меті подання оцінок експертів, відповідної інфографіки, інтерв'ю, повідомлень насичених аналітичним змістом.

Зауважимо стосовно двох останніх, що такі матеріали можуть посилатись на одну та більше подій відносно певного об'єкту, процесу, явища, тощо, проводити порівняння, висвітлювати динаміку подій та надавати їм відповідну оцінку в позитиві або негативі. Підбір фактів та відносність шкали оцінки потенційно криють певні механізми спрямування впливу. З метою досягнення найбільшого ефекту впливу за більш короткий проміжок часу можуть використовуватись не правдиві (або фейкові) факти.

Крім того, важливою особливістю інформаційного простору, на нашу думку, є наявність в ньому тематичних трендів (вузькому розумінні, як послідовність певних «тем дня» та за визначенням у [2]), що іманентно підпорядковуються деякій цільовій функції (часто прихованому задуму) у випадку свідомо спрямованого впливу («PR-впливу»). Емпіричні дослідження показують, що такі тематичні тренди (а особливо ті, що складають модель «PR-впливу») не є корпускулярними, тобто замкнутими на собі. Тренди перетинаються на тематичному рівні, що надає спрямованому впливові мережевих ознак. Відповідно, подібних ознак набувають усі наступні похідні (інформаційна операція (ІО), агресія, війна, що окремо відзначено у [2]). Тому проблема побудови методу виявлення інформаційних потоків типу «PR-впливу» на якісному рівні, який ми будемо розглядати в цій роботі, в умовах сьогодення набуває особливої *актуальності*.

Аналіз останніх досліджень і публікацій

Дослідженню ліній трендів динаміки публікацій в інформаційному просторі, що відповідають ІО присвячені роботи багатьох вітчизняних науковців. В деяких з них тренди розглядаються як динаміка кількісної оцінки наповнення ІП повідомленнями та застосовуються аналітичні моделі (логістична модель, модель дифузії інформації з метою врахування старіння інформації) для побудови шаблону виявлення інформаційних операцій [2]. У [1,2] глибоко опрацьовані сучасні підходи, принципи, механізми

та моделі інформаційних операцій і впливів, втім якісна оцінка та критерії тематичного визначення спрямованого інформаційного впливу не надаються. У [1,3,4] з метою дослідження закономірностей формування та розвитку сукупності документів, що відповідають певній тематиці пропонується мультіагентна модель (інструмент імітаційного моделювання), де поняття агент асоціюється з повідомленням корпусу тематичного інформаційного потоку (ТІП). Підхід дозволяє отримати різні типи еволюції ІП, але не враховує можливі перетини кількох тем в рамках тренду. Окремі роботи присвячені феномену інформаційно-мережевої війни в сучасному інформаційному просторі. Однак, більшість з них розглядають її загальну природу, характерні риси, методи та цілі [5], оставляючи поза увагою зміст прикладних моделей виявлення складових елементів цього явища. У попередніх роботах нами було розглянуто завдання, методи та механізми роботи PR-служб, що дозволило сформулювати відповідне поняття «PR-вплив» як засіб реалізації певних інтенцій через публікації в Інтернет ЗМІ. Крім того ми визначили його особливості, критерії та оціночні характеристики. Сформульована гіпотеза про відмінність процесу «PR-впливу» від «неупередженого» інформування. Запропонована формальна модель виявлення цих впливів на множині інформаційних повідомлень отриманих від Інтернет ЗМІ [6]. Крім того, нами визначено поняття «теми дня» як конструкції, що враховує лексичні та змістовно-сміслові ознаки групи пов'язаних документів з колекції повідомлень ІП, та запропоновано алгоритм їхньої побудови [7]. Послідовності «тем дня» ранжовані у часі дають інформаційні тренди («тематичні тренди»), яким притаманні певні характеристики. *Дослідження трендів ТІП у характеристичному просторі лінгвістичних ознак, на нашу думку, дозволить отримати метод виявлення «PR-впливу».*

Метою цієї роботи є побудова методу виявлення «PR-впливу» через пошук способу розподілу потоку інформаційних повідомлень які представляють колекції відповідних документів з лексикою властивою «PR-впливу» та відмінних від таких на основі формування відповідних трендів, що і буде *завданням* нашої роботи.

Об'єктом дослідження обрано інформаційний простір, утворений впорядкованою у часі множиною інформаційних повідомлень електронних Інтернет ЗМІ.

Предмет наукової роботи – метод виявлення «PR-впливу» через інформаційні повідомлення електронних Інтернет ЗМІ.

Виклад основного матеріалу

Як ми вже зазначали, тренди інформаційних повідомлень можуть утворювати мережеву модель «PR-впливу». Втім така модель у процесі створення не є заздалегідь відомою. Наперед можливо сформували лише загальні принципи або певні правила формування та поведінки окремих елементів (повідомлень), тобто загальну модель. У синергетиці доводиться, що ніякими зовнішніми

впливами неможливо «нав'язати» системі бажану кому-небудь поведінку - можна лише вибрати найбільш підходящу траєкторію з потенційно закладених. При плануванні та моделюванні Ю одним з основних завдань є пошук точок біфуркації соціальних процесів і формування флуктуацій, що приводять до вибору необхідної траєкторії еволюції [1]. Іншими словами підбір реальних (або формування фейкових) подій, використання лексичного апарату для опису подій (фактажний/оціночний), та надання їм необхідних характеристик (емоційність, маніпулятивність, позитив/негатив) складає суть технологічного процесу реалізації «PR-впливу».

Дослідники інформаційного простору, утвореного електронними ЗМІ звертають увагу на його значну неоднорідність (виникнення певної структури), яка з'являється завдяки групуванню інформаційних елементів (документів) пов'язаних між собою за тематикою [3]. Ми називаємо такі структури «тематичними трендами». З урахуванням достатньої складності системи та процесів інформаційного простору [1], для вирішення проблем виявлення аномалій, трендів та прогнозування динаміки ІП однією з найбільш прогресивних імітаційних моделей вважається мультиагентна модель ІП.

Сутність її полягає у співставленні поняття агент - повідомлення з корпусу ТПП. В початковий момент часу ($t=0$) тематику S_i представляє єдиний агент a_1 з енергією E_0 . Подальша еволюція агента складається із дій, що символічно позначаються певними індексами та відповідають правилам зміни енергії агента (повідомлення): *like* - збільшення енергії a_i на 1; *dislike* - «помирати», зменшення енергії на 1; *repost* - «самозародження» нового агента (копії a_i), та збільшення енергії на 2; *link* - «самозародження», «посилання» на агента a_i , та збільшення енергії на 1. Кожна з перелічених подій може відбутися з певною ймовірністю

$$P_k^n = P_{k0} \phi(E_n), \quad (1)$$

де $k=\{like, dislike, repost, link\}$, а

$$\phi(E_n) = \begin{cases} \frac{E}{2(2E_0 - E)}, & E \in \left[0; \frac{4E_0}{3}\right] \\ 1, & E > \frac{4E_0}{3} \end{cases}. \quad (2)$$

Зазначені ймовірностей P_{like0} , $P_{dislike0}$, $P_{repost0}$, P_{link0} та E_0 є параметрами системи і за певних значень дозволяють змодельовати числові ряди подібні до реальних, які б відповідали ТПП [3].

Представлена модель ІП за результатами аналізу самоподібності часових рядів (R/S-аналіз [1]) показує високий ступінь автокореляції ряду, що характерно і для реальних ТПП. З іншого боку, на нашу думку, створена таким чином мультиагентна модель динаміки проведення інформаційної атаки могла б використовуватись у якості шаблону для тестування ТПП на відповідність реальному трендові представленому у вигляді числового ряду, аналогічно наведеному у [2].

Зазвичай подальший аналіз інформаційних потоків зводиться до задач виявлення та візуалізації

трендів, виявлення гармонійних складових, трендів, локальних особливостей часових рядів, фільтрації шуму, які на сьогодні досить ґрунтовно теоретично опрацьовані та практично вирішуються методами фрактального, вейвлет, Фур'є-аналізу та інших [1]. Втім, відкритим залишається низка питань прикладного технологічного характеру:

- як позначати виявлену тематику реального ІП, яка породжує тренд;
- як саме формується тематичний переріз інформаційного потоку (ТІП);
- що є основою побудови ТІП та тренду.
- визначення якісної оцінки тренду;
- критерії тематичного визначення спрямованого інформаційного впливу.

Наша робота покликана конкретизувати ці питання та запропонувати метод виявлення «PR-впливу». Отже зупинимось більш формально на визначенні деяких термінів, на яких технологічно ґрунтується наш метод.

У попередній роботі ми визначили «теми дня» як структуроване поєднання сюжетних ланцюжків інформаційних повідомлень, отримане протягом певного періоду (добі). На якісному рівні «тема дня» позначає тематично пов'язану групу повідомлень, які подають за змістом: - фактичний опис події; - розгортку події стосовно різних контекстів; - розвиток події; - оцінку події; - аналіз події та її наслідків [7].

Будемо називати *паспортом «теми дня»* - список ключових слів корпусу повідомлень «теми дня». Паспорт «теми дня» є базою (утворює) лінгвістичного фільтру «теми дня», застосування якого до бази повідомлень протягом певного періоду дає тематичний переріз інформаційного потоку.

Вектор ІП - це часовий ряд (де кожному значенню вектора відповідає кількість повідомлень за проміжок часу Δt - зазвичай доба), загального ІП.

Тематичний вектор - це часовий ряд (де кожному значенню вектора відповідає кількість повідомлень за проміжок часу Δt - зазвичай доба), отриманий шляхом застосування фільтру «теми дня».

Тематичний тренд - сукупність пов'язаних, через повідомлення тематичного вектору, «тем дня».

Для виявлення «PR-впливу» пропонується метод, що полягає у наступному.

1. У якості вхідної інформації маємо репрезентативно достатню базу інформаційних повідомлень відібраних з новинних Інтернет ресурсів ЗМІ, яка постійно оновлюється. Тобто маємо злічену множину повідомлень $M = \{m_1, m_2, \dots, m_i, \dots\}$ інформаційного потоку $I = (M, \gamma)$, де відношення квазіпорядку $\gamma \in M \times M$, впорядковано за $T = \{t_1, t_2, t_3, \dots\}$. Розбиття T на проміжки Δt (доба) дозволяє побудувати вектор ІП $M_{\Delta t} = \{|M_{\Delta t1}|, |M_{\Delta t2}|, \dots, |M_{\Delta ti}|\}$, де $|M_{\Delta ti}|$ - потужність добової множини повідомлень $M_{\Delta ti}$, а $i = \overline{1, n}$ - кількість днів обраного для дослідження часового проміжку.

2. 3 метою отримання характеристичного простору лінгвістичних ознак індексуємо кожне повідомлення бази за допомогою алгоритму лінгвостатистичного оцінювання за наступними ознаками: наявність лексики що характеризує констатацію фактів, подій чи стосунків - «подієвість» ($P_{EV}(m_i)$) - показник, що повідомлення m_i є неподієвим; наявність лексики що характеризує спрямованість на довільний опис та оцінювання відношення до події або стосунків - «не подієвість» ($P_{NEV}(m_i)$) - показник, що повідомлення m_i є неподієвим, при чому потужність множини повідомлень ПП M :

$$|M| = \sum_{i=1}^n P_{EV}(m_i) + \sum_{i=1}^n P_{NEV}(m_i), \text{ де } n - \text{ загальна кількість}$$

повідомлень бази на поточний час; наявність «позитивної»/«негативної» лексики ($P_{PZ}(m_i)/P_{NPZ}(m_i)$) - показник, що повідомлення m_i є позитивним/негативним, при чому

$$|M| \neq \sum_{i=1}^n P_{PZ}(m_i) + \sum_{i=1}^n P_{NPZ}(m_i); \text{ наявність «емоційної»}$$

лексики ($P_{EM}(m_i)$) - показник, що повідомлення m_i є емоційним; наявність «маніпулятивність» лексики ($P_{MA}(m_i)$) - показник, що повідомлення m_i є маніпулятивним). При цьому:

$$P_r(m_i) = \begin{cases} 1, & p_r(m_i) \geq 0,85 \\ 0, & p_r(m_i) < 0,85 \end{cases} \quad (3)$$

де $r=\{EV, NEV, PZ, NPZ, EM, MA\}$, а $p_r(m_i)$ - ймовірність що m_i -те повідомлення за наявною r -тою лексикою з достовірністю 85% (результат емпіричних досліджень) не відрізняється від r -тої навчальної вибірки повідомлень.

3. Для кожної доби (множина повідомлень $M_{\Delta i}$ за проміжок часу Δt_i) за допомогою алгоритму побудови «тем дня» [7] формується добова множина «тем дня» $S_{\Delta i} = \{s_{\Delta i,1}, s_{\Delta i,2}, \dots, s_{\Delta i,j}, \dots, s_{\Delta i,h}\}$, ранжована за зменшенням потужності корпусу згрупованих повідомлень M_{TDj} (де $j = \overline{1,h}$, а h - кількість «тем дня»). Кожна «тема дня» $s_{\Delta i,j}$ отримує назву, відповідну назві центроїдного повідомлення групи M_{TDj} [7].

4. У процесі поточного моніторингу «тем дня» (множина $S_{\Delta i}$) виявляємо такі, що мають ознаки: новизна, гучність (привертає увагу), важливість (геополітична, екологічна, економічна, соціальна). Такі теми зазвичай посідають перші позиції загального списку [7].

5. Для відібраної «теми дня» $s_{\Delta i,j}$ формуємо паспорт, тобто список ключових слів W_{TDj} усього корпусу M_{TDj} , який є основою семантичного фільтру $L_{TDj}=f(W_{TDj})$.

6. Застосування L_{TDj} до інформаційного потоку M дозволяє відокремити ТПП асоційований з «темою дня» $s_{\Delta i,j}$, що визначена у п. 4, тобто отримати злічену множину повідомлень $M_{TTP} = \{m_1, m_2, \dots, m_i, \dots\}$ потоку $TIP = (M_{TTP}, \gamma)$, де відношення квазіпорядку $\gamma \in M_{TTP} \times M_{TTP}$, впорядковано за $T = \{t_1, t_2, t_3, \dots\}$. Розбиття T на проміжки Δt (доба) дозволяє побудувати вектор ТПП

$M_{TTP\Delta t} = \{|M_{TTP\Delta t,1}|, |M_{TTP\Delta t,2}|, \dots, |M_{TTP\Delta t,n}|\}$, де $|M_{TTP\Delta t,i}|$ - потужність добової множини повідомлень $M_{TTP\Delta t,i}$, а $i = \overline{1,n}$ - кількість днів обраного для дослідження часового проміжку.

7. Проводимо дослідження часового ряду $M_{TTP\Delta t}$ за тривалий період для встановлення часових характеристик теми (початкова фаза), тривалості її обговорення, динаміки та тонального окрасу. За результатами дослідження корегуємо проміжок часу n для обмеження $M_{TTP} = \{m_1, m_2, \dots, m_i, \dots, m_k\}$ (де k - кількість повідомлень корпусу ТПП за n днів) та побудови вектору $M_{TTP\Delta t}$.

8. Для відібраної у п. 7. множини повідомлень тематичного вектору $M_{TTP\Delta t}$ розраховуємо показник наявності «PR-впливу» ТПП. Відповідно до положень описаних у [6] (модифікація виразу (6)), будемо розраховувати математичне очікування щодобової оцінки показника наявності «PR-впливу» ТПП:

$$PRI(M_{TTP\Delta t}) = M[PRI(M_{TTP\Delta t,i})], \quad (4)$$

де показник наявності «PR-впливу» ТПП для i -тої доби дорівнює $1/3$ відношення кількості повідомлень з ознаками «не подієвість», «емоційність», «маніпулятивність» до загальної кількості повідомлень тематичного вектора МТІР:

$$PRI(M_{TTP\Delta t,i}) = \frac{(|M_{TTP\Delta t,i}| - P_{EV}(M_{TTP\Delta t,i})) + P_{EM}(M_{TTP\Delta t,i}) + P_{MA}(M_{TTP\Delta t,i})}{3 \times |M_{TTP\Delta t,i}|}, \quad (5)$$

$$P_r(M_{TTP\Delta t,i}) = \frac{|M_{TTP\Delta t,i}|}{\sum_{\ell=1}^{|M_{TTP\Delta t,i}|} P_r(m_\ell)}, \quad (6)$$

де $r=\{EV, EM, MA\}$ - індекс позначення характеристики з лінгвістичними ознаками відповідно {«подієвість», «емоційність», «маніпулятивність»}; $|M_{TTP\Delta t,i}|$ - потужність множини повідомлень ТПП за i -ту добу. Виходячи з побудови, значення $PRI(M_{TTP\Delta t})$ завжди є позитивним та знаходиться в межах від 0 до 1.

9. У якості критерію прийняття рішення щодо наявності «PR-впливу» у ТПП асоційованого з обраною у п. 3 «темою дня» встановлюється граничне значення α . Тобто, якщо

$$PRI(M_{TTP\Delta t}) > \alpha, \quad (7)$$

то такий ТПП, представлений тематичним вектором $M_{TTP\Delta t}$ містить ознаки «PR-впливу».

При цьому, п. 3-9 фактично являють собою спосіб розподілу потоку інформаційних повідомлень які представляють колекції відповідних документів з лексикою властивою «PR-впливу» та відмінних від таких.

Таким чином, запропонований метод дозволяє розподілити реальний ІП на складові ТПП, та дослідити відповідні числові ряди за ознакою наявності лексики властивої «PR-впливу» та дослідити динаміку тренду.

З метою оцінювання технологічних можливостей розробленої нами Системи аналізу інформаційного простору (САІП) [7] для практичного вирішення завдань виявлення «PR-впливу» та оцінювання числових значень виразів запропонованого методу, нами були проведені

попередні емпіричні дослідження, результатами яких стало наступне.

По-перше, стосовно граничного значення α , виразу (7) зауважимо, що емпіричні дослідження (результати наведені у Таблиці 1) показали відмінність $PRI(M_{TIP\Delta})$ на 21% його значення для тем, яким природно не властива лексика «PR-впливу» (наприклад для теми «Галузь інформаційних технологій» $PRI(M_{TIP\Delta})=0,50$ - рис. 1) від тем з ознаками новизни та гучності (наприклад для теми «Правий сектор - Мукачево» у період з 11.07.2015 р.

по 24.07.2015 р. $PRI(M_{TIP\Delta})=0,71$ - Рис. 2). Втім, як видно з виразу (5) п. 6 методу, на даний час, розрахунок показника наявності «PR-впливу» не враховує оцінку наявності «позитивної»/ «негативної» лексики, хоча розвідні експериментальні досліді дають позитивні результати відповіді на питання про можливість розподілу потоку інформаційних повідомлень яким властива лексика «PR-впливу» та відмінних від них. Тобто, для попередніх досліджень α можливо встановити на рівні 0,6.

Таблиця 1

Значення обчислення показника наявності «PR-впливу» за виразом (4,5,6) для тем: «Галузь інформаційних технологій» та «Правий сектор - Мукачево» у період з 11.07.2015 р. по 24.07.2015 р. та відповідні значення математичного очікування $PRI(M_{TIP\Delta})$ для них

Тема	11.07.2015	12.07.2015	13.07.2015	14.07.2015	15.07.2015	16.07.2015	17.07.2015	18.07.2015	19.07.2015	20.07.2015	21.07.2015	22.07.2015	23.07.2015	24.07.2015	25.07.2015	$PRI(M_{TIP\Delta})$
Галузь ІТ	0,50	0,40	0,61	0,51	0,45	0,55	0,50	0,37	0,61	0,55	0,51	0,51	0,56	0,44	0,47	0,50
Пр.Сект. Мукачево	0,66	0,74	0,72	0,73	0,73	0,68	0,66	0,71	0,73	0,69	0,78	0,59	0,81	0,66	0,78	0,71

Діаграма значень показника Наявності "PR-впливу" для теми "Галузь інформаційних технологій"

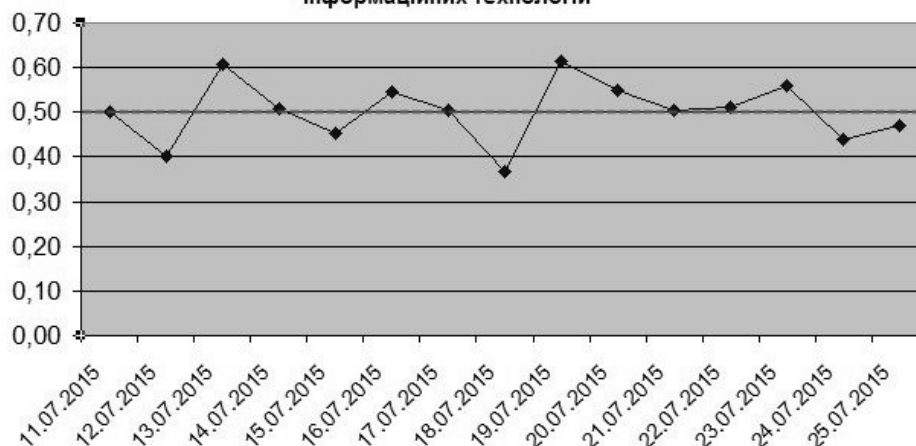


Рис. 1. Діаграма щодобових значень показника наявності «PR-впливу» розрахованих для теми «Галузь інформаційних технологій» у період з 11.07.2015 р. по 24.07.2015 р. (пунктирна лінія відповідає середньому значенню $PRI(M_{TIP\Delta})=0,50$)

Діаграма значень показника Наявності "PR-впливу" для теми "Правий сектор - Мукачево"

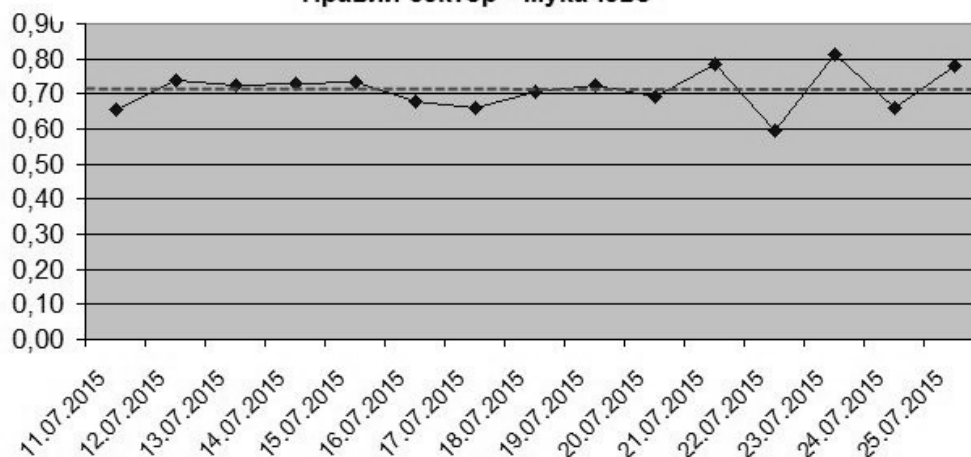


Рис. 2. Діаграма щодобових значень показника наявності «PR-впливу» розрахованих для теми «Правий сектор - Мукачево» у період з 11.07.2015 р. по 24.07.2015 р. (пунктирна лінія відповідає середньому значенню $PRI(M_{TIP\Delta})=0,71$)

По-друге, як особливість запропонованого показника наявності «PR-впливу» (4) відзначимо його нечутливість до тижневого коливання кількості повідомлень (на Рис.1,2: 11,18,25 липня – субота, 12,19 липня – неділя). Це виключає необхідність

застосування апроксимації для часового ряду при проведенні аналізу на тривалих проміжках часу.

По-третє, при дослідженні теми «Правий сектор - Мукачево» з'ясувалось що вона перетинається на рівні повідомлень з іншими темами (рис. 3.).

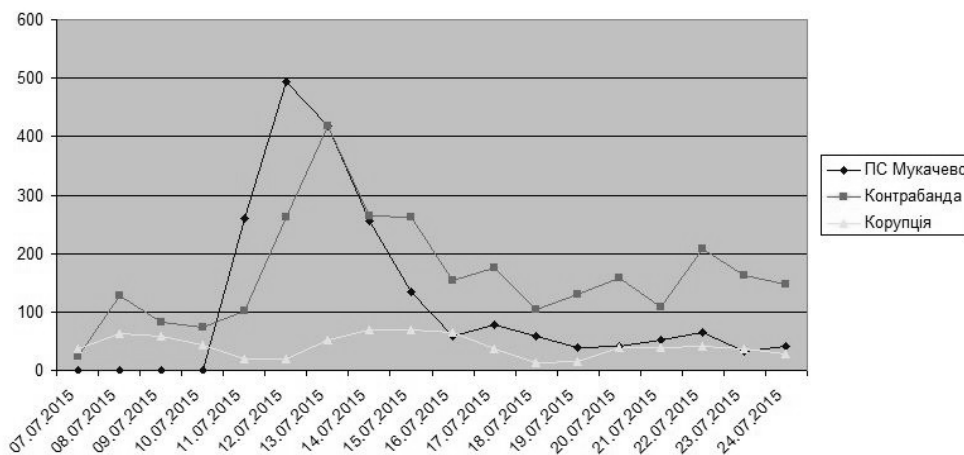


Рис. 3. Діаграма щодобової кількості повідомлень за темами «Правий сектор - Мукачево», «Контрабанда», «Корупція» у період з 11.07.2015 р. по 24.07.2015 р. які перетинаються за рахунок однакових повідомлень

Висновки

Таким чином, наше дослідження підтвердило гіпотезу яка полягає у тому, що дослідження трендів ТП у характеристичному просторі лінгвістичних ознак дозволить отримати метод виявлення «PR-впливу».

Наведені дані попередніх емпіричних досліджень спонукають до наступних міркувань. По-перше, гучні події збагачують потужність ТП із властивостями «PR-впливу» (для теми «Контрабанда» $PRI(M_{TPA})=0,62$). По-друге, паралельні у часі тематики за рахунок перетину через певні повідомлення утворюють тематичний тренд. По-третє, на нашу думку для дослідження тематичного тренду можливо застосовувати методологічний апарат багатоагентного моделювання, якщо спів ставляти поняттю агент визначену нами «тему дня». Останнє, на сам кінець, напевне дозволить перейти до використання математичних моделей для прогнозування можливих сценаріїв динаміки соціальних процесів на якісному рівні.

Друга теза потребує окремого зауваження. Як ми вже зазначали, деякі ТП в динаміці формування реального інформаційного простору набувають мережевих ознак. Завдяки нашому методу розподілу загального ІП на складові отримуємо певний механізм відтворення (реконструкції) його мережевої структури. А саме, лінгвістичний фільтр, побудований на базі паспорту «теми дня» - W_{TDj} . Але, кожне повідомлення має особистий перелік ключових слів W_{mi} який у загальному випадку може не бути підмножиною W_{TDj} . За рахунок цього деякі повідомлення при поточному розгляді мають увійти до інших «тем дня» та відповідних до них ТП, які є паралельні у часі, але далекі за тематикою. За рахунок таких перетинів різних ТП через окремі

повідомлення, отримуємо певну мережу. Де вузлами є ТП, а ребрами повідомлення що їх пов'язують. Тоді можемо визначити наступні мережеві ознаки: потужність зв'язку двох ТП (*ступінь близькості ТП*) – це кількість зв'язуючих повідомлень; *ступінь свободи ТП* – це кількість пов'язаних ТП (або кількість ребер що виходять з вузла ТП); *вага ТП* – це потужність множини корпусу повідомлень ТП. Таким чином, застосування нашого методу розподілу загального ІП на складові ТП відкриває додатково методологічний апарат теорії складних мереж для реконструкції топології реального інформаційного простору в «суб'єкт-вплив-об'єктному» вимірі та проведення якісного аналізу ІП.

Слід зазначити, що наш підхід по суті є індикативним і досліджує потенційну можливість інформаційного впливу на особу та суспільство, яка властива ІП у певний проміжок часу, через встановлення лінгво-статистичних оцінок складових елементів (повідомлень) корпусу відповідних структурних елементів інформаційного простору. Тобто ми не досліджуємо реакцію суспільства на ті чи інші інформаційні впливи, а обмежуємось тестуванням ІП на їхню наявність.

Втім, запропонований нами метод виявлення факту наявності контенту, спрямованого на впровадження «PR-впливу» фактично є інструментом інформаційно-психологічної боротьби.

Перспективи дослідження

Тож, в подальших експериментально-прикладних дослідженнях плануємо зосередитись на:

- детальному обґрунтуванні показника наявності «PR-впливу» (4,5);
- моделі визначення критерію α прийняття рішення щодо наявності «PR-впливу» у ТП вираз (7);

- посиленні показника наявності «PR-впливу» у ТПІ за рахунок оцінки наявності «позитивної»/«негативної» лексики;
- опрацюванню ефективності запропонованого методу;
- застосуванню інструментарію та методологічного апарату теорії складних мереж для проведення якісного аналізу ПП.

Література

[1] Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. – К.: Інтертехнологія, 2009. – 164 с.

[2] Ланде Д.В. Тренди відображення інформаційних операцій в інформаційному просторі / Д.В. Ланде // Інформація і право, 2013. – № 1 (7). – С. 82-88.

[3] Грайворонська А.М. Дослідження інформаційних потоків, як динамічних мультиагентних систем. / А.М. Грайворонська, Д.В. Ланде // Системний аналіз і інформаційні технології: матер. 17-й Міжнарод. науч.-техн. конф. SAIT 2015, Київ, 22-25 июня 2015 г. / УНК «ИПСА» НТУУ

«КПИ». – К.: УНК «ИПСА» НТУУ «КПИ», 2015. – С. 62-63.

[4] Додонов О.Г. Мультиагентная модель поведения тематических информационных потоков / Додонов О.Г., Ланде Д.В. // Материалы 6-ой Всероссийской мультиконференции – 2013. – Т. 4, № 6. – С. 102-107.

[5] Шумка А.В. Інформаційно-мережева війна – нова форма міждержавного протидержавства початку ХХІ ст. [Електронний ресурс] / А.В. Шумка, П.П. Черник // Військово-науковий вісник. – 2013. – Вип. 19. – С. 243-255. – Режим доступу: http://nbuv.gov.ua/j-pdf/vnv_2013_19_20.pdf.

[6] Рябий М. Модель виявлення PR-впливу через публікації в інтернет ЗМІ. [Електронний ресурс] / М. Рябий, О. Хатян, С. Багацький // Безпека інформації. – 2015. – Т.21, № 2. – С.131-139. – Режим доступу: <http://jrn1.nau.edu.ua/index.php/Infosecurity/article/view/8731/10703>.

[7] Хатян О.А. Алгоритм побудови «тем дня» Інтернет ЗМІ / О.А. Хатян // Інформаційна безпека людини, суспільства, держави. – № 2 (18), 2015 – С. 110-123.

УДК 004.912 (045)

Рябий М.А., Хатян О.А., Багацький С.П. Метод виявлення PR-впливу через Інтернет-СМІ

Анотація. Рассматривается информационное пространство, образованное информационным потоком сообщений Интернет СМІ. В основе работы лежит гипотеза о различии трендов тематических информационных потоков типа беспристрастного информирования и «PR-влияние» в характеристическом пространстве лингвистических признаков. Приведен метод распределения общего информационного потока на составляющие тематические потоки, и способ отделения тематических потоков с лингвистическими признаками «PR-влияния». Приведенные данные предварительных эмпирических исследований позволяют оценить граничные условия для принятия решения относительно показателя наличия «PR-влияния» тематического потока. Также анализ приведенных результатов показывает наличие связи отдельных тематических трендов, полученных в результате применения нашего метода. Этот факт открывает возможности теории сложных сетей для реконструкции топологии реального информационного пространства и проведения анализа информационных потоков.

Ключевые слова: информационное пространство, информационный поток, информационное воздействие, информационная операция, метод выявления «PR-влияния», тематический тренд, тематический вектор, сложные сети.

Ryabyy M., Hatyany O., Bagatskiy S. Model of detection PR-impact by means of Internet mass media

Abstract. The information space formed by the information thread of publications in the Internet of mass media is examined. The basis of work is hypothesis of about distinction of trends of thematic informative threads of type of the impartial informing and «PR-impact» in characteristics space of linguistic attributes. Showed a method for the distribution of the total flow of information on the components of the thematic threads, and a method of defining thematic streams with linguistic attributes «PR-influence». Cited data of preliminary empirical researches allow to estimate border terms for a decision-making in relation to an index presence of «PR-impact» of thematic thread. An analysis the results shows the presence of connection of the separate thematic trends got as a result of application of our method. This fact opens possibilities of theory of difficult networks for the reconstruction of topology of the real informative space and realization of analysis of informative threads.

Key words: information space, information flow, informational impact, information operation, detection method PR-impact, thematic trend, thematic vector, complex networks.

Отримано 23 вересня 2015 року, затверджено редколегією 9 жовтня 2015 року