

МОДЕЛЬ ВИЯВЛЕННЯ PR-ВПЛИВУ ЧЕРЕЗ ПУБЛІКАЦІЇ В ІНТЕРНЕТ ЗМІ

Мирослав Рябий¹, Олег Хатян¹, Сергій Багацький²

¹Європейський університет, Україна

²ТОВ «Миронівський хлібопродукт», 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: завідувач кафедри організації комплексного захисту інформації з 2013 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: Київський національний технічний університет (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. Властивість інформаційної загрози полягає у здатності до запуску механізмів самознищення складної інформаційної системи, якою є окрема особа, громадські об'єднання та суспільство взагалі. Провідником зазначеного типу загроз є акт інформаційного впливу через інформаційні джерела ЗМІ, який ми називаємо спрямованим або «PR-впливом». Ефективна модель виявлення інформаційних впливів є основою механізму виявлення та протидії інформаційним загрозам. У роботі свідомо спрямований вплив характеризується як засіб зміни якісної оцінки або відношення до певних фактів або об'єктів цільовою аудиторією. Проведений аналіз завдань, методів та механізмів роботи PR-служб дозволив сформулювати відповідне поняття «PR-вплив» як засіб реалізації певних інтенцій через публікації в Інтернет ЗМІ, а крім того визначити його особливості, критерії та оціночні характеристики. Сформульована гіпотеза про відмінність процесу «PR-впливу» від «неупередженого» інформування. Запропонована формальна модель виявлення цих спрямованих впливів на множині інформаційних повідомлень отриманих від Інтернет ЗМІ.

Ключові слова: спрямований вплив, критерії оцінки впливу, Інтернет ЗМІ, алгоритм класифікації колекції документів, сюжети новин, теми дня, об'єктно-подієві патерни, модель виявлення PR-впливу.

Вступ

В умовах сьогодення, коли якість протистояння на міждержавному рівні та градус суспільного настрою визначається насамперед станом інформаційної сфери держави, важливість

завчасного виявлення інформаційної загрози навряд чи можна переоцінити. Російський дослідник інформаційних війн Расторгуєв С.П. визначає принципову особливість інформаційної загрози як певного механізму, що потенційно може привести в дію процес, спрямований на самознищення складної

інформаційної системи, здатної до самонавчання [1], якою є окрема особа, громадські об'єднання та суспільство взагалі. Всебічний аналіз наукових праць, присвячених зазначеній проблематиці дозволяє зробити висновок, що так чи інакше, провідником інформаційної загрози є акт впливу, впроваджений як правило через інформаційні джерела, серед яких Інтернет ЗМІ займають не останню роль. Тому проблема побудови формальної моделі виявлення інформаційних впливів як основи виявлення інформаційних загроз є *актуальною*.

Аналіз останніх досліджень і публікацій

Загально доступних наукових робіт, присвячених моделюванню інформаційних впливів в аналітичному вигляді, за певних обставин, не дуже багато. Вітчизняний науковець Залкин С.В. опрацьовує модель інформаційно-психологічних впливів на бойову готовність підрозділів Повітряних Сил Збройних Сил України [3]. Однак ці дослідження стосуються лише окремих аспектів прояву емоційної реакції інформаційного впливу на свідомість осіб, оціненої часом виконання окремої операції. Безпосередньо проблемі моделювання інформаційних впливів через ЗМІ у контексті ведення інформаційних війн присвячені роботи вже згаданого нами російського дослідника Расторгуєва С.П. [1,2]. Вчений визначає поняття загрози як множини взаємозалежних та взаємообумовлених подій, реалізація яких здатна заподіяти шкоду, звідки випливає, що *виявлення загрози передбачає виявлення відповідної послідовності подій, які не завжди відомі заздалегідь* [1]. На відміну від цього, наш підхід базується на припущенні, що інформаційна загроза безпосередньо складається і є похідною актів навмисного (штучно спрямованого) інформаційного впливу, що може бути прогнозованою від моменту виявлення. У попередніх роботах ми розглянули формальну модель та загальні принципи процесу класифікації інформаційних повідомлень на базі оцінювання лінгвостатистичних показників [4], дослідили ознаки та технологічні складові суспільної комунікації з точки зору аналізу потоків інформаційних повідомлень [5]. Отримані результати дозволили створити прикладну інформаційно-комунікативну модель суспільної комунікації та апробувати її в процесі дослідження соціально-економічних явищ.

Метою цієї роботи є побудова формальної моделі виявлення штучних інформаційних впливів, шляхом вирішення *завдання* аналізу особливостей свідомо спрямованого впливу («PR-впливу») через публікації в ЗМІ певних інформаційних матеріалів на загальному фоні інформаційного контенту.

Виклад основного матеріалу

Проблема виявлення штучного впливу через ЗМІ, на наш погляд, тотожна проблемі розуміння або встановлення смислу окремо взятого акту комунікації (в окремому випадку інформаційного повідомлення) в рамках певного контексту (тобто проблемі розпізнавання контексту повідомлення / комунікації). Відмінними особливостями реалізації

впливу є - хронологічна протяжність (у часовій шкалі - $T = \{t_1, t_2, \dots, t_n\}$), змістовна роздрібненість спрямованого на вплив корпусу повідомлень (як правило, обов'язкова в силу вимоги латентності), тобто загальний смисл впливу диференційований та рознесений у множині інформаційних повідомлень. Смисл окремого повідомлення не обов'язково співпадає із метою впливу. При цьому низька надмірність окремого повідомлення (що обумовлено форматом подання контенту в ЗМІ), компенсується значною надмірністю загального корпусу тематично об'єднаного потоку (у якому синтетично для посилення ефекту впливу може поєднуватися комплексом ключових тем). Тому зупинимось предметно на понятті «смисл».

В одній з робіт, присвячених, серед іншого, теорії комунікації, Бейтсон Г. дає таку характеристику поняття «смисл» - «смисл» можна розглядати як приблизний синонім слів «патерн», «надмірність», «інформація» та як обмеження усередині наступної парадигми.

Слід вважати, що деякий конгломерат подій або об'єктів (наприклад, послідовність фонем, картина, «жаба» або культура) містить «надмірність» («патерн»), якщо цей конгломерат деяким способом може бути розділений «межею» таким чином, що спостерігач, який сприймає те, що є з одного боку цієї межі, може здогадатися (з успіхом, що перевищує випадковий), що знаходиться з іншого боку межі. Ми можемо казати, що те, що знаходиться з одного боку межі, містить інформацію (смисл) того, що є з іншого боку. На інженерній мові можна сказати, що конгломерат містить «надмірність». З точки зору спостерігача-кібернетика, інформація, доступна з одного боку межі, обмежуватиме помилкове відгадування (тобто знижувати його вірогідність) [6].

Так Г. Бейтсон визначає поняття «смисл» в контексті вивчення творів мистецтв (архітектури, живопису, скульптури, музичних творів, театральних постановок, літератури тощо). В контексті ж проблематики виявлення штучного впливу «межею», на нашу думку, можна вважати певний момент t_i , що поділяє відповідну впливу послідовність інформаційних повідомлень на ті, що вже опубліковані, та ті що будуть опубліковані в майбутньому. За таких умов «надмірності» контенту на момент t_i може вистачити для визначення «смислу» впливу (тобто відповіді на питання, а чи є він навмисним).

З метою забезпечення латентності, на думку автора, при реалізації акту впливу, надмірність патерну у первинних фазах знижується, але нестримно зростає у наступних. Суб'єктом впливу (комунікативного процесу) є соціальна група або цільова аудиторія, яка відносно об'єкту впливу проявляє деякі однорідні характеристики або наявні соціальні прояви, що дозволяє умовно називати таку аудиторію «соціальною масою». Вплив в такому випадку має масовий характер. Ми вважаємо, що як правило, *кінцевою метою впливу* є зміна умовної або фактичної соціальної (громадської, яка іноді штучно подається як персональна) цінності тієї або іншої

характеристики об'єкту, відносно якого здійснюється вплив. Прикладами можуть бути: очікування «поліпшень» при здійсненні вибору – де об'єктом є кандидат для вибору; відношення до події або очікуваної події (хтось/щось постраждав(о), перемістився(лось) або збагатився(лось) і так далі) – фактична інформація про об'єкт або символічне представлення об'єкту в інформаційному просторі; реалізація комплексу потреб – об'єктом є товари і послуги. У такому трактуванні поняття «вплив», у соціальному плані, еквівалентно поняттю *латентної загрози* [7].

Отже, основою наших міркувань щодо побудови формальної моделі виявлення інформаційних впливів є розуміння про відмінність процесу навмисного впливу на процес ухвалення рішення (тобто такого впливу, що має на меті досягнення певного стану або контексту суб'єктом ухвалення рішення, при якому рішення вибирається із заздалегідь визначеної обмеженої підмножини з безлічі усіх можливих рішень) від «неупередженого» інформування (тобто *своєчасного* швидкого, досить *повного* і *достовірного* надання фактів, які подаються з усебічним, об'єктивним і неупередженим аналізом, при якому *не відбувається зміщення акцентів контексту* ухвалення рішення). Самі цілі впливу зумовлюються постановкою завдання деяким «замовником» перед спеціалізованою організацією або підрозділом організації (умовно PR-служба). Такі завдання технічно реалізуються через публікації в ЗМІ (часто підконтрольних «замовнику») певних інформаційних матеріалів, сукупність яких ми будемо називати «*PR-впливом*». ЗМІ, натомість, мають можливість міняти пропорції реальної події, трансформують їх в більш потужні або слабкіші [8]. Термін «*PR-вплив*» визначається нами з метою фіксації фокусу уваги на особливих характеристиках певних елементів «штучної» (або синтетичної) організації неоднорідностей загального інформаційного потоку. Серед таких неоднорідностей, зокрема, при ретроспективному аналізі інформаційного простору нами неодноразово виявлялися інформаційні сплески (значне збільшення інформаційних повідомлень у контексті певної теми), що у сукупності мали характер «інформаційної атаки» (так звані «газова», «сирна», «м'ясна» війни; актуалізація мовних та/або етнічних питань; «інформаційні атаки» на певні фінансові установи). Проте, термін «інформаційна атака» дуже специфічний і не здатний охопити широкий клас інтенсивних явищ цілеспрямованого маніпулювання узагальненим контентом, який формує досить «тонкі» патерни об'єктно-подієвого простору. Тому охарактеризовані вище загальні явища надалі ми називатимемо «*PR-вплив*».

З урахуванням того, що об'єктом наших досліджень є інформаційний простір, породжений узагальненим контентом інформаційних повідомлень електронних ресурсів ЗМІ (ці джерела за об'єктивних обставин випереджують публікацію контенту, що транслюється у телевізійних та радіо новинах), а метою дослідження – побудова формальної моделі, спробуємо відповісти на

питання «Де шукати?» (які предметні області мають найбільш характерні риси неоднорідностей і значимість результатів), «Що саме шукати?» (які оцінки неоднорідностей доцільно вивчати) і «Як шукати?» (яка формальна гіпотеза стане основою для моделі свідомого інформаційного впливу).

«*Де шукати?*»

Генеральною метою «паблік рілейшнз» (PR) є «формування ситуації успіху в суспільстві» [9] (розуміємо, що поняття «формування» має прямий семантичний зв'язок з поняттям «вплив»). Перелік областей інтересу, де переважним чином можуть бути поставлені завдання і цілі для PR-служб, найкращим чином формулюють самі PR-практики [10]. Серед загального переліку пріоритетних напрямів застосування PR (предметні точки) виділимо:

1. Система цінностей.

2. Особисті стосунки (питання етики і моралі, гендерні питання, визначення професіоналізму, проблематика вибору).

3. Громадські стосунки.

4. Урядові стосунки (структурні питання).

5. Міжнародні і міжнаціональні відносини (формування іміджу).

6. Стосунки в промисловості і фінансах (корекція попиту-пропозиції, питання якості товарів і послуг, корекція бренду, формування потреб).

7. Діяльність в засобах масової інформації.

Серед різних видів PR в літературі виділяють термін *чорний PR* для означення використання так званих «чорних технологій» (обман, фальсифікація) з метою очорнення, знищення конкуруючої партії, групи тощо, поширення образливих або економічно небезпечних заяв та інше [9]. У сучасній дійсності чорний PR особливо яскраво проявляється в політичних комунікаціях, зокрема в партійній та опозиційній пресі. Чорний PR використовує *маніпулятивні форми дії*. Серед них [9]:

1. *Заміна імен, або наклеювання ярликів*

(наприклад, «обличчя кавказької національності», «джамшуті») – становлять небезпеку завдяки швидкому входженню через ЗМІ в суспільний ужиток, стають звичними, повсякденними словами, заміщують суміжні, але менш агресивні поняття. Маніпулятивні терміни, або «ярлики», є одним з видів психологічного програмування масової свідомості, та є одним із способів насадження масових стереотипів, протистояти якому дуже складно.

2. *Повтор інформації* – сприймається цільовою аудиторією та фіксується у масовій свідомості після чого стає звичайною «істиною». Варіантом техніки «повтору» є вживання гасел і ключових слів типу «Рівні права для усіх», «Народ і партія єдині» і тому подібне. Такі, часто безглузді, фрази відіграють велику роль в політиці та рекламі.

3. *Твердження*. Особливість цього прийому полягає в перевазі супереччі *голих тверджень* на підтримку певної тези, чим обмежують плюралізм думок та представляють тільки одну, найбільш вигідну сторону для підтримки тези.

4. *Постановка риторичних питань*, або *розпускання чуток*. ЗМІ ставлять перед аудиторією питання та залишають його без відповіді, але наділяють певним контекстом, що примушує глядача/слухача/читача мимоволі мислити та розвивати «підкинуту ідею» у потрібному для маніпулятора напрямі. Чутки вносять відчутний вклад у формування громадської думки. На цьому рівні, наприклад, запускається негативна інформація про опонента під час передвиборної кампанії. Результатом є укорінені в масовій свідомості негативна думка та, як наслідок, зниження рейтингу політика.

5. *Осміяння* – досить часто використовується в ЗМІ з метою інформаційно-психологічного впливу на масову свідомість. При осміянні окремих висловлювань, елементів поведінки людини, до нього ініціюється жартівливе та несерйозне відношення, яке надалі поширюється на інші його висловлювання та погляди, шкодить іміджу особи. Таким чином, якщо психологічними бар'єрами блокується пряма критика і негативна дія, гумористична форма знижує захисні реакції особи. Особливий ефект впливу цього методу має на молодіжну аудиторію.

6. *Використання образливих назв*. Необхідно відмітити тенденцію до насичення гумором політичних статей і широке використання в них всіляких образливих назв політиків і політичних партій.

7. *Створення негативних асоціацій*, завданням яких є активізувати в пам'яті людини який-небудь негативний образ і перенести його на конкретного політика.

Як видно з наведеного, на практиці PR не обмежуються тільки інформаційною складовою. При цьому, типовими *агресивними технологіями інформаційної роботи PR-служб* є:

- створення інформаційного приводу (освітлення деякої нібито «актуальної» для усіх події або факту);

- інформаційні викиди («утка» - неперевірена або навмисно неправдива інформація або факти);

- концентрація уваги на неправдивому об'єкті (реальний, але не значимий фактаж або штучний об'єкт, імітуючий дійсний для введення в оману).

Отже, проаналізувавши особливості впровадження PR-впливу, можемо сформулювати критерії, що його характеризують:

1. Найбільш оптимальним для реалізації «PR-впливу» є матеріали, що орієнтовані не на констатацію фактів, подій чи стосунків, а спрямовані на опис та оцінювання відношення до події або стосунків, при чому вибір конкретної сфери є питанням суто прикладним і залежить від умов завдання. Умовно назвемо цю характеристику «ступінь подієвості» (подання факту чи його оцінки) і будемо позначати множину таких «подієвих» повідомлень – EV , альтернативну множину – NEV . Певну сферу відображену корпусом повідомлень окремої рубрики будемо позначати – R (множина інформаційних повідомлень певної тематичної / галузевої рубрики).

2. Найбільш ефективним механізмом здійснення «PR-впливу» є використання маніпулятивних технік, що базуються на використанні певної лексики (якщо мова йде про контент у текстовому вигляді). Серед інших, а на нашу думку найбільш значимих і технологічно опрацьованих, ознаками такого впливу є:

- «позитив» / «негатив» (позначимо відповідні множини PZ та NPZ);

- «емоційність» / «не емоційність» (позначимо відповідні множини EM та NEM);

- «маніпулятивність» / «не маніпулятивність» (позначимо відповідні множини MA та NMA).

«Що шукати?»

У певному значенні суть цього питання проходить лейтмотивом нашої первинної гіпотези про відмінність процесу навмисного впливу на процес ухвалення рішення від «неупередженого» інформування. Тобто, якщо такі відмінності існують, тоді вважаємо можливим і необхідним сформулювати характеристики та оцінки і критерії достовірності (наприклад, статистичні) відмінності згаданих вище процесів для підтвердження або спростування нашої гіпотези.

Оціночний інструментарій (характеристики)

Наша попередня робота, присвячена опрацьованню прикладної інформаційно-комунікативної моделі суспільної комунікації, та дослідженням соціально-економічних явищ демонструвала особливості технологічних складових аналізу інформаційного потоку на прикладі інтегратора новин Wwell (wwell.com.ua) [5]. В свою чергу, інформаційною базою інтегратора є розроблена автором система аналізу інформаційного простору (САІП). Її основу, за класичною схемою організації, становить підсистема збору, впорядкування і накопичення у хронологічному порядку контенту з електронних ресурсів Інтернет ЗМІ. Тоді, фактично маємо можливість сформулювати наступні загальні оціночні характеристики інформаційного потоку.

Добове число повідомлень (M_i – множина повідомлень за добу). У САІП ця характеристика обчислюється у процесі побудови діаграми «Місячної активності ЗМІ» на сторінці стрічки новин (Рис.1). При цьому позначимо $I = (M, \gamma)$ – інформаційний потік новин, де відношення квазіпорядку $\gamma \in M \times M$ становить час опублікування повідомлень [11].

З іншого боку загальний інформаційний потік охоплює усі сфери та галузі життєдіяльності суспільства. Тому, і достовірно встановлено, значні інформаційні сплески мають місце тільки у випадках резонансу подій загальнонаціонального масштабу («Балтійський циклон і небували весняні снігопади», «газова війна Росії з Україною», «Вибори Верховної ради», «Вибори Президента»). В основному на фоні загального потоку частка окремих галузей носить адитивний характер, тому логічно визначити у якості оціночної характеристики – *галузеве добове число повідомлень* (R_i – множина повідомлень певної рубрики, що пов'язана з окремою галуззю).

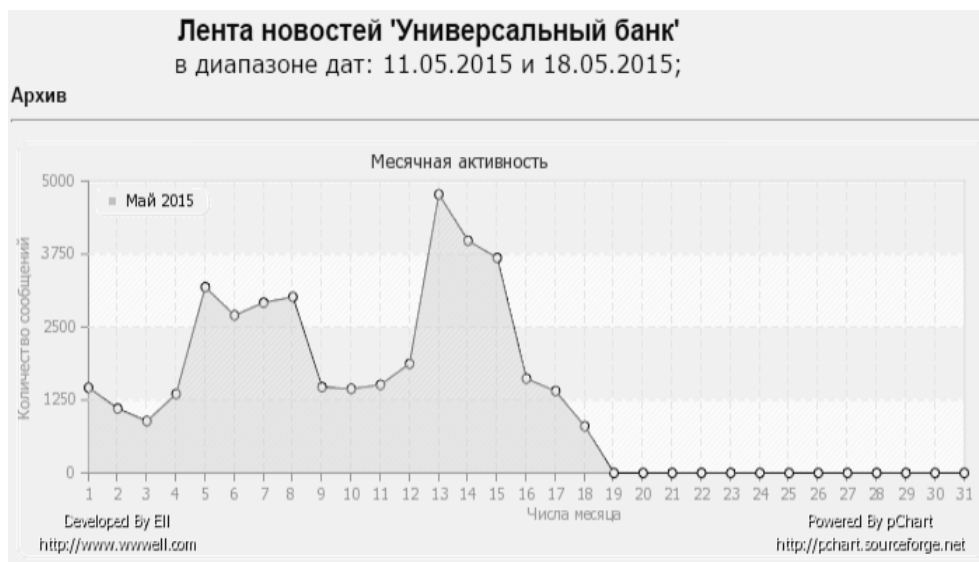


Рис. 1. Діаграма «Місячної активності ЗМІ» побудована САПІ. Спостерігаємо сплеск кількості повідомлень загального інформаційного потоку протягом 13 – 15 травня

Відповідний *галузевий інформаційний потік* позначимо як $I_{R_i} = (R_i, \gamma)$. Базою для відбору таких повідомлень є відповідний семантичний фільтр [5], що виділяє галузевий інформаційний потік із загального. Такий показник дозволить встановити випадки резонансу подій галузевого масштабу

(«відсутність запасів гречаної крупи», «незадовільний стан дорожнього покриття», «заборгованість із заробітної плати працівникам комунального господарства»). Дані для цієї характеристики також доступні на ресурсі САПІ (Рис. 2.).



Рис. 2. Діаграма «Місячної активності ЗМІ» рубрики «Фінанси/банки/інвестиції/страхування ризиків» побудована САПІ. Інформаційний сплеск 13 – 15 травня, обумовлений обговоренням теми переговорів на міжнародному рівні щодо інвестування української економіки та повернення активів «сім'ї»

Вказані показники мають макроскопічний оглядовий характер і дають можливість встановити наявність неоднорідностей на великому об'ємі інформації. В той же час, вони є показовими при ретроспективному аналізі, тобто дозволяють констатувати факт впливу у фазі глибокого резонансу (коли подія отримала достатнє освітлення, «обросла» об'єктно-подієвим матеріалом, виникли відповідні *патерни*, спричинила громадський резонанс, реалізувався *позитивний зворотний зв'язок* і численні ЗМІ публікують повідомлення на резонансну тему в піковій фазі її розвитку). Проте за характером сплеску складно оцінити його першопричину. Іншими словами не можна бути

упевненим, що сплеск синтетичний і викликаний саме завдяки цілеспрямованому впливу. Вказані характеристики носять дуже абстрактний характер і є показовими при ретроспективному аналізі на значних часових інтервалах, а також дозволяють фокусувати увагу дослідника на характерних або особливих фрагментах інформаційного потоку.

Значно більший практичний інтерес являє встановлення ранніх стадій, підготовки впливу (бажано момент зародження), які максимально наближені до моменту створення інформаційного приводу, здійснення «інформаційного викиду». Так, нами встановлено, що початкові фази проведення «PR-впливу» надає можливість виявити порівняльний

щоденний аналіз динаміки тематичного інформаційного потоку $I_{x(t)}$, де $x(t_i)$ – кількість повідомлень визначеної тематики x в момент часу t_i інформаційного потоку у порівнянні не лише з попередньою тенденцією, яка в ньому спостерігалась, але й у співвідношенні до загального інформаційного потоку $I = \sum_{R_i} I_{R_i}$, причому елементи рядів для порівняння тематики x отримуємо із співвідношення:

$$x'(t_i) = \frac{x(t_i) - \bar{x}(t)}{\sigma_{x(t)}}, \quad (1)$$

де середнє $\bar{x}(t)$ і стандартне відхилення $\sigma_{x(t)}$ отримують за оцінками інформаційного потоку [11]. Тому поряд із зазначеними вище макроскопічними показниками, що характеризують «PR-вплив» на високому рівні абстракції, доцільно визначити більш деталізовані оцінки інформаційних потоків.

Подальше міркування по зниженню рівня абстракції [12] спонукає сконцентруватися на розгляді прикладних аспектів галузі, де вже можуть бути виділені окремі *об'єктно-подієві патерни*. Для реалізації такої деталізації в САПІ створений інструментарій «Теми дня» на сторінці стрічки новин. Суттю цього інструменту є набір алгоритмів дворівневої кластеризації (класифікації) множини повідомлень (колекції документів), відібраних протягом доби.

Значимо, що з точки зору загального підходу рішення щодо автоматизованої класифікації колекції документів складаються з чотирьох етапів і вони полягають у наступному:

1. Перехід від множини документів до множини векторів значимих термів (множина таких термів вихідної колекції документів задають простір ознак).
2. Розрахунок ваги кожного терму у кожному документі.
3. Побудова матриці близькості між документами.
4. Виконання процедури класифікації.

Детальний аналіз усього комплексу алгоритмів потребує окремого розгляду і не є завданням цієї роботи. Тому зупинимось лише на окремих відмінностях нашого підходу до класифікації колекції документів з метою формування «тем дня».

З метою підвищення сукупної продуктивності методу ми пропонуємо технологічно розділити етапи класичної схеми на дві групи:

1. «Перехід від множини документів до множини векторів значимих термів» і «Розрахунок ваги кожного терму у кожному документі»
2. «Побудова матриці близькості між документами» і «Виконання процедури класифікації».

Такий розподіл на нашу думку дозволить виконувати обчислення для першої групи етапів у процесі поточного збору інформації, а для другої групи безпосередньо при виконанні запиту на класифікацію певної колекції документів.

Отже, результатом першого рівня класифікації є побудова *сюжетних ланцюжків* інформаційних повідомлень на основі ключових слів самих повідомлень (різновид методу кластеризації К-найближчих сусідів). Полягає він у наступному. Для кожного повідомлення m_i відібраної для аналізу колекції документів $\hat{M} = \{m_1, m_2, \dots, m_i, \dots, m_n\}$ формується, відповідно критерію запропонованою автором [13], вектори ключових термів, які інтегрально утворюють простір ознак (загальний вектор ключових термів) $K = \{k_1, k_2, \dots, k_j, \dots, k_m\}$, і формується матриця зв'язків «повідомлення-ключові слова»:

$$D_{MK} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{m1} & w_{m2} & \dots & w_{mn} \end{pmatrix}, \quad (2)$$

де $w_{ij}, (i = \overline{1, n}; j = \overline{1, m})$ – кількість певного ключового слова i в повідомленні j . Після цього розраховується коефіцієнт кореляції повідомлень за ключовими словами (матриця векторного добутку D_{MM}), що є еквівалентом матриці близькості між документами:

$$D_{MM} = \begin{pmatrix} \bar{w}_{11} & \bar{w}_{12} & \dots & \bar{w}_{1n} \\ \bar{w}_{21} & \bar{w}_{22} & \dots & \bar{w}_{2n} \\ \dots & \dots & \dots & \dots \\ \bar{w}_{n1} & \bar{w}_{n2} & \dots & \bar{w}_{nn} \end{pmatrix}, \quad (3)$$

по рядкам матриці D_{MK} , за формулою яка фактично забезпечує співставлення кожному терму документу його статистичної міри \bar{w}_{ij} :

$$\bar{w}_{ij} = \frac{\left(\sum_{j=1}^m w_{ij} \right)}{\lambda_i}, \quad i = \overline{1, n}, \quad (4)$$

де $\lambda_i = \max(w_{ij}), (j = \overline{1, m})$ – максимальне значення входження i -того ключового слова серед усіх j -тих повідомлень.

Тоді, остаточне розбиття множини \hat{M} повідомлень за класами $\Omega = \{\omega_1, \omega_2, \dots, \omega_l\}$, (де $\omega_i = m_{\omega_i}$ – є повідомлення-центроїд, на роль якого послідовно тестуються повідомлення із множини \hat{M}) відбувається за правилом:

$$m_i \in \omega_k, (k = \overline{1, l}) \Big|_{\bar{w}_{ik} - \bar{w}_{ok} < \alpha},$$

де \bar{w}_{ok} – коефіцієнт кореляції центроїда, а α – граничне значення відстані повідомлення від центроїду класу, або коефіцієнт подібності повідомлення до певного класу тобто *сюжетного ланцюжку*.

Подальші міркування стосовно формування *теми дня* виходять з ідеї, що отримані з колекції документів сюжети фактично дають перелік описів подій та оцінок стосовно окремих фактів та об'єктів реального світу. З іншого боку, окреме повідомлення змістовно стосується більш ширшого семантичного кола у певних контекстах, ніж з яких лексично

складається сюжет. А тому, на нашу думку, класифікація сюжетів дозволить отримати шукане змістовно-сміслову групування повідомлень.

Тоді, на другому рівні сюжетні ланцюжки об'єднуються в *теми дня* на основі кластеризації повідомлень (процедура побудови аналогічна наведеній вище), які увійшли до різних сюжетних ланцюжків. Теми дня становлять перелік подій, які сталися протягом певного періоду і містять контент що складає об'єктно-подієві патерни та журналістську/редакторську оцінку цих подій. Назва теми визначається навою повідомлення-центроїду, яке об'єднало інші повідомлення кластера. Зауважимо, що отриманий за нашою процедурою перелік тем дня змінюється протягом доби і стає сталим тільки по її закінченню. Така особливість дає змогу, в решті, проведення ретроспективного аналізу контенту за цією ознакою.

За побудовою наведеного алгоритму визначення *тем дня* відповідні корпуси у середньому мають значну потужність кожної з тем. Звідси природно випливає думка про значну надмірність узагальненого контенту сформованої теми, що відзначалось нами в загальних міркуваннях. Таку надмірність у медіа сфері визначають поняттям «інформаційний шум» як множини тем, які «на слуху», актуальні, знаходяться «в топах таблоїдів», а так само «замовні статті» і інший інструментарій «PR-впливу» [14], тобто позначені як важливі, на відміну від прийнятого в теорії інформації *information noise* – дані, що не відповідають інформаційній потребі або такі що не становлять предмету новизни для інформування суб'єкта.

Слід зазначити, що за результатом кластеризації Ω множини повідомлень \hat{M} певна множина $\bar{\Omega}$ повідомлень *не увійшла до тем або сюжетів* дня. При цьому нами емпірично встановлено, що саме ця множина містить комунікаційні одиниці (інформаційні повідомлення), які можуть перспективно стати темами дня в тих випадках, коли вони мають потенціал для громадського резонансу (наприклад матеріал для штучної сенсації).

Підсумовуючи викладене зазначимо:

1. Визначені нами макроскопічні оціночні характеристики інформаційного потоку дозволяють шляхом проведення порівняльного щоденного аналізу динаміки потужності контентного наповнення на початкових фазах встановити факт імовірного акту «PR-впливу».

2. Сюжетні ланцюжки та теми дня в поєднанні із вектором ключових слів та зазначеними вище лінгво-статистичними оцінками окремого повідомлення («подієвість», «емоційність», «позитив/негатив», «маніпулятивність») формують простір мікроскопічних характеристик (впорядковують множину інформаційних повідомлень) який, на нашу думку, є необхідним для побудови моделі виявлення «PR-впливів».

3. На підставі викладеного робимо висновок про можливість виявлення фактів «PR-впливу» шляхом ретроспективного аналізу множини повідомлень загального та галузевого потоку, тем

дня і тих, що не увійшли до них (для чого у подальшому передбачається механізм автоматичної побудови лінгвістичних фільтрів на основі корпусу повідомлень тем дня) у просторі мікроскопічних характеристик. Часове вікно такої ретроспекції імовірно обчислюється днями, що в масштабі проведення акції «PR-впливу» також носить мікроскопічний характер.

4. Фактично теми дня з одного боку відповідають найменуванню найбільш важливих подій що висвітлюються ЗМІ, а з іншого – потенційно містять т.з. інформаційний шум, який є первісною функцією «PR-впливу».

5. Механізм побудови тем дня поряд з їх безпосереднім отриманням забезпечує надмірність контенту (у розумінні визначення «смыслу» за Г. Бейтсоном) певної теми і за умови визначення переліку об'єктів, що мають прикладний інтерес дозволяє на заключних етапах аналізу «PR-впливу» формувати об'єктно-подієві патерни з метою іменування головних «дієвих осіб» проведення акту впливу.

«Як шукати?»

Для побудови моделі виявлення «PR-впливів» використаємо модельний приклад. Будемо відрізняти діяльність PR-служб, що мають на меті серед іншого впровадження навмисного впливу, та незалежного журналістського корпусу (НЖК), цілями якого умовно є «об'єктивне» інформування. При цьому вважається, що при використанні PR методів і технологій, журналіст перестає відноситися до НЖК і стає агентом PR-служб. Цікавим для нас є те, що інформаційним продуктом діяльності PR-служб є множина PR-повідомлень або PR-послідовностей (позначимо як $PR = \{pr_1, pr_2, \dots, pr_i, \dots, pr_n\}$), а результатом діяльності НЖК є множина НЖК-повідомлень або НЖК-послідовностей (позначимо як $IJC = \{ijc_1, ijc_2, \dots, ijc_i, \dots, ijc_m\}$). Таким чином, загальний (галузевий) інформаційний простір (тобто деяка множина повідомлень що висвітлює в інформаційному просторі подієвий план) складається з об'єднання множин НЖК- та PR-повідомлень ($M = PR \cup IJC$).

За цих умов можемо сформулювати гіпотезу щодо можливості розпізнавання PR – послідовностей, покликаних впровадити акт навмисного штучного впливу у загальному інформаційному потоці:

Для деякої множини повідомлень $\hat{M} = \{m_1, m_2, \dots, m_i, \dots, m_n\}$ інформаційного потоку $I = (M, \gamma)$, де відношення квазіпорядку $\gamma \in M \times M$, впорядкованою за $T = \{t_1, t_2, \dots, t_n\}$ існує можливість виділення множини $PR = \{pr_1, pr_2, \dots, pr_i, \dots, pr_n\}$ для якої:

$$P(\hat{M}, A, T) = \begin{cases} pr_i \geq \alpha, m_i \in PR; \\ pr_i < \alpha, m_i \notin PR; \end{cases} \Big|_{|PR| \rightarrow \max; \Delta t \rightarrow \min} \quad (5)$$

із статистичною достовірністю, що не перевищує α .

При цьому

$$P(\hat{M}, A, T) = \{p_i, i = \overline{1, n}\},$$

є ймовірність події $A = \{a_i, i = \overline{1, n}\}$, яка в свою чергу обчислюється як нормалізована інтегральна складова лінгво-статистичних ознак:

- $P_{NEV}(m_i)$ – ймовірність, що повідомлення m_i є не подієвим;

- $P_{NPZ}(m_i)$ – ймовірність, що повідомлення m_i є негативним;

- $P_{EM}(m_i)$ – ймовірність, що повідомлення m_i є емоційним;

- $P_{MA}(m_i)$ – ймовірність, що повідомлення m_i є маніпулятивним.

Тобто,

$$P(a_i) = \sum \{P_{NEV}(m_i), P_{NPZ}(m_i), P_{EM}(m_i), P_{MA}(m_i)\}. \quad (6)$$

Отже, (5,6) формалізують модель виявлення загальної PR-послідовності на множині M – загального інформаційного потоку, покликаної до впровадження певної множини впливів (багатовекторність) у різних сферах життя суспільства. Тоді, галузеве розмежування та/або сюжетно-об'єктові патерни отримуємо шляхом «накладення» на отриману множину повідомлень відповідної топології галузевих рубрик та/або тем дня. За умови звуження потоку до галузевого \hat{M} , отримуємо окрему PR-послідовність ймовірного акту впливу.

Висновки

Таким чином, шляхом аналізу особливостей свідомо спрямованого PR-впливу через публікації в ЗМІ інформаційних матеріалів стосовно певних предметних областей та типових форм дії з метою навіювання певних станів та настроїв суспільної свідомості, ми встановили, що найбільш характерними рисами відповідного контенту є зміщення фокусу уваги з самої події на її оцінку («подієвість»), наявність «негативної», «емоційної» та «маніпулятивної» лексики. Потім ми розглянули певні макроскопічні (добова галузева та загальна активність ЗМІ) та мікроскопічні (сюжетні ланцюжки та теми дня) ознаки контенту загального інформаційного потоку, що генерується ЗМІ, та охарактеризували оцінки неоднорідностей, які доцільно вивчати практично для встановлення часових меж (1) проведення акту впливу. Це дало нам змогу в решті сформулювати формальну гіпотезу та безпосередньо саму модель (5,6) виявлення факту наявності контенту, спрямованого на впровадження свідомого інформаційного PR-впливу на множині інформаційних повідомлень, на експериментальну апробацію якої будуть спрямовані подальші дослідження.

Література

[1] Расторгуев С.П. Математические модели в информационном противоборстве. Экзистенциальная математика. / С.П. Расторгуев; Центр стратег. оценок и прогнозов. – М.: АНО ЦСОиП, 2014. – 260 с.

[2] Расторгуев С.П. Информационная война. / С.П. Расторгуев – М.: Радио и связь, 1999. – 221 с.

[3] Залкин С.В. Влияние информационно-психологических воздействий на боевую готовность подразделений // С.В. Залкин / Наука і техніка Повітряних Сил Збройних Сил України. – 2012, № 3(9). – С. 155-158.

[4] Хатян О.А. Технологічні засади побудови моделі аналізу потоків інформаційних повідомлень. / О.А. Хатян // Сучасні інформаційні технології у сфері безпеки та оборони. – № 2(5). – С.43–49.

[5] Хатян О.А. Інформаційно-комунікативна модель суспільної комунікації як основа дослідження соціально-економічних явищ. / О.А. Хатян // Сучасні інформаційні технології у сфері безпеки та оборони. – 2011. – № 3(12). – С.66–71.

[6] Bateson G. Style, Grace and Information in Primitive Art // A Study on Primitive Art / Ed. by A.Forge. Oxford, 1971.

[7] Хатян О.А. Історичний розвиток поняття «загроза» у сфері забезпечення державою інформаційної безпеки. / О.А. Хатян // Інформаційна безпека людини, суспільства, держави. Науково-практичний журнал. – 2009. – № 2(2). – С. 38-42.

[8] Почепцов Г.Г. Паблик рилейшнз для професіоналов. / Г.Г. Почепцов. – Издание 2-е, исправленное. – М.: «Рефл-бук»; К.: «Ваклер», 2000. – 624 с.

[9] Кравченко Н.П. Использование PR-технологий в информационной работе пресс-служб. / Н.П. Кравченко / Методическое пособие для работников пресс-служб, руководителей средств массовой информации, органов государственной власти и местного самоуправления. Выпуск №21 – Краснодар, 2011 г. – 153 с. – С. 17., Иванченко Г.В. Реальность Паблик Рилейшнз. / Г.В. Иванченко – М: Смысл, 1999.

[10] White J., Blamphin J. What we need to know // Journal. The Institute of Public Relations. - 1995. - Vol. 13. - N 8.

[11] Панченко В.М. Методика визначення часових меж спеціальної інформаційної операції шляхом моніторингу потоків новин / В.М. Панченко, О.А. Хатян // Науковий вісник НА СБ України. – №33. – 2010. – С. 107–118.

[12] Хатян О.А. Методологічні засади дослідження інформаційно-суспільних явищ. / О.А. Хатян // Сучасні інформаційні технології у сфері безпеки та оборони. – 2010. – № 3(7) – С. 90-95.

[13] Хатян О.А. Потужність множини ключових слів як критерій інформативності новинного повідомлення. / О.А. Хатян, М.О. Рябий // Актуальні питання забезпечення кібернетичної безпеки та захисту інформації: Зб. наук. праць науково-практичної конференції; м. Київ, 25-28 лютого 2015 р., Європейський університет / Редкол.: О.І. Тимошенко (голова) та ін. – К.: Вид-во Європейського університету, 2015. – С. 117-119.

[14] Інформаційний шум українських ЗМІ. Замовні статті і гонитва за сенсаціями. [Електронний ресурс] / режим доступу: <http://job-sbu.org/informatsionnyiy-shum-ukrainskih-smi-zakaznyie-statii-pogonyia-za-sensatsiyami.html?height=400&width=600>.

УДК 004.912 (045)

Рябий М.А., Хатян О.А., Багацкий С.П. Модель обнаружения PR-влияния посредством Интернет-СМИ

Аннотация. Свойство информационной угрозы заключается в способности запустить механизм самоуничтожения сложной информационной системы, которой является отдельная личность, общественные объединения и общество вообще. Угроза такого типа реализуется посредством акта информационного влияния через информационные источники СМИ, который мы называем направленным или «PR-влиянием». Эффективная модель выявления информационных влияний лежит в основе механизма выявления и противодействия информационным угрозам. В работе характеризуется сознательно направленное влияние как средство изменения качественной оценки или отношения к определенным фактам или объектам целевой аудиторией. Проведенный анализ заданий, методов и механизмов работы PR-служб позволил сформулировать соответствующее понятие «PR-влияние» как средство реализации определенных интенций через публикации в Интернет СМИ, а кроме того определить его особенности, критерии и оценочные характеристики. Сформулирована гипотеза об отличии процесса «PR-влияния» от «объективного» информирования. Предложена формальная модель выявления направленных влияний на множестве информационных сообщений полученных из Интернет СМИ.

Ключевые слова: направленное влияние, критерии оценки влияния, Интернет СМИ, алгоритм классификации коллекции документов, сюжеты новостей, темы дня, объектно-событийные паттерны, модель обнаружения PR-влияния.

Ryabyy M., Hatyay O., Bagatskiy S. The model of PR-impact detection by means of Internet mass-media

Abstract. The properties of information threat consists in abilities to start the mechanism of self-destruction of the complicated informative system (personality, public associations and society in general). Threat of such type realized by means of act of information influence through the informative sources of mass-media, that we name «PR-impact». An effective model of identification information influences is underlying of the mechanism of detection and counteraction of information threat. In the paper is defined consciously directed influence as means of change of quality estimation or attitude toward certain facts or objects of target audience. Conducted analysis of tasks, of methods and mechanisms of work of PR-services allowed to set forth a corresponding concept «PR-impact» as means of realization of certain intensions through publications in the Internet of mass-media, and in addition to define his features, criteria and evaluation of descriptions. Is set forth a hypothesis about the difference of process of «PR-impact» from the «non preconceived» supply of information. Is offered the formal model of finding «PR-impact» from the collection of information messages obtained from the Internet of mass-media.

Key words: the directed influence, criteria of estimation of influence, Internet mass-media, algorithm of documents collection classification, plot of news, pick of day, objects-event pattern, model of searching PR-impact.

Отримано 20 квітня 2015 року, затверджено редколегією 19 травня 2015 року
