

Ю.О. Колос, к.т.н., доц., пров.н.с.

Ю.І. Міхєєв, к.т.н., н.с.

І.О. Орищук, с.н.с.

*Житомирський військовий інститут ім. С.П. Корольова
Національного авіаційного університету*

СПОСІБ ВІДБОРУ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ В АВТОМАТИЗОВАНІЙ СИСТЕМІ МОНІТОРИНГУ РАДІОТЕЛЕВІЗІЙНОГО ПРОСТОРУ

У роботі розглянуто порядок моніторингу інформаційних джерел. Визначено основні параметри радіотелевізійного простору, які дають змогу встановити показники та критерії для відбору каналів та програм. Запропоновано математичний апарат для розрахунку важливості інформаційних джерел.

Вступ. Постановка проблеми та її актуальність. На сьогоднішній день питання автоматизації пошукової діяльності в інформаційній сфері набуває все більшої актуальності. Розробка автоматизованих систем моніторингу пов'язана із необхідністю проведення оперативного аналізу інформаційної обстановки у різних сферах людської діяльності. Постійне оновлення ринку спеціалізованого програмного забезпечення, призначеного для моніторингу радіотелевізійного простору, свідчить про те, що на сьогоднішній день засоби масової інформації традиційно виступають одними з найінформативніших серед каналів отримання даних про останні світові події [1–4].

Моніторинг радіотелевізійного простору в інтересах замовника передбачає виконання таких основних етапів: відбір та спостереження за рядом потенційних джерел, збереження знайденого матеріалу, його фільтрація, категоризація та [5, 6]. Постійне дублювання інформації на різних радіотелевізійних каналах призводить до ускладнення аналітичної роботи оператора, перевантажує засоби запису та збереження даних, що в цілому зменшує оперативність моніторингу. Таким чином, актуальними є питання, що присвячені підвищенню оперативності роботи автоматизованих систем моніторингу радіотелевізійного простору.

Аналіз останніх досліджень та публікацій. Аналіз існуючих систем автоматизованого моніторингу свідчить про те, що більшість з них мають широкі можливості щодо пошуку інформації у радіотелевізійному просторі. У публікаціях розглядається архітектура побудови таких систем, однак, недостатньо опрацьовані питання відбору інформаційних джерел [1–4, 7, 8]. У поданих системах зменшення часу обробки інформації досягається за рахунок збільшення кількості робочих місць та організації цілодобового чергування, що призводить до ускладнення організації роботи системи.

Одним з можливих шляхів підвищення оперативності моніторингу радіотелевізійного простору є вдосконалення процесу відбору інформаційних джерел. Тому **метою статті** є розробка способу відбору радіотелевізійних каналів для подальшого їх запису в автоматизованих системах моніторингу.

Викладення основного матеріалу. Враховуючи можливості сучасних технічних засобів та особливості радіотелевізійного простору організацію та порядок ведення моніторингу можна зобразити у такій послідовності (рис. 1). У результаті сканування радіотелевізійного простору можна встановити його параметри:

- кількість доступних телерадіоканалів (супутникових, наземних);
- параметри доступу (частота або номер каналу, орієнтація антени для приймання, коефіцієнт підсилення та поляризація антени, вид сигналу аналоговий чи цифровий, вид кодування сигналу);
- належність каналу, радіостанції до певної держави;
- інформаційна спрямованість каналів, радіостанцій;
- мова, на якій проводяться передачі.

Далі необхідно провести розподіл телерадіоканалів за інформаційною спрямованістю (новини, інформаційно-аналітичні, розважальні та ін.).

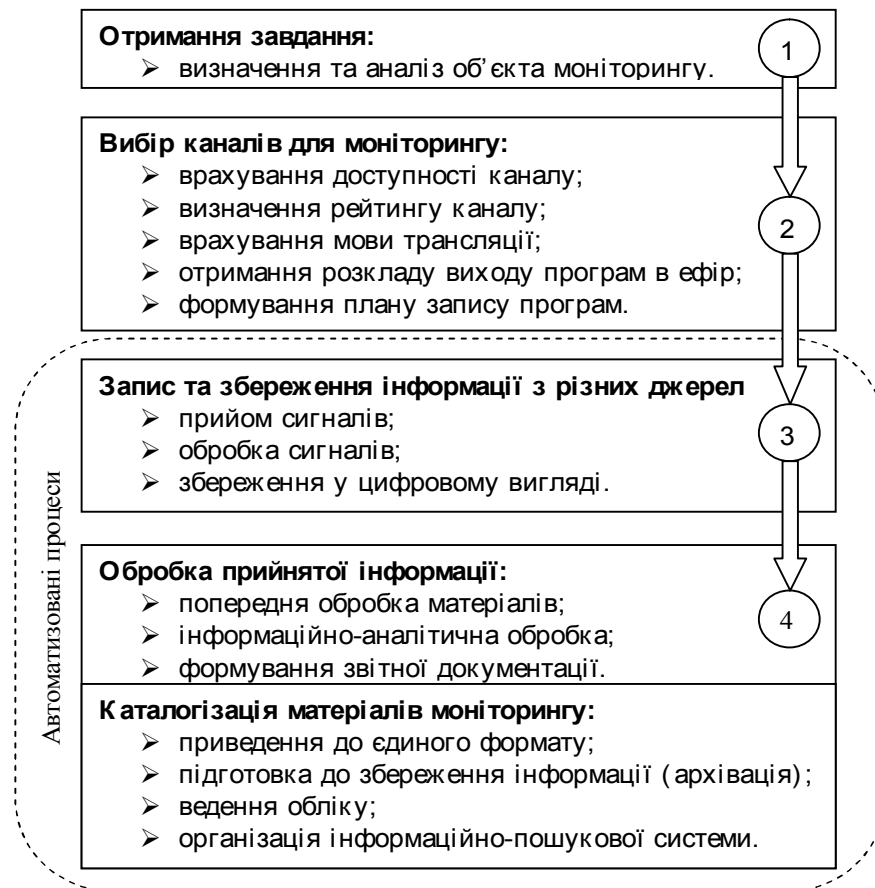


Рис. 1. Етапи моніторингу радіотелевізійного простору

Встановлення відповідності каналів завданню за державною належністю, мовою, інформаційною спрямованістю позначається “1” – якщо відповідає, “0” – якщо не відповідає. Результати сканування та вивчення доступних каналів оформлюються у вигляді таблиць (табл. 1).

Таблиця 1

Приклад оформлення результатів сканування та вивчення доступних каналів

Супутникові канали							
№ з/п	Назва каналу	Належність	Відповідність завданню за належністю	Мова	Відповідність завданню за мовою	Інформаційна спрямованість	Відповідність завданню за інформаційною спрямованістю
1	Poland1	Польща	1	Польська	1	Науково-популярні	1
2	BBC	Англія	0	Англійська	1	Новини	0

Крім того, необхідно врахувати, що канали можуть бути універсальні, програми, що містять практично всі види передач, та спеціалізовані (тематичні), абсолютна більшість часу проводяться передачі певної спрямованості. Завданням моніторингу найбільш відповідають канали, на яких проводяться новини та інформаційно-аналітичні передачі.

У подальшому шляхом спостереження, вивчення програм передач, додаткової інформації встановлюються:

- належність каналу державі, організації, особі;
- у чий інтересах подається інформація (провладні, опозиційні, незалежні);
- масштабність (центральні, загальнодержавні, місцеві) і формується реєстр телерадіоканалів.

Можлива форма реєстру показана у таблиці 2.

Таблиця 2

Форма каталога для зберігання інформації про телеорганізації

Назва та адреса телеорганізації, позивні, логотип	Керівник	Засновник (власник)	Вид діяльності та мовлення	Територія розповсюдження програм	№ каналу або частота	Обсяг мовлення, мова
1	2	3	4	5	6	7

З точки зору наявності та подачі інформації певної спрямованості, наприклад, новин, телевізійний канал можна охарактеризувати такими параметрами: оперативність та часова щільність повідомлень про певні об’єкти, певної тематичної спрямованості – термін подачі інформації про події, явища після їх початку; скільки часу протягом доби здійснюються повідомлення певної тематики. З урахуванням зазначеного вище необхідно встановити показники та критерії для відбору I каналів та програм на них з метою подальшого моніторингу.

Множину завдань моніторингу позначимо $L = \{l_1, l_2, l_n, \dots, l_N\}$, де N – загальна кількість завдань.

На основі аналізу кожного n -го завдання l_n створюється множина тем і об’єктів, що підлягають моніторингу на його виконання $Y = \{y_{n1}, y_{n2}, y_{n3}, \dots, y_{nT_n}\}$, де T_n – загальна кількість тем, об’єктів відповідно до n -го завдання.

Тоді множина Y включає всі теми та об’єкти моніторингу. Множини L, Y визначають предметну галузь моніторингу. Далі необхідно визначити множину каналів для вивчення.

На першому етапі проводиться запис запланованих передач, на другому – виявляється та відбирається корисна інформація для подальшої обробки, що відповідає предметній галузі. У спрощеному варіанті необхідно встановити наявність таких тем і об’єктів, як і в множині Y . У загальному випадку до корисної також повинна бути віднесена інформація за темами про об’єкти, що завчасно не встановлені, але відповідають завданням. Якщо оператор не може прийняти рішення щодо корисності інформації, то вона також відбирається і записується для подальшої експертизи. Нові теми й об’єкти можуть бути додані до множини Y . За наявності анонса оператор може зразу виконувати другу операцію і записувати не всю передачу, а відібрані частини. Відібрана корисна інформація, а також інформація про канал, передачу, дату, час початку і кінця повідомлення записується до бази повідомлень.

Більшість опитаних експертів вважають, що важливість телерадіопередачі (каналу) для моніторингу, в першу чергу, визначається достовірністю, повнотою, оперативністю, частотою подачі в них інформації, що відповідає предметній області моніторингу. Крім того, показники слід вибрати так, щоб їх можна було визначати в процесі моніторингу і вони повинні вказувати на можливу зміну важливості. У результаті аналізу різних підходів вибрані такі показники:

– P_i – повнота охоплення тематики на i -му каналі (частка заданих для моніторингу тем, об’єктів, що висвітлюються на каналі). Вона характеризується відношенням кількості тем, об’єктів K_{oi} із визначеного переліку, які висвітлюються в передачах каналу, до їх загальної кількості в переліку K_{zo} :

$$P_i = \frac{K_{oi}}{K_{zo}}; \tag{1}$$

– λ_i – відносна середня частота висвітлення на каналі визначених для моніторингу тем, об’єктів (відношення кількості повідомлень про визначені об’єкти чи відповідно до визначеної тематики протягом $m \gg 1$ діб на i -му каналі K_{zi} до сумарної кількості таких повідомлень на всіх каналах K_{zm}):

$$\lambda_i = \frac{K_{zi}}{K_{zm}}; \tag{2}$$

– t_{ci} – середня тривалість висвітлення на i -му каналі визначених для моніторингу тем, об’єктів протягом доби (відношення сумарної тривалості повідомлень протягом $m \gg 1$ діб до кількості діб спостереження):

$$t_{ci} = \frac{1}{m} \sum_{j=1}^{K_{ci}} (t_{ej} - t_{bj}), \tag{3}$$

де t_{bj} – час початку трансляції повідомлення; t_{ej} – час закінчення трансляції повідомлення;
– відносна середня тривалість:

$$t'_{ci} = \frac{t_{ci}}{\sum_{i=1}^I t_{cp}}, \tag{4}$$

де t_{cp} – середня тривалість інформаційних повідомлень на i -му каналі;
– оперативність подачі інформації на каналі:

$$O_i = \sum_{j=1}^{K_{ci}} \frac{1440 - t_{bj}}{1440}, \tag{5}$$

де 1440 – тривалість доби у хвиликах.

Даний показник буде тим більший, чим раніше відносно початку доби на каналі передаються інформаційні повідомлення.

Розглянуті вище показники змінюються від 0 до 1, причому їх збільшення означає покращення інформативності та оперативності. Тому важливість каналу для моніторингу тим більша, чим більші повнота, частота, тривалість, оперативність повідомлень:

$$B_i = P_i \lambda_i t'_{ci} O_i. \tag{6}$$

У даному випадку всі показники рівноцінні. За рішенням експертів можна змінювати вплив на важливість того чи іншого показника шляхом введення вагових коефіцієнтів. Такі ж показники можна застосувати для визначення важливості передач. Для визначення значень вибраних показників необхідно сформувати предметну галузь, провести спостереження та набрати статистичні дані про наявність інформативних для виконання завдань моніторингу повідомлень та їх параметри.

Подальша обробка полягає в ідентифікації, формуванні бази повідомлень, встановленні параметрів повідомлень і формуванні бази параметрів. Ідентифікацію повідомлень пропонується здійснювати за: датою; номером завдання n у множині L , на виконання якого відібрана інформація; номером об'єкта, теми в множині Y , про які є інформація в повідомленні; порядковим номером повідомлення, яке містить інформацію про визначений об'єкт, тему. Тоді структуру ідентифікаційного номера доцільно обрати такою, як в таблиці 3.

Таблиця 3

Структура ідентифікаційного номера

№ завдання		№ об'єкта, теми			№ з/п повідомлення		
0	1	0	3	1	2	3	4

Ідентифікаційний код містить всього 8 розрядів, що поділяються для номера завдання – 2 (максимальна кількість завдань 99), номера об'єкта, теми – 3 (максимальна кількість об'єктів, тем 999); порядкового номера повідомлення за певною темою, про певний об'єкт – 3 (максимальна кількість повідомлень 999).

База повідомлень повинна включати такі розділи: дату; ідентифікаційний номер; зміст повідомлення. До параметрів повідомлень необхідно включити такі, які б дозволили розраховувати показники важливості каналів (1)–(5). У результаті накопичуються інформаційні повідомлення і створюється база даних про них у вигляді таблиці 4.

Таблиця 4

Таблиця параметрів повідомлень

Дата	№ каналу	Програма	№ завдання	№ теми (об'єкта)	№ з/п	Час початку	Час кінця
1	2	3	4	5	6	7	8

За даними таблиці можна визначити важливість каналу шляхом проведення відповідних розрахунків за виразами (1)–(6). За результатами вивчення для подальшого спостереження відбираються найбільш важливі канали. Але цього може бути недостатньо. Використовуючи дані з таблиці 4, а також показник (1), необхідно додатково перевірити і скорегувати сукупність відібраних каналів так, щоб найбільш повно охопити визначені теми та об'єкти моніторингу. Аналогічно визначається сукупність передач для моніторингу.

Висновки. Таким чином, запропонований спосіб відбору інформаційних джерел дає змогу формувати план запису радіотелевізійних каналів з урахуванням відомостей про об'єкт моніторингу, що зменшує час подальшої їх обробки. Аналіз результатів випробування запропонованого способу відбору телевізійних каналів у пересувній системі моніторингу радіотелевізійного простору ПРТК-1 показав, що час відпрацювання оперативних завдань зменшився на 15 %.

ЛІТЕРАТУРА:

1. Tv monitoring – система моніторингу телевизионного эфира. – [Электронный ресурс]. – Режим доступа : http://www.hopsteam.ru/index.php?item=tv_monitoring&lang=ru.
2. Автоматизированная система мониторинга и анализа СМИ. – [Электронный ресурс]. – Режим доступа : <http://s-graph.ru/web/ru/50>.
3. Медиалогия – первая система анализа СМИ в реальном времени. – [Электронный ресурс]. – Режим доступа : <http://www.mlq.ru/about>.
4. Система автоматизированного мониторинга телевизионных программ (САМ-ТВ). – [Электронный ресурс]. – Режим доступа : <http://www.niir.ru/rus/page.php?trid=94>.
5. Колос Ю.О. Автоматизована система комплексного моніторингу інформаційних джерел / Ю.О. Колос, О.В. Манько, Ю.І. Міхєєв // Вісник ЖДТУ / Технічні науки. – 2010. – № 3 (54). – С. 59–65.
6. Тренинги: журналистика. Тема “Методика мониторинга электронных и печатных СМИ”. – [Электронный ресурс]. – Режим доступа : <http://islamtat.ru>.
7. Системы мониторинга и анализа СМИ. – [Электронный ресурс]. – Режим доступа : http://www.newart.ru/oparin/smi_oparin.htm.
8. Система автоматизированного мониторинга телевизионной рекламы. – [Электронный ресурс]. – Режим доступа : <http://broadcasting.ru/articles2/econandmen>.

КОЛОС Юрій Олександрович – кандидат технічних наук, доцент, провідний науковий співробітник науково-дослідної лабораторії інформаційних технологій наукового центру Житомирського військового інституту ім. С.П. Корольова Національного авіаційного університету.

Наукові інтереси:

– інформаційні технології.

МІХЄЄВ Юрій Іванович – кандидат технічних наук, науковий співробітник науково-дослідної лабораторії інформаційних технологій наукового центру Житомирського військового інституту ім. С.П. Корольова Національного авіаційного університету.

Наукові інтереси:

– інтернет-розвідка.

E-mail: yuramiheev@ukr.net

ОРИЦУК Ігор Олександрович – старший науковий співробітник науково-дослідної лабораторії інформаційних технологій наукового центру Житомирського військового інституту ім. С.П. Корольова Національного авіаційного університету.

Наукові інтереси:

– інформаційні технології.

Подано 03.05.2011

