

Забезпечення вибіркового доступу до контенту мережі Інтернет як засобу оцінки ситуації та прийняття рішень

Проблематика

Україна, як і весь розвинутий світ, намагається побудувати нове, постіндустріальне суспільство - суспільство інформаційне, яке характеризується підвищенням ролі інформації і знань, підвищенням частки інформаційних комунікацій, інформаційних продуктів і послуг у валовому внутрішньому продукті, створенням глобального суспільства, що забезпечує реалізацію прав і свобод людини, ефективну інформаційну взаємодію своїх суб'єктів, їхній доступ до національних та світових інформаційних ресурсів та задоволення соціальних, економічних та особистих потреб та інтересів в інформаційних продуктах і послугах.

Це прагнення потребує постійної оцінки, особливо з урахуванням суспільної думки, поточної ситуації та корегування, за необхідністю, прийнятого курсу та, відповідно, конкретних дій. Викладені положення стосуються майже всіх сторін діяльності суспільства – зовнішньо- та внутрішньополітичного курсу України, організації і проведення виборів, питань національної безпеки тощо.

Наразі, у часи постійного поповнення і вдосконалення інформаційного наповнення (контенту) світової мережі Інтернет, у часи, коли інформація не має кордонів і цензури, у часи, коли інформаційний простір вражає своїми обсягами і розмаїтістю, спрямованістю (точніше – її відсутністю) відкритої інформації у багатьох випадках навіть відпала необхідність у організації та проведенні розвідницької агентурної роботи однієї країни у інших країнах. Досить мати доступ до Інтернет та необхідну кількість висококваліфікованих спеціалістів-аналітиків, а також відповідний програмно-алгоритмічний інструментарій. І не треба ризикувати життям людей, ризикувати політичною репутацією та коштами платників податків.

Сьогодні Інтернет - це величезне сховище інформації, інтегрований доступ до найбільш динамічної складової якого - новинних ресурсів – ускладнений [1]. Розмаїтість інформації, у тому числі, наприклад, повідомлень Мережі не може бути корисна на практиці при відсутності ефективного доступу [2].

Для вирішення цієї проблеми розроблено чисельні концепції, серед яких слід зазначити концепції Семантичного Web та Web другого покоління [3], у рамках яких йдеться щодо надання універсального доступу саме до знань.

Але сьогодні замало отримати доступ до загальносвітової інформації, замало зробити її первісну обробку та здійснити інші необхідні процедури. Сьогодні необхідно спромогтися зберегти свою власну інформацію. Недаремно, стандартом де-факто, став принцип фізичного розділення зовнішньої та внутрішньої телекомунікаційних мереж.

Крім того, величезні обсяги інформації в Інтернет щонайменше обумовлюють завдання вибору найважливіших відомостей з новинних інформаційних потоків, щоб надати ці дані користувачам шляхами розсилки електронною поштою (найпростіший варіант), публікації на веб-сайтах та публікації у ЗМІ після аналітичного узагальнення.

Служби синдикації новин

Оптимальне рішення, здатне допомогти орієнтуватися у новинній інформації Інтернет, сьогодні надають інформаційні служби нового типу - системи синдикації новин [4]. Під синдикацією у цьому випадку розуміється збір інформації в Інтернет і подальше поширення її фрагментів відповідно до потреб користувачів. Крім того, служби синдикації забезпечують публікацію однакових даних на різних сайтах (у тому числі, призначених для кишенькових комп'ютерів і мобільних телефонів).

Вирішують ці завдання шляхами збору, сканування інформації з мережі Інтернет; приведення її до єдиного формату (у останній час – до XML, RSS); вибіркового розподілу, формування тематичних каналів (фідів) та подальшого розповсюдження [5].

Формування та розповсюдження фідів здійснюється за попередньо встановленими темами (найчастіше) або за запитами користувачів (у розвинених, професійних системах контент-моніторингу).

Технологія синдикації Інтернет-новин містить у собі "навчання" програм структурі обраних джерел (Web-сайтів), безпосереднє сканування інформації, її приведення до загального формату (останнім часом - до XML), класифікацію й доставку користувачам різними шляхами (e-mail, Web, WAP, SMS і т.д.).

Для рішення завдання синдикації новин було створено кілька форматів опису даних на основі XML. Найпоширеніший формат одержав назву RSS, що означає Really Simple Syndication, Rich Site Summary, хоча спочатку він називався RDF Site Summary. Зміст всіх цих аббревіатур укладається в простому способі узагальнення й розподілу інформаційного наповнення Web-сайтів - синдикації контенту.

Спочатку RSS створювався компанією Netscape для порталу Netcenter як один з перших XML-додатків, але потім став використовуватися на багатьох інших сайтах. Сьогодні практично всі провідні новинні сайти, "живі журнали", що працюють в Інтернет, використовують RSS як інструмент оперативного представлення своїх поновлень. Наприклад, сьогодні експорт в RSS здійснюють всі найбільші портали, включаючи CNN, BBC News, Amazon, CNet News, MSNBC, The Register, Wired тощо.

Існує багато версій RSS (0.90, 0.91, 0.92, 0.93, 0.94, 2.0), при цьому, багато хто вважають всі версії, крім 2.0, застарілими та «скасованими», але це далеко не так, поки ще самим популярним є RSS 0.91.

У всіх версіях RSS є деякі особливості, але поєднує їхня орієнтація на один тип інформації, внаслідок чого вони включають загальні базові поля: основний блок даних (channel), що містить як атрибути заголовки каналів (title), посилання (link), дані про мову повідомлень (language) і логотип (image), після яких іде список самих повідомлень, де в кожному пункті (item) вказується заголовок (title), короткий опис (description) і посилання на новину (link). Крім того, кожен RSS-файл починається обов'язковими елементами xml й rss. Перший із цих елементів містить атрибути version (версія) і encoding (кодування).

Крім заголовка блоку даних у форматі RSS передбачений опис окремих інформаційних елементів (item). Кожен елемент <item> - це окрема стаття або коротка анотація й посилання на повну версію статті. Канал (channel) може включати будь-яке число елементів <item>, що містять тільки два обов'язкові вкладені елементи - назва (title) і опис (description). Крім того, часто

використовуються такі вкладені елементи: посилання на першоджерело (link), категорія (category), коментар (comments) і автор (author).

Сучасні підходи до одержання новинної інформації з мережі Інтернет передбачають застосування RSS-агрегаторів - спеціальних програм для персональних комп'ютерів, мобільних пристроїв (подібних програм чимало, наприклад: Atilon, FeedReader, Syndirella, FeedExpress, HotSheet, Beaver, NewsDesk тощо), плагінів для веб-браузерів (на цей час існує багато спеціальних RSS-плагінів) та звичайних засобів для перегляду веб-сайтів, до яких вбудовано фіди як новинні рядки або колонки [5].

В останньому випадку можливий інформаційний шум, особливо коли фіди формуються за запитами користувачів автоматично (перший режим). Тому на допомогу адміністраторам веб-сайтів приходять два інших режими: постмодерації, коли адміністратор може вилучати посилання на новини після їх публікації, або модерації, коли публікуються лише ті новини, які відбирає адміністратор (цей режим особливо важливий для веб-сайтів органів державної влади) [6].

Система контент-моніторингу

Саме з метою використання інформаційних масивів мережі Інтернет для оцінки ситуації та прийняття рішень створюються та впроваджуються системи контент-моніторингу. Як приклад однієї з таких систем надалі будемо розглядати систему InfoStream (<http://infostream.ua>) та сервіси на її основі. Ці сервіси, зокрема, дозволяють користувачеві одержувати новинну інформацію з тематики, що його цікавить, електронною поштою, SMS або вмонтувати постійну добірку у свою веб-сторінку тощо. Персоналізація інтерфейсу користувачів системи InfoStream, що працюють у режимі онлайн, тобто збереження їхніх постійних запитів й організація підписки, реалізується на основі сучасної технології RSS 0.91.

Для отримання тематичного каналу (RSS-фіду) від комплексу контент-моніторингу у відповідне поле RSS-агрегатора треба ввести адресу в форматі такого типу:

[http://\[ДОМЕН\]/rss.php\[?<ЗАПИТ>\]](http://[ДОМЕН]/rss.php[?<ЗАПИТ>])

де *ДОМЕН* - це адреса сервера контент-моніторингу, а як *ЗАПИТ* можна ввести слово або словосполучення мовою запитів інформаційно-пошукової системи InfoRes.

На основі технології InfoStream створені численні фіксовані новинні фіди, серед яких можна назвати, наприклад:

Банки - <http://uaport.net/cgi-bin/infostream.rss?rubr02>

Економіка України - <http://uaport.net/cgi-bin/infostream.rss?rubr04>

Нормативні акти - <http://uaport.net/cgi-bin/infostream.rss?rubr09>

Оборона, конверсія - <http://uaport.net/cgi-bin/infostream.rss?rubr10>

Енергетика - <http://uaport.net/cgi-bin/infostream.rss?rubr16>

Освіта - <http://uaport.net/cgi-bin/infostream.rss?rubr23>

Політика - <http://uaport.net/cgi-bin/infostream.rss?rubr34>

Застосування RSS-технології дозволяє користувачам, які мають доступ до сервера контент-моніторингу в режимі онлайн, зберігати свої постійні запити та організувати підписку на них за допомогою різноманітних програмних рішень.

Для організації підписки на тематичний канал (RSS-фід) треба в основному вікні інформаційно-пошукової системи налаштувати запит, після чого запустити

програму-агрегатор (наприклад, FeedReader), активізувати опцію New та ввести таку інформацію:

- адресу RSS-фіду, включаючи запит інформаційно-пошуковою мовою у наведеному вище форматі;
- назву інформаційного каналу, яка може визначатися користувачем;
- періодичність оновлення.

Для управління підпискою у цьому режимі існують додаткові опції, що активуються для кожного RSS-фіду окремо:

- поновлення фіду (переліку активних повідомлень);
- відмітка повідомлень, як вже прочитаних;
- вилучення переліку повідомлень;
- зміна властивостей підписки, включаючи тему, періодичність тощо.

Для отримання повного тексту повідомлення, заголовок і анотація якого зацікавила користувача, треба:

- здійснити подвійне натискання клавішею миші на заголовок або
- активізувати гіперпосилання "Read on" в полі анотації, або
- натиснути на відповідну кнопку, що стоїть перед заголовком, або
- натиснути на праву клавішу миші, знаходячись курсором на заголовку, при цьому можна відкрити текст повідомлення у новому вікні Інтернет-браузера.

Організація стрічок новин на веб-сайтах користувачів

Застосування RSS-технологій та автоматично отримуваних фідів від систем контент-моніторингу для встановлення їх на веб-сайтах має одну велику ваду. Серед формально релевантних тематичному запиту повідомлень можуть з'являтися:

- не достовірні;
- дублікати (змістовні дублі);
- застарілі;
- літературно недопустимі;
- релевантні, але не пертинентні (що не відповідають суті проблематики).

Названа вада унеможлиблює автоматичну публікацію таких новинних блоків, особливо на серйозних веб-сайтах органів державної влади.

Тому для вирішення цієї проблеми було створено два допоміжних режими: постмодерації, коли адміністратор може вилучати посилання на новини після їх публікації, та модерації, коли публікуються лише ті новини, які відбирає адміністратор [6].

На Рис. 1. схематично показано, як новинна інформація з системи контент-моніторингу поступає на веб-сайт, проходячи при цьому через автоматизоване робоче місце адміністратора (АРМ), який модерує цей потік новин.

Відкритим залишається питання обсягів інформації, які має переглядати адміністратор, працюючи у режимі відбору-модерації. Ці обсяги залежать від запиту. Відомо, що відгуки інформаційно-пошукової системи з точки зору користувача характеризуються повнотою та пертинентністю (відповідністю отриманої вибірки його інформаційним потребам), а також те, що ці характеристики є антагоністичними. У ідеалі максимальна повнота – це повний обсяг бази даних, який найчастіше неможливо переглянути у режимі реального часу, а максимальній пертинентності відповідає один або декілька документів, яких недостатньо для представлення у колонці новин. З цього випливає, що

мистецтво формування точних запитів є необхідною складовою роботи адміністраторів новинних ресурсів. Зокрема у сервісі WebStream (про який буде йтися далі) передбачається надання адміністратору для модерування одночасно не більше 50 документів. Щоб досягти цього було створено спеціальний інтерфейс налагодження запитів (Рис. 2).

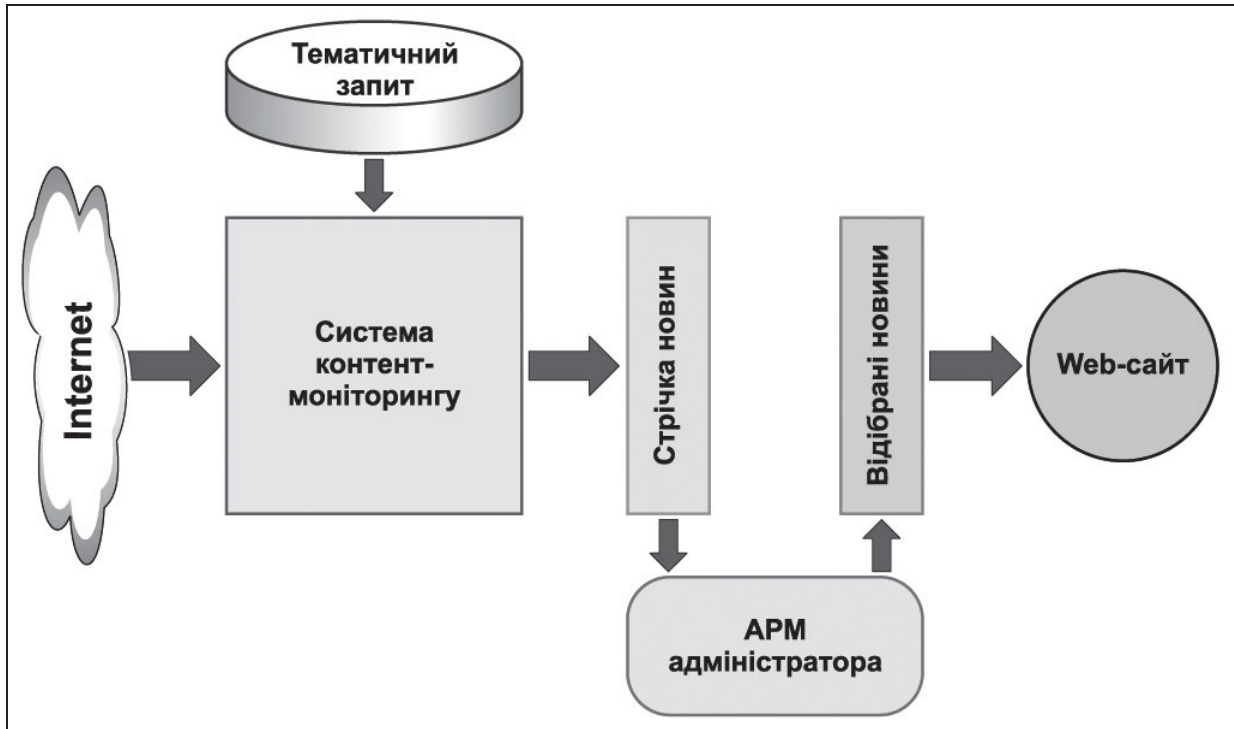


Рис. 1. Схема проходження новин у режимі модерації

Редактирование запросов

Имя пользователя:

E-mail:

URL:

| No | Тема | Запрос | Режим |
|----|-----------------------------------|---|---|
| 1 | <input type="text" value="nato"/> | <input type="text" value="(нато) евроатлантическ евроатлантич)&(украин україн)"/> | <input type="text" value="Модерируемый"/> |
| 2 | <input type="text"/> | <input type="text"/> | <input type="text" value="Модерируемый"/> |
| 3 | <input type="text"/> | <input type="text"/> | <input type="text" value="Постмодерируемый"/> |
| 4 | <input type="text"/> | <input type="text"/> | <input type="text" value="Модерируемый"/> |
| 5 | <input type="text"/> | <input type="text"/> | <input type="text" value="Модерируемый"/> |

Рис. 2. Інтерфейс з конфігурування новинної стрічки

WebStream - спеціальний сервіс для Інтернет-проектів, за допомогою якого здійснюється розміщення на Web-сайті користувача автоматично обновлюваної стрічки новин. Новинна інформація у вигляді заголовків й анонсів статей розміщується безпосередньо на сайті користувача за допомогою вбудованого JavaScript-застосування. Існуючий пакет WebStream дозволяє користувачеві самостійно управляти стрічкою новин, що спочатку формується автоматично. Одержавши логін і пароль, користувач одержує можливість управління відображенням тієї або іншої новини в стрічці новин через Web-інтерфейс. Включення або виключення новини зі стрічки здійснюється за допомогою простого поля вибору (CheckBox), тобто установкою або зняттям ознаки відображення для конкретної окремої новини по розсуду користувача. Причому управління може здійснюватися у двох режимах (Рис. 2). У режимі включення новини в стрічку за замовчуванням, коли користувач вилучає зі стрічки ті новини, відображення яких він вважає недоцільним (режим постмодерації). В іншому режимі новина буде відображена в стрічці тільки після того, як користувач установить ознаку відображення для цієї новини (режим модерації).

Додатково зручний і простий веб-інтерфейс дозволяє легко змінювати настроювання відображення новинних повідомлень та автоматично генерувати код, який можна розташовувати на веб-сторінках (Рис. 3).

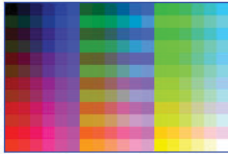
| | |
|---|---|
| 1. Ширина таблиці: (в точках или процентах) | <input type="text"/> |
| 2. Цвет фона: (шестнадцатиричный) | <input type="text"/> |
| |  |
| 3. Формат новостей: | <input type="text" value="1"/> |
| 4. Название стиля для заголовка статьи: | <input type="text"/> |
| 5. Показывать дату: Название стиля для даты: | <input checked="" type="radio"/> Да <input type="radio"/> Нет <input type="text"/> |
| 6. Показывать время: Название стиля для времени: | <input checked="" type="radio"/> Да <input type="radio"/> Нет <input type="text"/> |
| 7. Показывать анонс статьи: Название стиля для анонса статьи: | <input checked="" type="radio"/> Да <input type="radio"/> Нет <input type="text"/> |
| 8. Показывать название источника: Название стиля для названия источника: | <input checked="" type="radio"/> Да <input type="radio"/> Нет <input type="text"/> |
| 9. Имя окна для открытия новостей: | <input type="text"/> |
| 10. Количество новостей: | <input type="text" value="5"/> |
| <input type="button" value="По умолчанию"/> <input type="button" value="Генерировать код"/> | |
| Ваш Код: <input type="text"/> | |

Рис. 3. Фрагмент інтерфейсу з вибору параметрів відображення

Таким чином, роботу із впровадження сервісу можна розділити на два незалежні етапи: налагодження запиту, відповідно до якого буде формуватися новинна стрічка; налаштування параметрів новинної стрічки (вибір дизайнерського рішення щодо новинної стрічки, вибір шрифтів і формату для відображення новин у стрічці), що здійснюється за допомогою веб-інтерфейсу та розміщення готового коду на веб-сторінці.

Висновки

З метою підтвердження ефективності запропонованих і розроблених механізмів та інструментарію використання інформації, яка циркулює у мережі Інтернет, для оцінки ситуації та прийняття відповідних рішень, а також - підвищення рівня обізнаності населення країни щодо співробітництва України з НАТО розроблено нову версію веб-сайту апарату Національного центру з питань євроатлантичної інтеграції України, який розміщено за адресою www.nceai.gov.ua (Рис. 4).



Рис. 4. Головна сторінка веб-сайту апарату Національного центру з питань євроатлантичної інтеграції України

Інформаційна структура цього сайту виконана в контексті Плану дій Україна-НАТО і, серед іншого, забезпечує доступ користувачів до ретроспективної та оперативної баз даних з євроатлантичної інтеграції України на основі технології

InfoStream [6]. Однією з "родзинок" сайту є "онлайнова стрічка" актуальної інформації електронних ЗМІ з євроатлантичної тематики, яку реалізовано на базі сервісу WebStream.

Системи синдикації Інтернет-новин вирішують проблему знаходження необхідної інформації, але залишають без уваги такі завдання, як узагальнення даних - їхню обробку й аналіз. Одним із самих перспективних напрямків узагальнення новинних інформаційних потоків на цей час визначено концепцією «глибинного аналізу текстів» (Text Mining). По відношенню до новинних потоків її ідеологію можна сформулювати як постійне відтворене в часі виконання змістовного аналізу [7]. Саме безперервна аналітична обробка повідомлень є самою характерною рисою цієї концепції, що дозволяє формувати автоматичні дайджести, виявляти нові поняття і їхні взаємозв'язки, будувати різноманітні рейтинги. Саме системи такого типу зможуть позбавити користувачів від інформації, що дублюється, інформаційного шуму, дозволять виявляти головні тенденції, знаходити ланцюжки подій, що корелюють.

Темп життя, що постійно прискорюється, вимагає оперативного одержання повної та актуальної інформації, яка має враховуватися при прийнятті управлінських, маркетингових, виробничих рішень. Сучасні технології синдикації контенту будуть кращим помічником у цьому процесі - потрібно тільки вибрати відповідний інструментарій.

Список літератури

1. Ландэ Д.В. Поиск знаний в Internet. Профессиональная работа -М.: "Диалектика", 2005. - 272 с.
2. Брайчевский С.М., Ландэ Д.В. Современные информационные потоки: актуальная проблематика // Научно-техническая информация. Сер. 1. –М., 2005. - №11, - С. 21-33.
3. Ландэ Д.В. 2GW – Майбутнє Інтернету // Правова інформатика. –К., 2006. - № 1 (9). –С. 34-41.
4. Фурашев В.М., Ландэ Д.В. Інформаційне середовище нового суспільства. // ДНІ НАУКИ: Зб. тез. доповідей: В 3 т. / Гуманітарний університет "ЗІДМУ", 27-28 жовтня 2005. - Запоріжжя: ГУ "ЗІДМУ", 2005. - Т.1. - С. 300-303.
5. Д.В. Ландэ, Морозов А.Ю. Новостной Интернет. «Телеком», - № 11, 2004. – С. 54-57. - № 1-2, 2005. – С 58-62.
6. Ландэ Д.В., Фурашев В.М., Григор'єв О.М. Програмно-апаратний комплекс інформаційної підтримки прийняття рішень: Науково-методичний посібник. - Київ: Інжиніринг, 2006. – 48 с.
7. Фурашев В.М., Ландэ Д.В., Григор'єв О.М., Фурашев О.В. Електронне інформаційне суспільство України: погляд у сьогодення і майбутнє. – К.: Преса України, 2005. – 166 с.

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
НАЦИОНАЛЬНЫЙ АЭРОКОСМИЧЕСКИЙ УНИВЕРСИТЕТ
им. Н.Е. ЖУКОВСКОГО
"ХАРЬКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ"**

**ОТКРЫТЫЕ ИНФОРМАЦИОННЫЕ
И КОМПЬЮТЕРНЫЕ ИНТЕГРИРОВАННЫЕ
ТЕХНОЛОГИИ**

Сборник научных трудов

Выпуск 32

**Харьков
«ХАИ»
2006**

ББК 32.973.3
О 83
УДК 629.7.01; 681.327.11

Утверждено к печати ученым советом Национального аэрокосмического университета им. Н. Е. Жуковского «ХАИ», протокол № 1 от 20.09.2006 г.

Редакционная коллегия: А. И. Бабушкин, П. В. Балабуев, В. К. Борисевич, В. Е. Гайдачук, А. Г. Гребеников, Н. В. Ефремова, В. В. Кириченко, Б. М. Конорев, В. С. Кривцов, А. К. Мялица, В. И. Рябков, И. Б. Сироджа, В. Г. Сухоребрий, О. Е. Федорович, В. Н. Фурашев

Открытые информационные и компьютерные интегрированные технологии: Сб. науч. трудов. Вып. 32. - Харьков: Нац. аэрокосм. ун-т "ХАИ", 2006. – 150 с.

ISBN 966-662-143-6

Освещены вопросы аэродинамики, прочности, конструкции и технологии производства авиационной техники, методы интегрированного проектирования, подготовки производства, инженерного анализа с помощью систем CAD/CAM/CAE. Представлены методы обеспечения безопасности конструкций при длительной эксплуатации, системы информационного обеспечения жизненного цикла самолетов, системы обеспечения безопасности жизнедеятельности для летательных аппаратов, вопросы экономического анализа рисков, методы поддержки, принятия управленческих решений.

Для специалистов промышленных и научных организаций, преподавателей и студентов.

**Адрес редакционной коллегии: 61070, г. Харьков, ул. Чкалова, 17,
Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ»,
тел. (057) 707-43-63**

**Свидетельство о государственной регистрации КВ №7975 от 09.10.2003 г.
В сборнике публикуются статьи на украинском и русском языках.**

ISBN 966-662-143-6

**© Национальный аэрокосмический университет
им. Н.Е. Жуковского «ХАИ», 2006**

Содержание

| | |
|--|-----|
| Д.С. Кива, Б.П. Маслов, А.Н. Климюк. Концептуальный подход к определению параметров надежности агрегатов планера самолета из КМ | 5 |
| В.Д. Белый, А.Г. Гребеников, П.О. Науменко, В.В. Парфенюк, С.В. Удовиченко. Формирование мастер-геометрии беспилотных летательных аппаратов «Пчелка» | 14 |
| А.В. Наумов. Реализация концепции «Free Flight» на пассажирском региональном самолете | 26 |
| А. Н. Нарожный. Компьютерная поддержка процесса определения летно-технических характеристик самолета. Часть 4 | 32 |
| О.А. Коршець. Автоматизоване управління літаками в групі у просторі станів різницевих параметрів | 38 |
| Д.В. Ланде, В.М. Фурашев. Забезпечення вибіркового доступу до контенту мережі Інтернет як засобу оцінки ситуації та прийняття рішень | 47 |
| А.Г. Гребеников, В.Н. Клименко, А.Ю. Ефремов. Анализ влияния осевого и радиального натягов на податливость одноболтового односрезного соединения и распределение усилий между рядами многоболтового соединения из сплава ВТ6 | 55 |
| В.П. Максимов, В.Д. Белый. Система автоматизированного управления и связи беспилотных автоматизированных авиационных платформ семейства Пчелка» | 65 |
| С.М. Мельников. Дефекты формы ячейки сотового заполнителя, возникающие в процессе формообразования сотового пакета, и их регламентация | 69 |
| В.Ф. Сорокин. Создание штамповой оснастки в машиноориентированном производстве лопаток | 76 |
| А.А. Цирюк, М.А. Яровой. Алгоритм проектирования рациональных силовых конструкций многостеночных крыльев малого удлинения | 85 |
| В.В. Баранник, А.К. Юдин. Двухпризнаковое структурное двоичное полиадическое декодирование | 96 |
| М.С. Мельников. Автоматизация поиска технологического процесса с применением конструкторско-технологического кода детали | 101 |
| В. М. Илюшко, Д. В. Павлюк. Разработка методов построения командно-измерительного радиоканала для беспилотного летательного аппарата | 105 |
| А.Е. Агеев. Моделирование организационных структур и процессов управления рисками проекта | 110 |
| И.Б. Сироджа, И.Е. Россоха, С.В. Россоха. Интеллектуальная информационная технология поддержки принятия диспетчерских решений | 114 |