

УДК 621.941

**СИСТЕМА ПУБЛІКУВАННЯ/ПЕРЕДПЛАТИ
НА НАУКОВО-ТЕХНІЧНІ ІНФОРМАЦІЙНІ РЕСУРСИ**

*М.О. Алексєєв, доцент;
Ю.М. Молчанов, аспірант;
О.М. Алексєєв*, доцент*

*Інститут телекомунікаційних систем НТУУ “КПІ”, м. Київ
Сумський державний університет, м. Суми

В статье авторы рассматривают и обосновывают преимущества системы публикаций/подписки на научно-технические информационные ресурсы как наиболее перспективной для использования в корпоративных Интранет-сетях предприятий и в сети Интернет, которая не требует дополнительного программного обеспечения, специальных протоколов со сменным источником подписки.

У статті автори розглядають і обґрунтовують переваги системи публікації/передплати на науково-технічні інформаційні ресурси як найбільш перспективної для використання у корпоративних Інтранет-мережах підприємств і в мережі Інтернет, що не вимагає додаткового програмного забезпечення, спеціальних протоколів зі змінним джерелом передплати.

Із розвитком автоматизованих систем керування підприємством невтримно збільшується кількість інженерної та управлінської інформації, що перебуває в обігу як в таких системах, так і за їхніми межами. Крім того, все більше найновішої технічної інформації та її постачальників знаходиться у мережі Internet. Інформаційні системи автоматизованих систем керування підприємством на сьогодні можуть включати тисячі компонентів, бути розподіленими по всьому світу. При цьому розташування та призначення цих компонентів може значно змінюватися впродовж існування системи. Тому формування їхніх зв'язків у вигляді точка-точка чи синхронних комунікацій приводять до появи негнучких і статичних реалізацій, роблять розробку динамічних широкомасштабних систем громіздкою.

Для зменшення труднощів проєктувальників розподілених інформаційних систем керування підприємством зв'язок між компонентами в таких системах також повинен забезпечуватися спеціалізованою проміжною інфраструктурою, що базується на відповідній схемі зв'язку. Завдяки цьому схема автоматизованого отримання інженерної та науково-технічної інформації у вигляді взаємодії публікація/передплата звертає на себе все більше уваги. Вона становить схему автономної комунікації між учасниками обміну подіями в інформаційних системах, що надає потужний механізм для розподілених та мобільних архітектур [1]. Передплатники можуть виражати їх зацікавленість в події або зразку події і згодом бути оповіщеними про будь-яку подію, що була згенерована видавцем

(публікатором), яка відповідає їх певному інтересу. Подія асинхронно пересилається всім передплатникам, які заявили інтерес в цій події. Сильна сторона цього подіє-орієнтованого стилю полягає в повному розділенні в часі, просторі і синхронізації між тим, хто публікує, і передплатником. Багато промислових систем і дослідницьких прототипів підтримують даний стиль взаємодії, і відомі деякі дослідницькі розробки в галузі оригінальних схем взаємодії постачальник/передплатник.

Принцип взаємодії публікація/передплата

Схема взаємодії публікація/передплата, наведена на рисунку 1, складається з трьох учасників: постачальника або сторони, що публікує, передплатника або споживача і посередника.

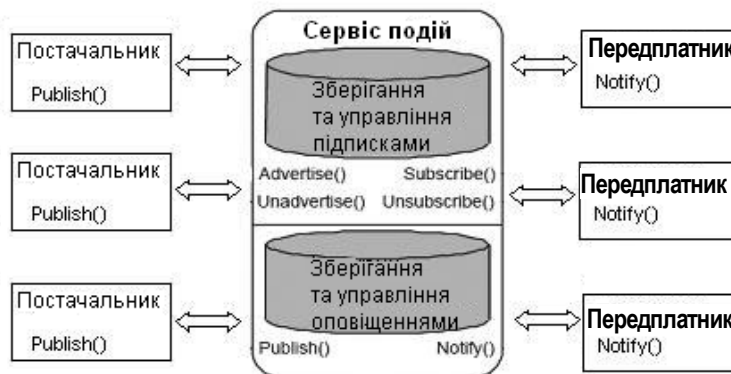


Рисунок 1 - Проста об'єктно-орієнтована система публікації/передплата

Постачальники публікують оповіщення, а споживачі підписуються на них за допомогою виданих підписок, які, по суті, становлять фільтри повідомлень без запам'ятовування стану. Споживачі можуть мати декілька активних підписок, а після того як передплатник видасть підписку, сервіс оповіщення доставить всі майбутні співпадаючі оповіщення, які були опубліковані будь-яким постачальником.

У загальному випадку передплатники реєструють інтерес у події, викликаючи операцію subscribe() в сервісі подій. Об'єкт їх інтересу передається як параметр. Зворотна операція unsubscribe() закінчує підписку. Для генерування події постачальник викликає операцію publish(). Сервіс подій спрямовує цю подію всім передплатникам, які проявили інтерес в ній. Постачальники також, як правило, можуть афішувати природу майбутніх подій за допомогою операції advertise().

На відміну від класичної моделі взаємодії запит/відповідь модель взаємодії публікація/підписка надає три рівні розділення [1]:

1 Просторове розділення наведено на рис. 2 – учасники взаємодії не потребують інформації один про одного. Публікація здійснюється через сервіс подій, а передплатники отримують ці події не безпосередньо, а через сервіс подій. Постачальники, як правило, не мають посилання на передплатників і не знають, скільки передплатників беруть участь у взаємодії. Також передплатники, як правило, не підтримують посилання на постачальників.

2 Часове розділення наведено на рис. 3 – учасники взаємодії не мають бути активні протягом взаємодії одночасно. Зокрема, постачальник може опублікувати деякі події тоді, коли передплатник відключений і навпаки.

3 Розділення синхронізації наведено на рис. 4 – постачальник може створювати подію у будь-який час, а передплатник може отримувати повідомлення асинхронно (через зворотний виклик).



Рисунок 2 - Просторове розділення

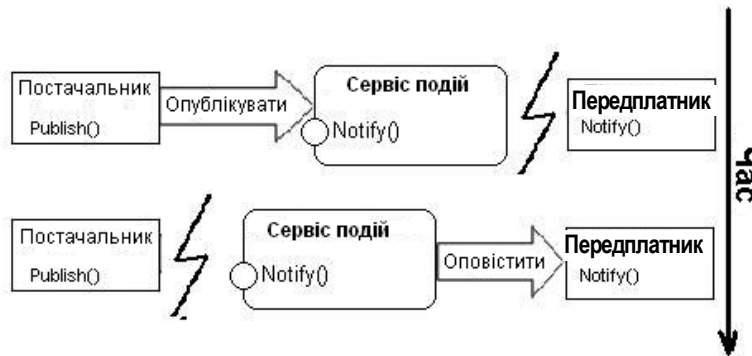


Рисунок 3 - Часове розділення

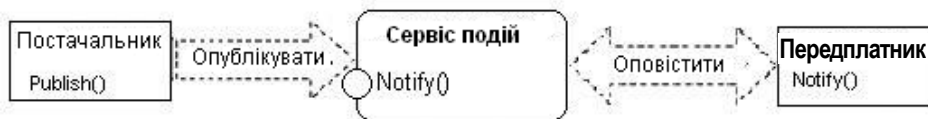


Рисунок 4 - Розділення синхронізації

Існуюча класифікація систем публікації/передплати

Серед систем публікації/передплати (СП/П) можна виділити 3 типи.

1 Темо-орієнтовані – успадковують концепцію каналів, які використовуються для зв'язку комунікаційних вузлів із методами класифікації вмісту подій [2]. Це означає, що постачальники можуть публікувати події у відповідних темах, а споживачі - підписуватися на них. Теми ідентифікуються за ключовими словами. Підписка на певні теми означає участь у групі з назвою даних тим. З'єднання з темою вимагає вибору імені у форматі URL. Теми організуються в дереві тем за допомогою точкової нотації (наприклад, wifi.радіозв'язок.телекомунікації), і передплатники можуть підписатися на окрему тему або на тему з її підтемами. Кожна тема ідентифікується в сервісі подій унікальним ім'ям. Коли в системі з'являється подія, вона асинхронно пересилається всім передплатникам по цьому URL. До таких систем належать Corona (Cornell Online News Aggregator), SCRIBE.

Основними перевагами темо-орієнтованих систем є простота реалізації, малі затримки, висока пропускна спроможність і простота переадресації

підписок великій кількості користувачів, що складаються з груп, через просту відповідність ключових слів іменам груп.

Серед недоліків можна виділити обмежені виразні здібності семантики, неможливість передплати на частину теми і, як наслідок, неефективне використання пропускової спроможності.

2 Вміст-орієнтовані – повідомлення про події порівнюються з тими, що зберігаються в базі даних передплати і відповідають зумовленому критерію відповідно до вмісту подій, а не ключового слова [3]. Кожна підписка ідентифікується визначенням зразка передплати, що необхідне для розповсюдження події. Зразок передплати є основою для визначення фільтру передплати. Фільтри вибирають відповідні події з опублікованої інформації з використанням мови передплати і реалізуються у формі пар ім'я-значення властивостей і базових операторів порівняння. Зразок передплати визначається одним із таких способів: рядки, об'єкти шаблонів, виконуваний код. Якщо система використовує для визначення зразків передплати рядки, фільтри повинні узгоджуватися з семантикою передплати. Можливе використання SQL, XPath як мови передплати.

Основною перевагою є підвищення виразності і, як наслідок, підвищена гнучкість механізму вибору сповіщень.

Недоліками є низька продуктивність, складна реалізація масштабованої системи. Найбільш характерним представником даної групи СП/П є система SIENA.

3 Тип-орієнтовані – оповіщення діляться відповідно до певних типів, які можуть інкапсулювати атрибути так само як і методи [4]. Передплатник указує тип події і визначає вираз фільтру, який оперує з атрибутами, що надаються цим типом події. Такий підхід робить тіснішою інтеграцію мови програмування і проміжного ПЗ. Головними перевагами таких систем є простота і гнучкість розподіленої реалізації за допомогою використання видаленої фільтрації вмісту, що дозволяє збільшити кількість можливих клієнтів і трафіку.

Класифікація, що пропонується

Розглянута вище класифікація може бути використана для опису СП/П з точки зору їх призначення, завдань виконуваними системами. Проте для опису СП/П з точки зору особливостей функціонування систем в глобальних телекомунікаційних мережах доцільно використовувати критерії:

- використовуюваного проміжного ПЗ - клієнтська або серверна частина, наприклад додаткові модулі броузера або сервера;
- використовуваних мережевих протоколів - HTTP, TCP/IP, WebDAV, SOAP, NNTP;
- джерел передплати, наприклад, БД, служби новин, RSS.

XPath базована система публікації/передплати

Авторами статті було розроблено дослідний макет системи публікування/передплати. Було прийнято рішення використати модульну побудову системи, що, як було зазначено вище, надасть системі переваги масштабованості та інтегрованості.

Таким чином, система складається з шести модулів:

1 *Форма*. Цей модуль розміщений на сервері у вигляді веб-форми та може бути доступний як з мережі, так і локально через URL. На цій формі користувач може обрати:

- URL документа, на який проводиться підписка (локальний або віддалений);
- вираз XPath, який ідентифікує інтерес користувача в певній частині документа;

– тип передплати – вибір передплати на зміни в документі або появу в документі деякої ключової фрази.

2 *Тригер*. Тригер є центральним модулем системи публікування/передплати та контролює та об'єднує роботу всіх інших модулів, при цьому постійно перебуваючи в оперативній пам'яті операційної системи.

3 *База даних*. База даних є стороннім модулем системи. Відповідає за зберігання підписок користувачів системи.

4 *Виявлення змін*. Як зазначалося вище, модуль відповідає за порівняння поточної версії документа з версією документа (або частини документа), що зберігається в базі даних.

5 *Адаптивний таймер*. Цей модуль відповідає за визначення інтервалу порівняння між поточною та версією документа, що зберігається.

6 *Зберігання XML*. Модуль, що відповідає за визначення частини XML документа, яку потрібно занести до бази даних.

Схема взаємодії модулів системи наведена на рисунку 5.

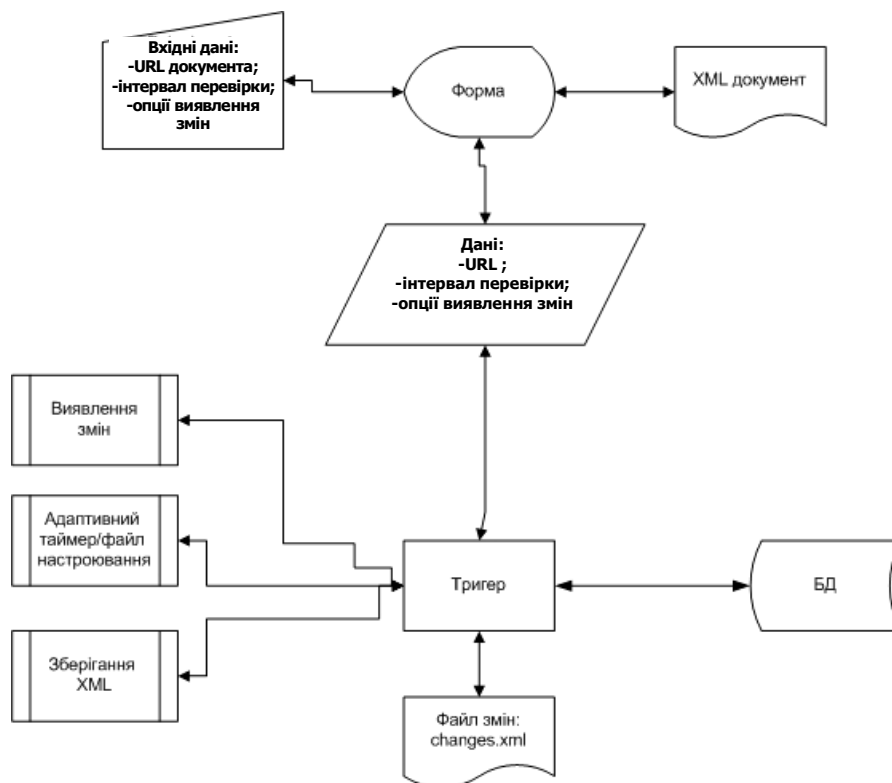


Рисунок 5 - Схема взаємодії модулів системи

Алгоритм роботи програми складається з таких етапів:

- 1 Після заповнення користувачем форми дані передаються у тригер.
- 2 Частина або весь оригінальний XML-файл заноситься до БД згідно з рішенням, прийнятим модулем зберігання XML.
- 3 До БД заноситься інформація про час наступного порівняння оригінального та поточного станів XML-документів.
- 4 Коли час порівняння збігається з поточним системним часом, з БД добувається оригінальний файл для порівняння та передається разом з поточним станом файлу в модуль виявлення змін.

5 Якщо зміни виявлено, модуль змін передає виявлені зміни як потік у тригеру, якщо ні, то в БД оновлюється інформація про наступний час перевірки.

6 Тригер формує файл змін та записує його на жорсткий диск.

Програмна реалізація системи публікування/передплати

Для програмної реалізації системи публікації/передплати було обрано середовище розробки Microsoft Visual Studio 2005, а як мови програмування С#. Оскільки дана мова програмування є об'єктно-орієнтованою, її використання дозволяє розробити модулі системи у вигляді окремих класів, що, у свою чергу, забезпечує модульність та інтегрованість даної системи. Як сторонній модуль БД було використано СУБД MS SQL Server 2005.

Процес передплати користувача на необхідний документ є дуже простим та складається з таких етапів:

1 Обирається необхідний локальний або віддалений ресурс у полі URL передплати та інтервал порівняння (у секундах) у полі інтервал перевірки основного вікна програми, що зображено на рисунку 6.

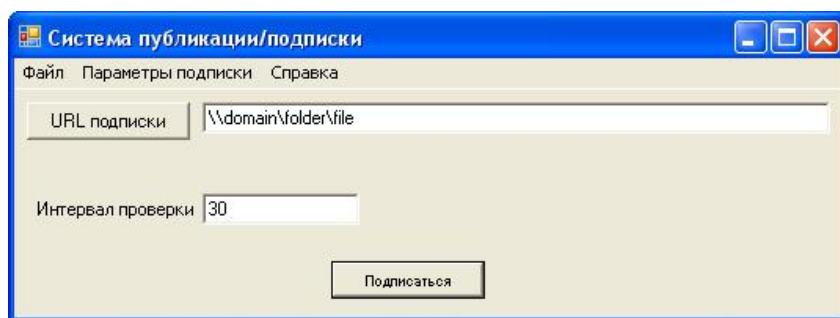


Рисунок 6 - Обрання необхідного ресурсу та інтервалу перевірки

2 Обираються необхідні параметри передплати та натискається кнопка «Подписаться», що зображено на рисунку 7. Серед обраних параметрів можуть бути: ігнорування порядку нащадків, інструкцій обробки, коментарів, декларацій XML, пробілів, DTD, простору імен, префіксів; порівняння фрагментів файлів; вибір необхідного алгоритму.

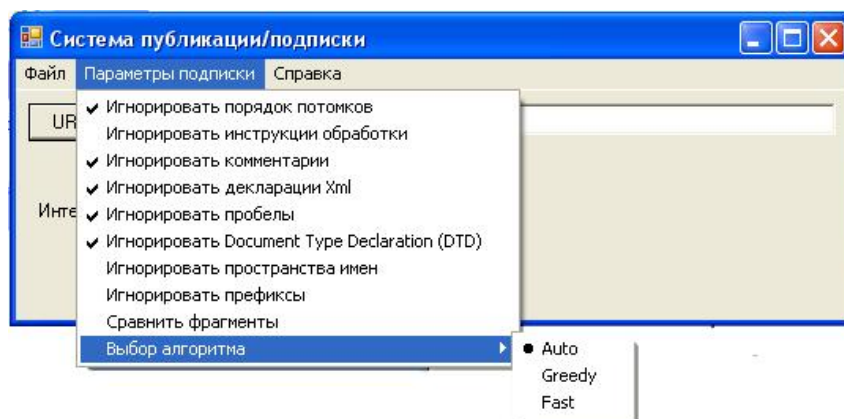


Рисунок 7-- Вибір параметрів передплати

3 Перегляд результатів роботи програми у файлі changes.xml.

Лістинг оригінального файлу origin.xml.

```
<examples>
  <example id="1">
    <name>change</name>
  </example>
  <example id="2">
    <name>add</name>
  </example>
  <example id="3">
    <name>remove</name>
  </example>
  <example id="4">
    <name>change id into 5</name>
  </example>
</examples>
```

Лістинг файлу changes.xml.

Идентификатор передплати:1260448572

URL передплати: <http://neytrin/Export/origin.xml>

Обнаруженные изменения:

```
<xd:xmldiff version="1.0" srcDocHash="1829959912340266426"
options="IgnoreChildOrder IgnoreComments IgnoreWhitespaces
IgnoreXmlDecl IgnoreDtd" fragments="no"
xmlns:xd="http://schemas.microsoft.com/xmltools/2002/xmldiff">
  <xd:node match="1">
    <xd:node match="1">
      <xd:node match="1">
        <xd:change match="1">Changed text</xd:change>
      </xd:node>
    </xd:node>
  <xd:node match="2">
    <xd:add>
      <title>added</title>
    </xd:add>
  </xd:node>
  <xd:node match="3">
    <xd:remove match="1" />
  </xd:node>
  <xd:node match="4">
    <xd:change match="@id">5</xd:change>
  </xd:node>
</xd:xmldiff>
```

Таким чином, можна зробити висновки:

– схема комунікацій публікування/передплата є перспективною для розвитку розподілених інформаційно-телекомунікаційних систем керування підприємством, оскільки, на відміну від класичної моделі взаємодії запит/відповідь надає три рівні розділення: у часі, у просторі та у синхронізації;

– існуюча класифікація систем публікування/передплати є недостатньою для класифікації веб-орієнтованих систем, отже розробникам доцільно використовувати розширену класифікацію;

– найбільш перспективною системою для використання в корпоративних Інтранет-мережах підприємств та мережі Інтернет є вміст-орієнтована система публікування/передплати, що не вимагає

додаткового ПЗ, не вимагає спеціальних протоколів із змінним джерелом передплати.

SUMMARY

PUBLICATION/SUBSCRIPTION SYSTEM ON SCIENTIFIC AND TECHNICAL INFORMATION RESOURCES

*M.O. Alekseyev, Yu.M. Molchanov, O.M. Alekseyev **

Sumy State University;

Institute of telecommunication systems NTUU “KPI”

In the article authors consider and prove advantages of system of publications/subscriptions on scientific and technical information resources as the most perspective for use in corporate Intranet-networks of enterprises and in the Internet, which does not demand the additional software, special protocols with a replaceable source of a subscription.

СПИСОК ЛІТЕРАТУРИ

1. Patrick Eugster, Pascal A. Felber, Rachid Guerraoui, Anne-Marie Kermarrec. The many faces of Publish / Subscribe // ACM Computing Surveys. - 2003. - Vol. 35, No 2. - P. 114-131.
2. Tova Millo, Tal Zur, Elad Verbin. Boosting topic-based publish-subscribe systems with dynamic clustering // International Conference on Management of Data. Proceedings of the 2007 ACM SIGMOD international conference on Management of data. - 2007. - P. 749 - 760.
3. Gianpaolo Cugola, Davide Frey, Amy L. Murphy, Gian Pietro Picco. Content-Based Routing for Publish-Subscribe on a Dynamic Topology: Concepts, Protocols, and Evaluation // ACM Journal Name. -2005. - Vol.V. - P. 1-51.
4. Patrick Eugster. Type-based publish/subscribe: Concepts and experiences // ACM, New York, USA. - January 2007. -Vol.29. - P. 32-46.

Надійшла до редакції 2 грудня 2008 р.