Comprehensive Study of RSS: A Simple Web Technology

Priyanka Bane[1], Jayesh Kudase[2]

B.E. Student, Dept. of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Maharashtra, India[1]

B.E. Student, Dept. of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Maharashtra, India[2]

ABSTRACT: As all of us as web users, we are interested in many websites whose content may be changing on an unpredictable schedule. It is very difficult to keep track of it. Because checking each website separately again and again to see for updates can be a very complex task. These websites could be news sites, community and religious information pages, medical websites and weblogs. This paper provides a comprehensive study and analysis of RSS web technology. RSS functions like a customizable, digital newspaper. RSS is used for simplifying, organizing and delivering web content. Various aspects of RSS web technology, its working and applications have been discussed.

KEYWORDS: Really Simple Syndication (RSS), Aggregator, Feed, Weblogs, EXtensible Markup Language (XML)

I. INTRODUCTION

Web surfing is considered as fun but it wastes a lot of time if one wishes to visit every single website or blog he enjoys. Wouldn't it be amazing if we could get notified with the headlines of latest stories which comes from our favourite websites, blogs or even channels. For these, RSS comes into picture. It provides a way by which we get online notifications of a list of current story headlines of our interested sites.

It stands for "Really simple syndication". It allows personalized views for various sites/blogs. It is one of the technologies which is being used by billions of web users around the globe to keep track of their favourite websites. Basically it is that technology which gives you up-to-date information about sites of your choice so that you can read it in your free time or whenever possible.

RSS are kind of live bookmarks. In earlier days of internet, when we wish to see updates of our favorite website we need to bookmark that particular website on our browser and then manually go to those websites on a regular basis to check what new has been added to it or uploaded on it. Here you as a web user had to do all these work. If you have bookmarked too many websites then it gets complicated to keep track of each and every website manually. There are chances of you missing some useful information or update when you do not remember to check the sites you have bookmarked previously. This leads to going through the same information again and again on sites that do not update frequently.

RSS changesthese scenarios completely. Suppose, what if you can tell your favorite website to ping you every time when an update comes. In that case, RSS will come to your rescue. The limitation with using RSS as live bookmarks is that you can only have access to your RSS list from the computer/laptop you created the bookmarks on. If we want the flexibility of accessibility then we can make use of RSS Reader route.

II. BACKGROUND

A. *What is RSS?*

RSS – a technical abbreviation of "Really Simple Syndication" is a XML formatted plain text to organize news and one of the ways to share the web content. It ensures compatibility and reliability which means the format is easy to read and understand both by automated machines and by humans. It is a mechanism for delivering frequently updated web content. As we have described above, it follows XML format because as soon as the updated content is retrieved, RSS reader makes use of XML structure to give an organized display for the end users. Hence we can say that it is a lightweight, multipurpose, extensible metadata description and syndication format.

B. *History of RSS*
- 1997: ScriptingNews was developed by Dave Winer at UserLand. RSS was born.
- 1999: Netscape develops RSS 0.90. It supported ScriptingNews.
- 1999: ScriptingNews 2.0b1 was developed by Dave Winer. It included RSS 0.90 features.
- 1999: Netscape developed RSS 0.91. Many features of ScriptingNews 2.0b1 were included in RSS 0.91.
- 1999: UserLand uses only RSS 0.91.
- 1999:Netscape stops their RSS development.
- 2000: UserLand releases the official RSS 0.91 specifications.
- 2000: O'Reily develops RSS 1.0. This format uses RDF and namespaces.
- 2000: Dave Winer at UserLand develops RSS 0.92.
- 2000: After leaving UserLand, Dave Winer develops RSS 2.0.
- 2003: The official RSS 2.0 specification is released.

RSS was developed using World Wide Web Consortium RDF (Resource Description Framework) standard. There is no official standard for RSS.

III. OVERVIEW OF RSS

A. *How RSS Works?*

RSS system to syndicate news and information over the web is very simple. It is as follows:
1) There are some web pages user wants to get displayed by other website. This group of web pages is known as the RSS feed.
2) This RSS feeds will be defined by XML file. This file contains title, URL and description for each page that will be shown to the user.
3) A web user who is interested in reading those feeds on its computer will make use of an RSS reader or a browser.
4) Other website that wants to display the feed has to load the RSS file from the provider to extract URL of pages and display title and summary. This can be done with the help of PHP script.
5) At the time user visits the website of the receiver, the script is launched, it brings the RSS file from the providers website and displays a list of news from the information that is been extracted.
6) By clicking on the line of the list, respective page from the provider will be shown to the user.

B. *What Information does RSS provide?*

RSS is made up of a list of items presented in an order from starting with newest to oldest. Each of these items is composed of a title along with description of that item plus a link to a web page with the actual content described over there. This description can be summary or detailed information. It is not necessary that web page will be always linked with the item. Sometimes title and description may contain all the requested information.

C. *What does RSS Include?*

RSS includes the following:

1) *RSS Reader*: It is a software which is called as "RSS aggregators/RSS channels/Feed reader" which can be desktop based or mobile-device based, provides RSS feed information to respective users. It reacts to feed changes in an appropriate way. It is all about simplifying the way we view and display the information on web. Hence it is a program that gathers and sorts out RSS feeds. Users can subscribe to feeds of their choice by putting an URL of a feed into the RSS Reader or just by clicking on the feed icon of the browser.RSS Reader runs in the background and regularly checks and monitors the feeds of a particular website or blog that the user has subscribed to for update of any new content. When something new pops up on that site, it instantaneously sends you. It aggregates the content for users so that they can view it in one place hence the name RSS aggregators.

RSS readers are categorized as follows:
- Could be Browser based interface.

Example: Tiny Tiny RSS: It is suitable if one wants to host his own feeds.

- Can be dedicated apps or soft wares.
  Example:PRESS: It is a paid app. It gives best visual environment for its users as compared to all other readers.
- Can be accessible from both browser and software
  Example:FEEDLY: It provides both browser and a software based interface to users. It has hosting services and many hard-to-find features in paid version.

2) *RSS Modules:*In the context of RSS, it is a strongly bonded vocabulary established via XML namespaces to give names to concepts and relationships among the concepts. It is similar to XML schema.As it is similar to XML schema, primary objective of RSS modules is to extend the basic XML schema which could provide more powerful content organization.

3) *RSS feed:*RSS works on a website in which author of that website maintains a list of notifications in a structured way. This list of changes is known as "RSS feed". It has been automatically accessed by the RSS readers or we can say RSS aggregators on user's behalf and it syndicates the result for the user. Basically RSS feed of a page cuts the information down to essentials like main text content, main images.This feature is been used by thousands of websites now including major organizations like The New York Times, The BBC, Reuters and several weblogs.
   Types of RSS feed:
   - Convertible URLs: In URL based RSS, it is essential that you manipulate awebsite's URL which can get converted into an RSS feed. By appending "/RSS" or ".RSS" at the end of a URL will make it into an RSS feed.
     Example: One of the recognized sites Reddit allow its users to modify URL into RSS feed.On the other hand, our well-known sites YouTube and Facebook requires more advanced modifications to a URL in order to function it as an RSS feed.
   - Two more types are FeedBurner based and Direct XML links.

D. *RSS Document*
   A sample RSS document is as follows:

```
<? xml version="1.0"?>
<rss version="2.0">
      <channel>
              <title>…</title>
              <link>…</link>
              <description>…</description>
              <item>
                      <title>…</title>
                      <link>…</link>
                      <description>…</description>
              </item>
      </channel>
</rss>
```

The first line defines the XML version. Second lines identifies that this is an RSS document.
<channel>: describes the RSS feed and has three child elements as follows:

- <title>: defines title of the channel
- <link>: defines hyperlink to the channel
- <description>: describes the channel , summary of the article in about 200 characters

Additional child elements, parent element "channel" can have:

- <category>: It is used to specify category of the RSS feed. RSS aggregator makes use of it to group websites based on category.
- <copyright>: It is used to notify about copyrighted material.
- <image>: It is used to display an image while presenting a RSS feed and is optional.Its sub child elements include:
  <url> : It gives address of the image
  <title> : Text to be shown if the image is not displayed.
  <link> : gives hyperlink to a site that provides the channel
- <language>: It is used to mention the language in which RSS document will be written. It can be used to group the sites on the basis of the language.

Additional child elements, parent element "item" can have:

- <author>: It is used to provide e-mail address of the author of an item.
- <comments>: It is used to give comments about a particular item.
- <enclosure>: It is used to include a media file associated with the item.Its sub child elements include:
  <url>: specifies URL to media file
  <length>: gives length of media file
  <type>: gives file type of media

E. *Relationship of RSS and XML*

As we know, XML is a standard that helps in displaying textual information in a structured way. It is a text based format using tags used to store information in an organized way in a document.Whereas RSS follows that XML standard but goes further by defining names for those specific tags and what type of information should be stored in them. Therefore, we can say that RSS is rendered with the help of XML. This meant whenever you will request for an RSS feed, you will be provided with an XML document which conforms to RSS structure.

IV. ADVANTAGES AND APPLICATIONS

A. *Benefits of Using RSS*

The benefits of using RSS are as follows:

- Resolves problem with Email notifications: Notifying the users via an email, when any new information gets available on your favourite website was an earlier solution. But when you receive email notifications from several websites that information will be in an unorganized form and can be often mistaken as a spam. But RSS is better solution as it notifies the user with changes in an organized way and it handles the content easily. It gives 100% control over the situation as it does not let us reveal our email address, since there is no email address involved. There is no way a publisher can sell, rent or give away the means to contact you. This avoids virtues, phishing or any identity theft. Thus, RSS assures privacy.
- RSS is like you are subscribing to your favourite magazine that comes to your house each month. Instead of this, it is delivered to your RSS reader every time your favourite website gets added with new information. It lets you know about the updated parts from your website. This removes the need for the user to manually go to a website and check for any new content on it. Instead of this, browser having RSS enabled, monitors the sites on a regular basis and informs the user if there are any new updates. The web browser can also give the orders to automatically download the new data or information for that user.
- RSS enables publishers to combine data automatically. It is useful for those users who wish to receive timely updates from their favourite websites. It easily distributes latest news headlines about social affairs, politics, entertainment, education etc. sometimes to a group of people.

- One of the main advantage is that user can create their own channels and publish them on the internet.

B. *Applications of RSS*
- Notifying the user about the arrival of a new product in a particular store.
- Alerts about changing weather conditions.
- Notifications of the insertions of new item to a database or addition of new members to a group.
- Listing and notifying email newsletter issues.
- Used for the propagation of software update lists.

V. CONCLUSION

RSS is a way to syndicate and consume the web with increased efficiency and ease of use. It was supposed to be the very popular and useful means for communication at earlier times. You can view it as a newspaper with the advantage to view only the headlines. Also it allows busy people to receive updates to web-based content of their interest. It provides fast browsing news and offers instantaneous updates. Further it helps us to enhance user's information gathering capabilities. Unfortunately, the big data companies such as Google don't like RSS. They invented and then killed the best browser-based RSS client of the time – Google Reader. While some sites such as Reddit embrace RSS, others are trying to deprecate it. Taking everything into consideration we can conclude that it's a must-have software for web readers.

REFERENCES

[1] ProBlogger Contributors, Available at http://www.problogger.net/what-is-rss/
[2] Blog "CravingIdeas" Contributors, Available at http://cravingideas.blogs.com/backinskinnyjeans/2006/09/how_to_explain_.html
[3] Wikipedia contributors, RSS.
[4] SoftwareGarden Contributors, Available at http://rss.softwaregarden.com/aboutrss.html
[5] DigitalTrends Contributors, Available at http://www.digitaltrends.com/computing/how-to-use-rss/#:YN2Nw_tFFacebook
[6] W3Schools Contributors, Available at http://www.w3schools.com/xml/xml_rss.asp
[7] MakeUsOf Contributors, Available at http://www.makeuseof.com/tag/what-is-rss-and-how-can-it-improve-your-life/
[8] https://tools.ietf.org/id/draft-nottingham-rss-media-type-00.txt