Concepts and Implementation of a Semantic Web Archiving and Simulation System for RF Propagation Measurements

Vinodh K. Rajendran, James N. Murdock, Adrian Duran, Theodore S. Rappaport
Wireless Networking and Communications Group (WNCG)
University of Texas at Austin, Austin, TX, USA 78712
vinodh@cs.utexas.edu, jmurdock@mail.utexas.edu, aduran@mail.utexas.edu, wireless@mail.utexas.edu

Abstract

In this paper, we present an Open-Source web-based archiving system to organize and share wireless RF propagation measurement data, models, and simulation software in a centralized, standardized archive. This archiving system is based on Semantic Web ideas that will enable the wireless research community to easily share and access measured data and simulators provided by researchers across the globe. To begin development of the web-based archiving environment, we use a previously developed RF propagation simulator, SIRCIM [6], to represent the range of values, types of measurements, and file format types that would be needed to properly archive measurements from the research community at large. This paper also explores development issues and considerations required to build a Semantic Web on-line propagation channel measurement and modeling archiving system for global use.

Keywords- Semantic Web, Wireless Measurements, SIRCIM, SMRCIM, Simulation, Propagation models, Cloud Computing

I. Introduction

Many fields of research, including astronomy, biology, medicine and power systems, make extensive use of online archiving technologies to organize and share critical data [1],[2],[3]. A semantic web is a linked data-structure that enables machines to understand the semantics, or meaning, of information on the World Wide Web. It extends the network of hyperlinked human-readable web pages by inserting machine-readable metadata and how they are related to each other. This enables cross-referencing and clearer data representation, with high searchability. With continued growth of wireless technologies, and the growing interest in unexplored frequencies and applications (such as for emerging vehicle-to-vehicle, or sensor communication systems, and new propagation measurements at or above 6 GHz), the wireless research community would greatly benefit from the use of semantic web technologies to organize and share wireless measurement and modeling data in a collaborative on-line repository. These online collaborative tools will help researchers and practitioners standardize models that can be easily transported across platforms, to aid measurement model development and comparison.

II. Propagation Measurements, Models, and Sharing

The wireless communications research community requires accurate measurement data and channel models to develop and standardize future wireless networks. For future technologies to be successful, researchers and product developers must understand wireless propagation at particular frequencies, and in many new environments. Today, the state-of-the-art in wireless device and infrastructure design requires separate research teams to each make their own measurements of these new frequencies and environments. Tremendous amounts of money and time are spent by researchers replicating many of the same kinds of measurements, in order to develop specific knowledge about a prospective new wireless opportunity.

Unfortunately, most researchers in wireless propagation use a “show and throw” approach, in which the carefully measured channel data are discarded after analysis and publication. Despite the enormous amount of time required to design and conduct propagation measurements, once analyzed, the raw data are often discarded or left in the heap of unorganized university or corporate storage and are not available to the research community at large, let alone to the same or other researchers at the originating institution. There are many reasons why measurement data are often discarded. First, measurements are often specifically designed for a particular application or frequency band, and once the desired knowledge is derived, the measurements are no longer needed or perceived as valuable by the researcher. Unfortunately, this approach precludes other researchers from using the data to develop new knowledge or test new concepts not related to the specific goals or application of the original research. Secondly, there are a wide range of file formats and processing techniques used by propagation researchers, and standards do not exist for such measurements, making it impossible for third parties to verify the accuracy or utility of the data, even if it were available. Thirdly, significant time and effort is required to setup and maintain file systems and servers to make these data available online. The IT support required to maintain a database is often expensive, and researchers are often not able to afford additional costs to share raw data.

978-1-4244-8327-3/11/$26.00 ©2011 IEEE

Existing standardization organizations, such as the IEEE 802 wireless community and the COoperation européenne dans le domaine de la recherche Scientifique et Technique (COST), spend tremendous amounts of time and money debating and determining channel models for use in standardization activities. The COST Forum already offers some measurement data online, through projects such as COST273 [4] and COST2100[5]. The measurement and model data from COST projects are available on request as separate files in different formats. Presently, however, COST does not offer a single on-line location where all of its data are available to researchers at large. Furthermore, the wireless community does not have a web-based environment that allows various measurement databases to be archived and processed in a uniform manner, with a standardized file format to allow new measurements to be contributed to the archive. This would allow data to be processed for representation with various propagation models, or checked for suitability or accuracy for various applications.

III. Archiving Framework and Standardized Data Formats – Starting With SIRCIM And SMRCIM

This paper develops a framework that will allow a wide range of propagation measurement data to be systematically indexed, sorted, archived, processed, and made available to everyone using an open-source web portal. The raw measurement data will be processed to be available in standardized data formats for easy portability across different simulation platforms. The website will permit file downloads and uploads, and will use integrated visualization tools to render high-quality plots and graphs to allow users to visualize and compare measurements and models.

To develop a prototype of the semantic web environment described above, we started with existing channel simulation tools SIRCIM [6], [7] and SMRCIM [8]. These simulators are used as data generators to emulate the raw data that could be produced by various measurement campaigns and which might be submitted to the open-source web archiving system. These open-source computer simulation programs run under MATLAB and generate Channel Impulse responses and a myriad of other channel data that are typically measured or modeled.

SIRCIM (Simulation of Indoor Radio Channel Impulse Response Models) and SMRCIM (Simulation of Mobile Radio Channel Impulse Response Models) are open-source RF propagation simulators developed by Virginia Tech Intellectual Properties, Inc. and Wireless Valley Communications, Inc., that take into account different physical channel scenarios. They provide narrowband and wideband simulated data of realistic channel impulse responses based on underlying statistical models collected empirically and through the literature [10] The channel statistics generated by the simulators have been found to be consistent with measured data [7]. SIRCIM and SMRCIM have been in use for about 20 years by over 100 companies and universities, for a wide range of indoor and outdoor products, standards, and investigations [9],[10],[11].

SIRCIM simulates channel responses for a wideband signal received by a stationary or mobile receiver operating from 10 MHz to 60 GHz in a user specified environment. In a typical SIRCIM simulation run, the user selects the frequency, Environment (amount of environmental clutter and floor plan), transmitter receiver separation distance, and velocity of the receiver. SIRCIM then computes an impulse response for 64 equally spaced positions on a track that is 10 wavelengths long at the receiver, in addition to narrowband pathloss for each point between the transmitter and receiver. The user may specify the type of simulation output, such as raw impulse responses or narrowband (CW) fading. Simulation runs generate wideband data for each multipath component, (e.g. Angle of Arrival or AOA data), and also provide narrowband (CW) data including fading voltages, fading distributions, and CDFs of RMS delay spread.

SIRCIM and SMRCIM have powerful visualizing tools based on MATLAB. This integrated visualization tool is capable of generating high-quality customized plots like those in Figures 1, 2, and 3. Any of the plots can be regenerated without the need to rerun the entire simulation. Figure 1 shows a typical SIRCIM simulation window with the impulse response plot for a 60GHz channel in a soft partitioned building. The plot shows the impulse response for a receiver as it is moved over the ten wavelength long track. SIRCIM and SMRCIM provide data files that represent the amplitude, phase, time delay, and angle of arrival of each multipath component within each impulse response shown in Figure 1. Figure 2 is a scatter-plot of path loss vs. transmitter-receiver separation distance for the case of 1000 simulation runs in the same 60 GHz channel. Figure 3 shows the angle-of-arrival of different multipath components for the simulated channel.

The attraction of using SIRCIM and SMRCIM as a first step in the development of a web-archive is that they produce simulated measurements in common pre-defined file formats shown in Table 1. Each of these types of files is generated for every simulation. SIRCIM and SMRCIM are strong starting points for the archive project because other researchers could use these formats to provide data to the archive, and well-defined file formats are critical for any application in the semantic web. For example, for each simulation SIRCIM creates HD1, HD2, BIN, AOA, PHS, CDF, WPL, and NPL files. The Header HD1 and HD2 files are short ASCII files used to encapsulate important simulation parameters. BIN files store impulse responses represented with 64 equally spaced time bins (i.e. data points) for excess delays from 0 to 500 nSec for wideband simulations. Each local maximum in an impulse response in a BIN file corresponds to a multipath component in the simulation. AOA files contain angle-of-arrival results (in radians) for each of these multipath components, and PHS files provide phases information (in radians) for each multipath component. Continuous Wave (CW) files are ASCII files that contain narrowband signal strengths over the ten-wavelength long simulation track. Cumulative distribution (CDF) files contain information about RMS delay spreads of the impulse response data. WPL and NPL files provide wideband and narrowband path loss data, respectively.

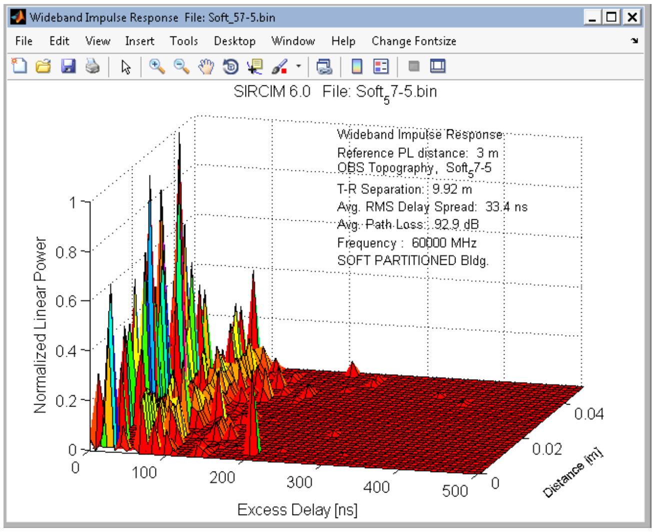

Figure 1 Typical SIRCIM simulation window – with impulse response

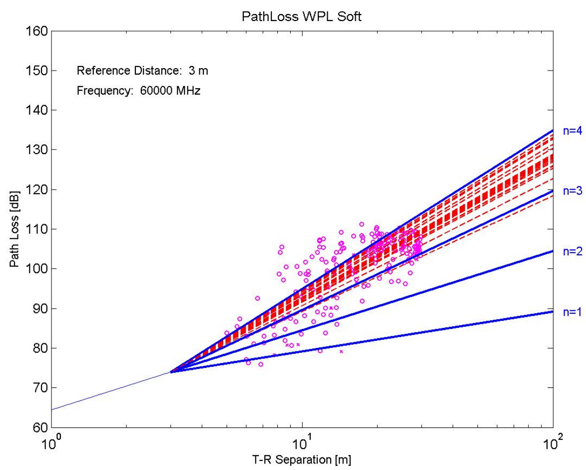

Figure 2 Path loss scatter plot vs T-R separation for channels with soft partitions.

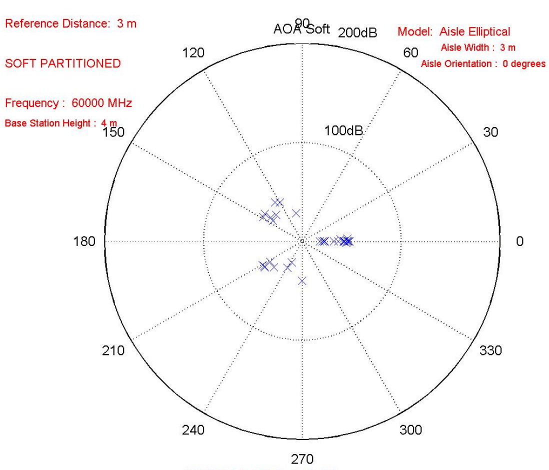

Figure 3 Angle of arrival for soft partitioned channels.

File Type	Extension	Storage
Wideband channel impulse responses	BIN	Binary
Multipath component phases	PHS	Binary
Angle-of-arrival of multipath	AOA	Binary
Small scale narrowband signals	CW	Ascii
Channel header1 files	HD1	Ascii
Channel header2 files	HD2	Ascii
RMS delay spread values	RMS	Ascii
CDF of RMS delay spread	CDF	Ascii
Path loss values of wideband signals	WPL	Ascii
Path loss values of CW signals	NPL	Ascii
Impulse noise data	NSE	Binary
Impulse noise data	DIS	Ascii
Impulse noise header3 files	HD3	Ascii
Impulse noise data not below threshold	DAT	Ascii
Random impulse noise data	RND	Ascii
Repetitive impulse noise data	REP	Ascii
Program data	MAT	Binary

TABLE I SIRCIM/SMRCIM OUTPUT FILE TYPES AND EXTENSIONS (FROM [7][8])

PLOT	Filetype
Wideband impulse response screen	.BIN
Angle-of-arrival plots	.AOA
Polar plot of aoa Vs path loss	.AOA
Aoa Vs rms delay	.AOA
Rms delay spread cumulative distribution	.CDF
Narrowband envelope and phase	.CW
Narrowband cumulative distribution	.CW
Wideband path loss scatter plot	.WPL
Narrowband path loss scatter plot	.NPL
Noise waveform plot	.NSE
Noise probability distributions plot	.NSE

TABLE II PLOTS GENERATED BY SIRCIM AND SMRCIM AND FILE TYPES

We will use the SIRCIM file types to create the first version of the open-source, Semantic Web, propagation measurement and model archive. Raw measurement data uploaded by any user will be processed offline to generate similar separate file types as shown in Table 1. These files will then be associated and linked to raw-measurement data in the archive.

The separation of different parameters in separate files serves multiple purposes. First, if a user is interested only in some specific information (eg. angle-of-arrival data) from a particular measurement data set, he or she does not need to download the entire measurement archive. Second, the availability of preprocessed information, such as angle-of-arrival and phase components, enables the user to consume the data directly without the need for any post-processing. Third, these files are self-contained and can be ported across multiple simulation platforms without additional control or header files. Seeding the web-archive with simulation data generated by SIRCIM and SMRCIM will provide a good starting point for building the foundation of a more extensive channel model and measurement archive application.

IV. Online Simulation and Visualization

SIRCIM and SMRCIM simulators will be made accessible online as part of the open-source web-archive. Users will be able to generate MATLAB-quality plots by accessing the archive without any specialized or proprietary software other than a web browser. The on-line archive will make it possible to compare simulation results with existing models, and new simulation or measurement results may be stored as new data in the archive.

Open-source online simulation can be of great use in educating students about wireless technology by giving them access to real-world simulation or measurement data. With SIRCIM and SMRCIM, students will be able to tweak different parameters in the simulation model and visualize the effects in graphical plots.

V. Semantic web, Scalability and Cloud Computing

One of the major technologies in building the semantic web is Resource Description Frameworks (RDF) encoded in XML. Formulating an RDF begins with a map of the structure to be described, such as shown in Figure 4, to describe a wireless channel (based on the explanation in [12]). The map describes a general propagation model with statements that relate a subject (e.g. pathloss) to an object (e.g. value of 2) using predicate logic (e.g. pathloss has a value of 2). These statements are written into XML for interpretation and automatic computer management. Describing channel measurements or models in a way that can be easily and automatically managed is the first step to creating an on-line semantic web application for wireless researchers. Once this is available, these software representations can be queried for specific information by the research community.

As a first step the archive and the simulator will be hosted in a university server with limited access to interested participants. As the project gains more popularity and exceeds the bandwidth and computational limits of the server, the archive will be transitioned to a shared cloud-computing environment such as Amazon's EC2 [13] and Simple Storage Service (S3) platforms. Cloud computing solutions offer practically unlimited computing resources and unbounded storage space with periodic backups. A user could be directed to the cloud-based web repository through a link on the IEEE Vehicular Technology Society (VTS) website to access models or measurement data. One can simulate using SIRCIM with a specifically designed interface, which invokes the simulation binaries remotely in the cloud (S3). The simulation results or the measurement data can be downloaded along with the post-processed files (as in Table 1). Graphical plots and other visualizations can be generated for these measurement or simulated data and downloaded as image files (in jpeg or png format). Dynamic plots with zoom, pan and rotate features can be viewed online.

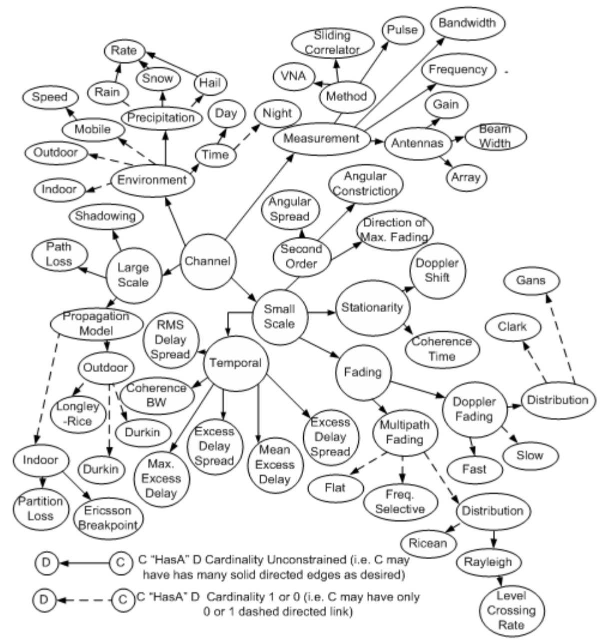

Figure 4 This map serves as the basis for creating a Resource Description Framework (RDF) to model a channel. [14]

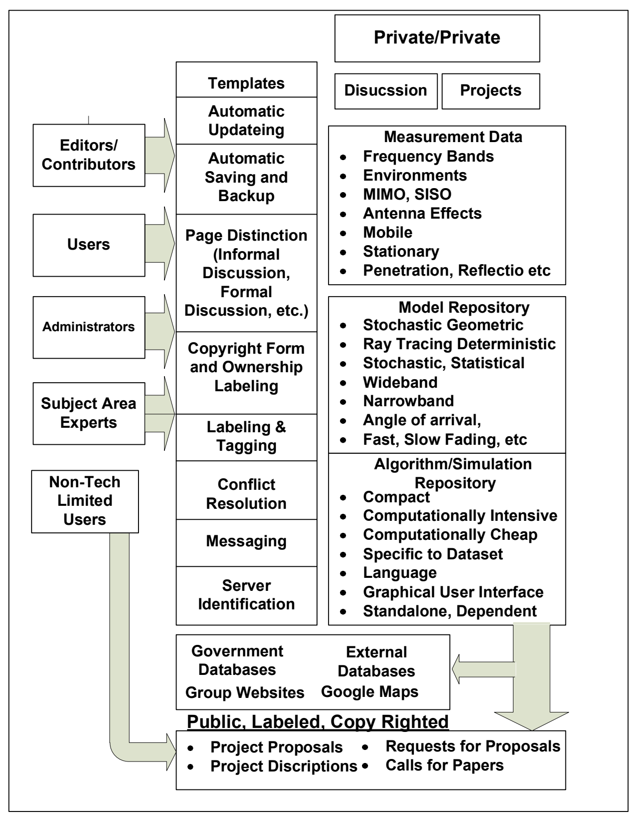

Figure 5 This figure illustrates the structure of the proposed on-line archiving tool.

VI. Challenges

The primary challenge in this kind of data management solution is to ensure data authenticity and to validate every user input. To guarantee secure data transfer into and out of the web-archive, we propose the hierarchical structure shown in Figure 5. In this collaborative environment, users would have different roles according to their requirements and responsibilities. Simple users would have read-only access to the tool, giving them ability to access models and download simulated data and visualizations. Editors and contributors would have higher privileges to edit and add new measurement data sets and simulation code. Administrators will monitor the archive for any security holes or malicious code and maintain the general health of the archive and oversee the functioning of the tool.

A simple way to authenticate the data is to have other users review the proposed data and approve it to be valid and authentic. A more comprehensive authenticating scheme is to have a group of subject matter experts, reviewing each measurement data set that is uploaded, or to have experts serve as editors or administrators who mediate or offer suggestions. Having a set of computer-verified rules and test simulation runs, that demonstrate data set correctness would reduce human error and ensure correctness of new contributions.

VII. Conclusion

This open-source web repository will enable researchers to pool the data sets and share measurements effectively. One can extract additional value from data that may have been collected for other purposes. It will also aid in the standardization of the formats of the measured data and allow the comparison of models derived from certain data sets. With developments in semantic web technologies, the capabilities of such repositories will greatly improve in coming years. There are challenges in creating the repository: enticing researchers to use and contribute to the archive, maintaining data integrity with authentication, and developing a versatile and useful portal that is credible and useful. However, despite these challenges, the likely value to the research community from this central repository would be great, and could justify the effort required for setup and maintenance.

VIII. Acknowledgements

This project is sponsored by the IEEE Vehicular Technology Society, and the IEEE VTS Propagation Committee.

References

[1] Detter, R.; Mooney, M.; Fatig, C.C.; , "XML - James Webb space telescope database issues, lessons, and status," Aerospace Conference, 2004. Proceedings. 2004 IEEE , vol.5, no., pp. 3312 Vol.5, 6-13 March 2004

[2] Hunter, P.J.; , "Modeling Human Physiology: The IUPS/EMBS Physiome Project," Proceedings of the IEEE , vol.94, no.4, pp.678-691, April 2006

[3] Field, L.; Andreozzi, S.; Konya, B.; , "Grid Information System Interoperability: The Need For A Common Information Model," eScience, 2008. eScience '08. IEEE Fourth International Conference on , vol., no., pp.501-507, 7-12 Dec. 2008

[4] COST273, http://www.lx.it.pt/cost273/

[5] COST2100, http://www.cost2100.org

[6] Theodore S. Rappaport, Scott Y. Seidel, and Koichiro Takamizawa. Statistical channel impulse response models for factory and open plan building radio communication system design. IEEE Transactions on Communications, 39(5):794-807, May 1991.

[7] S. Seidel and T.S. Rappaport, A User's Manual for SIRCIM (Version 1.0). Virginia Polytechnic Institute and State University, Blacksburg, VA, odober 19,1990.

[8] Wireless Valley Communications, Inc., SMRCIM Plus 4.0 (Simulation of Mobile Radio Channel Impulse Response Models) User's Manual, Aug. 1999.

[9] Xiaolei Shi, Mario Hernan Castaneda Garcia, Guido Stromberg: Converting SIRCIM Indoor Channel Model into SNR-Based Channel Model. ICN (1) 2005: 231-238.

[10] J. E. Nuckols, Implementation of Geometrically Based Single-Bounce Models for Simulation of Angle-of-Arrival of Multipath Delay Components in the Wireless Channel Simulation Tools, SMRCIM and SIRCIM,Master thesis in Electrical & Computer Engineering, Virginia Polytechnic Institute and State University, 1999.

[11] Benvenuto, N.; Salloum, A.; Tomba, L.; , "An all digital receiver for DECT radios," Communications, 1996. ICC 96, Conference Record, Converging Technologies for Tomorrow's Applications. 1996 IEEE International Conference on , vol.2, no., pp.862-866 vol.2, 23-27 Jun 1996

[12] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall. 2002

[13] Amazon Web Services http://aws.amazon.com/

[14] T.S. Rappaport, J. N. Murdock, D. G. Michelson, and R.Shapiro, "An open-source archiving system for RF propagation measurements and radio channel modeling and simulation," to appear in *IEEE Vehicular Technology Magazine,* June 2011.

[15] MATLAB application deployment http://www.mathworks.com/help/toolbox/compiler/f12-999353.html