JournalofCurrentDevelopmentin Artificial Intelligence

September – December 2025 Vol – 13 Issue – 3

ENRICHED PUBLICATIONS PVT. LTD

S-9, IIndFLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR-5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-47026006

JournalofCurrentDevelopmentin Artificial Intelligence

Aims and Scope

Journal of Current Development in Artificial telligence is a Journal addresses concerns in applied research and applications of artificial intelligence (AI). the journal also acts as a medium for exchanging ideas and thoughts about impacts of AI research. Articles highlig t advances in uses of AI systems for solving tasks in management, industry, engineering, administration, and edu aationvaluauations oxisxisting AI systems and tools, emphasizing comparati e studies and user experiences and the economic, social, and cultural impacts of AI. Papers on key applications, highlig hting methods, time schedules, son months needed, and other relevant material are welcome.

JournalofCurrentDevelopmentin Artificial Intelligence

JournalofCurrentDevelopmentin Artificial Intelligence

(Volume - 5, Issue -1, 2017)

Content

Sep-Dec2025Vol–13Issue-3

AReview of Artificial NeuralNetworks
TheNextMajor"Advanced"Technology

KautilyaPachorieandAbhinavParashar
Electronics & Communication Engineering
Faculty of Engineering & Technology, Agra College, Agra.

Abstract
This document provides us the information about h ow the Artificial Neural Networks are acting as a brain model, which promises a less echnical way to develop machine solutions and pr vides a more gracceful degradation duringystem overload than its more traditional ounterparts.

Keywords—*ANN(Artificial Neural tworks), Neurons, P Pern RecRecognition, Neural System, Feedback Network, Perceptrons, MCP(More Complicated Network).*

1. Introduction

It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. [1]The firsrst artificial neuron was produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits. An ANN is configured for a specific applicacation, such as patern reccognition or ta classificaation, through a learning process. Learning in biological systems involves adjustments to the synaptic connections th t exist between the neurons. Even simple animal brains are capable of functions that are currently impossible for computers. Computers do rote things well, like keeping ledgers or performing complex math. But computers have trouble recognizing even simple paterns much less generalizing those paterns of the past intto actions of the future. Brains store information as paterns. Some of these paaerns are very complicated and allow us the ability to recognize individual faces from many different angles. This process of storing information as patterns, utilizing those perns, and then solving problems encompasses a new field in ccomputing. This field, as mentioned b forre, does not utile traditional pogramming but involves the creation of massively parallel networks and the training of those networks to solve specific problems.

2. Human Neuron V/s Artificial Neurons

The fundamental processing element of a neural network is a neuron. Basically, a biological neuron receives inputs from other sources, combines them in some way, performs a generally nonlinear operation on the result, and then outputs the final result. These components are known by their biological names - dendrites, soma, axon, and synapses(Figure 1). Dendrites are hair-like extensions of the soma which act like input channels. These input channels receive their input through the synapses of other neurons. The soma then processes these incoming signals over time. The soma then turns th t processed value into an output which is sent out to other neurons through the axon and the synapses. Recent experimental data has provided further evidence that biological neurons are structurally more complex than the simplistic explanation abo e.

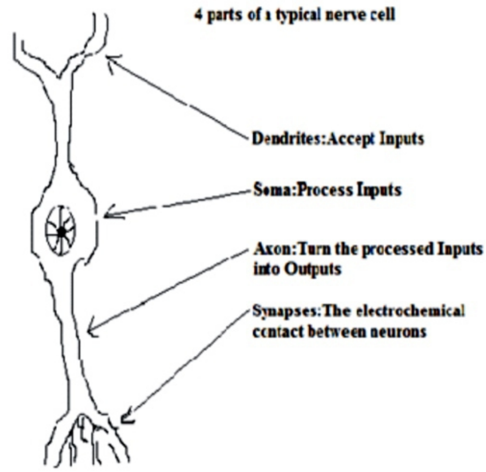


Figure 1: A Simple Neuron.

Neural network researchers are seeking an understanding of nature's capabilities for which people can engineer solutions o problems that have not been solved by traditional computing[2]. o do this, the basic unit of neural networks, the artificial neurons, simulate the four basic functions of n tural neurons. Various inputs to the network are represented by the mathemati al symbol, x(n)(Figure 2). Each of these inputs are multiplied   y a cconnection eight. These weights are represented by w(n). In the simplest case, these products are simply summed, fed through a transfer function to generate a result, and then output. This process lends itself to physical implementation on a lage scale in a small package. This electronic implementation is still possible with other network structures which utilie differennt summing functions asell as different transffer functions.

The applications including, the recognition of ext, the idennti aication of speech, and the image deciphering of scenes are required to turn real-world inputs into discrete values. These potential alues are limited to some known set, like the ASCII characters or the most common 50,000 English words. Because of this limitation of output options, these applications don't always utilie networks composed of neurons that simply sum up, and thereby smooth, inputs. These networks may utilie the binary properties of O Ring and AN Ding of inputs. These functions, and many others, can be built into the summation and tansffer functions of a network.

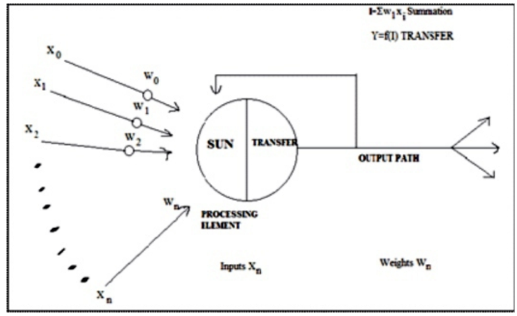


Figure 2: A Basic Artificial Ne on.

Approach

A. A Simple Neuron

An artificial ne on is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the using mode. In the training mode, the neuron can be trained to fire (or not), for particular input p aerns. In the using mode, when a taught input patern is detected at the input, its associated output becomes the current output. If the input patern does not belong in the taught list of input paterns, the firing rule is used to determine whether to fire or not.

B. Firing Rules

The firing rule[3] is an important concept in neural networks and accounts for their high flexibility. A firing rule determines how one calculates whether a neuron should fire for any input paaern. It relates to all the input paterns, not only the ones on which the node was trained. A simple firing rule can be implemented by using Hamming distance technique. The Rule Goes As Follows: “Take a collection of taining paaerns for a node, some of which cause it to fire (the 1-taught set of paaerns) and others which prevent it from doing so (the 0-taught set). Then the paterns not in the collection ause the node to fire if, on comparison, they have more input elements in common with the 'nearest' patern in the 1-taught set than with the 'nearest' patern in the 0-taught set. If there is a tie, then the patern remains in the undefined state.”

For example: A 3-input neuron is taught to output 1 when the input (X1,X2 and X3) is 111 or 101 and to output 0 when the input is 000 or 001. Then, before applying the firing rule, the truth table is:

X1	0	0	0	0	1	1	1	1
X2	0	0	1	1	0	0	1	1
X3	0	1	0	1	0	1	0	1
OUT	0	0	0/1	0/1	0/1	1	0/1	1

As an example of the way the firing rule is applied, take the patern 010. It differs from 000 in 1 element, from 001 in 2 elements, from 101 in 3 elements and from 111 in 2 elements. Therefore, the 'nearest' patern is 000 which belongs in the 0-taught set. Thus the firing rule requires that the neuron should not fire when the input is 001. On the other hand, 011 is equally distant from two taught paterns that have different outputs and thus the output stays undefined (0/1).

By applying the firing in every column the following truth table is obtained:

X1	0	0	0	0	1	1	1	1
X2	0	0	1	1	0	0	1	1
X3	0	1	0	1	0	1	0	1
OUT	0	0	0	0/1	0/1	1	1	1

The difference between the two truth tables is called the *generalization of the neuron.* Therefore the firing rule gives the neuron a sense of similarity and enables it to respond 'sensibly' to paterns not seen during training.

A. Patern Reccognition

An important application of neual networks is patern reccognition[4]. Patern recognition can be implemented by using a feed-forward neural network(Figure 3) that has been trained accordingly. During training, the network is trained to associate outputs with input paterns. When the network is used, it identifies the input patern and tries to output the associated output patern. The power of neural networks comes to life when a patern that has no output associated with it, is given as an input. In this case, the network gives the output that corresponds to a taught input patern that is least different from the given patern.

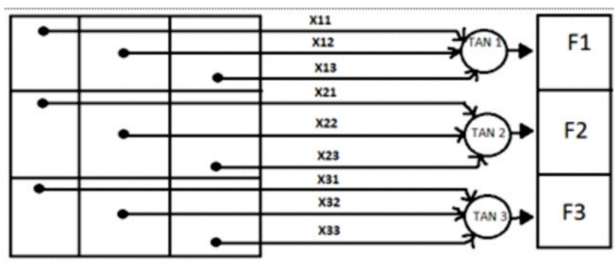


Figure 3: A Feed Forward Neural Network.

4 Architecture of Neural Networks[5]

A. Feed –Forward Networks

Feed-forward ANNs allow signals to travel one way only; from input to output. There is no feedback i.e. the output of any layer does not affect that same layer. Feed-forward ANNs (Figure 4) tend to be straight forward networks that associate inputs with outputs. They are extensively used in patern recognition. This type of organization is also eferred to as botom-up or top-down.

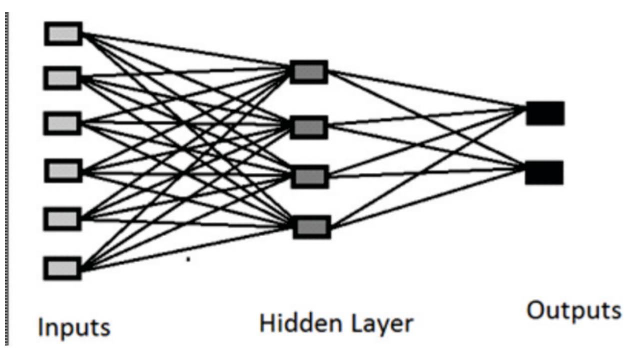


Figure 4: An Example of a Simple Feed Forward Network.

B. Feedback Networks

This is where the output of one layer routes back to a previous layer(Figure 5). The way that the neurons are connected to each other has a significant impact on the operaation of the network. In the larger, more professional soft are development packages the user is allowed to add, delete, and control these connections at will. By "tweaking" parameters these connections an be made to either excite or inhibit.

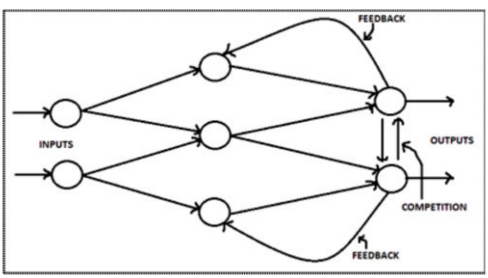


Figure 5: Simple Network with Feedback and Competition.

C. Perceptrons

The most influenntialork on neural nets in the 60's went under the heading of 'perceptrons '[6] a term coined by Frank Rosenblatt. The perceptron (Figure 6) turns out to be an MCP model (neuron with weighted inputs) with some additional, f ed, pre--processing. Units labeled A1, A2, Aj , Ap are called association units and their task is to extract specific, localised featured from the input images. Perceptrons mimic the basic idea behind the mammalian visual system. They were mainly used in patern reccognitionven though their capabilities xtended a lot more.

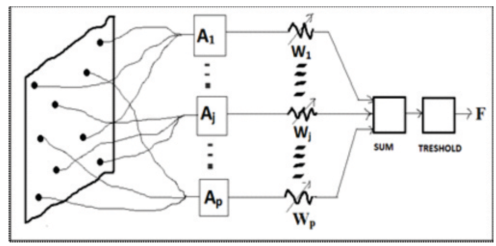


Figure 6: A Perceptron.

5. Electronic Implementation of Artificial Ne ons

These artificial neurons are called "processing elements" and have many more capabilities than the simple artificial ne on. In Figure 7, inputs enter into the processing element from the upper left.

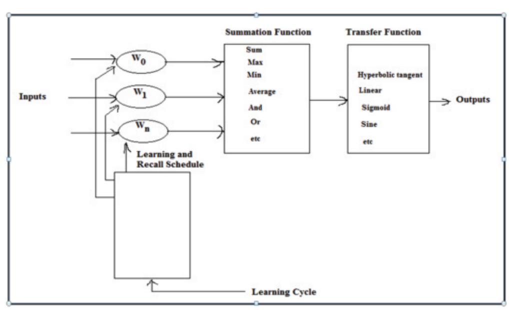


Figure 7: A Model of a "Processing Element”

The first step is for each of these inputs to be multiplied y their rrespecte weighting actor (w(n)). Then these modified inputs are fed into the summing function, which usually jut sums these products. These operations ould produce a number of different values which are then propagated forward; values such as the average, the largest, the smallest, the ORed values, the ANDed values, etc. Furthermore, most commercial development products allow soft are engineers to create their own summing functions via routinesoded in a higher level language. Sometimes the summing function is further complicated by the addition of an acti aation function which enables the summing function o operatte in a time sensiti e way.

Either way, the output of the summing function is then se t into a transfer function. This function then turns this number into a real output via some algorithm. It is this algorithm that takes the input and turns it into a zero or a one, a minus one or a one, or some other number. The transfer functions that are commonly supported are sigmoid, sine, hyperbolic tangent, etc. This transfer function also an scale the output or control its value via thresholds. The result of the transfer function is usually the direct output of the processing element. An example of how a transfer function orks (Figure 8).

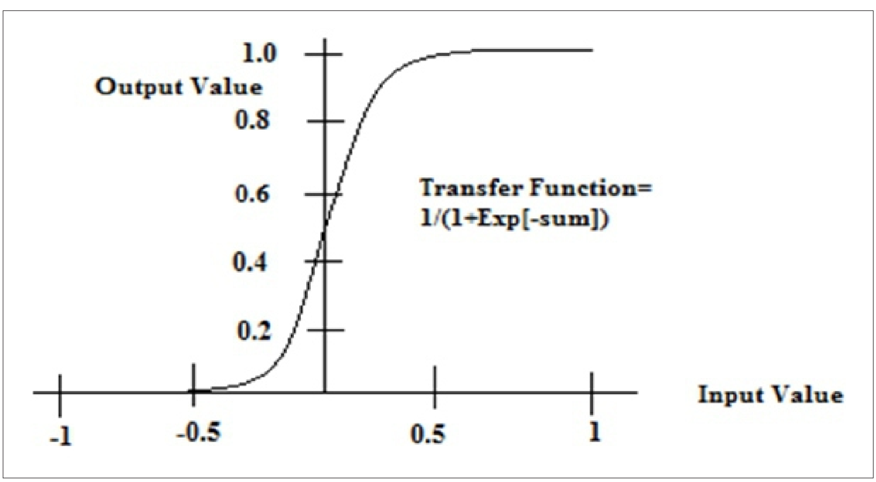


Figure 8: Sigmoid Transfer Function

This sigmoid transfer function akes the value from the summation function, alled sum (Figure 8), and turns it into a value between zero and one. Finally, the processing element is ready to output the result of its transfer function. This output is then input into other processing elements, or to an outside connection, as dictated by the structure of the network.

All artificial neural networks are constructed from this basic building block - the processing element or the artificial ne on. It is variety and the fundamental differences in these building blocks which partially ause the implementing of neual networks to be an "art.”

6. Application of Neual Networks

Neural networks are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identiy the disease[7]. Neural networks learn by example so the details of how to recognize the disease are not needed. What is needed is a set of examples that are representati e of all the variations of the disease. The qua ntity of examples is not as important as the 'quantity'. The xamples need to be selected very carefully if the system is to perform reliably and efficiently.

Neural Networks are used experimentally to model the human cardiovascular system[8]. Diagnosis can be achieved by building a model of the cardiovascular system of an individual and comparing it with the real time p ysiological measurements taken from the patie t. If this rroutine is carried out regularly, potential harmful medical conditions an be detected at an early stage and thus make the process of combating the disease much easier.

ANNs are used experimentally to implement electronic noses[9]. Electronic noses have several potential appli aations inelemedicine. Telemedicine is the practice of medicine over long distances via a communication link. The electonic nose would identiy odors in the remote surgical environment. These identified odors would then be electronically transmited to another site where a door generation ystem would recreate them. Because the sense of smell can be an important sense to the surgeon, telesmell would enhance telepresent surgery.

An application d veloped in the mid-1980s called the "instant physician" trained an auto associati e memory neural network to store a large number of medical records, each of which includes information on symptoms, diagnosis, and treatment for a particular case. Aer training, the net can be presented with input consisting of a s t of symptoms; it will then find the full stored paaern that represents the "best" diagnosis and treatment.

Conclusion

The computing orld has a lot to gain from neural networks. Their ability to learn by example makes them very flexible and powerful. Furthermore there is no need to devise an algorithm in order to perform a specific task; i.e. there is no need to understand the internal mechanisms of that task. They are also very well suited for real time systems because of their fast response and computational times which e due to their parallel architecture.

Neural networks also contribute to other areas of research such as neurology and psychology. They are regularly used to model parts of living organisms and to investi ate the internal mechanisms of the brain.

Perhaps the most exciting aspect of neural networks is the possibility that some day 'conscious' networks might be produced. There is a number of scientits arguing that consciousness is a 'mechanical' property and that 'conscious' neural networks are a realistic possibilit.

Finally, I would like to state that even though neural networks have a huge potential e will only get the best of them when they are integrated with computin , AI, fuzzy logic and related subjects.

References

[1] An introduction to neural computing. Ale sander, I. and Morton, H. 2nd edition

[2] DARPA Neural Network Study (October, 1987-February, 1989). MIT Lincoln Lab. Neural Networks, Eric Davalo and Patrick Naim

[3] An Introduction to Computing with Neural Nets (Richard P. Lipmann, IEEE ASSP Magazine, April 1987)

[4] Patern Recognition of Pathology Images http://opernik-

eth.npac.syr.edu:1200/Task4/patern.html

[5] Alkon, D.L 1989, Memory Storage and Neural Systems, Scientific Amer an, July, 42-50

[6] Minsky and Papert (1969) Perceptrons, An introduction o computational eometry, MIT press, expanded edition.

[7] Artificial Ne al Networks in Medicine http://ww .emsl.pnl.gov:2080/docs/cie/techbrief/NN.techbrief.ht

[8] A Novel Approach to Modelling and Diagnosing the Cardiovascular System http://ww .emsl.pnl.gov:2080/docs/cie/neural/papers2/keller.wcnn95.abs.html

[9] Electronic Noses for Telemedicine http://ww .emsl.pnl.gov:2080/docs/cie/neural/papers2/keller.ccc95.abs.html

SoftComputingMethodsforWebIntelligence

NishatZulquernainandMdWaseemAkram
M.Tech (Scholar)
Department of computer science
Al-Falah School of Engineering & Technology
Dhauj Faridabad.

Abstract
Web has become the primary means for information distribution. It is being usedor educational purposes and thus its popularity resuled in heavy traffic in the Internet. Web Intelligence (WI) deals with the scientific exploration of the w territories of the Web. As a new field of computer science, it ccombines artificial telligence and advanced information echnology in the context of the Web, and goes beyond each of them. Data mining has a lot of scope in e-Applications. The ey problem is how to find useful hidden pa erns for beber appliclication. Probl em to address soft computing echniques like Neural networks, Fuzzy Logic, Support Vector Machines, Genetic Algorithms in vvolution y Compututation. In this pa , we ee explore soft computing echniques use to achieve web intelligences.

Keywords*—Soft Computing, Neural twork, Data Mining, Web Intelligence*

Introduction
Data mining has useful business applications such as finding u ful hidden information fom databases, predicting futue trends, and making good business decisions [1,6,7]. Soft computing echniques such as fuzzy logic, genetic al orithm and neural networks are useful in data mining [l,2,l0]. Web intelligence, a term that was coined in the late 1999's, concerns about research and application of machine learning and information echnology with a specific focus on the Web platorms. Typical Web Intelligence applications include but not limited to online text classificaation, Web document

clustering, Web recommender for e-commerce, Web usage profiling and similar knowledge discovery tasks are drawing atenntion om c communities of global researchers. The data, in the context of data that are originated from the Web, called Web Intelligence data pose certain challenges to knowledge discovery tasks and Web mining. WI (Web Intelligence) is studied carefully from different aspects [3]. WI exploits Artificial Intelligence (AI) and advanced Information echnology (IT) on the Web and Internet [3].

1)Web- The World Wide Web (abbreviated as WWW or W3, commonly known as the web) is a system of interlinked hypertext documents accessed via the internet With a we, one can view web pages that may contain text, images, videos, and other multimedia and n vigate between them via.hyper link On March 12,1989,Tim berners lee a british omputer scienntt and former cern (European Organization for Nuclear Research) employee, wrote a proposal for what *would eventually become the World Wide Web.[5]*

2)Web Intelligence- Web intelligence is the area of study and research of the application of artificial intelligence and information echnology on the web in order to create the next generation of poducts, services and frameworks based on The internet. The term was born in a paper writen by Ning Zhong, Jiming Liu Yao and Y.Y.Ohsuga in the Computer Soft are and Applications Co ference in 2000. [11]

3) Methods of Data Mining –

.Artificial ne al networks - Non-linear predicti e models that learn through training and resemble biological neural networks in structure. Warren McCulloch and Walter Pitts [19] (1943) created a computational model for neural networks based on mathematics and algorithms. They called this model threshold logic. Neural network is used in data mining for patern reccognition.

Decision trees - Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classificaation of a taset. Although decision trees have been in development and use for over 60 years (one of the earliest uses of decision trees was in the study of television broadcasting y Belson in

1956).Decision tree is used in data mining for the classificaation.

A decision tree consists of 3 types of nodes:

a. Decision nodes - commonly represented by squares.

b. Chance nodes - represented by circles.

c. End nodes - represented by triangles.

- Rule induction - The xtrraction of useful if-then rules from data based on statistal significance. The rule induction al orithm was first used by Hunt in his CLS system in 1962.
- Genetic al gorithms - Optimazation techniques based on the concepts of genetic ombinaation, matation, and natural selection. It was introduced by John Holland in 1975.In 1989, Axcelis, Inc. released Evolver, the world's first commercial GA product for desktop computers.[20]
- Nearest neighbor - A classificaation technique that classifies each record based on the records most similar to it in an historical database. Donald Knuth in vol. 3 of The Art of Computer Programming (1973) called it the postoffice problem, referring to an application of assigning to a residence the nearest post office. Nearest is used in data mining for clustering.

Literature Review

Soft Computing-

Soft computing is an emerging approach to computing which parallels remarkable ability of human mind to reason and learn in an environment of uncertainty and imprecision[12].Soft Computing consists of several computing paradigms like Neural Networks, Fuzzy Logic, and Genetic al gorithms. Soft Computing uses ybridizaation of these techniques. A hybrid technique would inherit all the advantages of constitue t techniques. Thus the components of Soft Computing e complementary, not competit e, offering their own advantages and techniques to partnerships to allow solutions to otherwise unsolvable problems [13].

1. Difference betwween soft and hd c computing-

Hard computing	Soft computing
Conventional computing Requires a precisely stated analytical model.	Soft computing is tolerant to imprecision.
Often requires a lot of computation time.	Can solve some real world problems in reasonably less time.
Not suited for real world problems for which ideal model is not present.	Suitable for real world problems
It requires full truth.	Can work with partial truth.
It is precise and accurate.	Imprecise.
High cost for solution.	Low cost for solution

2. Soft Computing thodsa)

a) Fuzzy Logic- As one of the principal constitue nts of soft comput , fuzzy logic is playing a key role in what might be called high MIQ (machine intelligence quotie t) systems. Two concepts within fuzzy logic play a central role in its applications. The f st is a linguistic variable; that is, a variable whose values are words or sentences in a natural or synthetic language[14]. The other is a fuzzy if-then rule, in which the antecedent and consequents are propositions containing linguistic variables [14]. While variables in mathematics usually take numerical values, in fuzzy logic applications, the non-numeric linguistic ariables arre oen used to facilitate the expression of rules and facts [15]. For example, a simple temperature regulator that uses a fan might look like this:

IF temperature IS very cold THEN stop fan

IF temperature IS cold THEN turn down fan

IF temperature IS normal THEN maintain level

IF temperature IS hot THEN speed up fan

There is no "ELSE" – all of the rules are evaluated, because the temperature might be "cold" and "normal" at the same time o different degrees.

b) Neural networks- Based on the computational simplicity Artificial Neural Network (ANN) based classifier are used. In this proposed system, a feed forward multil yer network is used. Back propagation (BPN) Algorithm is used for training. There must be input layer, at least one hidden layer and output layer. The hidden and output layer nodes adjust the weights value depending on the error in classificaation. In BPN the signal flow will be in feed forward dirrection, but the error is back propagated and weights are up dated to reduce error. The modificaation of the eights is according to the gradient of the error curve, which points in the direction o the local minimum. Thus making it much reliable in prediction as ell as classifying tasks.

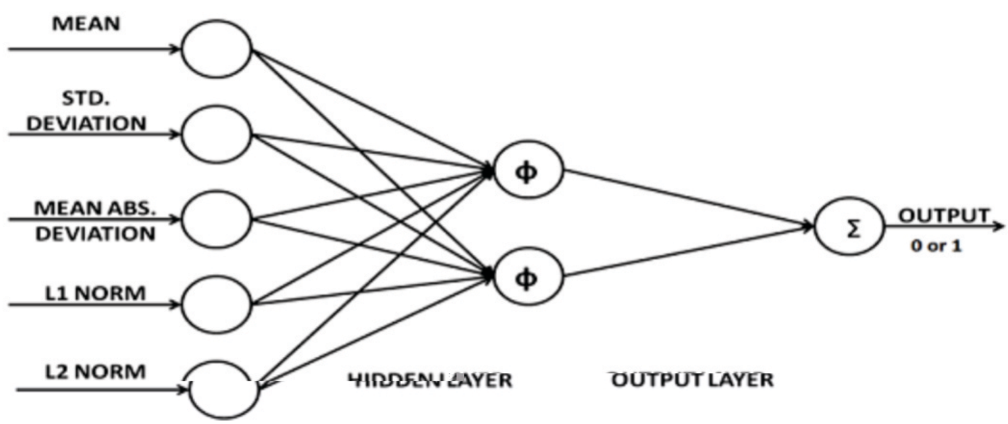


Fig Artificial Ne al Network Structure

In propose Methodology, We train an ANN with back propagation as eror tracing method and weight updated with biological genetic algorithm. Proposed Method consist of two phase.

• c) Support Vector Machines- Support Vector Machine (SVM) is a classificaation technique based on statistal learning theory. It is based on the idea of a hyper plane classifier. The goal of SVM is to find a linear optimal hyper plane so that the margin of separation b tween the two classes is maximized [17].The figure 2 below presents an overview of the SVM process

Genetic Algorithm- A genetic algorithm (GA) is a search heuristic that mimics the process of natural selection. This heuristic (also sometimes alled a metaheurisstic)

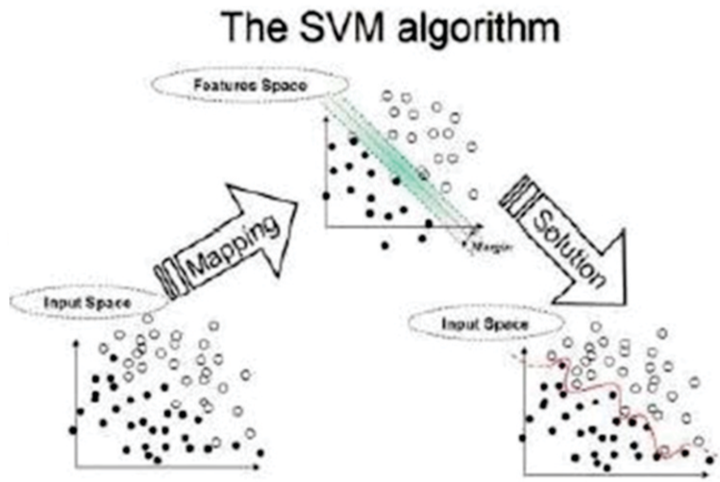


Fig SVM process

is routinely used o generate useful solutions to optimiaation and search problems.[16]

1. In proposed Methodology, GA work as

Parent selection Previous weight applied to error occurred in NN layer are the parent chromosome.

3. Split parent into two parents.

4. Select Crossover indices from parent then swap the element to generate new offspring.

5. Combine this two offspring and train NN with newly generated child.

6. Set weights to layer at which the minimum error occurred.

7. Muted selected indices

8. Repeat from step no 1 to step no 6 for every error occur NN Layers.

9. Stop.

2.3 Web Intelligence- This definition has the following implications. The basis of WI is AI and IT. The "I" happens to be shared by both "AI" and "IT", although with different meanings in them, and "W" defines the plaaorm on which WI research is carried out, [18]. The goal of WI is the joint goals of AI and IT on the new platorm of the Web.

1. *Web Intelligence Ralated Topics-* An overview of WI related topics as shown in Figure 3 and list several major subtopics in each topic below [11].

a. Web Human-Media Engineering:

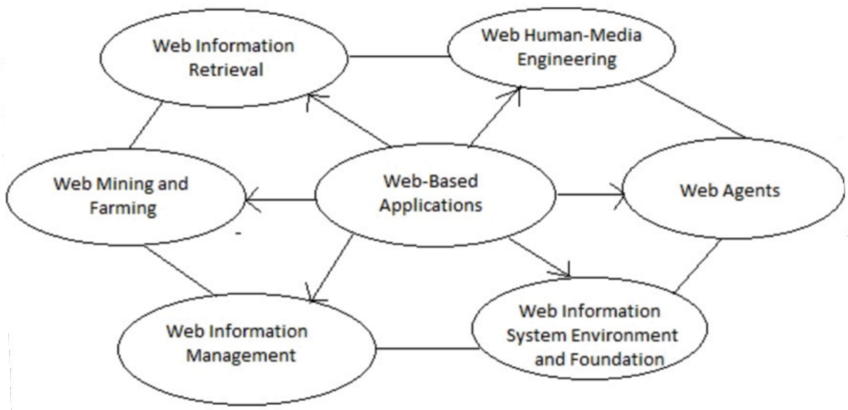


Fig- WI topics

The art of Web page design, multimedia information representation, multimedia information pocessing, visualizaation of Web information, and eb-based human computer interface.

b. Web Information Management: Data quality management, information transformation, Internet and Web-based data management, multidimensional Web databases and OLAP (on-line analyti al prrocessing), multimedia information management, new data models for the Web, object oriented Web information management, personalized information mana ement, semi-structured data management, use and management of metadata, Web knowledge management, Web page automatic eneraation and updatin , as well as Web security, integrity, privacy and trust.

c. Web Information Retrieval: Approximate retrieval, conceptual information extraction, image retrieval, multilinguistic information etrievval, multimedia retrieval, new retrieval models, ontology-based information retrieval, as well as automatic Web content cataloging and indexing.

d. Web Agents: Dynamics of information sources, e-mail filtering, email semi-automatic reply, global information collectin , information filtering, navigation guides, ecommender systems, remembrance agents, reputation mechanisms, esource intermediary and coordination mechanisms, as well as Web-based cooperati e problem solving.

e. Web Mining and Farming: Data mining and knowledge discovery, hypertext analysis and transformation, lerning user profiles, multimedia ta mining, regularities in eb surfing and Internet congestion, ext mining, Web-based ontology engineering, Web-based reverse engineering, Web farming, Web-log mining, and Web warehousing.

f. Web Information System Environment and Foundations: Competit e dynamics of Web sites, emerging Web technology, network community formation and support, new Web information description and qu y languages, theories of small world Web, Web information system development tools, and Web protocols.

Conclusions

Current research in soft computing mainly focuses on the discovery algorithm and visualization techniques. Soft computing methodologies, like fuzzy logic, neural networks, Support Vector Machines and genetic al orithms have recently been used to solve data mining problems can also be in web intelligence. This paper suggest the soft computing techniques used in soft computing. ac ording to the new characteristic, so to achieve intelligence of the web.

References

[1] K. Cios, W. Pedrycz, and R. Swiniarski, Data Miming methodr for Knowledge Discovery, Kluwer Academic Publishers, 1998.

[2] Meijuan Gao, Jingwen Tian, and Shiru Zhou ,"Research of Web Classificaation Mining Based on Classify Support Vector Machine" 2009 ISECS International Colloquium on Computin , Communication, Control, and Management

[3] Y.Y. Yao, Zhong, N., Liu, J. and Ohsuga, S., "Web Intelligence (WO:Research challenges and trends in the new information

age", Proc. Of WI2001, pp. 1-17, Springer, 2001.

[4] Y.-Q. Zhang and T.Y. Lin, "Computational eb Intelligence (CWI): Synergy of Computational Intelligence and Web Technology," Proc. of FUZZIEEE2002 of World Congress on Computational Intelligence 2002: Special Session on Computational Web Intelligence, pp.1104-1107, May 2002.

[5] "World Wide Web Consortium". "The World Wide Web Consortium (W3C)..."

[6] K. Thearling, "Increasing customer value by integrating D ta Mining and Campaign Management SoAware," Direct Marketing Magazine, February 1999.

[7] C. Westpbal and T. Blaxton, Data Mining Solulions -Methods and Tools for Solving Real- World Problems, John Wiley \& Sons, 1998.E. Gelenbe, Y. Feng, K.R.R. Krishnan, Neural Network Methods for Volumetric Magnetic

[8] V.Vapnik, Statistal Learning Theory, Wiley, 1998.

[9] J.Weston and C. Watkins, Support vector machines for multi-class patern reccognition, In oceeding of the 6th European symposium on Artificial Neural Network (ESANN), 1999.

[10] Y,-Q. Zhang, M. D. Fraser, R. **A.** Gagliano and **A.** Kandel, "Granular Neural Networks for Numerical-Linguistic Data Fusion and Knowledge Discovery,"Special Issue on Neural Networks for Data Mining and Knowledge Discovery, IEEE Transactions on Neural Nehvorks, Vol. 11, No. 3, pp.658-667, May, 2000.

[11] Zhong, Ning; Liu Yao, Jiming; Yao, Y.Y.; Ohsuga, S. (2000), "Web Intelligence (WI)", Web Intelligence, Computer Soft are and Applications Conference, 2000. COMPSAC 2000. The 24th Annual International, p. 469, doi:10.1109/CMPSAC.2000.884768, ISBN 0-7695-0792-1.

[12] J. S. R. Jang, C. T. Sun and E. Mizutani, Neuro-Fuzzy and Soft Computing :A Computational Approach to Learning & Machine Intelligence, Prentice Hall, Saddle River, New Jersey, 1997.

[13] S.N. Sivanandam, S.N. Deepa, Principles of Soft Computin , Wiley India (P) Ltd, 2007.

[14] L.A. Zadeh, “Outline of a New Approach to the Analysis of Complex Systems and Dccision Processes,” IEEE Trans. Systems, Man and Cybmzetics, pp. 28-44.

[15] Zadeh, L. A. et al. 1996 Fuzzy Sets, Fuzzy Logic, Fuzzy Systems, World Scientific Press, ISBN 981-02-2421-4.

[16] Mitchell, Melanie (1996). An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press. ISBN 9780585030944.

[17] C.J.C. Burges, “ A tutorial on support vector machines for patern recognitation “ , Data Mining and Knowledge Discovery , vol.2 , 1998, pp. 121-167.

[18] Y.Y. Yao, “Web Intelligence (WI): New Fronties Exploration” University of Regina, Saskatchewan, Canada 2003.

[19] McCulloch, Warren; Walter Pitts (1943). "A Logical Calculus of Ideas Immanent in Nervous Activity". Bulletin of Mathemati al Biophysics 5 (4): 115–133. doi:10.1007/BF02478259.

[20] Markoff, John (29 August 1990). "What's the Best Answer? It's Survival of the Fitest". New York Times. Retrieved 9 August 2009.

[21] Yuehui Chen, “Hybrid Soft Computing Approach to Identif acation and Control Of Nonlinear Systems”.

[22] Xiaozhe Wang, Ajith Abraham and Kate A. Smith, “Soft Computing Paradigms for Web Access Patern Analysis”

DischargePredictioninStraightCompoundChannels Using Artificial NeuralNetwork

PreetiPatel,[1] KishanjitK.Khatua,[2] ArpanPradhan[3]

[1]M.Tech. Student, Dept. of Civil Engineering, National Institute of Technology, Rourkela-769008, India

[2]Associate Professor, Dept. of Civil Engineering, National Institute of Technology, Rourkela-769008, India

[2]Ph.D. Scholar, Dept. of Civil Engineering, National Institute of Technology, Rourkela-769008, India

Abstract

Reliable estimation of discharge apacity helplps practitione o provide essenssential information regarding flood mitigatio onsconstructionydraulic structures and prediction of sedime t loads so as to plan for efficient prevenventive measures. Conventional m thods for esstimating the discharg apacity of a channel have been based on standard uniform flow formulas, such as Chezy's, Manning's, and Darcy-Weisbach's equations, by either treating the cross-section as a whole o y dividing it horizontally, verti ally, or diagonally into non-intteracting subareas. Another standard method for discharge estimation is y considering the intnteraction ects between the main channel and the flood plains, otherwise known as the method of coherence (COHM). Coherence method is observed to give better prediction of discharge with respect to the other methods but involves a variety of difficult formulae for calculating discharge undergoing a vey complicated procedure. Therefore these standard discharge prediction m thods are not always suited to be used in all sorts of data sets. Hence in this paper an atempt has been made to predict the discharge capacity in straight compound channels by Artificial Neural Network which does not involve complicated calculations. In this paper both the conventional models and ANN are analyed with the experimental data sets reported by other investigaors, such as the large scale channel data of FCF (Flood Channel Facility) (1991), straight channel data of Knight and Demetriou (1983), *Myers (1984) and the data observed at NIT Rourkela by Khatua (2008) and Mohanty (2013).The paper provides with a comparison among the discharge predictions and the percentage of error with respect to the actual discharges.*

Keywords: *Discharge Prediction, Straig t Channel, Coherence Method, ANN, Error Analysis*

1. Introduction

There are numerous hydrologic processes such as rainfall, snowfall, flood and draughts etc. which are usually investi ated by analyzing their records of observations. And ma y characteristics of these processes represent definite relaation le flood in river due to excess rainfall. In river channel analysis, the accurate prediction of discharge carrying capacity is a major problem. A flood is an overflow of water that submerges land area adjacent to a river or stream know as floodplain and this type of channel is commonly known as compound channel. The accurate prediction of discharge carrying capacity of straight compound channel is important for flood warning, determination of flood risk areas, flood controlling work in river, and in designing artificial aterways. The conventional m thod of discharge calculation in compound open channel flow, like Single Channel Method (SCM), where the whole compound section is onsidered as a single section. In SCM or a compound open channel, it is difficult to assign a single value manning's *n* for the whole section; this problem can be overcome by Divided Channel Method (DCM) by assigning different manning's *n* values for main channel and floodplains. The selection of interface plane for the separation of the compound section o sub-areas can be made by using the value of apparent shear at assumed interface plane (Knight and Demetriou 1983). The problem associated with DCM is turbulent interaction of flow tween main channel and floodplains which leads to momentum transfer. The Coherence Method (COHM) of Ackers (1992) for a compound open channel is a function of eometry only and along with geometrical weted parameters of the channel it uses the Darcy-Weisbach friction actor *f* and the discharge deficit factor (DISDEF)

for discharge calculations.
In this paper work, an atempt has been made to predict discharge capacity for straight compound channels. Along with the experimental studies at NIT Rourkela, experimental data sets of other researchers are also investi ated. Comparison among the various discharge predictions and pecentage of error with respect to actual discharges have been made.
A simple but reliable prediction method for estim ating dischge of flood channel is highly desirable for field engineers. And easily implement able technique ANN has been adopted in this paper. Typically, the architecture of ANNs consist of a series of nodes that are usually arranged in layers an input layer, an output layer and one or more hidden layers. The number of hidden layers required generally depends on the complexity of the relationship between input parameter and output values. Back propagation al orithm of Artificial Neural Network (ANN) is used for prediction of dischage in compound open channel and result obtained by ANN is compared with Single Channel Method (SCM), Divide Channel Method (DCM) and the Coherence Method (COHM).

2. Methodologies

2.1 Single Channel Method (SCM)

A compound channel cross-section is comprised of the deep main channel along with a shallow flood plain on one or both sides of the main channel. Chow (1959) suggested that Manning's, or Chezy's or Darcy–Weischbach's equations shown in equations 1-3 respectely can be used to predict discharge capacity at low depths, i.e. when the flow is only in main channel. But during over-bank flows such as floods, these traditional formulae either overestim tes or underessti tes the discharge capacity. By applying a composite roughness value in these methods, the prediction of dischage is basically inadequate for compound channels as they are considered as a single unit. The discharge carrying capacity is underestim ted as the single channel method suffers from a sudden reduction in ydraulic radius as the main channel overtops to the floodplains.

$$Q = \frac{1}{n} A R^{2/3} \sqrt{S_o} \quad (1)$$

$$Q = AC\sqrt{RS_o} \quad (2)$$

$$Q = \sqrt{\frac{8g}{f}} A\sqrt{RS_o} \quad (3)$$

here Q is the overall discharge of the compound channel, A is the cross-sectional area of the compound channel, R is the hydraulic radius for the compound channel, S_o is the slope of the channel, n is the composite Manning's coefficient for the compound channel while C and f are the Chezy's constant and Darcy-Weischbach's friction actor for the compound channel respecti ely.

2.2 Divided Channel Method (DCM)

The separation of compound section into sub-section y using 'interface plane' between the junctions of main channel and floodplains is known as divided channel method (DCM). Assumed interface planes, verti ally at the junction of the main channel and floodplains and horizontally at the bank-full depth of the main channel.

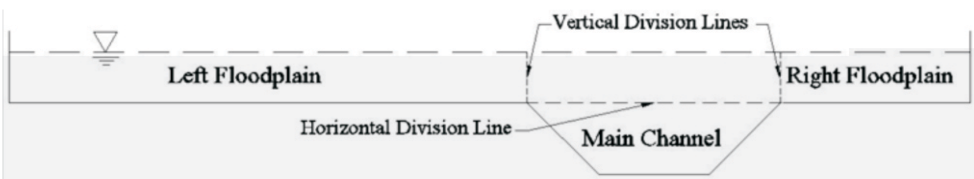


Figure 1 Typical Division of a Straight Compound Channel

Divided channel methods solve the issue of composite roughness as distinct oughness values of the main channel and floodplains can be used in their individual calculations. With the help of these division interfaces, the compound channel section is divided into subsections whose individual discharges are calculated by using Manning's or Chezy's equations and summed up to give the total discharge carried by the compound channel section. Geneally, Manning's formula is used in the calculations and is gi en by the equation 4.

$$Q_{total} = \Sigma Q_i = \Sigma \frac{1}{n_i} A_i R_i^{2/3} \sqrt{S_o} \quad (4)$$

where i refers to the subsections of the compound channel and the other variables have their usual denotation. When the verti al interfaces are considered, two types of methods can be perceived, VDM-I and VDM-II. The verti al division lines are not considered in the computation of the weted perimeter for the main channel for the case of VDM-I while it is considered for VDM-II. Similarly in the case of division of compound section y the horizontal interface, two approaches are observed, HDM-I and HDM-II where the division line is not included in calculations or HDM-I and included in HDM-II. The verti al division methods take into consideration the i terractionffect between the main channel and floodplains, while the horizontal methods do not. However by simply altering the weted perimeter by the verti al interfaces or the horizontal interfaces, does not completely reflect the interaction ffect as a simple function and in eneral overessti tes the flow rate. The discharge prrediction methods by these approaches do not provide with a particular tend. For the present study, by applying all these division method to the current dataset, HDM-I gave beter results for which only this approach is considered in the present investi aation and is denoted as DCM (Divided Channel Method).

2.3 Coherence Method (COHM)

Acker (1992 and 1993) established a one dimensional approach for dealing with overbank flows and the related difficulties of heterogeneous roughness and shape effects. Coherence denoted as COH, is defined as the ratio of the basic onveyance calculated by treating the channel as a single unit with perimeter weighting of the friction actor to that calculated by summing the basic conveyances of the separate zones which is given by equation 5.

$$COH = \frac{\sum A_i \sqrt{\frac{\sum A_i}{\sum f_i P_i}}}{\sqrt{A_i \sqrt{\frac{A_i}{f_i P_i}}}} \quad (5)$$

here *i* identifies each of th *n* flow zones, *A* is the sub-sectional aea, *P* is the weted perimeter of the subsection and *f* is the Darcy–Weisbach friction actor for the subsection.

As COH approaches unity, it is appropriate to treat the channel as a single unit using the overall geometry and discharge is estim ted as per single channel method. In extreme cases, COH may be as low as 0.5. When coherence is much less than unity, discharge deficit factors are required to correct the individual discharges in each sub-area and calculations ae similar to divided channel method. The coherence method was developed by Ackers by basically taking into consideration the experimental investi aations of the flood channel facility (FCF) phase A, data sets. Four distinct l vels of flow regions have been suggested by the study which occurs above the bank-full level of the main channel in the case of straight compound channel flows. Different discharge deficit factors are to be estim ted by methodologies given by Ackers (1993) for each region to estim te the overall discharge of the compound channel.

Region 1:

The depth of flow is quite low in this region; hence, the velocities in floodplains and the main channel are very dissimilar. This region is characterized by relati e depth, H_r<0.2.

$$H_r = \frac{H-h}{H} \tag{6}$$

where *H* is the water level above channel botom and *h* is the bank-full level of the main channel.

$$Q_{R1} = Q_{basic} - DISDEF \tag{7}$$

Region 2:

This zone is of a higher depth where interaction ffect is s still not domi t and the flow computaation depends on discharge adjustment factor, *DISADF* in the channel under consideration.

$$Q_{R2} = Q_{basic}\ X\ DISADF_2 \tag{8}$$

where $DISADF_2$ is the discharge adjustment factor for region 2.

Region 3:

This zone appears when the relati e depth is around 0.5 and the interference effect affects the discharge capacity and hence the discharge adjustment factor is different in this region and is termed as $DISADF_3$

$$Q_{R3} = Q_{basic}\ X\ DISADF_3 \tag{9}$$

Region 4:

This zone has a greater value of relati e depth i.e. above 0.6 and behaves as a single unit due to the coherence character that obeys both the main channel and the floodplains. The discharge adjustment factor for this region $DISADF_4$ is dependent on the aforesaid coherence character of flow.

$$Q_{R4} = Q_{basic}\ X\ DISADF_4 \quad (10)$$

here Q_{basic} is the basic total discharge calculated using zones separated by verti al divisions. The discharge deficit factor for zone 1 and the discharge adjustment factors for the others zones are calculated by individual methodologies and then the predicted discharge is estim ted from the choice of region.

3. Sources of Data Sets and Selection of Hydraulic Parameters

3.1 Sources of Data

The data used for analysis are;

1. Experimental findings from the Flood Channel Facility (FCF) at HR Wallingford, UK (1987) which is served as a high quality benchmark data used by many researchers to compare and validate their research work. FCF Phase A consists of experimental data sets for straight compound channels with different cross-sectional geometries having symmetric and asymmetric sections. Six such d ta series have been considered in this paper with symmetric cross-sections.

2. Experimental results from Knight (1983) are considered comprising of one rectangular main channel and two symmetrically disposed floodplains. Three such data sets with different floodplain widths are presented, demonstrating its effect on the discharge capacity of the compound channel with other variables remaining constant.

3. Smooth channel boundary data are collected from Myers (1984) relating o asymmetrical as well as symmetrical compound cross sections from which only the symmetric data set is considered.

4. Data sets from similar experimentations onducted at NIT, Rourkela by Khatua (2008) and Mohanty (2013) have been considered in the study.

The detailed parameters of all the datasets are given in table 1.

NOP represents the number of data points, *h* refers to the depth of main channel while *b* is the base width of main channel. *B* is the cross-sectional top width of the two-stage channel S_o refers to the longitudinal bed slope or the valley slope and *n* is the Manning's roughness coefficient. *β* refer to the relati e depth of the rrespecte data points.

Table 1 Geometric Parameters of Data Sets in Literature

Test Series		NOP	Main Channel C/S Shape	*h* (in m)	*b* (in m)	*B* (in m)	S_o, Bed Slope	Manning's *n*	*β* = (*H*-*h*)/*H*
FCF A (1987)	1	8	Trapezoidal	0.15	1.5	10	0.001027	0.01	0.056 - 0.400
	2	10	Trapezoidal	0.15	1.5	6.3	0.001027	0.01	0.042 - 0.479
	3	10	Trapezoidal	0.15	1.5	3.3	0.001027	0.01	0.050 - 0.500
	7	7	Trapezoidal	0.15	1.5	6.3	0.001027	0.01415	0.038 - 0.504
	8	8	Rectangle	0.15	1.5	6	0.001027	0.01	0.050 - 0.499
	10	8	Trapezoidal	0.15	1.5	6.6	0.001027	0.01	0.051 - 0.464
Knight (1983)	2	6	Rectangle	0.076	0.076	0.152	0.000966	0.01	0.108 - 0.493
	3	6	Rectangle	0.076	0.076	0.228	0.000966	0.01	0.131 - 0.490
	4	6	Rectangle	0.076	0.076	0.305	0.000966	0.01	0.106 - 0.506
Myers (1984)		10	Rectangle	0.08	0.16	0.75	0.00093	0.01	0.180 - 0.476
NITR (2008, 2013)	Khatua	10	Rectangle	0.12	0.12	0.44	0.0019	0.01096	0.119 - 0.461
	Mohanty	6	Trapezoidal	0.065	0.33	3.95	0.0011	0.01	0.109 - 0.435

3.2 Selection of Hydraulic Parameters

Discharge in a straight compound channels found to be substantially dependent on geometrical as well as hydraulic factors deduced from the study of literatures (Yang 2005). Therefore in the current study, the parameters considered in the neural network analysis are (i) the width ratio α=B/b, (ii) the ratio of main channel area to that of floodplain denoted as $A' = A_M / A_{FP}$, (iii) ratio of hydraulic radius of main channel to that of the floodplain referred as $R' = R_M / R_{FP}$, (iv) slope of the compound channel S_o (v) Darcy's friction factor *f* for the respecti e data points and (vi) the relati ve depth of the channel section for each data point denoted as *β*=(*H*-*h*)/*h*. the above non-dimensional variables are normalized in the range of [-1,1] for beter disstribution of ta sets and are considered as input values to the artificial ne al network code with the actual discharge data as the output range.

4. Back Propagation Neual Network (BPNN)

4.1 Back Propagation Neural Network Architecture

The architecture network of a back propagation neual network model consist of three layers namely input layer, hidden layer and output layer. Generally, input layer is considered a distributor of information fom the external sources. Hidden layer are considered to be feature detectors of such information. The output l yer is considered a collector of the features detected and producer of the network. The input signals of input layer and output signals of hidden layer are modified by intercconnection weight, known as weight factor. The sum of signals from input layer (total acti aation) is then modified by a sigmoidal transfer function f). And the sum of the modified signal from hidden layer is then modified by a pure linear transfer (f) function and output is collected at output layer.

Let $I_p = (I_{p1}, I_{p2}, \ldots I_{pl})$, p= 1, 2, . . . ,N be the pth patern among N input paaern.

Output from a neuron in the input layer is,

$O_{pi} = I_{pi}$, I = 1, 2, . .l (11)

Output from a neuron in the hidden layer is

$$O_{pj} = f\,(\text{NET})_{pj} = \sum_{i=0}^{l} W_{ji} O_{pi}, \; j=1, 2, \ldots m \quad (12)$$

Output from a neuron in output layer is

$$O_{pk} = f(\text{NET}_{pk}) = f \sum_{i=0}^{l} W_{kj} O_{pj} \quad k=1, 2, \ldots n$$

where W_{ji} represents connection weights between i^{th} input neuron to j^{th} hidden neuron, and W_{kj} represents connection weights between j^{th} hidden neuron to k^{th} output neuron.

4.2 Learning or Training in Back Propagation Neural N twork

Batch mode type of supervised learning has been used in the present case in which interconnection eights are adjusted using delta rule algorithm afer sending the entie training sample to the network. During training, the predicted output is compared with the desired output, and the mean square error is calculated. If the mean square error is more than a prescribed limiting alue, it is back propagated from output to input, and weights are further modified till the e or or number of iterations is within a pescribed limit.

Mean square error, E_p for patern p is defined as

$$E_p = \sum_{i=1}^{n} \frac{1}{2} \left(D_{pi} - O_{pi}\right)^2 \quad (14)$$

where, D_{pi} is the target output, and O_{pi} is the computed output for the i^{th} patern.

Weight change at any time t, is given by

$$\Delta W(t) = -\eta E_p(t) + \alpha \Delta W(t-1) \quad (15)$$

here η is the learning rate i.e. 0 < η <1 and α is the momentum coefficient i.e.

0 < α <1.

5. Result and Analysis

The collected data sets being 95 in number are divided into 70% of training data and the rest as testing data. The number of hidden layer neuron are fixed by experimentaation to find the ANN architecture with minimum value of mean squared error as shown in figure 2. The neural network structure architecture is chosen to be three layered structure denoted as 6-8-1 i.e. 6 input layers, 8 hidden layers and 1 output layer. Analysis on the residual value of the training data sets are carried out as shown in the figure 3, charted against the test series samples. The analysis indicates that the training data is well trained as the residuals are evenly distributed across the centerline. Figure 4 specifies the comparison among the actual and predicted values of discharge in the training and testing d ta ranges which are found to be quite synchronized.

The correlation plots b tween the actual and predicted discharges is presented in figure 5 for all the prediction m thods. ANN is seen to have the highest coefficient of determination of 0.99 which is quite high, indicating th t the model is potentially suitable for discharge computations. able 2 illustrates the different error values for each of the models indicated in the paper. Mean absolute error (MAE), the mean square error (MSE) and the root mean square error (RMSE) all show minimum values for the discharge predictions y ANN which are again incidentally much less in comparison to the other methods. MAPE or the mean absolute percentage error for ANN model is around 7% whereas that for coherence model and divided channel method is around 16% and 20% respecti ely, with 22% in the case of single channel method. The coefficient of determinanation, i.e. the R^2 values are also specified in the table with an increasing order of preference from SCM to the ANN prediction model.

6. Conclusions

The following conclusions that can be presented in this paper are

1. Estim ation of discharge prediction in a ompound channel is studied in this paper with a proposition of using an artificial neural network model as a tool for discharge estim ation. ANN is able

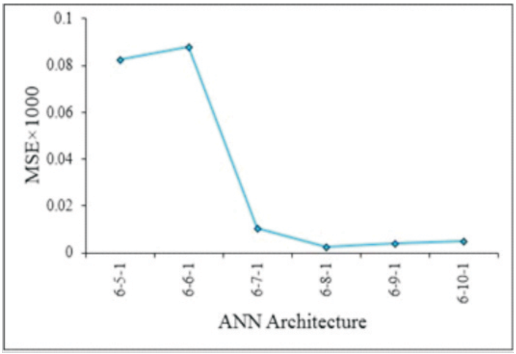


Figure 2 MSE for Different Ann Architectures

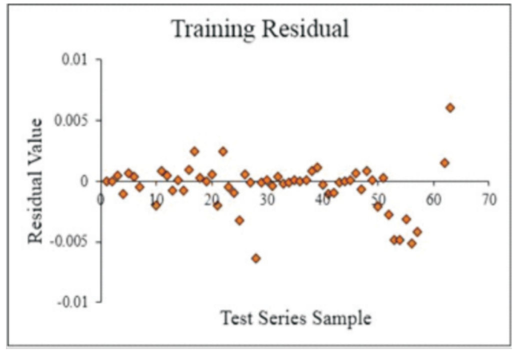


Figure 3 Residual Distribution or Training Data

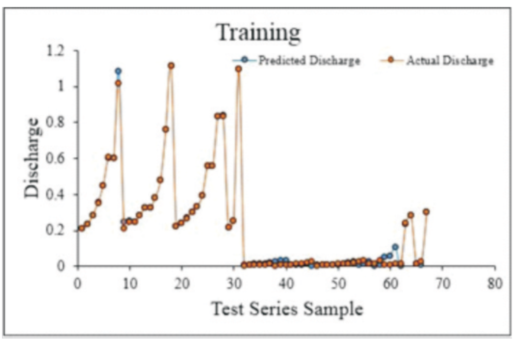


Figure 4.1 Comparision in Training Data

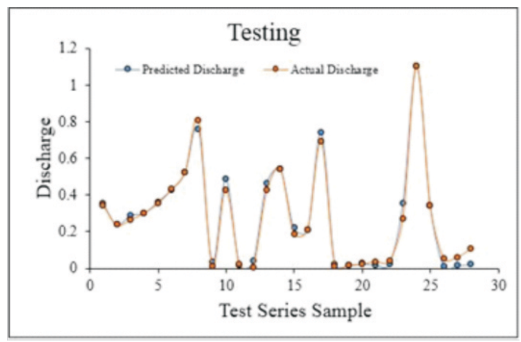


Figure 4.2 Comparision in Testing D ta

Figure 4 Comparision of Actual and Predicted Discharge Values in ANN

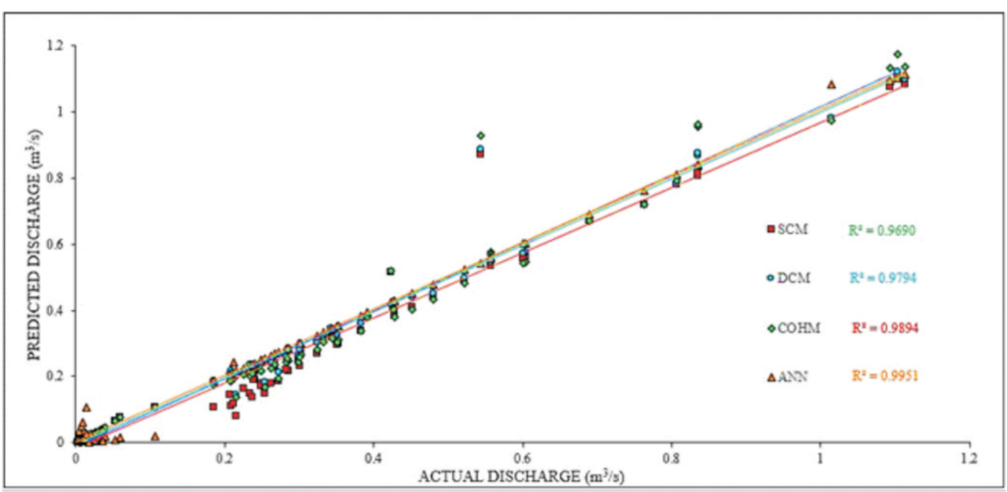


Figure 5 Correlation Plots b tween Actual and Predicted Discharges

Table 2 Error Analysis

Statistical Error Analysis	SCM	DCM	COHM	ANN
MAE	0.0320	0.0269	0.0276	0.0019
MSE	0.0045	0.0036	0.0027	0.00006
RMSE	0.0841	0.0711	0.0529	0.0081
MAPE	22.99	19.88	16.83	6.9962
R^2	0.9690	0.9794	0.9894	0.9951

to provide with decisive results with high degree of accuracy for a wide variety of data sets having nonlinear relationshi s, due to the nonlinear mapping of the input and output values.

2. Traditional m thods for discharge prediction in ompound channels such as single channel method and divided channel methods are investi ated where the SCM underestim tes while the DCMs overestim te the actual value primarily because the interaction effects between the main channel and floodplains are not taken into consideration.
3. The coherence method which takes the momentum transfer mechanism into account, gives reasonable prediction alues. The issue of overbank flow and composite roughness is also taken care of. COHM gives reliably beter result than the trraditional methods but ANN surpasses all the methods as far as accuracy is concerned.
4. The correlation plots or different methods show that the ANN model is fi ed with astounding accuracy whereas the coefficient of determination or the other methods are lower. The mean absolute percentage error for ANN is observed to be as low as 7% as compared to higher mean absolute percentage of errors for the other models.
5. The present study has immense scope where the data sets can be increased by applying the data of other investi ators to provide with a more decisive analysis, as the proposed ANN model can be examined with a wider variety of data ranges.

References

1. Ackers, P. (1991), "Hydraulic design of Straight compound channels Volume 1", *Report* SR 281, HR Wallingford, Wallingford, U.K.
2. Ackers, P. (1992),"Hydraulic Design of Two-Stage Channels", *Proceedings of the Institution of Civil Engineers, Water, Maritime and Energy*, Vol. 4, pp. 247-57.
3. Ackers, P. (1993), "Stage-Discharge Functions for Two-Stage Channels: The Impact of New Researches", Water and Environment Journal, Vol. 7, Issue 1, pp. 52-59.
4. Chow, V. T. (1959), "Open-Channel Hydraulics",McGraw-Hill Book Co.
5. Khatua, K. K. (2008), "Interaction of flow and estim ation of discharge in two stage meandering compound channels". Ph.D. Thesis, National Institue of Technology, Rourkela.
6. Knight, D. W, Demetriou J. D. (1983), "Flood Plain and Main Channel Flow InteractionJ Hydraulic Engineering, ASCE, Vol. 109, No. 8, pp. 1073-1092.
7. Knight, D.W. and Sellin, R. H. J. (1987), "The SERC Flood Channel Facility", Journal of The Institution of Water and Environmental Management, Vol. 1, No. 2, pp. 198-204.
8. Mohanty, P. K. (2013), "Flow analysis of compound channels with wide floodplains". Ph.D. Thesis, National Institue of Technology, Rourkela.
9. Myers, W.R.C. (1984), "Frictional Resistance in Channels with Floodplains", 1st International Conference on Channels and Channel Control Structures, Springer-Verlag, Heidelberg, Vol. 4, pp. 73-87.
10. Yang, K., Cao, S., Liu, X. (2005), "Study on Resistance Coefficient in Compound Channels.", Acta Mech Sin , Vol. 21, pp 353-61.

FrequencySelectivePropertiesofCoaxial TransmissionLinesLoadedwithCombined Artificial Inclusions

Ms.RiyaRai

(Research Scholar)

Abstract

The properties of a modified oaxial transmission line by periodic inclusions will be discussed. The introduction of split ring resonaors, conductor stubs, air gaps, and combination of these gives rise o neew frequency selective properties, such a top band or passband behavior, observable in planar as well as volumetric metamaterial structures. These results envisage new potential appli cations and implementation of d vices in coaxial transmission line technology.

1. Introduction

The inclusion of periodic elements loading a transmission line or a volumetric structure such as frequency selecti e surfaces has been of interest for over the past 60 years. The initial appli aations of artificial dielectrics and periodic structures, for the implementation of lenses to enhance radar technology [1], evolved to photonic bandgap and electromagnetic band ap structures from the Infrared frequency range [2] and scaled to the RF and microwave frequency range [3, 4]. In the late 90s, the split ring resonator (SRR) was proposed as an element with a strong magnetic response, implemented with conventional onductors [5]. The combination of these SRR eleme ts with thinwire media gave rise to a so called left handed meta material structure, characterized by exhibiting ntipaallel phase and group velocity and E-field, H-field, and k-vectors forming a lef-handed triad instead of a conventional rig t-handed one [6, 7]. The exploration of these tructures gave rise to the implementation of a broad range of devices in different configuraations, om free space to closed and open waveguides [8–15]. One of the most popular transmission lines employed in communication systems is coaxial cables. This is due to the fact that it is completely shielded, exhibiting high interference protection, and th t it supports a transverse electromagnetic mode as the fundamental propagating mode. This implies the lack of a cut-o ff frequency, which allows among other things the simultaneous transmission of information signals as ell as DC or AC power, which enables powering the remote end of the communication link, as applied, for example, in Power over Ethernet devices or in terrestrial microwave communication lin s. In this sense, the application of inclusions in order to tailor the frequency response of a coaxial transmission line can be of interest in this type of applications. This oncept has been extended in order to potentially impleme t devices such as power dividers, based on a coaxial transmission line loaded with SRR elements as well as with capaciti e loaded strips which exhibit frequency selecti e response [16].

In this work, an alternati e configuraation in der to implement lef-handed coaxial transmission lines is proposed. It consists in the superposition of Split Ring esonators with an array of concentric thin wires, emulating a thin wie media, which is in principle decoupled to the array of SRR elements. Due to decoupling of the electric response of the SRR elements and the thin wire media, it is possible to obtain an initial ejected frequency band that becomes a passband when both structures are combined. The proposed structure is schemati ally depicted in Figure 1. Initially, in Section 2, the equency response of the individual structures (Split Ring Resonators and thin wire media) will be described. Section 3 describes the combination of both in order to achieve an equivalent lef-handed medium, with a characteristic pass-band response, leading to the conclusions of the paper.

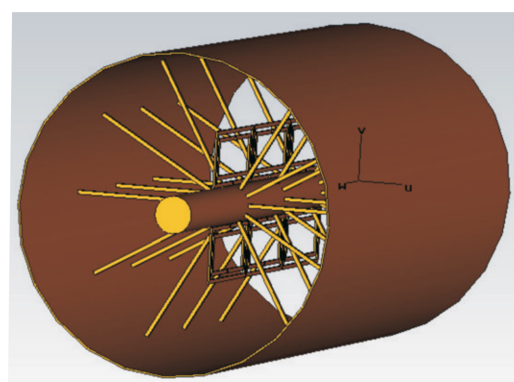


Figure 1: Schematic vi w of a coaxial transmission line loaded with radial stubs and Split Ring Resonators

2. Coaxial Transmission Line with Periodic Inclusions

Theinitial step in order to obtain frequency selecti e behavior is to introduce periodic inclusions in a conventional coaxial transmission line. The first step is to introduce split ring resonators embedded within the dielectric volume of the coaxial cable. The location of the SRR eleme ts will be given by physical restrictions o fit the inclusions as well as by the configuraation of the elecomagnenetic fields. It is a well-known fact that the coaxial transmission line supports a transverse electromagnetic mode, with E-field lines perpendicular to the central conductor spanning to the surrounding ground plane, whereas the H-field lines are solenoidal, closing around the central conductor strip.

The SRR element exhibits bianisotropic behaviour, being the main excitation omponent of H-field lines which are perpendicular to the plane that contains the metallic sloted rings. Taking this fact into account, initially the mot adequate position o obtain adequate excitaation of the SRR elements is to place them in an axial disposition aound the central conductor line, as depicted in Figure 1. The quasistatic esonance frequency can be determined by an equivalent circuit, in which the central frequency is given by an LC tank, as described in [8, 9], which is coupled to the equivalent circuit model of the transmission line. In the design proposed in this work, the dimensions of the SRR elements are given in Figure 2, which following the equivalent circuit model of the SRR exhibit a quasistatic resonance frequency in the vicinity of 2.9GHz [8, 9]. The Coaxial Transmission Line has an inner conductor of diameter 1mm, an outer diameter of 8.9mm, and a dielectric of ε_r= 2.43.

Full wave simulation esults have been obtained with the aid of CST Microwave Studio for the Coaxial Transmission Line loaded with SRR elements. The host transmission line is loaded with 6 SRR elements of the dimensions previously stated in Figure 2. The overall length of the proposed prototype is 33mm.The simulation results are shown in Figure 3, in which a rejection band in the vicinity of 2.9GHz can be observed, clearly given by the quasistatic resonance frequency of the SRR elements. Resonances at higher frequencies (beyond the obtained simulation result) are given either by dynamic resonances of the SRR elements or by Bragg resonances due to periodic loading. This can also be observed in the surface current density plots which are depicted in Figure 4. Excitation an be seen in the SRR elements in the quasistatic esonance frequency at 2.9GHz, whereas little or no response from the SRR elements is observed out of the resonance band. At this band, the structure can be viewed as equivalent media with a negati e value of magnetic permeability μ.

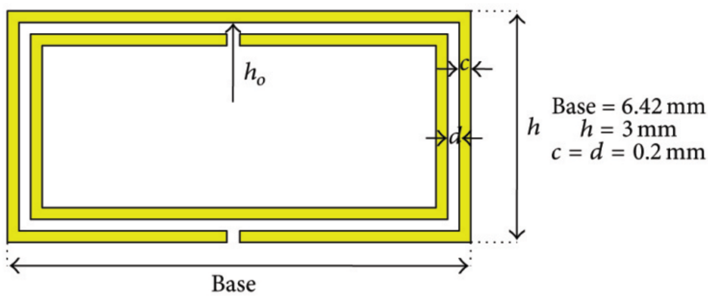


Figure 2: Dimensions of the SRR elements which load the host Coaxial Transmission Line. The conductor strips have a thickness of 0.1 mm.

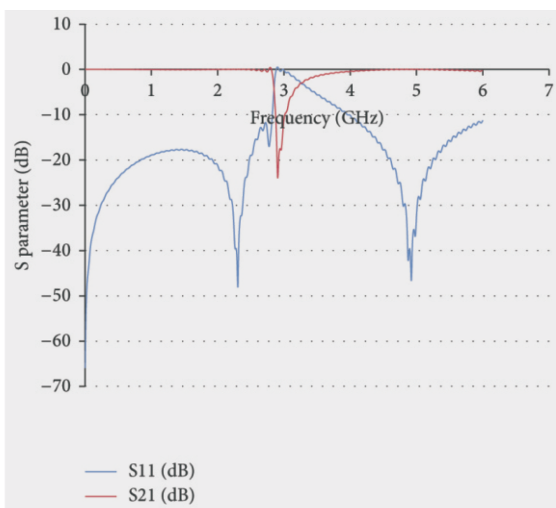
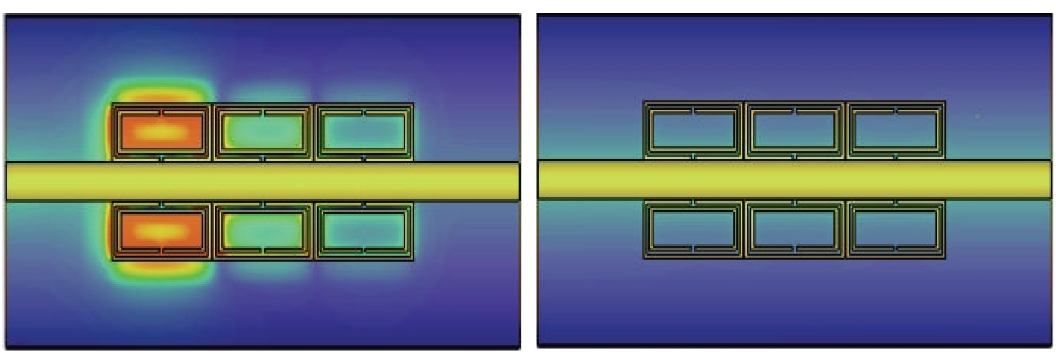


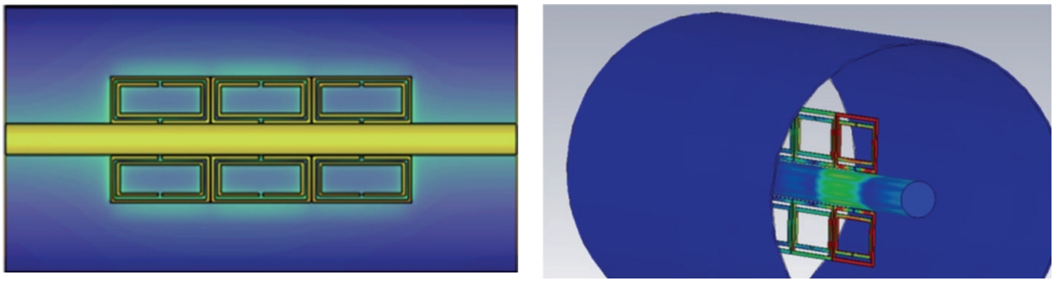
Figure 3: Full wave simulation esults for S21 and S11 parameters in a Coaxial Transmission Line loaded with Split Ring Resonators. The quasistatic esonance frequency is approximately at 2.9GHz, determined by the dimensions of the SRR elements embedded in the dielectric core of the Coaxial Transmission Line.

In order to achieve a Lef-Handed structure it is necessary to combine an equivalent negati e µ with the equivalent negati e value of µ which can be observed at the plasma frequency. A simple way to achieve this condition is y exxciting an aay of thin wires with a parallel electric field. This approach has been successfully applied in volumetric experiments, hollow metallic waveguides as well as in planar transmission lines. Following a transmission line analogy these radial stubs behave in a similar fashion as shunt inductors, which are connected from the central conductor to the ground plane. The number of stubs as well as their separation distance determines the overall requency response of the device, which intrinsically high pass. Full wave simulation esults have been obtained for a Coaxial Transmission Line loaded with 10 radial stubs, each one of them with a diameter of 0.1mm (Figure 6). Each set of radial stubs is separated 10 mm, in order to allocate additional SRR elements in the composite lef-handed configuraation that will be described in the following section. The results depicted in Figure 4 show strong rejection values at the operating frequency of the SRR elements, in the 2.9GHz frequency region. The rejection gi en in the 2.9GHz frequency region can be observed in the surface current plots presented in Figure 5.

As it can be seen, there is very low surface current content in the output port, due to high values in the reflection oefficient determined by the S11 parameter, in line with the inducti e nature of the radial stubs.



(a) (b)



(c) (d)

Figure 4: Surface current plots for the SRR loaded Coaxial Transmission Line. (a) Observation frequency of 2.9GHz, in which quasistatic esonances can be observed within the SRR elements, (b) observation fequency of 1GHz, (c) observaation equency of 4GHz, and (d) volumetric representatation of SRR excitation t the quasistaaticesonance frequency of 2.9GHz.

3. Combined Loading Elements

Once the equivalent value of negati e ε can be derived from the radial stubs loading the Coaxial Transmission Line and the equivalent value of negati e µis achieved by introducing SRR elements, both can be now combined in a single structure in order to obtain Lef-Handed behaviour. In order to enhance the frequency response of the resulting d vice, both the electric response (given by the radial stubs) and the magnetic esponse (given by the set of SRR elements) should be decoupled [8, 9]. Due to the fact that the fundamental mode in the Coaxial Transmission Line is a TEM mode, E-field components extend from the central conductor to the external ground plane conductor, perpendicular to the conductor surfaces, whereas the H-field lines close around the central conductor strip in a concentric manner. Therefore, the optimal onfigurarati on to achieve such decoupling is the arrangement depicted in Figure 1, in which the arrays of radial stubs are located within the normal plane in which the slits of the metallic rings of the SRR are located. The result of such a combined structure is a passband where previously a stopband was present, as can be seen from the full wave simulation esults depicted in Figure 7. This passband exists due to the combined effect of the quasistaatic resonance frequency of the SRR elements and the strong inducti e response of the radial stubs at this frequency. Once out of the bandwidth of operation of the quasistatic resonance frequency of the SRR element, rejection is once again observed, since only reacti e power can now be propagated. This can be clearly seen in the power flow plots depicted in Figure 8, as well as in the surface current density plots shown in Figure 9.

4. Conclusions

In this work, a novel configuraation in order to achieve Lef-Handed behavior in a Coaxial Transmission Line is presented. By combining a set of Split Ring Resonators and an array of radial conductor stubs, a passband where in principle a stopband should appear can be obtained. This is given by the adequate excitation of the combination of SRR eleme ts with the radial stubs, avoiding magnetoelectric coupling. The proposed approach can be employed in order to implement frequency selecti e devices in coaxial technology, which can be in principle scaled in frequency in order to cover a broad range of potential appli aations.

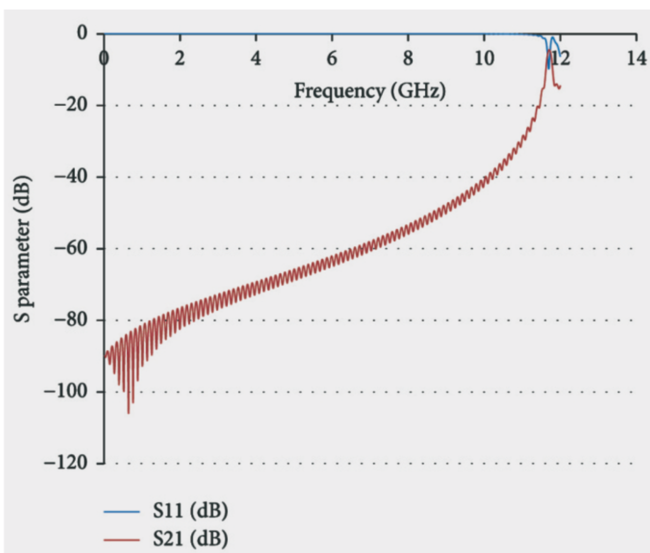
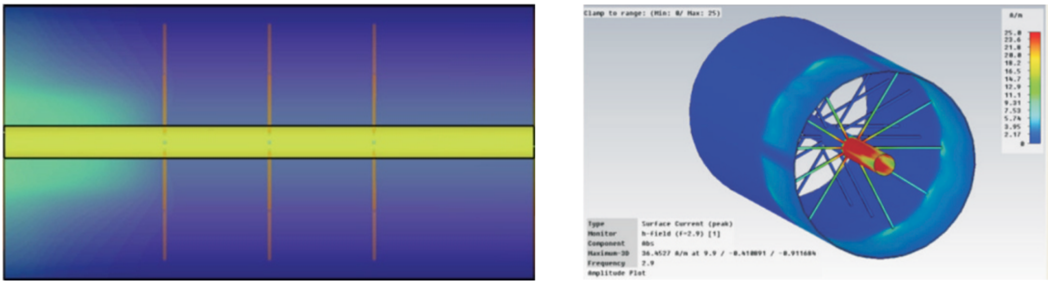


Figure 5: Full wave simulation esults for S21 and S11 parameters in a Coaxial Transmission Line loaded with shunt radial stubs, connecting the ce tral conductor strip with the ground plane. Due to the highly inducti e nature of the radial stubs, this composite structure exhibits high rejection t the previous quasistaaticesonance frequency of 2.9GHz given by the SRR elements.



(a) (b)

Figure 6: Surface current density plots for the Coaxial Transmission Line loaded with radial stubs connecting fom the central conductor to the ground plane. (a) Top view of the surface current plot, (b) volumetric view of the surface current, exhibiting strong reflection t the frequency of 2.9GHz.

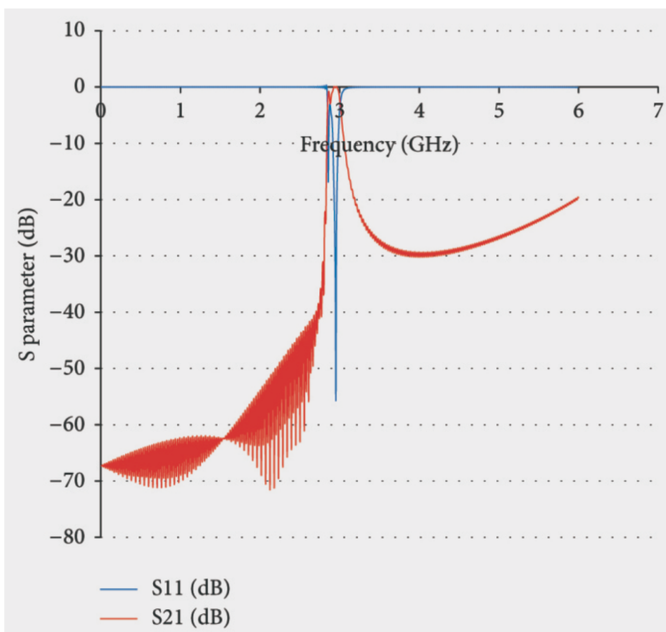


Figure 7: Full wave simulation esults for the combined SRR + radial stub loaded Coaxial Transmission Line. In this case a passband can be observed where initially a topband was given by each individual loading structure.

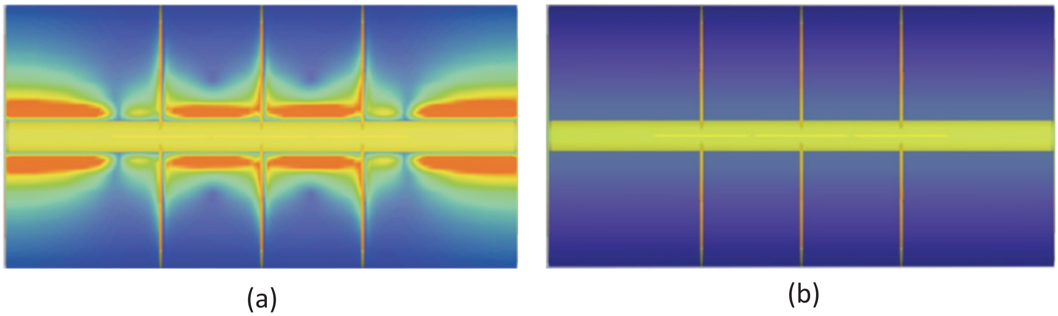


Figure 8: Power flow plots for the composite loaded Coaxial Transmission Line. The leeft image corresponds to a frequency of 2.9GHz,whereas the right image is for a frequency of 1GHz. As it can be seen there is high power transmission in the first case, whereas strong rejection an be observed in the second case.

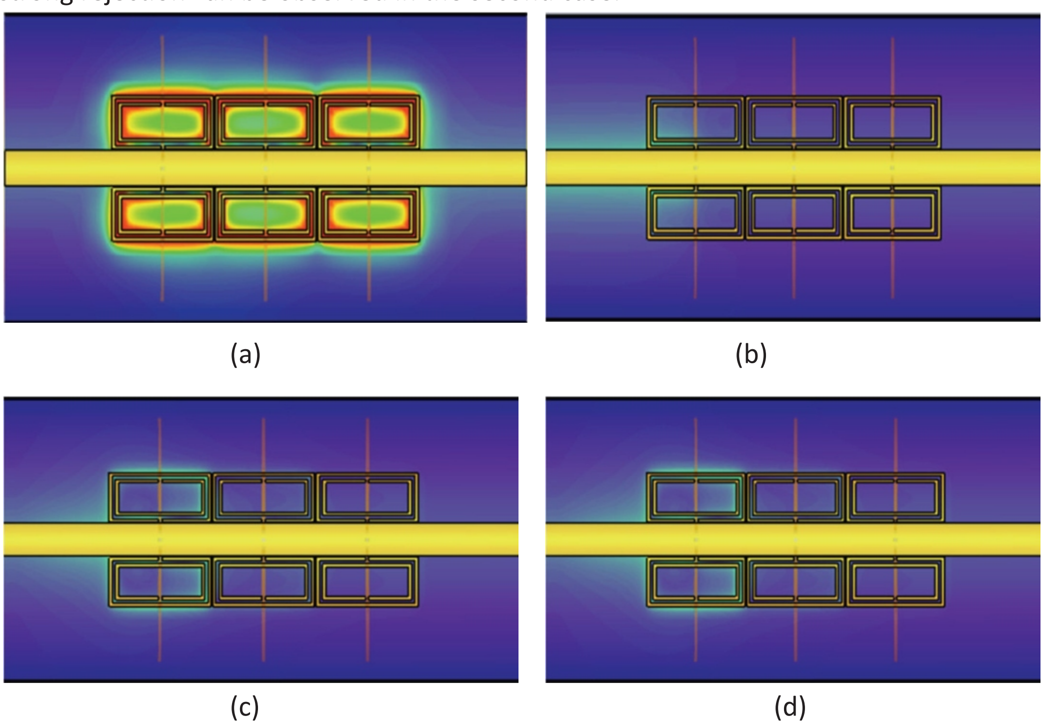


Figure 9: Surface current plots for the composite SRR + radial stub loaded Coaxial Transmission Line. (a) Observation fequency of 2.9GHz,in which quasistaaticesonances can be observed within the composite structure, (b) strong rejection pesent in the observation fequency of 1GHz, (c) observaation equency of 4GHz, and (d) volumetric representation of SRR xcitaationt the quasistatatiesonance frequency of 2.9GHz.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this pape.

References

[1] A. W. Lines, G. R. Nicoll, and A. M. Woodward, "Some properties of aveguides with periodic structure," Proceedings of the IEE, vol. 97, pp. 263–276, 1950.

[2] J. D. Joannopoulos, R. Meade, and J. Winn, Photonic Crystals. Molding the Flow of Light, Princeton University Press, Princeton, NJ, USA, 1995.

[3] V.Radisic,Y. Qian, R.Coccioli, andT. Itoh, "Novel 2-D photonic bandgap structure for microstrip lines," IEEE Microwave and GuidedWave Leters, vol. 8, no. 2, pp. 69–71, 1998.

[4] F. Mart´ın, F. Falcone, J. Bonache, T. Lopetegi, M. A. G. Laso, and M. Sorolla, "New periodic-loaded electromagnetic bandgap coplanar waveguide with complete spurious passband suppression," IEEEMicrowave and Wireless Components Leters, vol. 12, no. 11, pp. 435–437, 2002.

[5] J. B. Pendry, A. J. Holden, D. J. Robbins, andW. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," IEEE Transactions onMicrowave Theory and Techniques, vol. 47, no. 11, pp. 2075–2084, 1999.

[6] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negati e permeability and permittivi ," Physical Review Leters, vol. 84, no. 18, pp. 4184–4187, 2000.

[7] V. G. Veselago, "The electrodynamics of substances with simultaneously negati e values of e and ," Soviet PhysicsUspekhi, vol. 10, pp. 509–514, 1968.

[8] R. Marqu´es, F. Medina, and R. Rafii-El-Idrissi, "Role of bianisotropy in negati e, permeability and lef-handed metamaterials," Physical Review B, vol. 65, no. 14,Article ID144440, 6 pages, 2002.

[9] R. Marqu´es, F. Mesa, J. Martel, and F. Medina, "Comparati e

analysis of edge- and broadside-coupled split ring resonators for meta material design. Theory and Experiments," IEEE Transactions on A tennas and Propagation, ol. 51, no. 10, pp. 2572–2581, 2003.

[10] P. Gay-Balmaz and O. J. Martin, "Electromagnetic esonances in individual and coupled split-ring resonators," Journal of Applied Physics, vol. 92, no. 5, pp. 2929–2936, 2002.

[11] F. Falcone, T. Lopetegi, J. D. Baena, R. Marqu´es, F. Mart´ın, and M. Sorolla, "Efffecte negati ve- top band microstrip lines based on complementary split ring resonators," IEEEMicrowave and Wireless Components Leters, vol. 14, no. 6, pp. 280–282, 2004.

[12] J. Bonache, F.Mart´ın, F. Falcone et al., "Super compact split ring resonators CPW band pass filters," in Proceedings of the IEEE MITT-S International Microwave Symposium Digest, pp. 1483– 1486, Fort Worth, Tex, USA, June 2004.

[13] J. Bonache, I. Gil, J. Garc´ıa-Garc´ıa, and F. Mart´ın, "Novel microstrip bandpass filters based on complementary splitring resonators," IEEE Transactions on Microwave Theory and Techniques, vol. 54, no. 1, pp. 265–271, 2006.

[14] F. Falcone, T. Lopetegi, M. A. G. Laso et al., "Babinet principle applied to the design of metasurfaces and metamaterials," Physical Review Leters, vvol. 93, no. 19, Article ID 197401, 2004.

[15] J. D. Baena, J. Bonache, F. Mart´ın et al., "Equivalent-circuit models for split-ring resonators and complementary splitring resonators coupled to planar transmission lines," IEEE Transactions on Microwave Theory and Techniques, vol. 53, no. 4, pp. 1451–1460, 2005.

[16] H. Salehi and R. Mansour, "Analysis, modeling, and applications of coaxial waveguide-based lef-handed transmission lines," IEEE Transactions on Microwave Theory and Techniques, vol. 53, no. 11, pp. 3489–3497, 2005.

EmbeddingtheRoleofBigDataandWebIntelligence intheAreaofLibraryInformationScience

AnilKumarJharotia
Librarian, Tecnia Institute of Advanced Studies

Dr.JitenderRai
Associate Professor, Tecnia Institute of Advanced studies, Rohini,

Abstract
Libraries, Businesses, industries, governments, universities, scie ntts, consumers, and nonprofits are generating d a at unprecedented levels and at an incredible pace. We are not just talking about gigabytes of data. As we have entered an era of Big Data, processing large volumes of data has never been greater. Through beter Big Data analysis tools like Map Reduce over Hadoop and HDFS, guarantees faster advances in many scientific disciplines and imp ving the profitability and success of many enterprises. However, as the amount of data that need to be processed grows, many data processing methods have become not suitable or limited. This paper focuses on big data processing problems on library and large scale datasets in different domains. Moor's Law provides a simple way to scale your appliccation. It effortlessly scales from a single machine to thousands, providing fault tolerant and high performance.

1. Introduction

Big data is a popular term used to describe the exponential gowth and availability of data, both structured and unstructured. Big data may be important to business – and society – as the Internet has become. Big Data is so large that it's difficult to process using traditional database and soft are techniques. More data may lead to more accurate analyses. More accurate analyses may lead to more confident decision making. And beeer decisions can mean greater operational efficiencies, cost reductions and educed risk.

Analyzing big data is one of the challenges for researchers system and academicians that needs special analyzing techniques. Big data analytics is the pocess of examining big data to uncover hidden paterns, unknown correlations and other useful information th t can be used to make beter decisions.

Big data analytics efers to the process of collectin , organizing and analyzing large sets of data ("big data") to discover paterns and other useful information. Not only will big d ta analytics help ou to understand the information contained within the data, but it will also help identiy the data that is most important to the business and future business decisions. Big data analysts basically want the *knowledge* that comes from analyzing the data.

1.1 Web Intelligence

The study of Web intelligence was first introduced in several papers and books. Broadly speaking, Web intelligence is a new direction or scientific esearch and development that explores the fundamental roles as well as practi cal impacts of artificial intelligence, 1 such as knowledge representation, plannin , knowledge discovery and data mining, intelligent agents, and social network intelligence, as well as advanced information echnology (IT), such as wireless networks; ubiquitous devices; social networks; and data/knowledge grids; and the next generation of Web-empowered products, systems, services, and activities. The Web intelligence technologies revolutionie the way in which information is athered, stored, processed, presented, shared, and used through electoronization, virtualization, globalization, standardization, pesonalizaation, and portals. The new WI technologies will be determined precisely by human needs in a post-industrial era; namely (2):

- information empo erment,
- knowledge sharing,
- virtual social communities,
- service enrichment, and

• Practi al wisdom development. [1] The World Wide Wisdom Web (the Wisdom Web or W4) will become a tangible goal for WI research (1, 5, 6). The new generation of the www will enable humans to gain wisdom of living, working, and playing in addition to information seach and knowledge queries.

1.2 What Is Big Data?

There are many definitions ered for "big data,"one is that it means data that's so large, too fast, or too hard for existing ools to process. Here, "to large" means that organizations increasingly must deal with large scale collections of d ta that come from unstructured data system, transaction histories, sensors, and elsewhere."Too fast" means that not only is data big, but it must be processed quickly — for example, to perform fraud detection at a point of sale or determine which ad to show to a user on a webpage. "Too hard" is a catchall for data that doesn't fit neatly into an existing processing tool or that needs some kind of analysis that existing ools can't readily provide. A similar breakdown is being promulgated by Gartner (which is probably a sign that I'm oversimplifying things),citing the "three Vs" — volume, velocity, and variety (a catchall similar to "too hard"). A common example of big data usage and utiliaation is the eather forecasting of Indian m teorological department, which utilies the data of past 150 years to predict the trends of monsoon season in India.

1.3 Big Data Analytics

Big data is the next generation of d ta warehousing and business analytics. It has many deep roots and many branches. Big Data is defined as a large amount of data which require new technologies and architecture to make possible to extract value from it. Big data analytics [2] efers to the process of collectin , organizing and analyzing large sets of data "big data" to discover paterns and other useful informaation. Not only will big data analytics help you to understand the information contained within the data, but it will also help identiy the data that is most important to the business and future business decisions. Big data analysts basically want the knowledge that comes from analyzing the data.

"Big Data" refers to sets of data that are so large and complex, it is difficult

Figure 1: Big Data, Roots Branches

to use them efffectely and efficiently. In technology, the amount of available information inceases daily. In the next few years, we will be routinely trying to use petabytes of data stored in multiple ormats across different platorms. The volume and diversity of data make it extremely challenging to store, retrieve, analyze and utlize this information. Businesses, overnment agencies, and society needs experts with the skills and knowledge to design, develop and deploy complex information ystems and applicaations that deal with multi-erabyte data sets. The iSchool has created a specialization in the MSIS pogram that emphasizes big data analytics and provides students with the essential in-depth knowledge of techniques and technologies relevant for big data management. Coursework will cover the design and maintenance of infrastructure to efficiently store, easily access, and transfer over wide area networks, extremely large amounts of data.

The Big Data challenge involves three major dimensions: data size, data rate and data diversity: the iSchool's Big Data Analytics specialization will prepare students to address real-life problems along each of those dimensions. For instance, it is not uncommon for digital archives to store terabytes and even petabytes of data in hundreds of data repositories supporting thousands of appli aations. Maintaining such data repositories requires knowledge in ultra-large scale distributed systems, virtualization technologies, cloud computin , unstructured and semi-structured data management, optimiaation thods based on data replication and d ta migraation, as ell as in advanced data protection techniques. The exponential gowth of the amount of data calls for competence in advanced dynamic data processing techniques, including scalable data processing methods and technologies; data stream management; and large-scale process monitoring, modeling and mining. In order to comprehensively analyze such volumes of information fom disparate and various disciplines, information pofessionals will need to master advanced data integration techniques and business intelligence tools, crowd sourcing technologies, large-scale information fusion, d ta-intensive computation and sema ntic data management.

In a Compute rworld article published in November of 2012, the IT employment firm Gartner essti ted that 4.4 million IT jobs will be created in the area of big data between now and 2015. However, Gartner's head of research, Peter Sondergaard, notes a serious shortage of IT professionals with big-data skills: “There is not enough talent in the industry” and that only one-third of the new jobs will be filled. The new Big Data Analyytics specialization will pepare graduates to provide the much-needed expertise to advance this burgeoning field.

Why will there be such a significant need for Big Data Analysts and specialists? Because every industry sector and service entity has o deal with Big Data or can benefit from corralling the power of so much information. Obviously, those who work in data-rich disciplines such as astronomy or fields including online retail would depend on the tools and technologies in Big Data management. However, digital data is everywhere and employers from a wide range of sectors (healthcare, finance, place-based retail, manufacturing, and transportation, o name just a few) will be looking to build workforce capacity to enhance their productivity and competit e position in global markets. [4]

1.4 Data is Everywhere These Days

Libraries, Businesses, industries, governments, universities, scie ntts, consumers, and nonprofits are generating data at unprecedented levels and at an incredible pace. We are not just talking about gigabytes of data. IBM reports that every day we create "2.5 quintillion ytes of data" (www01.ibm.com/soft are/data/big data). In May 2012, Cisco forecast that "annual global IP traffic is forecast to be 1.3 zetabytes--(a zeeabyte is equal to a sextillion bytes, or a trillion gigabytes)." Cisco also predicts that, between 2015 and 2016, global IP traffic will grow to more than 330 exabytes(http://i vestor.cisco.com/re leasedetail .cfm? releaseid=678049). That's almost more data than the two previous years combined. To put it another way, think about McKinsey Global Institue's statement, in its "Big Data: The Next Frontier for Innovation, Competition, and Productivity" eport: "One exabyte of data is the equivalent of more than 4,000 times the i formaationtored in the US Library of Congress"(www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_d ta_The_next_fronntier_for_innovation). Now th t's a lot of data. Big Data is a big deal, and it is a big opportunity for librarians and information pofessionals to play a role in the Big Data universe. Why? Librarians have the skills, the knowledge, and the service mentality to help our businesses, governments, universities, and nonpofits capitalize on all that Big Data has to offer.

2.Data Explosion

Why now for this data explosion? One important reason is the widespread accessibility, affordability, and availability of new digital devices that make access to the internet easy and relati ely inexpensive. Billions of people and the billions of mobile phones, smartphones, tablets, computers, laptops, and other digital devices are able to access the internet and "contribute to the amount of big data available," as The Economist expressed in a 2010 special report, "Data, Data Everywhere" (www.economist.com/node/15557443).

The list of digital sources keeps growing. According to The Economist, the "trail of clicks that internet users leave behind from which value can be extracted-is becoming a mainstay of the internet economy." Consumers generate vast amounts of data every day through email, searching, browsing, blogging, tweetin , buying, sharing, and texting.

This "digital exhaust," data created as a byproduct of other activities, is contributing o the dramaatic inease in digital data. Combine that data with the data from the growing number of embedded networked sensors in cars, highways, and transportation networks; green buildings and the smart grid; the retail and manufacturing of RFID tags that track inventory; and healthcare services; and you can see why data is exploding like never before. It's not just increased amounts and types of data; it's also improved tools to store, aggregate, combine, analyze, and extract new insights. Put Big Data together with big analytics, and it becomes possible to spot business trends, uncover ways to prevent diseases, combat crime, add economic value, gain new insights in scientific research, and make government more transparent.

3. The Era of Big Data

McKinsey refers to Big Data as "datasets whose size is beyond the ability of typical database soft are tools to capture, store, manage, and analyze." Edd Dumbill from O'Reilly Media, Inc. defines it as "data that exceeds the processing capacity of conventional d tabase systems. The data is too big, moves too fast, or doesn't fit the strictures of your database architectures" (http://adar.oreilly.com/2012/01/what-is-big-data.html). Public and private sectors are feeling the pressure to take advantage of all that Big Data has to offer--with the expectation that it will spur new innovations and new product opportunities, achi ve cost savings and efficiencies, and use prpredice analytics. [5]

4. Today's Data Problems

Thanks to this accelerating trend, hospitals, schools, manufacturers, colleges, retailers, government agencies, and libraries have begun to collect and store truly enormous amounts of data. The goal in many cases is to make use of these data to provide valuable new services or to improve efficiency. The problem with reaching these goals is that as the amount of storage and processing has grown, the complexity of the data and the challenges of working with it have also accelerated.

In the good old days a programmer would write a program, a user would use the program, a statistician ould

Figure 2: Big Data Problem

analyze the data that the user produced with the program, and a librarian would archive the report that the statistician eated by analyzing the data. Those days are gone. The reason we now see lots of job advertiseme ts for “data scienntt” is that there is a pressing need for interdisciplinary bridge builders who understand all of the above: the Internet, databases, analytics, visualization, and d ta curaation. These professionals have their specialties – some are good at working with numbers, others are database experts, still othes have eexpertise in unstructured data (e.g., text) – but they also need generalist skills that let them blend the wide range of methods needed to manage today's data problems. [3]

5.Where Does the New Librarian Fit In?

Librarians have always been great at information management and organization. This is a ore skill in data science; it manifests most strongly in the data curation omponent of the big data problem. Many librarians are also outstanding communicators and have been trained in the art and science of transforming user information needs i to strategies and resources for investi aation and learning. So librarians clearly have roles at the start and the finish of the big data problem. But what about the middle of the equation, whee data transformation, analysis, and visualization ae the heart of the data science endeavor? This brings us back to our original question of how libary science and data science are connected.

The essential ask of the data science professional is to transform raw, messy data into actionable knowled e that can be used by decision makers. To paraphrase my astute colleague R. David Lankes, 'the mission of librarianship is to facilitate knowledge creation in communities.' It is e y to see the overlap here. A librarian does not need to become a programmer, but every librarian interested in knowledge creation should h ve some essential amiliarity with how various soft are tools can transform data. A librarian need not be a database engineer, but every librarian must understand the underpinnings of information etrieval tools. A librarian does not need to be a statistician, but every librarian should have a clear understanding of how descripti e summaries and basic tests of numeric data can be used and misused. Finally, a librarian does not need to be a graphic designer, but every librarian needs to recognize the features of efffecte data displays. In short, to fulfill their missions, librarians *can* exercise a range of sophisti ated skills that squarely occupy the central ground between understanding information user needs on one end and data curation on the othe.

When you consider some of the key values that drive librarianship, however, it becomes evident that librarians *must* take a leading role in working with big data lest this emerging specialty become the servant only of proprietary interests. Librarians stand for open access to information, or privacy rights, for serving the information needs of the community, for the importance of accurate information in a democaatic society, and for the necessity of preserving the legacy of historical information or future generaations. Public library users, students in school libraries, and faculty and students in university libraries all depend upon these bedrock values to support their missions of learning, exploration, and citienship. We've known for quite a while that fulfilling these missions requires much more than choosing, shelving, and lending books. In the near future, the ability to fulfill the roles of citienship will require finding, joining, examining, analyzing, and understanding diverse sources of data. For a citien to become an efffecte advocate tomorrow, she might need to "mash-up" map data, census data, health data, and environmental data to develop a meaningful understanding of a challenge that the community faces. Who but a librarian will stand ready to give the assistance needed, to make the resources accessible, and to provide a venue for knowledge creation when the community advocate arrive seeking answers? [3]

6.Moore's "Law"

Many people are familiar with the idea of Moore's Law that the amount of computer processing power per chip doubles every 18 months. A similar idea, promoted by Mark Kryder, the former chief technology officer of Seagate (manufacturer of hard disk), suggests that the amount of data storage one can fit on a given area of a magnetic medium*also* doubles every 18 months. What is little undestood about these “laws” is that doubling in a fixed amounnt of time eates an accelerating trend that starts o ffslow but eventually reaches a tipping poi t of massive growth. Between 2005 and the present we leaped from reasonably affordable disk drives that could comfortably hold all of your family photos to online “cloud” storage sufficient t to digie a whole floor of library books that is available *completely free* to anyone with a computer and an Internet connection. Today, for less than the price of a nice meal at a fancy restaurant one can buy a hard disk drive that has sufficient storage to hold the entie printed collection of the Libary of Congress.

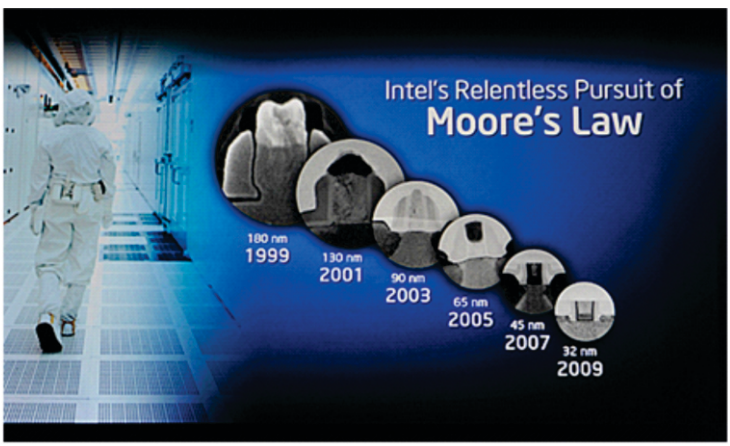


Figure. 3 : Moore's Law

7. Hadoop

Hadoop is open source soft are which made to handle big data. Hadoop was developed by Daug Cutting in 2005. Doug's son ve an yellow toy elephant he called it as Hadoop and Doug's like that name so Dougs name it Hadoop. It was developed by Apache Soft are Foundation. Its f st release was in 10 December 2011 and latest version 2.6.0 was release in 18 November

2014. It is writen in Java language. It has Apache License 2.0. The apache Hadoop project develops open source soft are for reliable, scalable, distributed computing. The apache Hadoop soft are library is a frame work that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single services to thousands of machines, each offering local computation and torage. Rather than rely on hardware to deliver high availability, the library itself is designed to detect and handle failures at the application l yer, so delivering a highly- available service on top of a cluster of computers, each of which may be prone to failures.

Hadoop includes following modules:

1. Hadoop Common – The common utilities that support the other hadoop modules.
2. Hadoop Distributed File System (HDFS) - A distributed file system that provides high throughput access to application d ta.
3. Hadoop Yarn – A framework for job scheduling and cluster resource management.
4. Hadoop Map Reduce: A yarn based system for parallel processing of large data sets. (hadoop)(3).

7.1 Need of Hadoop in Library- IBM claims 90% of today's stored data was generated in just the last two years. (hlideshare website)(1) If we think about library the situation will be same. In present era every library is on face book, twiter and on mobile also. In next few years maximum library resources will be published in electronic form. When the unstrcuchers data and structured data will be large then it is very difficult to find particular information in the seconds. So to handle this situation e required a big data manager in library which can retrieve information in seconds and can manage all unstrctuchered and structured data of the library. I think Hadoop is the best solution o overcome this problem.

7.2. Source of Big Data in Library

In library there are four sources of Big Data User Library data Sensors System.

7.3 Components of Hadoop

Hadoop have two major components – Distributed file system: It splits up large files intto multipleomputers.

Execution Engine (Map reduce framework): It is a framework used to process large data stored on the file system

7.4 Featurs of Hadoop

1. Large Data Sets –Map-reduce paired with HDFS is a successful solution or storing large volumes of unstructured data.
2. Scalable Algorithms –Any algorithm that can scale too many cores with minimal inter-process communication will be able to exploit the distributed processing capability of Hadoop.
3. Log Management –Hadoop is commonly used for storage and analysis of large sets of logs from diverse locations. Because of the distributed nature and scalability of Hadoop, it creates a solid platorm for managing, manipulatin , and analyzing diverse logs from a variety of sources within an organization.
4. Extract-Transform-Load (ETL) Platorm –Many companies today have a variety of data warehouse and diverse relational database management system (RDBMS) platorms in their IT environments. Keeping data up to date and synchronized between these separate platorms can be a struggle. Hadoop enables a single central location.

7.5 Hadoop is Work On

Hardware Failure one consequence of scale is that hardware failure is the norm rather than the exception. An HDFS instance may consist of hundreds or thousands of server machines, each storing part of the file system's data. The fact that there are a huge number of components and that each component has a non-trivial probability of failure means that some component of HDFS is almost always behaving badly. Even with RAID devices, failures will occur frequently. Therefore, detection of faults and quick, automatic ecovery from them is a core architectural goal of HDFS.

Streaming Data Access Applications that run on HDFS need streaming access to their data sets. They are not standard applications th t typically run on general purpose file systems. HDFS is designed more for batch processing rather than interacti e use by users. The emphasis is on high throughput of data access rather than low latency of data access. POSIX imposes many hard requirements that are not needed for applications th t are targeted for HDFS. POSIX semantics in a ew key areas have been relaxed to gain an increase in data throughput rates. Large Data Sets Applications th t run on HDFS have large data sets. A typical file in HDFS is gigabytes to terabytes in size. Thus, HDFS is tuned to support large files. It should provide high aggregate data bandwidth and scale to hundreds of nodes in a single cluster. It should support tens of millions of files in a single instance. Simple Coherency Model HDFS applications need a write-once-read-many access model for files. A file once creatated, wren, and closed need not be changed except for appends. This assumption simplifies data coherency issues and enables high throughput data access. A Map Reduce application or a eb crawler application fits pe ectly with this model.

Conclusion

Big Data is a big deal, and it is a big opportunity for librarians and information pofessionals to play a role in the Big Data universe. Librarians have the skills, the knowledge, and the service mentality to help our businesses, governments, universities, and nonpofits capitalize on all that Big Data has to offer. Big Data analysis tools like Map Reduce over Hadoop and HDFS, promises to help organizations b eer understand their customers and the marketplace, hopefully leading to beter business decisions and competit e advantages. For engineers building information processing tools and applications, lage and heterogeneous datasets which are generating continuous flow of ta, lead to more efffecte algorithms for a wide range of tasks, from machine translation o spam detection. The big d ta model has the flexibility to be expanded to incorporate more sophisti ated additional actors if needed.

References

[1]. J. Liu, N. Zhong, Y. Y. Yao, Z. W. Ras, The wisdom web: new challenges for web intelligence (WI), J. Intell. Inform. Sys.,(1): 5–9, 2003.

[2]. Jianqing Fan1, Fang Han and Han Liu, Challenges of Big Data analysis, National Science Review Advance Access published February, 2014.

[3]. http://i fospace.ischool.syr.edu /2012/07/16/data-science-whats-in-it-for-the-new-librarian/

[4] http://ww w.ischool.pitt.edu/t /degrees/specializations/big-data.php

[5]. htt s://www.quesstiaom/mag azine/1G1-301969491/big-data-a-big-opportunity-for-librarians

Instructions for Authors

Essentials for Publishing in this Journal

- Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.
- Conference papers may only be submitted if the paper has been completely re written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.
- All our articles are refereed through a double-blind process.
- All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page

Type of Articles

Classification of articles is a duty of the editorial sta ffand is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).
2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);
3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);
4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries.

   Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones).

All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250- Word abstract should be placed between the title and the keywords with the body text to follow. Besides a abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgments

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the tex and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.

The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.