David C. Wyld
Jan Zizka (Eds)

Computer Science & Information Technology

International Conference on Computer Science, Engineering and Information Technology (CSITY 2015)
Bangalore, India, February 14 ~ 15 - 2015

AIRCC

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Jan Zizka,
Mendel University in Brno, Czech Republic
E-mail: zizka.jan@gmail.com

ISSN: 2231 - 5403
ISBN: 978-1-921987-31-1
DOI : 10.5121/csit.2015.50301 - 10.5121/csit.2015.50304

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

International Conference on Computer Science, Engineering and Information Technology (CSITY-2015) was held in Bangalore, India, during February 14~15, 2015. International Conference on Signal and Image Processing (SIGPRO-2015) and International Conference on Data Mining (DTMN-2015) were collocated with the CSITY-2015. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The CSITY-2015, SIGPRO-2015, DTMN-2015 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, CSITY-2015, SIGPRO-2015, DTMN-2015 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the CSITY-2015, SIGPRO-2015, DTMN-2015.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld
Jan Zizka

Organization

General Chair

Natarajan Meghanathan	Jackson State University, USA
Dhinaharan Nagamalai	Wireilla Net Solutions PTY LTD, Australia

Program Committee Members

Abd El-Aziz Ahmed	Cairo University, Egypt
Abhishek Narkhede	Saraswati Vishwa, India
Adnan H. Ali	Institute of Technology, Iraq
Ahmed Y. Nada AQU	Al-Quds University, Palestine
Amir Khusru Akhtar	Cambridge Institute of Technology, India
Amit Choudhary	Maharaja Surajmal Institute, India
Anamika Ahirwar	Rajiv Gandhi Technical University, India
Anil Dubey	Government Engineering College, India
Archana Kakade	G.H.Raisoni College of Engineering, India
Asha.K	Higher College of Technology, UAE
Chanabasayya Vastrad	Mangalore University, India
Chetan S	Ambedkar Institute of Technology, India
Chiranjib Sur	University of Florida, US
Dac-Nhuong Le	Haiphong University, Vietnam
Debasis Giri	Haldia Institute of Technology, India
Deepak Dembla	JECRC University, India
Diptoneel Kayal	West Bengal University of Technology, India
Gaurang Panchal	Indian Institute of Technology, India
Giovanni Cordeiro Barroso	Federal University of Ceará, Brasil
Indrani Das	Assam University, India
Isa Maleki	Islamic Azad University, Iran
Islama	Alexandria university, Egypt
Jyoti Bajpai	GLA University, Mathura
Keneilwe Zuva	University of Botswana, Botswana
Laudson Souza	Integrated Faculties of Patos, Brazil
Long Chen	Lan zhou University of Technology, China
Marish Kr.Singla	Guru Nanak Institute of Technology, India
Narendra Kr Pareek	M L Sukhadia University, India
Mohamed Ben Ahmed	LIST laboratory, Morocco
Quan Yuan	University of Texas, USA
Raja Kumar Murugesan	Taylor's University, Malaysia
Ramayah T	Universiti Sains Malaysia, Malaysia
Ramesh C. Poonia	Amity University, India
Raveendra K Rao	University of Western Ontario, Canada
Reda Mohamed Hamou	Dr Tahar Unversity of Saïda, Algeria
Saad Mohamed Saad Darwish	University of Alexandria, Egypt
Sattar B.Sadkhan	IEEE Iraq Section, Iraq

Seyyed AmirReza Abedini	Islamic Azad University, Iran
Seyyed Mohammad Reza Farshchi	Tehran University, Iran
Shekoofeh Ghiam	Sharif University of Technology, Iran
Shuiyin Liu	Monash University, Australia
Sonia Pascua	University of the Philippines, Philippines
Subarna Shakya	Tribhuvan University, Nepal
Vasanth Ram Rajarathinam	Advanced Micro Devices, USA
William R Simpson	Institute for Defense Analyses, USA
Yasser Alginahi	Taibah University, Saudi Arabia
Yasser Hashemi	Islamic Azad University, Iran
Yung-Fa Huang	Chaoyang University of Technology, Taiwan

Technically Sponsored by

Networks & Communications Community (NCC)

Computer Science & Information Technology Community (CSITC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

International Conference on Computer Science, Engineering and Information Technology (CSITY 2015)

International Conference on Signal and Image Processing (SIGPRO 2015)

International Conference on Data Mining (DTMN 2015)

ADAPTIVE RESOURCE ALLOCATION AND INTERNET TRAFFIC ENGINEERING ON DATA NETWORK

Hatim Hussein

Department of Electrical and Computer Engineering,
George Mason University, Fairfax, Virginia, USA
hhussei1@gmu.edu

ABSTRACT

This research paper describes the issues of bandwidth allocation, optimum capacity allocation, network operational cost reduction, and improve Internet user experience. Traffic engineering (TE) is used to manipulate network traffic to achieve certain requirements and meets certain needs. TE becomes one of the most important building blocks in the design of the Internet backbone infrastructure. Research objective: efficient allocation of bandwidth across multiple paths. Optimum path selection. Minimize network traffic delays and maximize bandwidth utilization over multiple network paths. The bandwidth allocation is performed proportionally over multiple paths based on the path capacity.

KEYWORDS

Network Protocols, Internet, Network, Traffic Engineering, Network, Bandwidth, Allocation, Multipath

1. INTRODUCTION

In the last decade, business usage of the Internet has gone exponentially [1]. Taking North America as an example, 79% of the population is Internet users. North American usage growth was 152% between the years 2000 and 2011, compared with the world growth which was 32% during the same period. As result of these dramatic changes, it is no surprise that the Internet has become one of the major use and research area in the last decade. Research in the areas of hardware and application development, users are demanding expecting reliable throughput and trusted transport from their service providers. In addition to traditional Internet application, new multicast, multimedia and voice service applications are in the rise. New applications are developed constantly. These applications have increased the demand for bandwidth support and dictate the need for newer services. Along with the exponential growth of the Internet, these new services place ever-increasing strain on the existing resources. The Internet Protocol (IP) [2] has proven to be limited in scope of provide the functions necessary for today's Internet application and data demand. The network performance and resource allocation issues spawned the need for traffic engineering.

David C. Wyld et al. (Eds) : CSITY, SIGPRO, DTMN - 2015
pp. 01–07, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50301

2. ISSUES OF BANDWIDTH ALLOCATION

Network traffic delays and congestion are two of the main issues that faces today's Internet. Hybrid and adaptable bandwidth allocation can be seen as one solution to this type of problem. This issue will be addressed in this paper through traffic engineering based on link capacity, bandwidth availability, jitter, and other network configuration variables.

Equal and unequal traffic distribution across multi-paths should be based on the dynamic and changing network needs and recoverable network resources. For instance, a single MPLS network ingress LSR is connected an egress LSR across n parallel LSRs. Once network traffic enters at the ingress LSR, packets are typically distributed among available links and multiple paths based on the underlying link capacity, speed, throughput, and other network parameters.

As the network state changes, so does traffic distribution and bandwidth utilization. Bandwidth allocation can be discussed in two broad categories:

a) Allocation of network traffic with equal payload distributed across n multiple paths
b) Allocation of network traffic with unequal payload distribution across n multiple paths.

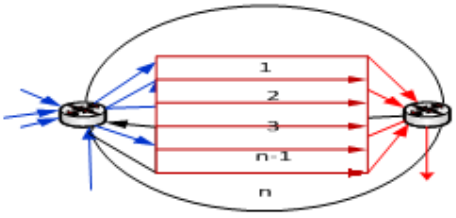

Figure 1: Bandwidth Allocation across an MPLS Network

3. RELATED WORK

3.1. Traffic Engineering:

Traffic Engineering (TE) has become the primary method of Internet traffic treatment [3]. Boarder Gateway Protocol (BGP) [4] as the main Internet Exterior Gateway Protocol (EGP) is responsible for transmitting packets across areas and autonomous systems (AS). BGP doesn't support traffic engineering as it becomes a necessity to optimize the Internet backbone. Traffic engineering can be a manual or automated process through a number of network resource controls [5] such as data resources, control mechanisms, and management tools. Initially, traffic engineering was implemented in IP networks, now it becomes part of the MPLS domain. TE is used to manipulate network traffic to achieve certain requirements. For instance, network utilization of a link is determined by the ratio of used bandwidth in relation to the allocated bandwidth. This can be accomplished by a uniform distribution of network traffic across the network. On the other hand, traffic engineering may be implemented to optimize scarce resources utilization. As one of traffic engineering objectives is to allocate available link bandwidth in relation to the required connection, avoiding congested and oversubscribed links. Load balancing will be handled in both resource allocation and optimal routing choices. [6].

Traffic engineering extensions were implemented in some of the link state routing protocols such as OSPF-TE and Resource Reservation Protocol (RSVP) [7] [8]. Bandwidth Allocation for Minimizing Resource Usage with Restoration (BAMRUR) introduced in [9] claims to obtain an optimal set of valid paths and traffic distribution by linear programming. BAMRUR objective

function is to minimize the bandwidth assigned to a path with certain constraints related to the bandwidth assigned to an active path and the total bandwidth needed.

3.2. Multi-protocol Label Switching (MPLS)

MPLS addresses network routing issues, scalability, and network performance [10]. MPLS works with heterogeneous network infrastructure such as IP backbone networks, Asynchronous Transfer Mode (ATM) networks, and other technologies. MPLS maps an IP address to a fixed length tag known as a label for network packet forwarding. It works with protocols at both the data link and the network layers of the OSI Model. An MPLS router is called a Label Switch Router (LSR); it inspects the label and the additional fields in forwarding the packet.

At the ingress LSR of an MPLS domain, IP packets are routed based on the information carried in the IP header.

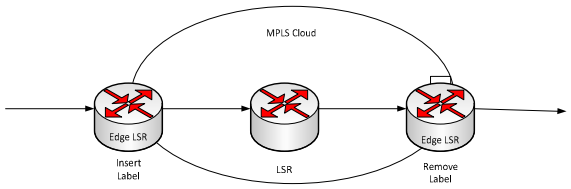

Figure 2: MPLS Network Framework

3.3. Equal Cost Multipath (ECMP)

ECMP, two well-known routing protocols namely; Open Shortest Path First (OSPF) [11] and Intermediate System to Intermediate System (IS-IS) [12] support Equal-Cost Multipath (ECMP) [13]. ECMP is a routing protocol for transmitting data packets across multi-paths with equal cost from source to destination.

Load balancing over multiple paths is one of the mechanisms that Internet Service Providers use to balance network payload. Load balancing has several benefits: it helps in capacity planning, it reduces traffic congestions, and it offers a reliable fault tolerance, since network traffic can pass across alternative multiple paths [14]. ISPs are striving to provide the maximum available bandwidth between networking nodes.

Load balancing requires the use of a key inserted in an available field in the packet. Finding the right field in a packet to use for load balancing is difficult. In the past, the extra encapsulation required fairly deep packet inspection to identify the right field at every hop that the packet traverses. The Entropy Label [14] concept was introduced to eliminate the need for deep packet inspection. The key information would be extracted once, at the entry of the MPLS LSP, and encoded within the label stack itself. The benefits of the introduction of the entropy label are discussed in [14]. Figure 3 shows load balancing of network payload between node *X* and node *Y*. The payload enters the network via the Ingress LSR and is load balanced to the next two LSRs, and then to the next three LSRs across the MPLS network until it reaches to the destination node *Y*.

For all the LSRs in the path, the ingress LER receives the incoming flow with the most contexts, for instance incoming traffic trunk for an IP or L2TP tunnel to be carried over an LSP. Packet data payload is beyond the TCP ports, which is where the deep inspection is needed.

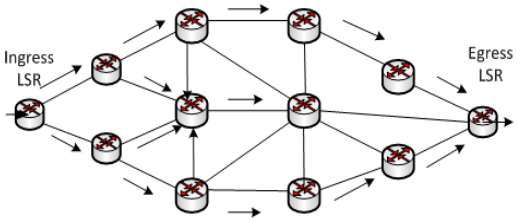

Figure 3, ISP network with load balancing from the Ingress to Egress LSR

3.4. Resource Allocation and Traffic Engineering

With today's Internet traffic demand and the need for efficient and optimize bandwidth utilization, multipath routing and optimal resource allocation should be considered for extensive research in the coming years. The Internet infrastructure should be the primary source of data for this type of investigation. The development and the implementation of multipath routing involve four major steps:

1. Find the shortest path or paths between two end-nodes,
2. Find the set of paths that meets certain requirements.
3. Find the capacity needed to handle network traffic flow between the two end nodes.
4. Allocate network traffic flow among the set of multiple paths.

This model aims to introduce proportional and efficient bandwidth allocation scheme to distribute network traffic flows across multiple paths with the objective to minimize network congestion and delays and maximize links bandwidth utilization.

Given available network capacity, in a traditional network setting traffic would take the shortest path from the source node to the destination node. With the development of the MPLS technologies in today's Internet demands and to support needed features such as load balancing and fault tolerance, many service providers have implement multipath technologies such as ECMP. Such techniques route traffic across multiple paths but doesn't take into account network congestions and network bottlenecks. Such problem may leads to underutilization and over unitization of link's capacities which leads network performance degradation and inefficient use of network resources.

This research will focus on resolving this problem by allocating network traffic flow across multiple paths based on the link's capacity and taking into account bandwidth availability and network congestion conditions.

1. A source *s* receives network traffic flow sourced from node *s* heading to destination *d*.
2. *s* has multiple paths to destination *d*.
3. Each of these links may have different bandwidth capacity *C*. This research assumes multiple links may have identical or different bandwidth capacities. More emphasis will be put on the latter.
4. The research environment assumes a packet-switching communication network such as a WAN network or an Internet backbone.
5. Two types of broadband links can be assumed here: (a) The WAN side connections can be configures as multilink PPP [15], to combine multiple T1 circuits (1.5 Mbps) bonded as one logical connection. Theoretically a single customer premise equipment (CPE) can handle up to 12N x T1. Which can build 3 set of multilinks - MU1, MU2, and MU3 - with 3MB, 6MB, and 8MB of bandwidth capacity respectively.

(b)The network backbone with high speed optical fiber network links carried on SONET fiber optic networks transmission rate based unit is 51.84 Mbit/s. The speed ranges from OC1 (51.84 Mbit/s), OC3 (155.52 Mbit/s, OC12 (622Mbit/s), OC24, etc.

6. Multiple links and multipath routing exists between source/destination (*s/d)* pair, where the data packets traverse the network via intermediate nodes that connect the *s/d* pair.
7. Network traffic control, bandwidth capacity allocation, and user request distribution is managed at the ingress of the network.
8. User data transmission requests enter the network through the source node. All requests are buffered in store-and-forward fashion.
9. Connection-oriented mechanism is used to support network connectivity between nodes.

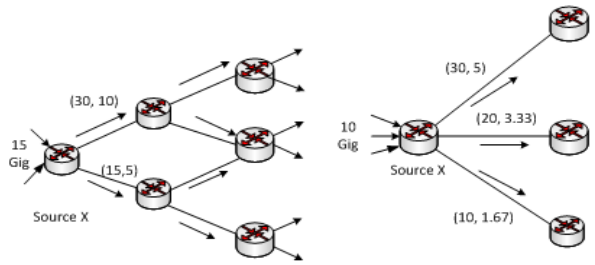

Figure 4: Network Traffic Flow Prototype

3.5. Network Optimization Construction

Capacity assignment problem and flow assignment problem. Both problems have constraints to satisfy and objective function to attain.

Capacity assignment problem (CA), chooses the optimum link capacity for an incoming given network traffic flow that would minimize cost and satisfy user's network payload request.

Traffic Flow assignment problem (TFA) objective is to match the user traffic flow with the available link's capacity. To minimize traffic delays for that user's network payload flow request. The question is how to select the optimum path [links] capacity that meets the user flow requirements for efficient traffic flow and minimize traffic delays.

Generally Link Available Capacity can be defined in the following simple equation:

$$LAC = TLC - LRC - LBC$$

Where:

LAC: Link Available and Usable Capacity.
TLC: Total Link Capacity
LRC: Link Reserved Capacity
LBC: Link Busy Capacity.

In a multipath environment the objective is to assign network traffic flow (User request) to the appropriate capacity. The user's traffic need to be optimized and mapped across multiple paths.
In general, network components are: source node *s*, destination node *d*, link *l* between *s* and *d* with intermediate nodes *i*. multiple paths may exist between every *s and d* pairs.

The average traffic flow between *s-d* is (Υ_{sd}/μ) should be more than link's capacity *C*.
Let π_{sd} denote the path between *s-d*

Υ_{sd}= User's traffic flow between nodes *s-d* (packet per second)

λ = Average number of messages flow per second traverse the i^{th} channel.

The sum of all the network traffic packets traverse the network is the summation of the average number of message flow traverse the network [16].

For simplicity we consider all delays equal *0*. In real network design all delays should be taken into account and factored into the capacity equation.

The total user traffic flow that originates from source node *s* to destination node *d*, traffic flows takes Poisson distribution process as presented in [16].

All network incoming traffic messages are assumed to be independently drawn and take exponential distribution function with the mean of $1/\mu$ (bits) X_{sd} represents expected network delay that affects network performance or message sourced from node *s* with node *d* as the destination.

Given the cost of the network *Q* (dollar), i.e. the cost of network links, equipment, and channel fees ... etc

$$Q = \sum_{i=1}^{M} q_i\,(C_i) \quad [16] \; (2)$$

Network delays represent the total time that takes network packet to traverse the network.

Network delays such as: transmission delay, propagation delay and queuing delay … etc.

Let assume that R_{sd} is the average for a network packet to traverse the network form source *s* to destination *d*.

Given the network capacity *C,* we are interested in finding the optimum message flow interval for each connection that would achieve the minimum average delay.

This research will take this optimization problem further by mapping network traffic flow based on network capacity and availability.

With a predetermined network channel capacities we need to find the optimum traffic flow to minimize the average network delay. So traffic flow sourced from *s* and to *d* needs to be adjusted and distributed across *n* paths. That implies the average available channel capacity C_{sd} must be more than the average traffic flow (user request) λ_{sd}.

To allocate network bandwidth across multiple paths, this paper proposed the following approach: First, calculate the k-shortest path from source node to the destination node.

Second, then calculate the Constraint Shortest Path first (CSPF).

Third, choose the paths that meet the bandwidth constraints requirements.

Forth, an algorithm will calculate the network traffic requirements.

Fifth, another algorithm will distribute the network traffic requirements across the chosen paths.

Sixth, the network traffic requirements will be distributed across the chosen paths proportionally.

Seventh, distribute network traffic payload across paths proportionally based on links bandwidth capacity.

4. CONCLUSIONS

The general overall objective of corporate network management is to minimize the cost function on network spending, seeks efficient traffic flow and minimizes network traffic delays, and optimized network performance. The goal of this research is to work toward achieving some of these novel objectives. In summary, this research uses the k-shortest path to identify the shortest paths. Implement the Constrained-base Shortest Path First (CSPF) to choose those shortest paths meet certain requirements. Allocate the network payload across the multiple paths proportionally base on each path capacity.

This framework can be implemented in network fault tolerance, load balancing, and MPLS traffic engineering with Entropy Label [14].

In future papers I will describe in details the framework building blocks and provide numerical examples of the framework implementation.

REFERENCES

[1] Internet World Stats, Usage, and Population Statistics http://www.internetworldstats.com/stats.htm
[2] DARPA Internet Program, Protocol Specification, Internet Protocol, RFC 791, September 1981.
[3] http://techtarget.com/definition/traffic-engineering
[4] Y. Rekhter, Ed, T. Li, Ed., S. Hares, Ed., A Border Gateway Protocol 4 (BGP-4), RFC 4271, January 2006
[5] Technical Report, UCAN-CL-TR, University of Cambridge, Computer Laboratory: http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-532.pdf
[6] I. Cidon, R. Rom, Y. Shavitt. Multi-Path Routing combined with Resource Reservation; INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE Volume: 1
[7] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE: Extensions to RSVP for LSP Tunnels, RFC 3209, December 2001.
[8] D. Katz, K. Kompella, and D. Yeung. Traffic Engineering (TE) Extension to OSPF version 2. RFC 3630, September 2003
[9] X. Yu, G. Geng, K. Gay, and C. Siew, An Integrated Design of Multipath Routing with Failure Survivability In MPLS Networks, IEEE Computer and Communications Societies. Proceedings IEEE Volume: 1, 2004
[10] A. Viswanathan E. Rosen and R. Callon. Multiprotocol Label Switching Architecture.RFC 3031, January 2001
[11] J. Moy, Open Shortest Path First (OSPF Version 2), RFC 1247, July 1981.
[12] JP. Vasseur, N. Shen, R. Aggarwal, Intermediate System to Intermediate System (IS-IS) Extensions, RFC 4971, July 2007.
[13] C. Hopps, Analysis of an Equal-Cost Multi-Path Algorithm, RFC 2992, November 2000.
[14] K. Kompella, J. Drake, S. Amante, W. Henderickx, and L. Yong, The Use of Entropy Labels in MPLS Forwarding , draft-ietf-mpls-entropy-label-00.
[15] The Multi-Class Extension to Multi-Link PPP, C. Bormann, RFC 2686, September 1999.
[16] L. Kleinorck, Queuing Systems, Volume II: Computer Application, Wiley-Interscience Publication, New York 1976.

INTENTIONAL BLANK

ADAPTIVE LIFTING BASED IMAGE COMPRESSION SCHEME USING INTERACTIVE ARTIFICIAL BEE COLONY ALGORITHM

Vrinda Shivashetty[1] and G.G Rajput[2]

[1]Department of Computer Science, Gulbarga University Gulbarga, India
mbvrinda2004@yahoo.co.in
[2]Department of Computer Science,
Rani Channamma University, Belagavi, India
ggrajput@yahoo.co.in

ABSTRACT

This paper presents image compression method using Interactive Artificial Bee Colony (IABC) optimization algorithm. The proposed method reduces storage and facilitates data transmission by reducing transmission costs. To get the finest quality of compressed image, utilizing local search, IABC determines different update coefficient, and the best update coefficient is chosen optimally. By using local search in the update step, we alter the center pixels with the co-efficient in 8-different directions with a considerable window size, to produce the compressed image, expressed in terms of both PSNR and compression ratio. The IABC brings in the idea of universal gravitation into the consideration of the affection between onlooker bees and the employed bees. By passing on different values of the control parameter, the universal gravitation involved in the IABC has various quantities of the single onlooker bee and employed bees. As a result when compared to existing methods, the proposed work gives better PSNR.

KEYWORDS

IABC, Image Compression, Wavelet Transform, Adaptive Lifting Scheme, PSNR.

1. INTRODUCTION

The wavelet coding method has been recognized as an efficient coding technique for lossy image compression. The wavelet transform decomposes a typical image data to a few coefficients with large magnitude and many coefficients with small magnitude. As most of the energy of the image concentrates on these coefficients with large magnitude, lossy compression systems just by using coefficients with large magnitude can realize the reconstructed image with good quality and high compression ratio. For wavelet transforms, Lifting scheme(LS) allows efficient construction of the filter banks. The restriction of this structure is that the filter structure is fixed over the entire signal. In many applications to shape itself to the signal it is very much desirable to design the filter banks. A number of such adaptive Lifting Schemes are proposed in the literature[12,14] which consider local characteristics of the signal for adapting. In this paper, image compression using IABC is proposed based on intelligent behavior of Honey bee swarms [8]. The paper is described as follows. In section II Compression techniques discussed, In section III a general lifting scheme is discussed and compared with the adaptive lifting scheme where the update step

David C. Wyld et al. (Eds) : CSITY, SIGPRO, DTMN - 2015
pp. 09–21, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50302

is modified with the IABC algorithm. Section IV discusses about the proposed work. Section V explains the IABC algorithm and section VI describes the proposed algorithm.

2. COMPRESSION TECHNIQUES

The image compression techniques are generally classified into two categories depending whether or not an exact replica of the original image could be reconstructed using the compressed image. These are:

1. Lossy technique
2. Lossless technique

1. Lossy Compression Techniques

Lossy schemes provide much higher compression ratios than lossless schemes. Lossy schemes are widely used since the quality of the reconstructed images is adequate for most applications. By this scheme, the decompressed image is not identical to the original image, but reasonably close to it. The most popular current lossy image compression methods use a transform-based scheme.

2. Lossless Compression Techniques

In lossless compression techniques, the original image can be perfectly recovered from the compressed image. These are also called noiseless since they do not add noise to the signal. It is also known as entropy coding since it use decomposition techniques to minimize redundancy.

3. LIFTING SCHEME

Lifting scheme is used to implement critically sampled filter banks which have integer output. The lifting scheme can custom design the filters, essential in the transform algorithms. Independent of translating and dilating, needless of frequency analysis lifting scheme is processed into space domain. An answer to the algebraic stage of wavelet construction is provided by lifting scheme, which leads to a fast in-place calculation of the wavelet transform, i.e. it does not require auxiliary memory. Different wavelets show different image compression effect; the compressed image quality and the compression rate is not only relational to the filter length, but also concerns with regularity and local frequency, vanishing moment, orthogonality, biorthogonality. In this paper, we implement adaptive lifting scheme based upon wavelet decomposition. Then, with the help of IABC algorithm, we find the best directional window size to get better compression ratio with considerable quality.

A. The Lifting Concept

Lifting is a spatial (or time) domain construction of bi-orthogonal wavelets. The lifting scheme procedure consists of three steps: Split, Predict and Update (Fig. 1) and inverse Lifting scheme is shown in Fig. 2.

Split:

Split the original data into two disjoint subsets. Though any disjoint split is possible, in the standard lifting scheme we split the original data set x[n] into the even indexed points, xe[n]-x[2n], and the odd indexed points xo[n]=x[2n+1]

Predict:

Generate the detail signals d[n] as the predicting error using prediction operator P

$$d[n] = xo[n] - P(xe[n]) \quad (1)$$

Update:

To obtain scaling coefficients c[n] that represent a coarse approximation to the original signal x[n] merge xe[n] and d[n]. This is accomplished by applying an U update operator to the wavelet coefficients and adding to xe[n].

$$c[n] = xe[n] + U(d[n]) \quad (2)$$

The above three steps is described as lifting stage. Iteration of the lifting stage on the output c[n] creates the complete set of DWT scaling and wavelet coefficients Cj[n] and dj[n]. At each step, we weight the Cj[n] with ke and dj[n] with ko respectively. The energy of the underlying scaling and wavelet functions is normalized.

The lifting stesp are inverted, even if P and U are nonlinear, non-invertible, or space-varyi.ng Rearranging (1) and (2), we have

xe[n]=c[n]-U(d[n]),

xo[n]=d[n]+P(xe[n]).

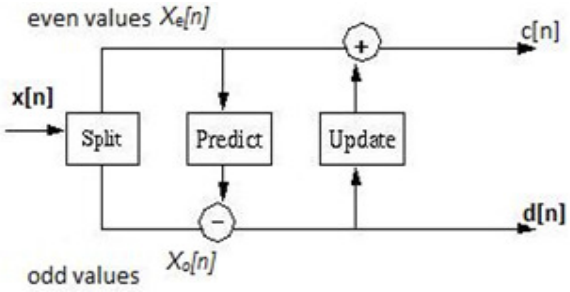

Fig.1. Lifting stage: Split, Predict, Update

As long as for the inverse and forward transforms U and P are chosen, the original signal will be perfectly reconstructed. The inverse lifting stage is shown in Fig.2.

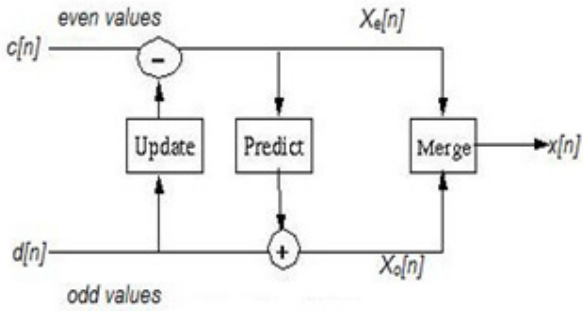

Fig.2. Inverse lifting steps: undo the Update, undo the Predict, and Merge the even and odd samples

B. Adaptive lifting scheme

The adaptive lifting scheme is classical lifting modification . The Fig.3 shows the adaptive update lifting scheme followed by a fixed prediction. At each sample n According to a decision function D(x[n],y) an update operator is chosen. As in the classical and space-varying lifting, the critical point is that D(x[n],y) depends on y, and it also depends on the sample being updated. The update operator and addition are fixed, in the standard lifting scheme. The choice of addition and the update operator depends on the information locally available within both the approximation signal and the detail signal in the adaptive lifting scheme.

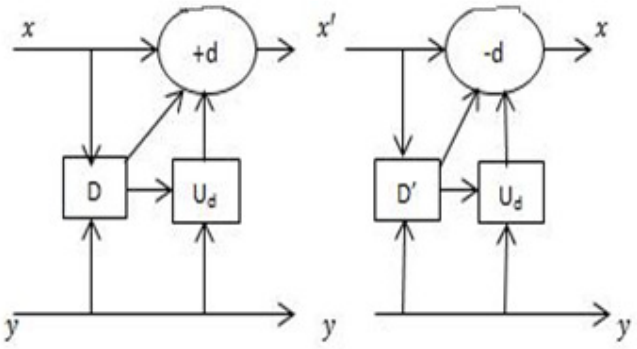

Fig.3. Adaptive update lifting scheme

According to the structure of lifting, Adaptive Lifting Scheme performs update first , and then performs prediction. Assume $x=x_o$ (2m,2n), where x_o is the input image, which is split into two signal one is x average signal and y detail signal . The y detail signal includes y_h horizontal signal , y_v vertical signal, and y_d diagonal signal.

The 2-D adaptive lifting formation is as follows:

Update: Coefficient y_h,y_v,y_d are used to update x:

$$x'= U(x,y_h,y_v,y_d) \tag{3}$$

Here, U is update operator, in which coefficients are chosen by D decidable factor.

Prediction: Updated low-frequency Coefficient x' is used to predict y_h,y_v,y_d :

$$y_h'= y_h - p_h(x, y_v,y_d) \tag{4}$$

$$y_v= y_v-p_v(x- y_d) \tag{5}$$

$$y_d= y_d-p_d(x') \tag{6}$$

The p_h,p_v,p_d, are prediction schemes for different frequency bands. According to the local feature adjacent to x, y_h,y_v, and y_d the scheme adaptively chooses U update operator and P prediction operator. Without recording any overhead information. the perfect reconstruction is ensured by the update and prediction scheme The choice of U update operator and the addition operator$\oplus$ in adaptive lifting scheme depends on the information locally available in the x approximation signal and the y detail signal. In reality, this choice will be triggered by the so called decision map $D:X\times Y \rightarrow D^Z$ where D is the decision set. We have a different U_d update operator and addition $\oplus_d$ for every possible decision d$\in$D of the decision map,. Thus the analysis step is given as

follows,

$$x'(n) = x(n) \oplus_{d_n} U_{d_n}(y)(n) \tag{7}$$

At location n d_n=D(x,y)(n) is the decision. Assuming that the reversibility condition $\oplus_d$ holds for every possible decision dЄD and it is given by

$$x(n) = x'(n) \ominus_{d_n} U_{d_n}(y)(n) \tag{8}$$

where $\ominus_{dn}$ denotes the subtraction that inverts $\oplus_d$.

The decision d_n = D(x,y)(n) depends on the x original signal. On the other hand, during synthesis, we do not know but "only" its update x'. In general, this prohibits the d_n computation and in such cases, perfect reconstruction is out of reach. However, it is still possible to recover d_n as there exist a number of situations from an posterior decision map.

4. PROPOSED BLOCK DIAGRAM

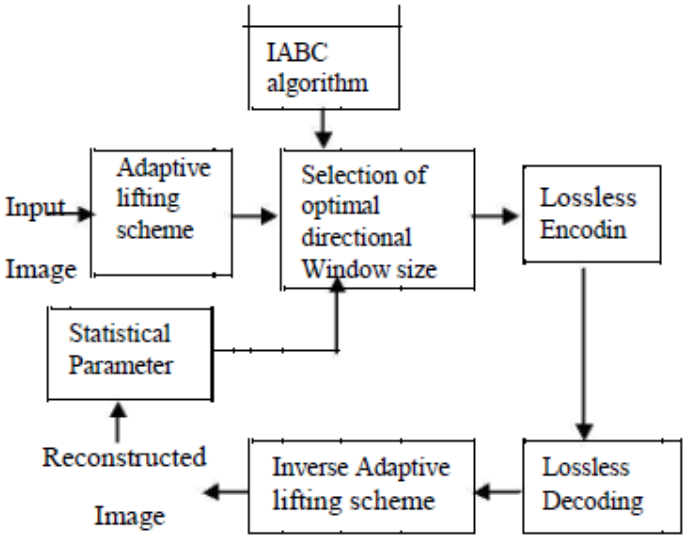

Fig. 4. Proposed Block diagram

In this method it is observed that, wavelet transform did not yield better quality for more detail texture image, so it gives a way for adaptive lifting scheme based decomposition. To determine the best directional window size and to produce the better quality for more detail texture image by local search process an Interactive Artificial Bee Colony algorithm, recent and successful optimization tool, is used. The lossless encoding technique is used to get a perfect compressed image. After the encoding process, data will be in digital form so that one can store or transmit the data to the long distance. For compressed data the image is reconstructed by applying decoding process followed by Inverse adaptive lifting scheme.

A. Need for Interactive Artificial Bee Colony algorithm

Choosing a global update coefficient does not give better compression ratio and quality. Interactive Artificial Bee Colony Algorithm by local search finds different update coefficient and helps to determine the best update coefficient optimally to get best quality of compressed image.

In update step, the center pixels are modified with the co-efficient in 8-different direction with a considerable window size and by using local search algorithm. To determine best directional window with an considerable size an IABC algorithm is used.

5. INTERACTIVE ARTIFICIAL BEE COLONY ALGORITHM

The most recently defined algorithm, motivated by the intelligent behavior of honey bees is Interactive Artificial Bee Colony Algorithm. It is as simple as Artificial Bee Colony Algorithm and differential Evolution(DE) algorithms Particle Swarm Optimization, and uses common control parameters such as maximum cycle number and colony size. As an optimization tool, IABC provides a population-based search procedure in which artificial bees with the time modifies the individuals called food positions and the aim of bee's is to discover the food sources placed with high nectar amount and at last the one with highest nectar.

In IABC algorithm, the solution space randomly spray percentage of the populations , fitness values called as nectar amounts is calculated, represents the ratio of employed bees to the total population. When these populations are positioned into the solution space they are called employee bees. The probability of selecting a food source is then calculated, select a food source to move by roulette wheel selection for every onlooker bees and then nectar amounts of them is determined. If the employed bees fitness values does not improve by predetermined number of iterations continuously, called "LIMIT", such food sources are abandoned, and these employed bees become the scouts. The scouts are moved. The position and the best fitness value found by the bees is memorized. We check whether the termination condition is satisfied by the total number of iterations. If the condition for termination is satisfied, terminate the program and output the results. The flow chart for IABC is shown in Fig. 5.

The process of the IABC can be described in 6 steps:

Step 1. Initialization phase:

In Initialization phase within the maximum boundaries of each pixels an window size is chosen

```
area=1+floor(maxarea*rand(1));

row=5+floor(r*rand(1));

col=5+floor(c*rand(1));
```

Step 2. Employed bees phase

In Employed bees phase of the algorithm, a local search xi, is conducted in the neighbourhood of each directional window, defined by using:

```
a1=img(row-area, col-area);

b1=img(row,col-area);

c1=img(row+area,col-area);

d1=img(row+area,col+area);
```

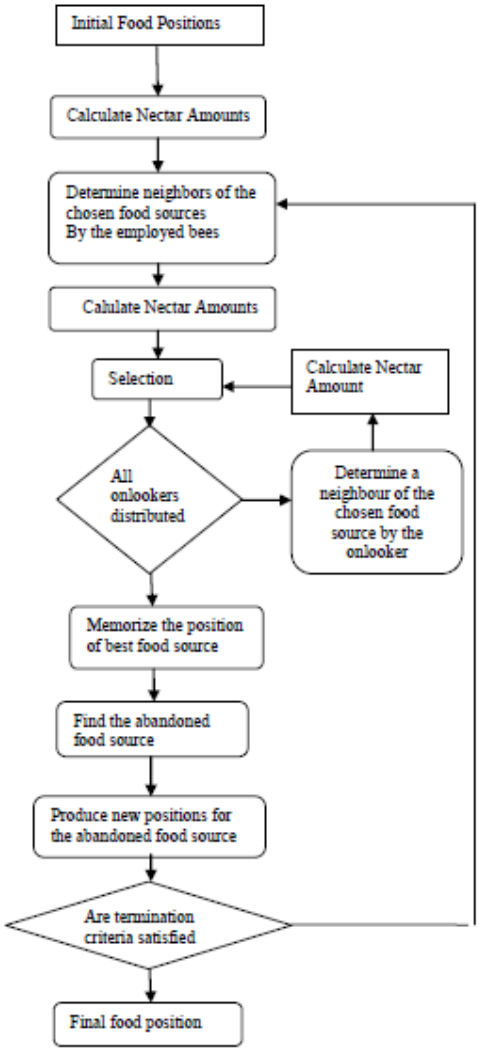

Fig. 5. Flow chart for IABC Algorithm

If we get fitness better than before, then memorize the current one.

If(localPSNR>prevPSNR)

prevPSNR=localPSNR;

bestimg=recconstimg;

After generating a new neighbour solution by local search, the new solution fitness (quality) is evaluated and better one is kept in the population. Now the counter is incremented for each local search up to 8 level.

Step 3. Onlooker bees phase:

In Onlooker bees phase of the algorithm, the probability of selecting a food source is calculated by using equation

$$P_i = \frac{F(\theta_i)}{\sum_{k=1}^{S} F(\theta_k)}$$

( ), is the employed bee fitness value that is picked by applying the roulette wheel selection.

P_i is the probability of selecting the th employed bee. ( ) randomly selected employed bee fitness value

By roulette wheel selection, select a food source to move for every onlooker bees and then the nectar amounts is determined. The onlookers movement follows the below equation

$$x_{ij}(t+1) = \theta_{ij}(t) + \sum_{k=1}^{n} \tilde{F}_{ikj} \cdot \left[\theta_{ij}(t) - \theta_{kj}(t)\right]$$

Step 4. Scout bees' phase

In the Scout bees' phase If the employed bees fitness values does not improve by predetermined number of iterations continuously, called "LIMIT", such food sources are abandoned, and these employed bees become the scouts. The scouts are moved by the equation

$$\theta_{ij} = \theta_{ijmin} + r.\left(\theta_{ijmax} - \theta_{ijmin}\right)$$

Step 5. Update the Best Food Source found so far:

In this step the best fitness value and the position is memorized, which are found by the bees.

Step 6. Termination Checking:

In this step check whether the termination condition is satisfied by the total number of iterations. If the condition for termination is satisfied, terminate the program and output the results.; otherwise go back to the Step 3.

6. PROPOSED ALGORITHM

In the proposed method, the input image is decomposed using wavelet lifting scheme and then the Interactive artificial bee colony algorithm is used in the update process to get considerable quality.

A. Algorithm steps:

Step 1: Input the Gray scale Image.

Step 2: split the image into odd and even pixel regions.

Step 3: Decompose the image as (odd-even) for next prediction step.

Step 4: Fix the maximum coverage size as 'M' and initialized

'K=0' for prediction co-efficient. Where M is maximum window size, upto which it will do local search for each center pixels maximum window size in our program is 5.

Step 5: Each pixel in the decomposed image is Scanned and its present fitness value and compression ratio is calculated.

Step 6: To predict 'a' and 'b' call direction finding algorithm co-efficient of all 8-direction combination. The 8- direction combinations are a1,b1,c1,d1,a2,b2,c2,d2

	DL		VT	DR	
		DL	VT		
HL	HL	HL	X	HR	HR
	DD	DD	VD	DB	
DD			VD		DB

Fig.6. Directional coefficient for center pixel 'x'

Where x is an center pixel to be update
HL is predicted coefficient in horizontal left direction
HR is predicted coefficient in horizontal right direction
VT is predicted coefficient in vertically top direction
VD is predicted coefficient in vertically down direction
DL is predicted coefficient in diagonally top left direction
DR is predicted coefficient in diagonally top right direction
DB is predicted coefficient in diagonally bottom left direction
DD is predicted coefficient in diagonally bottom right direction

Step 7: By using Update lifting formula for each direction prediction calculate update weight and find compression ratio and PSNR.

The Peak Signal to Noise Ratio(PSNR) represents a measure of the peak error and is expressed in decibels. PSNR is defined by

$$PSNR = 10.\log\_10\left(\frac{255^2}{MSE}\right)$$

Step 8: The best individual is memorized , CR and its direction using IABC local search.

Step 9: To predict and update the best value for different range of window size iterate K from (0 to M) .

Step 10: Using IABC local search, memorize the best window size in terms of its MSE and CR for each reference pixel.

7. EXPERIMENTAL RESULTS

The proposed algorithm is tested on standard images with different image formats. The Reconstructed images are shown in figure 7. The results are tabulated for various images in Table(1).Lena image is a JPEG image and the results obtained is better than existing methods.

Fig. 7. Reconstructed Images with lifting with IABC (a) Original Image (b) Output of lifting scheme with IABC.

Table 1: Comparison table for cameraman image

Compression Ratio	PSNR fro different images obtained using Lifting Scheme using IABC				
	Cameraman Image	Lena Image	Barbara Image	Pepper Image	Rice Image
30	41.78	43.4	42.39	40.83	40.27
40	38.92	39.4	38.95	38.10	37.82
50	38.63	38.2	38.31	37.85	36.74
60	34.26	33.9	33.78	34.57	33.74

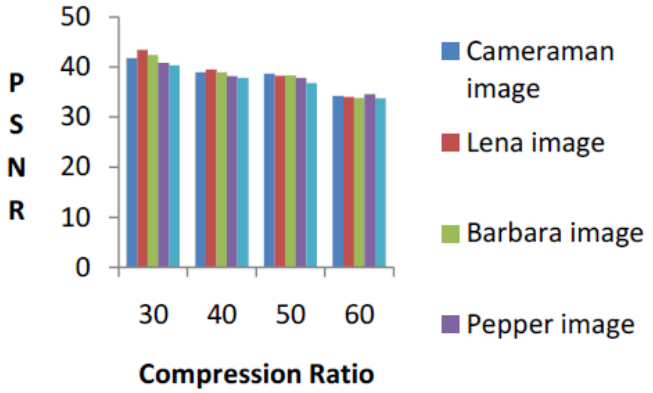

Fig. 8: Graph representing CR Vs PSNR for different images.

8. CONCLUSION AND FUTURE ENHANCEMENTS

In this paper, a method to optimize the prediction function used in lifting scheme using IABC algorithm for image compression is proposed. IABC algorithm is implemented in update process of lifting scheme to give better PSNR. From the experimental results, it is concluded that proposed method yields improved quality compare to existing methods in the literature. The proposed method gives the way to reduce the data to represent the image and thereby decreases transmission bandwidth. Hence, the transmission cost and memory cost is reduced.

In future work, Interactive Artificial Bee Colony algorithm shall be implemented in the thresholding process to reduce the number of coefficient representing the image by optimally choosing the thresholding value to get more better compression and quality.

REFERENCES

[1] Subramanya A. "Image Compression Technique, " potentials IEEE, Vol. 20, issue 1,pp19-23, Feb-March 2001.

[2] Rafael C. Gonzalez, Richard Eugene; " Digital image processing", Edition 3, 2008,page 466.

[3] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions", in Proc. SPIE, vol. 2569,1995,pp. 68-79.

[4] W. Sweldens, "The lifting scheme: A construction of second-generation wavelets", SIAM J. Math. Anal., vol. 29, no. 2, pp. 511-546,1997.

[5] A.R. Calderbank, I Daubechies, W. Sweldens, and B-L Yeo, "Wavelet transforms that map integers to integers", J. Appl. Comput. Harmon, Anal., vol.5, no. 3, 1998.

[6] M. Adams and F. Kossentini, "Reversible Integer-to-Integer Wavelet Transforms for Image Compression: Performance Evaluation and Analysis", IEEE Trans.On Image Processing, vol.9, no.6, pp. 1010-1024, Jun 2000.

[7] N.V Boulgouris, D. Tzovaras, and M.G. Strintzis, "Lossless image compression based on optimal prediction, adaptive lifting, and conditional arithmetic coding", IEEE, Trans. Image Process., vol. 10, no. 1,pp. 1-14, Jan. 2001.

[8] Pei-Wei Tsai, Jeng-Shyang Pan, Bin-Yih Liao, and Shu- Chuan Chu," Enhanced Artificial Bee Colony Optimization", International Journal of Innovative Computing, Information and control, vol.5, no.12, pp.1-12, Dec. 2009.

[9] J. Kennedy, R.C. Eberhart, and Y.Shi, "Swarm Intelligence", Morgan Kaufmann Publishers, San Francisco,2001.

[10] B. Akay and D. Karaboga, "Parameter tuning for the artificial bee colony algorithm," ICCCI 2009 (R.Kowalezyk, N.T. Nguyen and S.M.Chen,eds.), LNAI, vol. 5796, 2009, pp. 608-619.

[11] V . U Kale and N.N. Khalsa, "Performance evaluation of various wavelets for image compression of natural and artificial images", International Journal of Computer Science and communication 1(2010), no.1, pp. 179-184.

[12] D. Karaboga and B. Akay, A survey: "Algorithms simulating bee swarm intelligence", Artificial Intelligence Review 31(2009), no. 1, pp. 55-68.

[13] G. Piella and H.J.A.M. Heijmans, "Adaptive lifting schemes with perfect reconstruction", Research Report PNARO104,CWI, Amsterdam, Feb. 2001.

[14] F.W. Moore, "A genetic algorithm for optimized reconstruction of quantized signals", Evolutionary computation, 2005. The 2005 IEEE congress on, vol. 1,2005, pp. 105-111.

[15] R. Ramanathan, K. Kalaiarasi, D. Prabha, "Improved wavelet based compression with adaptivelifting scheme using Artificial Bee Colony algorithm", International Journal of Advanced Research in Computer Engineering & Technology, Vol 2 , Issue 4, April 2013.

[16] W. Trappe and K.J.R.Liu, "Adaptivity in the lifting scheme," in 33th Conference on Information science and systems, Baltimore, March 1999, pp. 950-958.

[17] D. Karaboga, " An Idea Based On Honey Bee Swarm For Numerical Optimization", Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

[18] E. Bonabeau, M. Dorigo, and G. Theraulaz, "Swarm Intelligence: From Natural to Artificial Intelligence", NY: Oxford University Press,1999.

[19] P.-W. Tsai, J.-S. Pan, S.-M. Chen, B. –Y. Liao, and S.-P. Hao, "Parallel Cat Swarm Optimization", Proc. Of 7th International Conference on Machine Learning and Cybernetics, pp. 3328-3333, Kunming, China, 2008.

[20] Y. Guo, X.Gao, H. Yin, and Z. Tang, "Coevolutionary Optimization Algorithm with Dynamic Sub-population Size", International Journal of Innovative Computing, Information and Control, vol.3, no.2, pp.435-448, 2007.

[21] M. Dorigo and L.M. Gambardella, "Ant Colony Optimization for Data Clustering", Proc. Of 8th Pacific Rim International Conference on Artificial Intelligence, Auckland, New Zealand, LNAI 3157, pp. 534-543,2004.

AUTHOR PROFILE

Vrinda Shiva Shetty received B.E from Gurbarga University and M.Tech degree from VTU in Computer Science and Engineering and presently pursuing Ph.D in Image Compression from the University of Gulbarga University, Gulbarga, and Karnataka. Field of Interest includes Intelligent Image Processing, Evolutionary Computation.

Dr. G. G. Rajput currently working as Associate Professor in the Department of Computer Science at Rani Channamma University Belagavi, Karnataka State, India. The Area of interest includes Image processing, Pattern recognition, Operations Research and Software Engineering.

INTENTIONAL BLANK

HINDI DIGITS RECOGNITION SYSTEM ON SPEECH DATA COLLECTED IN DIFFERENT NATURAL NOISE ENVIRONMENTS

Babita Saxena[1] and Charu Wahi[2]

Department of Computer Science, Birla Institute of Technology, Noida
babita.gs@gmail.com
charu@bitmesra.com

ABSTRACT

This paper presents a baseline digits speech recognizer for Hindi language. The recording environment is different for all speakers, since the data is collected in their respective homes. The different environment refers to vehicle horn noises in some road facing rooms, internal background noises in some rooms like opening doors, silence in some rooms etc. All these recordings are used for training acoustic model. The Acoustic Model is trained on 8 speakers' audio data. The vocabulary size of the recognizer is 10 words. HTK toolkit is used for building acoustic model and evaluating the recognition rate of the recognizer. The efficiency of the recognizer developed on recorded data, is shown at the end of the paper and possible directions for future research work are suggested.

KEYWORDS

HMM, Acoustic Model, Digit Speech Recognition, Grammar

1. INTRODUCTION

In the last few years, Hidden-Markov-Model based (HMM) algorithms have been the most successful techniques used for speech recognition systems. Using the same, the experiments are conducted for building a Digit Speech Recognition(DSR) for Hindi. Thus, for building a DSR, acoustic characteristics like pitch, formant frequencies etc have to be computed. These characteristics are captured and a model is built based on these. These models are further used for recognition purposes.

In this paper we present our work on building acoustic model for Hindi Digits. Hindi belongs to the Indo Aryan family of languages and is written in the devanagari script. There are 11 vowels and 35 consonants in standard Hindi. In addition, 5 Nukta consonants are also adopted from Farsi/Arabic sounds.

David C. Wyld et al. (Eds) : CSITY, SIGPRO, DTMN - 2015
pp. 23–30, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50303

This paper is organized as follows. Section 2 gives the related work. Section 3,4,5 describes database, feature extraction process the Acoustic model preparing procedure respectively. Section 6 describes grammar structure used for decoding. Section 7 tells how the utterances are recognized. Section 8 discusses the Observation and Results of speech recognition on this system. Conclusion and Future work is stated in Section 9 and Section 10 respectively.

2. RELATED WORK

Now-a-days research work is being carried out for Hindi Digits. Some speech recognition systems have been proposed for the isolated digit recognition in the Hindi language.

Sharmila et al. proposed hybrid features for speaker independent Hindi speech recognition system. In this paper Mel-frequency cepstral coefficients (MFCC), Perceptual linear prediction (PLP) coefficients along with two newly modified hybrid features are used for isolated Hindi digits recognition. Two modified hybrid features Bark frequency cepstral coefficients (BFCC) and Revised perceptual linear prediction (RPLP) coefficients were obtained from combination of MFCC and PLP. Experiments were performed for both clean as well as on noisy data. In this experiment six different noises: Car noise, F16 noise, Factory noise, Speech noise, LYNX noise and Operation room noise have been added to clean Hindi digits database at different SNR levels to get noisy database. The recognition performance with BFCC features was better than that with MFCC features. RPLP features have shown best recognition performance as compared to all other features for both noisy and clean databases.[2]

Dhandhania et al. proposed a speaker independent speech recognizer for isolated Hindi digits. They aimed to find the best combination of features which yields the highest recognition rate along with the optimal number of hidden states of the HMM. Using MFCC and delta-MFCC as the feature vectors and 8 hidden states, an average recognition rate of 75% is achieved on a dataset of 500 utterances.[3]

Mishra et al. proposed a connected Hindi digit recognition system using robust features such as Mel Frequency Perceptual Linear Prediction (MF-PLP), Bark Frequency Cepstral Coefficient (BFCC) and Revised Perceptual Linear Prediction (RPLP). A success of 99% was achieved using the MF-PLP feature extraction and training Hidden Markov Models (HMMs). Pre-defined 36 sets of 7 connected digits uttered by 35 speakers was used in training and the 5 other speakers for testing. The performance for this system might be high as predefined sets are used with a fix number of known digits in each set.[4]

Saxena et al. proposed a microprocessor based Speech Recognizer using a novel zero crossing frequency feature combined with a dynamic time warping algorithm. An overall success of 95.5% was reported with the implementation in MATLAB. The above systems involved training and testing on similar data leading to high performance. The number of speakers was limited to two in the experiments.[5]

Apart from English, successful results have been proposed in word digit recognition in other languages like Japanese, Thai, and Italian. Owing to their success we too evaluate the possibility of developing a robust system for Hindi digit recognition in natural noise environment.

3. DATABASE

The speech data is collected from 10 speakers in varying noise environments. The recording is done in every speakers' home with respective noise in their rooms. The respective noise refers to vehicle horn noises in some rooms, internal background noises in some rooms like opening doors, silence in some rooms, etc. All these recordings are used for training acoustic model. The audio data consists of Hindi digits recordings of 10 speakers. In all 10 digits in Hindi are recorded by the speakers from age group from 20 to 40 yrs. The utterances are recorded in 48 KHz, stereo 16 bit format. The recordings of 8 speakers are used for training acoustic model. The recording is done using Sony Xperia L headset on laptop model - Dell Inspiron 1440. It is then channel separated and down sampled to 16 KHz and then single channel is used to train acoustic models. The database of 10 speakers is divided into a training database of 8 persons and a testing database of 2 persons.

4. FEATURE EXTRACTION

The first step in any automatic speech recognition system is to extract features i.e. identify the components of the audio signal that are good for identifying the linguistic content and discarding all the other stuff which carries information like background noise, emotion etc.

The main point to understand about speech is that the sounds generated by a human are filtered by the shape of the vocal tract including tongue, teeth etc. This shape determines what sound comes out. If we can determine the shape accurately, this should give us an accurate representation of the phoneme being produced. The shape of the vocal tract manifests itself in the envelope of the short time power spectrum, and Mel Frequency Cepstral Coefficients (MFCC) accurately represent this envelope. MFCCs are a feature widely used in automatic speech and speaker recognition. They were introduced by Davis and Mermelstein in the 1980's, and have been state-of-the-art ever since.[6]

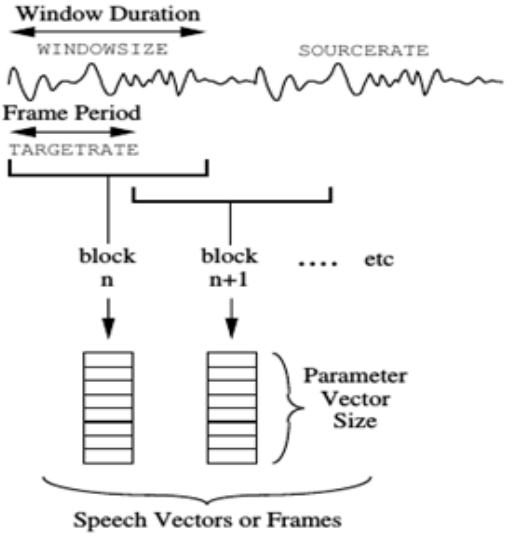

Fig.1: Feature Extraction Process

The primary feature vector size of 13 is considered along with their delta and delta-delta features. Thus in all 39 feature vectors are used.

5. ACOUSTIC MODEL

The acoustic model is built using HTK 3.4 toolkit. The acoustic model is HMM based since HMM based statistics for building acoustic model is very popular and have been shown to give better results than any other techniques. In a statistical framework for speech recognition, the problem is to find the most likely word sequence

$$\hat{W} = \arg\max_{w} p(W / A) \quad (1)$$

With a Bayesian approach to solving the above problem, we can write

$$\hat{W} = \arg\max_{w} p(A / W)\, p(W) \quad (2)$$

Equation 2 above gives two main components of a speech recognition system, the acoustic model and the language model. One type of language model is the grammar, which is a formal specification of the permissible structures for the language. The deterministic grammar gives the probability of one if the structure is permissible or of zero otherwise. Furthermore, the probabilistic relationship among a sequence of words can be directly derived & modeled from the corpora with the stochastic language model.[7] Language model is used for dictation type purposes systems & grammar is used for command & control systems or small vocabulary systems.[8] Since digit recognizer is a small vocabulary system, grammar is used for decoding.

The Acoustic Model preparation is described ahead with details of training steps. The acoustic model is made by training the recorded audio data out of which mfcc feature vectors are extracted and their delta and delta-delta features are considered. The hmm models are over 61 context independent phonemes. Each phone HMM definition file represents a single stream single-mixture diagonal covariance left-right HMM with five states.

Prototype models for 61 phonemes are built using flat start approach. These models were further refined by applying nine iterations of the standard Baum-Welch embedded training procedure. These models are then converted to triphone models and two iterations of Baum-Welch training procedure are applied, then there states are tied using decision tree based approach and two iterations of Baum-Welch training procedure are applied. Now the number of mixtures is incremented to 14 and seventeen iterations of the standard Baum-Welch training procedure were applied.

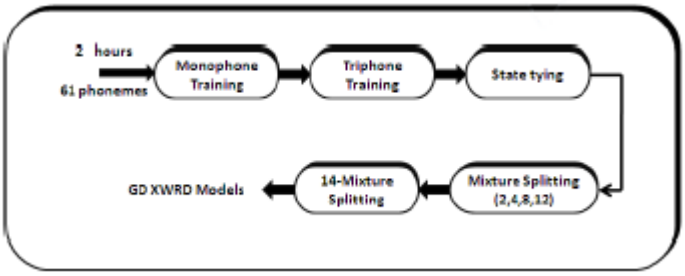

Fig.2: Acoustic model training procedure

The speech data is recorded in varying noise environment and pronunciation mistakes were taken care of, a few errors were found after recording. These errors were compensated by changing the transcription accordingly as per the pronunciation, by manually listening to the speech.

6. GRAMMAR

A speech recognition grammar contains one or more rules. Each rule defines a set of language constraints that a speech recognition engine uses to restrict the possible word or sentence choices during the speech recognition process. Speech recognition engines use grammar rules to control the elements of sentence construction using a predetermined list of recognized word or phrase choices.[9]

HTK provides a grammar definition language for specifying simple task grammars such as this. It consists of a set of variable definitions followed by a regular expression describing the words to recognize. For the digit recognizing application, a suitable grammar might be

$digit = एक | दो | तीन | चार | पाँच |
छः | सात | आठ | नौ | शून्य ;
({SENT-START} {$digit} {SENT-END})

where the vertical bars denote alternatives, the square brackets denote optional items and the angle braces denote one or more repetitions. The complete grammar can be depicted as a network as shown in Fig. 3.

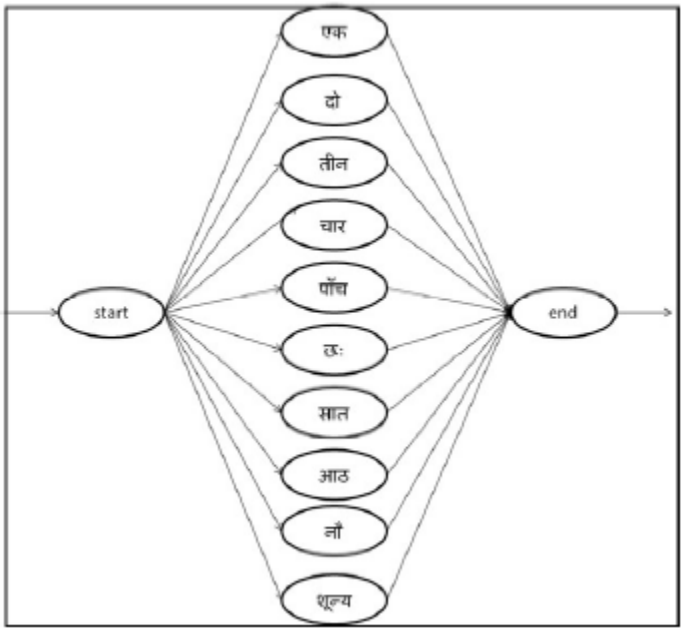

Fig.3: Grammar for Digit Recognition

7. DECODING

HTK's viterbi decoder Hvite is used for decoding the utterances. HVite is a general-purpose Viterbi word recogniser. It will match a speech file against a network of HMMs and output a transcription for each. When performing N-best recognition a word level lattice containing multiple hypotheses can also be produced.

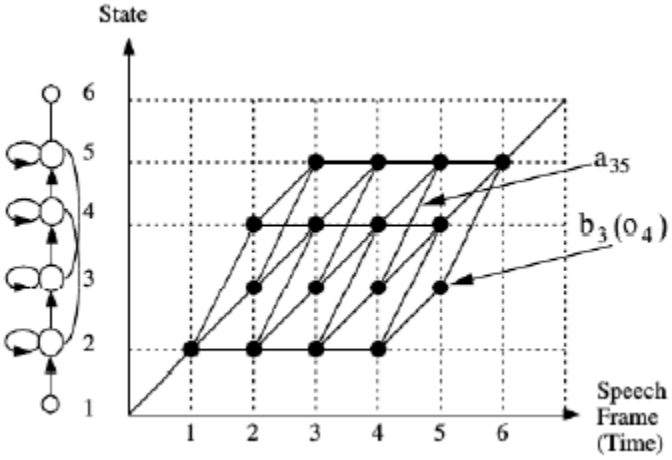

Fig.4: Viterby decoding process

8. OBSERVATION AND RESULT

The performance of DSR is measured in terms of recognition rate. The system is tested on

- 2 seen speakers
- 2 unseen speakers

where seen data means the data is taken from the training corpus itself and unseen data means the data is not from the training corpus. Recognition on the unseen data is carried out using the same acoustic model, lexicon and grammar. The performance is analyzed using HTK toolkit's HResult tool. The percentage number of correct labels recognized is given by:

$$\%\text{Correct} = \frac{H}{N} \times 100\%$$

where H and N are the number of correct labels and total labels respectively. In addition to the recognition rate, the tool also reports the number of insertions, deletions and substitutions for all the test data.

Table 1: Digits recognition rate for 2 speakers in training and test sets

Speaker	Phone level(%)	Word level(%)
Training	92.82	94.09
Test	86.17	85

The acoustic model and grammar is used for performing the recognition. The phone level recognition rate of 92.82% and 86.17% is observed on training (seen) data and test (unseen) data respectively. The word level recognition rate of 94.09% and 85% is observed on training (seen) data and test (unseen) data respectively.

9. CONCLUSION

In this paper, we described our experiments with the Hindi digits speech recognition. A baseline digits recognizer was developed and the results found are quite encouraging.

10. FUTURE WORK AND SCOPE

Our future work will be to further refinement of the word accuracy and supporting. A number of further experiments may be tried to achieve better accuracy. Some of them can be described as under.

- Continuous digits recognition experiments can be performed.
- Training corpus may be increased.
- Lexicon can be increased by adding 2-digit numbers.

REFERENCES

[1] Steve Young, Gunnar Ever, Thomas Hain, Dan Kershaw, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, Valtcho Vaitchev, Phil Woodland, "The HTK Book", copyright 2001-2002 Cambridge University Engineering Department

[2] Sharmila1, Dr. Achyuta, N. Mishra, Dr. Neeta, Awasthy, "Hybrid Features for Speaker Independent Hindi Speech Recognition", International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013

[3] Vedant Dahndhania, Jens Kofod Hansen, Shefali Jayanth Kandi & Arvind Ramesh, "A Robust Speaker Independent Speech Recognizer for Isolated Hindi Digits", International Journal of Computer & Communication Engineering, Vol 1, No. 4, November 2012

[4] A.N.Chandra, Mahesh Chandra, Astik Biswas, S.N.Sharan, "Robust Features for Connected Hindi Digits Recognition", International Journal of Signal Processing & Pattern Recognition, Vol. 4, No. 2, June 2011

[5] A. Saxena and A. Singh, "A Microprocessor based Speech Recognizer for Isolated Hindi Digits," in IEEE ACE, 2002.

[6] http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

[7] Xuedong Huang & Li Deng, Microsoft Corporation, "Chapter 15, An Overview of Modern speech Recognition, Handbook of Natural Language Processing" C5921_C012 Page 343, 2009-9-9

[8] http://en.wikipedia.org/wiki/Acoustic_model

[9] http://msdn.microsoft.com/en-us/library/hh378438%28v=office.14%29.aspx

AUTHORS

Ms. Babita has received her MCA from Uttar Pradesh Technical University, Lucknow (UP)-India in 2008. Presently she is pursuing M.Tech from Birla Institute of Technology, Noida -India. She has published more than 3 research papers in the area of Speech, Signal and Image Processing at National/International level. Her areas of interest are Speech and Signal Processing.

Ms. Charu Wahi is currently a Ph.D. candidate in the Department of Computer Science and Engineering, Birla Institute of Technology, Ranchi. She received her B.E. degree in 2000 and M.Tech - Computer Science in 2008. She is currently working as Assistant Professor in Department of Computer Science and Engineering, Birla Institute of Technology, Ranchi, Noida Campus. Her research areas include routing, security, quality of service especially in mobile ad-hoc networks

ASSOCIATING EVENTS WITH PEOPLE ON SOCIAL NETWORKS USING A-PRIORI

Srijan Khare, Vyankatesh Agrawal, Gaurav Tiwari, Gourav Arora and Bhaskar Biswas

Department of Computer Science & Engineering,
Indian Institute of Technology (BHU), Varanasi, India
{srijan.khare.cse11, v.agrawal.cse11, gaurav.tiwari.cse11,
gourav.arora.cse11, bhaskar.cse} @iitbhu.ac.in

ABSTRACT

In social media, same news or events are associated with two or more people, sometimes with different perspective. The representation of the news or events varies from person to person, perspective to perspective or time to time. In this paper, we present a simple model to associate events with different people (Personalities). To demonstrate our model, we have used real world social networks data (i.e. from Twitter) and results show the accuracy of the model.

KEYWORDS

Social Networks, Association rule, A-Priori.

1. INTRODUCTION

In the present times online social networks (social networks) plays an important role in modern lives. People not only use this powerful media as a tool for entertainment, news, sharing emotions and sentiments but also depend for news, political sentiments, and events. Most of the times, news or an event is associated with many people. Many times people express their views on an event based on their perspective. These expressions of views differ as they come from different people with different perspectives. People in social media or in general make their opinion about an event based on the views expressed by the others. This "word-of-mouth" is sometimes biased towards a particular aspect of the news or event and this is due to the difference in perspective. To get the holistic view of news or event, one may require getting many views/opinions associated with that news/event. In general, words from the famous personalities about an event build up the general opinion of that event. Another outlook is to judge or compare personalities or person based on their views on a particular event. So, to understand the events, on whole, different views are needed to be associated with views or personalities. In this paper, a model has been proposed to address this challenge.

The proposed model uses Twitter as a source of views of personalities and events. Association of different personalities/views with events has been done using A-Priori [17] algorithm. Twitter has been used by many researchers. Event detection has been a major challenge that was addressed by majority of researchers using twitter data. Using different approaches and methods, researchers have proposed to detect event from Twitter and applied for different applications and usage. A brief survey was conducted for the same ([2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13],

David C. Wyld et al. (Eds) : CSITY, SIGPRO, DTMN- 2015
pp. 31–34, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50304

[14] and [16]). A-priori algorithm is a basic association rule generation method which has been used by many researchers for different applications. One such application is given in [15].

In [10], Hila Becker et al. proposed a method to identify real event from non event using twitter messages. From the twitter stream, authors make clusters of similar messages to identify events. In this proposed model, similar approach was followed to identify the tweets from different personalities without distinguishing the events. Luca Cagliero and Alessandro Fiori [1] presented TweM Tweet Miner framework which discover generalized association rules from Twitter. [1] gives a generalized association among the content of the tweets, whereas the proposed model tries to associate different personalities with events. The authors of this paper acknowledge the work presented in [1] and [10] along with all the other references.

2. PROPOSED MODEL

The proposed model is shown in Figure 1. Using the Twitter API, raw tweets were collected. Based on the different 'hash-tags' and the inputs from the different personalities, the raw tweets were sorted. Pre-processing was done to different inputs. This step mainly includes word extraction and removal of symbol and words less than four letters, using Regex words extraction in our model. The Association Rule Algorithm, which in this case is A-Priori algorithm, requires unique words. So, duplicate items were removed. The pre-processed tweets were now stored in database in the form of transactions from different personalities. These transactions were given input to the algorithm. Based on the association rule generated, the events were associated with different personalities.

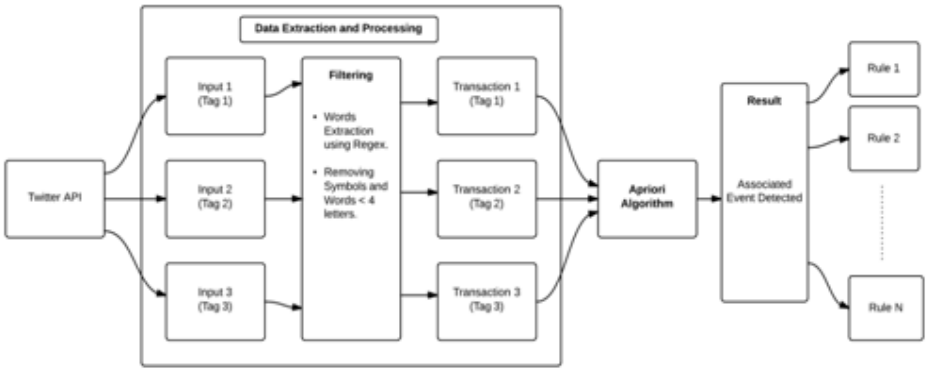

Figure 1. Proposed Model

3. EXPERIMENT AND RESULTS

To test the proposed model, tweets from famous personalities were extracted for a common time interval and limited to 100 tweets per day. Specifically, tweets about following were collected:

- Bill Gates
- Mark Zuckerberg
- Satya Nadella
- Jeff Bezos

These personalities represented 4 transactions. The tweets collected were broken into words using a simple regex pattern and words of length lesser than 4 were ignored. Moreover, special characters and words of non-English languages were discarded. The final input data comprised of

100 words within each. With Support 0.75 and Confidence: 0.5, a set of 666673 rules were identified. Similarly for Support 1.0 and Confidence: 0.5, a set of 223695 rules were found. The Ice Bucket Challenge which has been a recent trend was identified as a common link between all the transactions i.e. the personalities chosen.

Figure 2 shows a union of first 10 words of 4 transactions and their respective frequencies. The grey squares indicate the presence of a word in a particular transaction.

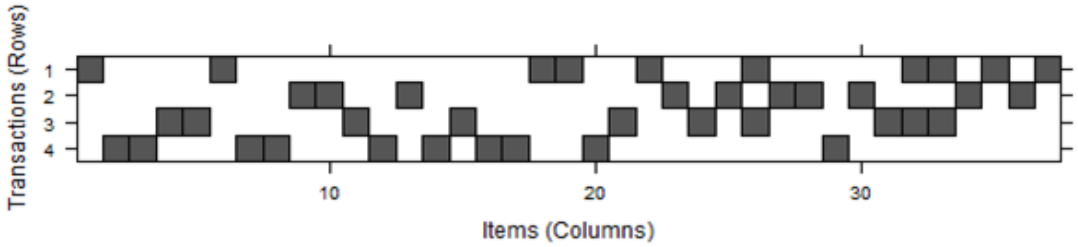

Figure 2. Items-Transaction Graph

Figure 3 shows some of the association rules found with support = 0.75 and confidence = 1.00.

```
87   {bucket,
      challenge,
      Elon,
      Gates,
      Jeff,
      Steve,
      watch}        => {take}
88   {challenge,
      Elon,
      Gates,
      Jeff,
      Steve,
      take,
      watch}        => {bucket}
89   {bucket,
      Elon,
      Gates,
      Jeff,
      Steve,
      take,
      watch}        => {challenge}
90   {bucket,
      challenge,
      Gates,
      Jeff,
      Steve,
      take,
      watch}        => {Elon}
91   {bucket,
      challenge,
      Elon,
      Jeff,
      Steve,
      take,
      watch}        => {Gates}
```

Figure 3. Association Rules

4. CONCLUSIONS

A simple model to associate personalities with events has been proposed in this paper. Initial experiments depict the success of the model. This idea can also be applied to a situation where tweets from different News agencies form different transactions in a selected time interval. Applying Apriori algorithm on it could help identify a trending topic. The news agencies and others could themselves use this trend to focus more on the identified topic.

This proposed model is a simple model, where sentiments analysis, opinion mining and machine learning techniques can be used to improve the efficiency. Methods can be proposed for the dynamic nature of the twitter data.

REFERENCES

[1] Luca Cagliero, Alessandro Fiori, " Discovering generalized association rules from twitter", Intelligent Data Analysis, Volume 17 Issue 4, July 2013, Pages 627-648.

[2] Salvatore Orlando, Francesco Pizzolon and Gabriele Tolomei, "SEED: A Framework for Extracting Social Events from Press News", WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.

[3] Hila Becker, Dan Iter, Mor Naaman and Luis Gravano," Identifying Content for Planned Events Across Social Media Sites", WSDM'12, February 8–12, 2012, Seattle, Washington, USA.

[4] Jeffrey Nichols, Jalal Mahmud and Clemens Drews," Summarizing Sporting Events Using Twitter", IUI'12, February 14-17, 2012, Lisbon, Portugal.

[5] Arkaitz Zubiaga, Damiano Spina, Enrique Amigó and Julio Gonzalo," Towards Real-Time Summarization of Scheduled Events from Twitter Streams", HT'12, June 25–28, 2012, Milwaukee, Wisconsin, USA.

[6] Rui Li, Kin Hou Lei, Ravi Khadiwala and Kevin Chen-Chuan Chang, "TEDAS : a Twitter-based Event Detection and Analysis System", 2012 IEEE 28th International Conference on Data Engineering.

[7] Ana-Maria Popescu , Marco Pennacchiotti and Deepa Arun Paranjpe," Extracting Events and Event Descriptions from Twitter", WWW 2011, March 28–April 1, 2011, Hyderabad, India.

[8] Adam Marcus, Michael S. Bernstein, Osama Badar,David R. Karger, Samuel Madden and Robert C. Miller," TwitInfo: Aggregating and Visualizing Microblogs for Event Exploration", CHI 2011,Session: Twitter Systems May 7–12, 2011,Vancouver, BC, Canada.

[9] Akshaya Iyengar, Tim Finin and Anupam Joshi," Content-based prediction of temporal boundaries for events in Twitter", 2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing.

[10] Hila Becker, Mor Naaman, and Luis Gravano," Beyond Trending Topics: Real-World Event Identification on Twitter", Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM '11), short paper.

[11] Yosuke Jimbo, Masaki Kohana and Shusuke Okamoto," A Twitter Service for School Event", 2011 Second International Conference on Networking and Computing.

[12] Ana-Maria Popescu and Marco Pennacchiotti, "Detecting Controversial Events from Twitter", CIKM'10, October 26–30, 2010, Toronto, Ontario, Canada.

[13] Makoto Okazaki and Yutaka Matsuo," Semantic Twitter: Analyzing Tweets for Real-Time Event Notification", BlogTalk 2008/2009, LNCS 6045, pp. 63–74, 2010.

[14] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo," Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development", WWW '10 Proceedings of the 19th international conference on World wide web.

[15] Alexander Toshev, Francois Br´emond and Monique Thonnat," An APRIORI-based Method for Frequent Composite Event Discovery in Videos", Computer Vision Systems, 2006 ICVS '06. IEEE International Conference on 4-7 Jan. 2006.

[16] Luca Cagliero and Naeem A. Mahoto," Visualization of High-Level Associations from Twitter Data", www.irma-international.org/viewtitle/80217/.

[17] http://en.wikipedia.org/wiki/Apriori_algorithm

AUTHOR INDEX