A CMOS Gm-C bandpass filter with a center frequency of 70 MHz and a bandwidth of 3 MHz is presented. The operational transconductance amplifier (OTA) presented in this paper along with the capacitors (Gm-C filter) is used in designing the biquad structure. The filter is realized in 0.18 µm CMOS technology using Virtuoso analog environment in Cadence tool. The simulation result shows the AC response of the filter with a center frequency of 69.98 MHz, bandwidth of 3.32 MHz and a gain of 18.46 dB.

Index Terms—Bandpass filter, biquad, operational transconductance amplifier (OTA), Super-heterodyne receivers.

1. INTRODUCTION

High performance cellular phones with low cost and small size have increasing demand in today’s market. The primary criterion in various cellular phone design approaches include lowering the complexity, cost, power, and number of external components. The super-heterodyne structure achieves good selectivity and avoids the problem of DC offset in homodyne (direct-down) receivers and this is mainly utilized in the wireless receivers for mobile phones. IF bandpass filters are then needed for the channel selection and filtering. In super-heterodyne receivers shown in figure 1, proper filtering is mandatory and is done by external surface acoustic wave (SAW) filters [1].

There are various designs of on-chip bandpass filters including simple RC filters, switched-capacitor (SC) filters, spiral inductor-capacitor tanks and Gm-C filters. The simple RC filters suffer from variation of resistors and capacitors. Each resistor has to be tuned after fabrication. SC filters solve the mismatch problem of the RC filters accurately in CMOS processes. However, SC filters are not suitable for frequencies above 10 MHz due to the strict requirement of the unity-gain frequency of operational amplifier (opamp) for fast settling. Spiral inductors are only suitable for high frequencies of GHz range. For operation in the IF band, which is from 10
MHz to 100 MHz, however, the Q of these inductors are too low (smaller than 1) to be compensated. A lot of chip area is required for the large inductors and the large capacitors. The resistance of the inductors dominates the reactance of the inductance. [6].

Filtering can be done by making use of active filters. The Gm-C circuits represent a popular technique of integrated realisation of high frequency continuous time filters [1]. The Gm-C filter offers many advantages in terms of low-power and high frequency capability. Gm-C filters can operate in a wide range of frequencies from several hundred of KHz to more than 100 MHz. Unlike the spiral inductors, the Q of Gm-C filters can be adjusted by controlling the output impedance even at lower frequencies. Therefore, for the IF band, Gm-C filter is a good choice for on-chip filtering. The main objective of this paper is to design a bandpass filter for super-heterodyne receivers. The filter design is presented in section II. Section III gives the implementation and simulation results. Finally, conclusion is drawn in section IV.

2. FILTER DESIGN

A. Operational Transconductance Amplifier (OTA)

The operational transconductance amplifier (OTA) is basically an op-amp without an output buffer. An Operational transconductance amplifier without buffer can only drive capacitive loads. An OTA can be defined as an amplifier where all nodes are low impedance except the input and output nodes. The transconductance of the OTA is given by

\[ g_m = \frac{i_{out}}{v^+ - v^-} \]  

(1)

where \( i_{out} \) is the output current of the OTA and \( v^+ \) and \( v^- \) is the differential input voltage to the OTA. The voltage gain of the OTA is given by

\[ A_v = \frac{v_{out}}{v^+ - v^-} \]  

(2)

where \( v_{out} \) is the output voltage of the OTA. The OTA used in designing the bandpass filter [2] is shown in figure 2. The circuit parameters for OTA are given in table I.

![Figure 2: Schematic of OTA](image)

Table I: Circuit Parameters for OTA

<table>
<thead>
<tr>
<th>Device</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMOS(W/L)</td>
<td>22µm/0.18µm</td>
</tr>
<tr>
<td>NMOS(W/L)</td>
<td>2µm/0.18µm</td>
</tr>
<tr>
<td>R</td>
<td>100KΩ</td>
</tr>
<tr>
<td>I bias</td>
<td>10µA</td>
</tr>
</tbody>
</table>

The symbol for OTA is shown in figure 3. A useful feature of OTA is that its transconductance can be adjusted by the bias current. Filters made using the OTA can be tuned by changing the bias current labeled I bias. Two practical concerns when designing an OTA for filter applications are the input signal amplitude and the parasitic input/output capacitances. Large signals cause the OTA gain to become non-linear. The external capacitance should be large compared to the input or output parasitic of the OTA. This limits the maximum frequency of a filter built with an OTA and causes amplitude or phase errors.
Both the OTAs are implemented in the same manner and the transconductance of each stage is same. The circuit parameters of the biquad (second order filter) are given in table II.

The transfer function of the second order bandpass filter is given by

$$H(s) = \frac{sC_1 g_m}{s^2 C_1 C_2 + sC_1 C_2 + g_m^2} \tag{6}$$

The filter design specification is given in table III.

### 3. Results

The proposed bandpass filter is designed and simulated using Virtuoso Analog Environment in Cadence tool and the following results are obtained.

#### A. Schematic and Results of OTA

The proposed OTA is implemented and simulated. The schematic diagram and the instance (symbol) for OTA are shown in figure 5 and figure 6 respectively. The bias current is set as $10 \mu A$. The control voltage is set as 1V. The resistance value is set to 100 KΩ. The non-inverting input voltage is set as 10 mV and the inverting input voltage is set as 5 mV.
Hence, the differential input voltage is

\[ \text{in differential} = (10 \text{ mV} - 5 \text{ mV}) \]

\[ = 5 \text{ mV} \]

The transient analysis for the OTA is done and the output current (\(i_{out}\)) and the input differential voltage (\(i_{in^+}\) and \(i_{in^-}\)) are plotted. The transient response is shown in figure 7 with output current and input differential voltage. The transconductance value of the OTA is calculated using equation (1).

\[ G_m = \frac{2.675 \mu A}{10 \text{ mV}} \]

\[ = 267.51 \text{ S} \]

B. Schematic and Results of bandpass filter

The OTA implemented is combined to form a biquad as shown in figure 8. The AC response of the biquad is shown in figure 9. The 3dB bandwidth obtained is 3.32 MHz and the maximum voltage gain is 18.46 dB. This shows an improvement in gain and bandwidth as compared to existing structures [7-9]. The simulation results are given in table IV.
### Table IV: Simulation Results

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Center frequency</strong></td>
<td>69.98 MHz</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>3.32 MHz</td>
</tr>
<tr>
<td><strong>Gain</strong></td>
<td>18.46 dB</td>
</tr>
<tr>
<td><strong>Transconductance of single OTA</strong></td>
<td>267.5 µS</td>
</tr>
<tr>
<td><strong>Power consumption</strong></td>
<td>0.578 mW</td>
</tr>
</tbody>
</table>

### 4. Conclusion

The bandpass filter is designed and implemented in Virtuoso analog environment in Cadence tool for a center frequency of 70 MHz and bandwidth 3 MHz. The filter is designed in 0.18 µCMOS process. The gain can be increased further by cascading the OTAs to design higher order filters. Also, common mode feedback block need to be introduced in order to fix the output drain voltages.

### References


Shunt Active Filter Based Voltage Regulation and Harmonic Damping

K. Suresh Kumar¹, P. Vijaya Kumar², L. Ganesh³

ABSTRACT

This paper discusses a shunt active filter intended for installation on a power distribution system. The active filter has an additional capability to regulate the distribution line voltage by means of adjusting reactive power. Theoretical analysis investigates the dynamic performance of combined harmonic damping and voltage regulation. As a result, harmonic damping makes it possible to improve the stability of the control loop for voltage regulation, and the combined harmonic damping. Experimental results are shown to verify the effectiveness of the combined harmonic damping and voltage regulation.

Index Terms – Digital control, distributed generators, power distribution systems, shunt active filters, voltage regulation.

I. INTRODUCTION

Harmonic problems have been serious in industrial and utility power distribution systems. “Harmonic amplification” is one of the most serious problems. It is caused by harmonic resonance between line inductance and power factor correction (PFC) capacitors installed by consumers. Active filters for damping out harmonic resonance in industrial and utility power distribution systems have been researched [1], [10]. The authors have proposed a voltage detecting shunt active filter for installation on a power distribution system [1], [2]. The main purpose of installing the active filter is to damp out the harmonic amplification. The active filter is characterized by behaving like a resistor for harmonic frequencies, resulting in damping out the harmonic amplification throughout a distribution line. Since a real distribution system consists of four to eight distribution lines, installing the active filter on the end bus of each line is effective in harmonic damping. On the other hand, distributed generators such as solar cells, wind turbines and soon have been developed and installed on utility power distribution systems. A distributed generator produces an amount of active power, and injects it into the distribution system. The active power influenced by weather and/or wind velocity causes voltage fluctuations in the distribution line, because it changes independently of power demand from loads. Moreover, distributed generators may produce the following problem in a utility distribution system: Some newly developed towns, apartments and buildings are being equipped with solar cells. This may make difficult voltage regulation throughout multiple distribution lines. A static synchronous compensator (STATCOM) [11]–[13] is one of the most effective solutions to regulate the line voltage. The STATCOM consists of a voltage source converter connected in shunt with the power system, and enables to control a leading or lagging reactive power by means of adjusting its ac voltage. A STATCOM for installation on a distribution system, or the so-called D-STATCOM,
has been researched and put into commission to solve voltage fluctuations and/or voltage flickers [14], [15]. It is possible to use a shunt active filter as a D-STATCOM, because they have the same circuit configuration. However, no literature has addressed the dynamic behavior of the active filter when it performs both harmonic damping and voltage regulation at the same time. This paper discusses a shunt active filter intended for installation on a power distribution system, with focus on voltage regulation capability. Theoretical analysis as well as computer simulation provides the dynamic performance of harmonic damping and voltage regulation. As a result, harmonic damping has the capability to improve the stability of voltage regulation. Thus, adjustment of the feedback gains makes it possible to reduce voltage fluctuation in transient states, when the active filter has the function of combined harmonic damping and voltage regulation. Experimental results are shown to verify the effectiveness of the active filter capable of both harmonic damping and voltage regulation.

II. ACTIVE FILTER FOR HARMONIC DAMPING

A. System Configuration

Fig.1 shows a 200 V simplified power distribution feeder under no-load conditions, where a shunt pure active filter for damping out harmonic propagation is connected by a LC filter. The purpose of the capacitor in LC filter is to impose a high impedance to the fundamental frequency so that the fundamental voltage appears exclusively across the capacitor. This means that no fundamental voltage is applied across the active filter. Fig.2. shows the detailed power circuit of the active filter, which consists of a three-phase voltage fed PWM inverter, and a series connection of L and C per phase. Note that the tuned frequency of L and C is not the 5th harmonic frequency but around the seventh - harmonic frequency. The reason is that the seventh - tuned LC filter is less bulky than the 5th - tuned LC filter as long as both filters have the same inductor as L.

Figure 1: Shunt Active Filter with LC filter.

Figure 2: Power circuit of the Active filter

A. Experimental System

Fig.3 shows the experimental system under no-load conditions. A three-phase power distribution feeder simulator rated at 200V, 50Hz, and 20kW is used for the laboratory experiments. In the feeder simulator harmonic is generated by a non-linear load connected at bus, which produces an amount of harmonic voltage. Table I summarizes the line simulator parameters. When a lossless line is assumed, the characteristic impedance of the feeder simulator, can be calculated as

\[ Z_0 = \frac{L}{\sqrt{LC}} = 0.8 \Omega \]  

(2)
A. Operating principle

The hybrid filter with current control is based on voltage detection. The compensating current of the Active filter consists of a fundamental component and harmonic components. The fundamental component is determined by the impedance of the LC filter while the harmonic components are controlled by the active filter. Three-phase voltages and currents are detected at the installation bus. The harmonic voltage in each phase is extracted from the detected three-phase voltage, and then the harmonic voltage is amplified by a control gain $K_v$. Thus the harmonic current reference $I^*_{ch}$ is given by

$$i^*_{ch} = K_v \cdot V_h$$  (1)

Figure 3: Simplified circuit of the distribution system.

The actual harmonic compensating current is extracted from the detected compensating current. Assuming that is equal to its reference, the hybrid filter behaves as a damping resistor of $1/K_v$ [V/A] for harmonic frequencies. The optimal value of is equal to the inverse of the characteristic impedance of the distribution feeder. With this value, the hybrid filter can damp out harmonic propagation effectively.

3. CONTROL SCHEME

A. Harmonic Damping

There are several methods to extract the harmonic components from the detected three-phase waveforms. Among them, the so-called p - q theory based on time domain has been widely applied to the harmonic extraction circuit of active filters. The detected three-phase voltage is transformed into the D - Q0 coordinates as shown in Fig.4. Two first order digital high pass filters (HPFs) with the same cut off frequency as 17 Hz extract the dc component $V_{d0}^*$, $V_{q0}^*$ and $V_0$ which corresponds to the fundamental frequency in the coordinates.

B. Voltage Detection

In line – voltage regulation part is performed by a feedback control. Two co-ordinates $V_d$ and $V_q$ is compared with harmonic extracted voltage $V_{d0}^*$ and $V_{q0}^*$. A gain $K_v$ amplifies and to produce current references for harmonic damping $I_{d0}$, $I_{q0}$, and $I_0$ as shown in (1). The current reference for the voltage – source inverter is the sum of the current references from the three parts, as follows:

$$I_{d0}^*(s) = K_v \cdot (G_h \cdot V_{d0}^* - V_d) + (V_{dc}^* - V_{dc})$$  (1)

$$I_{q0}^*(s) = K_v \cdot (G_h \cdot V_{q0}^* - V_q)$$  (2)

$$I_{0}^*(s) = 1/3 (V_a + V_b + V_c)$$  (3)

The obtained current reference is converted three phase current reference by inverse D – Q0 transformation $I_{a0}^*$, $I_{b0}^*$ and $I_{c0}^*$. The three three phase reference compensating current is compared with the active filter compensating current extracted from ac system. Thus
three phase compensating current $I_{ca}$, $I_{cb}$, and $I_{cc}$ are produced. The obtained reference current is given to a hysteresis current controller in order to generate controlled gate signal for shunt active filter.

C. DC – Bus Voltage Control

A critical issue in this hybrid active filter is the dc-bus voltage control. The dc bus consists of a single capacitor charged from the power supply. During operation, the active filter may absorb an amount of active power into, or release it from, the dc capacitor. Excessive active power absorption will increase the dc-bus voltage, and may damage the active filter. The strategy used to control the dc-bus voltage is based on active power control. According to the $D-q0$ theory, a dc component in the $D-q0$ coordinates corresponds to active power. No direct axis current on the $D-q0$ coordinates flows in the LC filter. Thus, the active power is controlled by adjusting the quadrature axis component. The direct axis is set to zero. Fig. 3 shows a block diagram for the dc bus voltage control. The dc-bus voltage is detected and compared with a reference, amplifying the error signal by a control gain of 0.12. A limiter is included in the dc-bus control loop. It is designed to ensure a smooth transient response and to avoid sudden increments or decrements in the dc-bus voltage. It is also designed to prevent the control loop from numerical saturation in the DSP signals. The limiter is set to ± 2.5 V in the digital controller which corresponds to 25% of the maximum control signal. For a 40-V dc-bus voltage, the maximum dc-bus control signal corresponds to a ± 10 V peak-to-peak fundamental voltage for the inverter. voltage for the inverter.

3. SIMULATION RESULTS

Table I

<table>
<thead>
<tr>
<th>Circuit Parameters used for the Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source voltage $V_a = V_b = V_c$</td>
</tr>
<tr>
<td>power $p$</td>
</tr>
<tr>
<td>Frequency $F$</td>
</tr>
<tr>
<td>Line inductance $L_a = L_b = L_c$</td>
</tr>
<tr>
<td>DC- Capacitor $c$</td>
</tr>
<tr>
<td>DC- voltage $V_{dc, ref}$</td>
</tr>
</tbody>
</table>

Table II shows the circuit parameters used in the simulation. The simulation is carried for distribution system with and without shunt active filter. Total harmonic distortion is calculated for the system voltage and current.

Table III

<table>
<thead>
<tr>
<th>Total Harmonic Distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>$I_a$, $I_b$, $I_c$</td>
</tr>
<tr>
<td>Without shunt active filter</td>
</tr>
<tr>
<td>With shunt active filter</td>
</tr>
</tbody>
</table>
Figure 5: Three Phase voltage of distribution system without Shunt Active Filter

Figure 6: Three Phase voltage of distribution system with Shunt Active Filter

Figure 7: Source current of distribution system without Shunt Active Filter

Figure 8: Source current of distribution system with Shunt Active Filter

Figure 9: Compensating current with Shunt Active Filter

Figure 10: DC – Voltage of Shunt Active Filter
5. CONCLUSION

This paper has dealt with a shunt active filter for installation on a power distribution system, with focus on harmonic damping and voltage regulation performance. Harmonics present in the distribution system is significantly reduced by shunt active filter. Simulated results have demonstrated good voltage – regulation performance, and have verified stable operation. Total harmonic distortion is also measured and compared. the main disadvantage of this P – Q theory is selective harmonics cannot be damped.

REFERENCES


Author's Biography

Mr. K. Sureshkumar received his B.E., degree in Electrical and Electronics Engineering from the Government College of Technology, Coimbatore, affiliated to Bharathiar University, in 1999 and M.E., degree in Applied Electronics from the Coimbatore Institute of Technology, Coimbatore, affiliated to Bharathiar University, in 2002. He is currently a research scholar at the Anna University, Coimbatore and presently working as Assistant Professor in the Department of Electrical & Electronics Engineering, Velammal Engineering College, Chennai. He has around
nine years of teaching experience. His primary areas of interest are in Embedded System Design, Power Quality and VLSI Design.

Dr. P. Vijayakumar born on July 6th 1971 at Coimbatore, Tamil Nadu state, India, graduated in Electrical and Electronics Engineering from PSG College of Technology (Bharathiar University) during 1992. He obtained his Post Graduation in Applied Electronics from PSG College of Technology (Bharathiar University) during 2002. He completed his Doctorate in the area of Low Power VLSI Design from Anna University, Chennai during 2007. At present he is working as Professor and Head in the Department of Electrical and Electronics Engineering at Karpagam College of Engineering, Coimbatore. He has around eleven years of teaching experience and about seven years of Industrial experience. His areas of interest are VLSI Design, Applied Electronics and Virtual Instrumentation. He has published seven papers in International Journals and six papers in National Journals besides more than 20 papers in various national and international conferences. He is a member of ISTE, ISSS, VSI and SSI.

Mr. L. Ganesh received his B.E., degree in Electrical and Electronics Engineering from Sri Venkateswara College of Engineering and Technology, Thiruvallur, affiliated to Anna University, Chennai in 2008 and M.E., degree in Power Systems from Velammal Engineering College, Chennai, affiliated to Anna University, Chennai in 2010. He currently applied for a Ph.D. His primary areas of interest are in Power Systems, Active filters, FACTS and Power Quality.
A Classification of Character Usage in Unique Addresses Employed for Accessing Yahoo! Groups Service

Jatinderkumar R. Saini, Apurva A. Desai

ABSTRACT

A tremendous increase in the use of the internet for online communication like message sending is witnessed worldwide. Yahoo! Inc. provides such a service in the form of Yahoo! Groups. Each such group is identified and accessed using a unique group address. The current paper presents an analysis of nearly 5000 Yahoo-group addresses. It presents a classification of characters employed by users in designing these addresses into 5 major sets. Our results show that around 90% characters used for designing the Yahoo-group addresses are alphabets whereas the remaining 10% constitute from the domain of digits and special characters. The paper also elaborates on the divisional values of these proportions highlighting the user’s preference for selecting a particular character. To the best of our knowledge this is the first attempt to study online user behavior based on the classification of character usage for designing a unique online identifier.

Keywords: Character Usage, Digits, Lower-case Letters, Upper-case Letters, Yahoo! Groups, Yahoo-group address, Yahoo-group Identifier

1. INTRODUCTION

The growth of internet has provided the users of 21st century with new means of communication. According to Jones and Fox [5], instant messaging, social networking, and blogging have gained ground as communications tools, but email remains the most popular online activity. Besides these, other online activities include sharing views and discussing various topics through groups and discussion forums. Yahoo! Inc. is one company that provides such internet services worldwide. The specific services provided by Yahoo! Inc. include mail, news, search, groups, video and maps, to name a few. Yahoo! Inc. was founded in 1995 and Yahoo! Groups which provides one of its services, came into existence in 2000 [9].

The Internet Marketing Definitions website [4] describes Yahoo! Groups as a service that operates as both electronic mailing list and Internet forum. Members can post and read messages either by receiving them in their email account or by going to the group’s homepage. Since the inception of Yahoo! Groups there has been a tremendous increase in its user bank. This service allows the user to create a group address and provide group name. Technically speaking, the group address is the one that is used to uniquely identify a Yahoo-group. In the current work, the concentration has been only on group addresses and not on group names which can consist of any character combination and are meant to provide a brief introductory description line of the group.
2. RELATED WORK

There are quite a few research instances available with research community, that have thrown light on statistical analysis of character usage. Even for those available, most of the works have concentrated on usage of characters as a stylometric parameter. Aaronson [1] has listed a set of 15 features which are important for performing data-driven stylometric analysis. These features are Ampersand Sign, Apostrophe, Colon, Comma, Dash, Dollar Sign, Exclamation Mark, Forward Slash, Left Parenthesis, Percent Sign, Period, Question Mark, Right Double Quote, Right Parenthesis and Semi-colon.

An instance of analysis of character usage is provided by the work of Saini [6]. In his work, he has presented a detailed discussion on the usage of characters by spammers for sending Unsolicited Bulk Emails (UBE) commonly known as spam emails. Calix et al. [3] in their work which was targeted towards e-mail author identification and authentication have also employed the statistical analysis of characters used in the emails. They have provided a platform to identify the author of a given e-mail based on writing-style features like number of words, number of commas, number of times “well” appears, etc. To the best of our knowledge and survey of related research literature, this is the first formal attempt aimed towards classification of usage of characters used for designing the Yahoo-group addresses.

3. METHODOLOGY

During this phase, the goal of following a methodological sequence of steps was to obtain a list of Yahoo-group addresses which could be subjected to further analysis. For better comprehension of the process, the sequence of steps followed towards this end is presented in the form of a flow-chart in Figure 1. Towards the first step, a corpus, consisting of various web pages retrieved from Yahoo website, is created. Specifically, the target area in the Yahoo website is Yahoo Groups, which follows a hierarchical structure. This can be visualized as an inverted tree, wherein Yahoo itself acts as the root node at top-most level and other nodes reside at various lower levels.
The basic model of free text consists of documents which are sequences of basic units called Tokens. In English language the tokens are words [11] and the act of breaking the text into tokens is called Tokenization. In order to make it easier for analysis and further processing, Tokenization is performed on the corpus of Yahoo web pages. Here, the problem was to identify the token delimiters. This means to say that there was a need to identify the beginning and ending of a token in a sentence. As a de-facto standard, it is known that space character is not allowed in web technologies for a number of instances like creation of email addresses and (Uniform Resource Locator) URL addresses. Besides, Yahoo also has explicitly specified that spaces are not to be allowed for creation of a Yahoo-group address. We tried the creation of a Yahoo-group address by using non-printable space (ASCII Code 32) character. The submission of this request yielded us with the expected and desired message. The message read as, “The group address may only contain letters, numbers, and - or _ characters. It may not have more than one - or _ in a row.” Consequently, it was possible to design the tokenization process to identify the words demarcated by any character other than letters, numbers, Dash (-) and Underscore (_). Tokenization, hence, was done in such a way that only unigrams could be focused.

The various text-processing activities completed so far yielded us with a text corpus in the Bag of Words (BOW) form. In BOW representation of a text document, terms or tokens in the document are identified with words in the document. Hence this representation is also called Set of Words (SOW) [7]. The BOW was further processed by removal of stop words from it. Sebastiani [8] has defined Stop Words as topic-neutral words such as articles and prepositions, which are eliminated in a pre-processing phase. Bharati et al. [2] have defined them as
few words which have high frequencies in all the categories, and hence are irrelevant for the classification exercise. The removal of stop words from the vector was helped by the fact that consecutive Dash (-) or consecutive Underscore (_) or consecutive combination of Dash and Underscore in any order was not allowed by Yahoo for the creation of a valid Yahoo-group address. Further, it was found that for creation of a Yahoo-group address, Yahoo explicitly specifies that it cannot have length more than 60 characters. Hence, tokens with length of more than 60 character length were also removed from the vector due to their irrelevance for the context of current work.

It is also to be noted that Yahoo allows the usage of only digits for creation of a Yahoo-group address. For instance an address formed of numbers like 12345 will be treated as a valid Yahoo Groups address by Yahoo website. Hence, the addresses consisting of only digits were prevented from removal as stop words. This processing of the BOW resulted in a vector consisting of 5129 word entries. As the vector was generated from the processing of a large text corpus, it was natural for the vector to contain duplicated entries. As a next step, the vector was refined by selecting only unique words. The removal of Stop Words was done before the removal of duplicated entries because the bulk of stop words was much more than the bulk of duplicated entries.

Table 1: Partial Snap-shots of Vector Containing Yahoo-group Addresses

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Vector Index No.</th>
<th>Yahoo-group Address</th>
<th>Sr. No.</th>
<th>Vector Index No.</th>
<th>Yahoo-group Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>2joel_kami</td>
<td>16</td>
<td>573</td>
<td>aykan8691</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>2ofy</td>
<td>17</td>
<td>574</td>
<td>aysegulunfistiklari</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>303kd_others</td>
<td>18</td>
<td>575</td>
<td>azalizan_ismail</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>317arazi27</td>
<td>19</td>
<td>576</td>
<td>azanimalrights</td>
</tr>
<tr>
<td>5</td>
<td>41</td>
<td>33_SSB</td>
<td>20</td>
<td>577</td>
<td>aziyar</td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>3bru</td>
<td>21</td>
<td>578</td>
<td>azncrew</td>
</tr>
<tr>
<td>7</td>
<td>43</td>
<td>3e6-o7</td>
<td>22</td>
<td>579</td>
<td>b_vikas1986</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>3lionsroaras1</td>
<td>23</td>
<td>580</td>
<td>b122</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
<td>407-thottigang</td>
<td>24</td>
<td>581</td>
<td>B29_pradan</td>
</tr>
<tr>
<td>10</td>
<td>46</td>
<td>43MirpurFriends</td>
<td>25</td>
<td>582</td>
<td>baarn</td>
</tr>
<tr>
<td>11</td>
<td>47</td>
<td>4asap</td>
<td>26</td>
<td>583</td>
<td>babajigroup</td>
</tr>
<tr>
<td>12</td>
<td>48</td>
<td>4-ever_friends</td>
<td>27</td>
<td>584</td>
<td>BABASEMENTgirls</td>
</tr>
<tr>
<td>13</td>
<td>49</td>
<td>4frenz</td>
<td>28</td>
<td>585</td>
<td>bablo_bablo40</td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td>4gats</td>
<td>29</td>
<td>586</td>
<td>baby-girl-4eva</td>
</tr>
<tr>
<td>15</td>
<td>51</td>
<td>4_i_green-mindedclub</td>
<td>30</td>
<td>587</td>
<td>babysugar</td>
</tr>
</tbody>
</table>
The reduced vector obtained after removal of duplicated entries consisted of 4940 entries and was ordered by sorting it in lexicographic ascending manner. The entries in the token vector of our experimental setup represented the Yahoo-group identifiers of the real world. Out of a total list of 4940 entries, two partial snap-shots of this vector from index positions 37 to 51 and 573 to 587 are selected randomly and presented in Table 1.

4. RESULTS AND FINDINGS

The end of various activities of processing the text-corpus was a one-dimensional vector consisting of 4940 unique Yahoo-group addresses. By the analysis of the data available in the vector, it was found that Yahoo allows upper-case, lower-case as well as mixed-case characters to be used for creation of Yahoo-group address. Moreover, by attempting to create multiple Yahoo-group addresses by changing the case of letters, it was found that Yahoo does not treat its group addresses as case-sensitive. This means to say that if a group address called ‘abc’ is already created, then one cannot create a group address called ‘ABC’ or one with mix-case letters like ‘Abc’. This is true also for addresses created in different categories. In other words, it can be said that a Yahoo-group address is unique irrespective of its belonging or categorization to any category in Yahoo groups. For the current work, the treatment of text corpus is considered to be case-sensitive. This was done to help meet our objective of studying the usage of lower-case and upper-case letters by users, in designing the Yahoo-group addresses. Each of the address in the vector partially depicted in Table 1, was a unigram and was made up of specific characters. These constituent characters used for creation of addresses into 5 categories were classified and the corresponding listing is presented in Table 2.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Character Type</th>
<th>No. of CCR</th>
<th>Usage of CCR</th>
<th>CCR Usage in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lower-case Letters</td>
<td>26</td>
<td>52764</td>
<td>83.46</td>
</tr>
<tr>
<td>2</td>
<td>Upper-case Letters</td>
<td>26</td>
<td>4048</td>
<td>6.40</td>
</tr>
<tr>
<td>3</td>
<td>Digits</td>
<td>10</td>
<td>3945</td>
<td>6.24</td>
</tr>
<tr>
<td>4</td>
<td>Underscore</td>
<td>1</td>
<td>2159</td>
<td>3.41</td>
</tr>
<tr>
<td>5</td>
<td>Dash</td>
<td>1</td>
<td>306</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>Others</td>
<td>1</td>
<td>9880</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>65</strong></td>
<td><strong>63222 (except others)</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Table 2: Classification and Usage of Constituent Character Representations (CCR) in Yahoo-group Addresses

From Table 2, it can be seen that the Yahoo-group addresses could be created by either exclusive usage or any combinational usage of 26 English lower-case letters, 26 English upper-case letters, 10 digits from 0 (zero) to 9 (nine) and two special characters Dash and Underscore. Here it is to be noted that only consecutive usage of special characters is not allowed by Yahoo. Finally, it was also tried to analyze the vector for any other characters that could have been missed otherwise or that could have admitted in unknowingly.
It is noteworthy to see that the total number of other characters found by us was 9880. This value is equal to multiplication of 2 with 4940 which is the total number of tokens. This value comes into play due to counting of carriage-return and new-line characters during the processing of the token vector. Also since no other character was found either omitted for analysis or admitted unknowingly, for further discussion, value of 9880 is not given any consideration. This value has been treated as statistically insignificant from the viewpoint of objectives of current work. Hence, in Table 2, the number of Constituent Character Representations (CCR) of Yahoo-group addresses for character type of 'others' is also depicted as 1. Table 2 also depicts the usage of various representations of each character type by the designers of Yahoo-group addresses.

The data corresponding to this is presented graphically in Figure 2. For simplicity and better comprehension, the values of character usage in Figure 2 have been represented in terms of percentage of usage. From Figure 2, it can be seen that users have made almost 90% use of the English alphabet characters for designing the Yahoo-group addresses. Further, the lower-case letters constitute 83.46% of the total character usage whereas upper-case English letters constitute 6.40% of the total value. An interesting result to note here was that almost 10% of the characters used for designing Yahoo-group addresses are not from the domain of English alphabets. This value of 10% is further divided into three categories for ten digits and two special characters. The digits contribute 6.24% of the total character usage for designing Yahoo-group addresses. The underscore special character constitutes 3.41% followed by the dash special character constituting a meager amount of 0.48% of the total character usage.

5. CONCLUSION

The current paper is an attempt to analyze the addresses used for uniquely identifying the access of Yahoo! Groups service. The classification of usage of various characters, employed towards designing nearly 5000 Yahoo-group addresses by users, has been presented. It is concluded that even though Yahoo allows the design of group addresses with lower-case, upper-case and mixed case characters, it does not treat them as case-sensitive. Hence, the usage of different case characters in Yahoo-group addresses is meant either to make them visually pleasing or to let users incorporate different words in the identifier, each word demarcated by change in case of the letter.

It is further concluded that the character set used for designing Yahoo-group addresses could be divided into five categories for lower-case letters, upper-case letters, digits, underscore and dash. Additionally, of the total character usage, lower-case letters constitute the maximum amount of nearly 84% share for designing the
Yahoo-group addresses. This is followed by an almost equal usage of upper-case letters and digits with value of around 6% for both. The use of underscore is very less with value of around 3%. The preference of usage of dash character by users for designing Yahoo-group addresses is minimum, with a value of around 1% only.

Our results are best reported on the dataset used and we do not promote or discourage the use of any specific character for designing of Yahoo-group address. We just present the classification of usage of characters based on the preferences of users for designing the Yahoo-group addresses. The current work is not only an insight into the usage of characters for designing Yahoo-group identifiers but is also having a wide range of general applicability to other domains. On the sidelines of the current study, it is advocated that the paper has also provided an insight into usage and behavior of selection of characters by users for designing passwords as well as other unique identifiers in the online and offline world.

REFERENCES


Author’s Biography

Jatinderkumar R. Saini is Ph.D. from Veer Narmad South Gujarat University, Surat, Gujarat, India. He secured First Rank in all three years of MCA at college and has been awarded Gold Medals for this. He is also recipient of Silver Medal for B.Sc. (Comp. Sci.). He is IBM Certified Database Associate – DB2 as well as IBM Certified Associate Developer – RAD. He has presented 12 papers in international and national conferences, all sponsored by either AICTE or ISTE. One of his papers has also won the ‘Best Paper Award’. His 5 papers have been accepted for publication at international level and 8 papers have been accepted for national level publication. He is a member of many academic committees. He is also a member of various international and national professional bodies and scientific research academies and organizations.

Apurva A. Desai completed his post graduation from Veer Narmad South Gujarat University, securing First Rank in the University. He earned his Ph.D. in the year 1997. He has got a long teaching and research experience since 1990. Many students have completed Ph.D. and M.Phil. under his supervision. He has delivered many lectures and invited talks as resource person in various national and international events. He has 3 books and 13 research papers to his credit. He is chairman of Board of Studies (Computer Science) for last 6 years. He has attended many International and National conferences. He has also visited Canada and Italy for various academic programmes.
Effective Utilization of Processor in a Distributed Environment
For PC Clusters Analysis

N. Kavitha 1  S. Karthikeyan 2

Abstract - Distributed Data mining is expected to perform partial analysis of data at clients and then to send the outcome, as results to the server where it is sometimes required to be aggregated to the global result. The primary issues to be considered for Distributed Data mining are Scalability, privacy of data and autonomy of data. These issues can be easily handled when we go for intelligent software agents for Distributed Data mining, because of its inherent features of being autonomous, capable of adaptive and deliberative reasoning. So by using software intelligent agents the average idle.

Key words: Distributed Data Mining, Software Intelligent Agent, and Scalability

1. INTRODUCTION

With the enormous amount of data stored in files, databases, and other repositories, it is increasingly important, if not necessary, to develop powerful means for analysis and perhaps interpretation of such data for the, extraction of interesting knowledge that could help in decision-making. Distributed data mining refers to the mining refers to the mining of distributed data sets. The data sets are stored in local data bases, hosted by local computers, which are connected through a computer network. Data mining takes place at a local level and at a global level where local data mining results are combined to gain global findings. In some applications, data are inherently distributed, but it is necessary to gain global insights from the distributed data sets. Distributed Data mining has emerged as a means for identifying patterns and trends from large quantities of data in a quick manner. The implementation of data mining in distributed computing has become crucial for ensuring system scalability and interactivity as data continues to grow inexorably in size and complexity time per node is kept low. To gain good understanding of the utilisation of the CPU time of identified workstations time to time, the software agent for the distributed environment was implemented. Also, this intelligent agent is capable of giving alert information to the server, when the CPU usage time of any identified workstations exceeds threshold value. Usage of intelligent agent approach to distributed data mining was applied and the expected performance was obtained successfully.

2. METHODOLOGY

PROPOSED SYSTEM MODEL

The Proposed System was aimed to develop the effective utilisation of the CPU time in workstations time to time; the software agent for the distributed environment was implemented. Also, this intelligent agent is capable of giving alert information to the server, when the CPU usage time of any identified workstations exceeds threshold value. The Apriori algorithm was implemented to mine the transaction data sets.

1Research Scholar, Department of CS, Karpagam University, Coimbatore. Email: nkavee1@yahoo.co.in
2Assistant Professor, Department of Information Technology, College of Applied Sciences, Sohar, Sultanate of Oman. Email: skarthi@gmail.com
Intelligent agent (IA): is an autonomous entity which observes and acts upon an environment (i.e. it is an agent) and directs its activity towards achieving goals (i.e. it is rational). Intelligent agents may also learn or use knowledge to achieve their goals. They may be very simple or very complex: a reflex machine such as a thermostat is an intelligent agent, as is a human being, as is a community of human beings working together towards a goal.

Primarily, the server interacts with the agent to know about the status of the clients in terms of idle time, periodically. Now, agent contacts client to know about its idle time. This in turn makes the client to response with its current idle time. This will be given as response to the server. Based on this information, now the server will start splitting the large data sets among various clients. After performing association rule mining on this given data set, clients send the processed results to the agent. Now, the agent collects the processed result and submits to the server to display in the required format.

3. IMPLEMENTATION OF APRIORI ALGORITHM IN A SINGLE MACHINE AND ANALYZE THE RESULTS

In this step, the implementation of Apriori algorithm for association rule mining in a single machine was considered and the results were analyzed. The apriori algorithm was implemented using JDK 1.4 and executed successfully to generate association rules for the given identified inputs in a single machine.

Different numbers of records were given as input and the results were analyzed and the same was presented in the following Table 1. The processing time required for mining 1, 00,000 records with single system is 1.1 secs.

<table>
<thead>
<tr>
<th>No Of Records</th>
<th>Duration (In Secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single System</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.03</td>
</tr>
<tr>
<td>20</td>
<td>0.06</td>
</tr>
<tr>
<td>30</td>
<td>0.08</td>
</tr>
<tr>
<td>40</td>
<td>0.11</td>
</tr>
<tr>
<td>60</td>
<td>0.19</td>
</tr>
<tr>
<td>75</td>
<td>0.27</td>
</tr>
<tr>
<td>85</td>
<td>0.58</td>
</tr>
<tr>
<td>100</td>
<td>1.1</td>
</tr>
</tbody>
</table>

4. IMPLEMENTATION OF APRIORI ALGORITHM IN THE NUMBER OF WORKSTATIONS USING INTELLIGENT AGENTS AND ANALYZE THE RESULTS

Apriori Algorithm in N (N=6) identified workstations were considered. This involves the calculation of CPU idle time in the workstations. Distribution of dataset among these workstations based on the calculated CPU idle time using intelligent agents were considered and the results were analyzed. The distribution of data sets among the identified workstations and the agent program that monitor the CPU idle has been implemented using JDK 1.4. Proposed approach consists of two logical components.

Client- There may be in N-number of clients connected in intranet. In contrast to the model proposed by, here the client receives data sets and mining operation to be performed as parameters from the server agent. Local agent running in the client machine takes care performing specified operation and storing the final results. The client side agent sends resource utilization information periodically to the server machine. Additionally, it is also having the Responsibility of alerting the server, if it is overloaded or can’t execute the task because of various reasons such as taking appropriate actions in the unwanted situation like overloaded message from the client agent, network failure etc.,
**Server** - In this proposed model, the server machine stores the large database or Data warehouse where the millions of records are stored and intelligent agent running in this machine. The System flowchart for the proposed work is illustrated in the Fig 1 and the frame work for the proposed system is illustrated in Fig 2.

**5. APRIORI ALGORITHM – IMPLEMENTATION**

The implementation of Apriori algorithm basically consists of the following steps:

- **Join Step**: $C_k$ is generated by joining $L_{k-1}$ with itself.
- **Prune Step**: Any $(k-1)$-itemset that is not frequent cannot be a subset of a frequent $k$-itemset.

**PSEUDO CODE**

- $C_k$: Candidate itemset of size $k$
- $L_k$: frequent itemset of size $k$
- $L1 = \{\text{frequent items}\}$

```plaintext
for (k = 1; Lk != Ø; k++) do begin
    Ck+1 = candidates generated from Lk;
    for each transaction t in database do
        increment the count of all candidates in Ck+1 that are contained in t
    Lk+1 = candidates in Ck+1 with min_support
end
return $\bigcup L_k$;
```

**Figure 1: System Flow Diagram for the Proposed Work.**

**Figure 2: Framework for the proposed work.**
GENERATION OF CANDIDATE SET

Suppose the items in L_{k-1} are listed in an order

Step 1: self-joining L_{k-1}
insert into C_k
select p.item1, p.item2, …, p.item{k-1}, q.item{k-1} from L_{k-1}
p, L_{k-1} q
where p.item1=q.item1, …, p.item{k-2}=q.item{k-2}, p.item{k-1} < q.item{k-1}

Step 2: pruning
for all item sets c in C_k do
for all (k-1)-subsets s of c do
if (s is not in L_{k-1}) then delete c from C_k

6. RESULTS AND DISCUSSION

This research work uses database from the Departmental Store of Nilgris. It consists of 1, 00,664 records with 15 different fields. The primary information in this item code module includes, Date of item purchased, item quantity, item weight, item cost. With this information, we are aiming at the following association and predictions.

• Number of items for each transaction.
• Regularity of Items sold
• Selection of items by the customer and the transaction for the particular items in terms of the days and months and year.

The original database consists of additional fields and information. It is essential to remove these fields and make transformations to improve the accurate results in the mining process and to improve the efficiency of the whole system.

The unnecessary information like Leave Type, Created User, Crated Date, Modified User, Modified Date and Batch were deleted from original database which are not related to the generation of association rules of the mining process.

All the entries marked with item code for each transaction, of the transformed database.

All the entries in AT_DAY, which represents the day in which the transaction was marked, were assigned a unique number in the transformed database to speed up the mining process.

AT_DATE, which represents the date in which the transaction was entered of the original database, is transformed to two fields AT_MONTHS, AT_DAYS contains Month Information and the Number of Days (calculated from the Difference between first record and existing date entry).

The Table 1 shows the original database considered and the Table 2 reflects the transformations performed on the original table.

<table>
<thead>
<tr>
<th>L_NAME</th>
<th>Item Name For Each Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>L_CODE</td>
<td>Purchased Items Code For Each Transaction</td>
</tr>
<tr>
<td>T_ID</td>
<td>Unique Value For Each Transaction: Transaction Id</td>
</tr>
<tr>
<td>TD_DATE</td>
<td>Corresponding Day Of The Item Purchased</td>
</tr>
<tr>
<td>I_QTY</td>
<td>Purchased Item Quantity</td>
</tr>
<tr>
<td>I_COST</td>
<td>Purchased: Item Cost</td>
</tr>
</tbody>
</table>
Effective Utilization of Processor in a Distributed Environment
For PC Clusters Analysis

Table 3 – Structure of Transformed Database

<table>
<thead>
<tr>
<th>I_CODE</th>
<th>Purchased Items Code For Each Transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>I_QTY</td>
<td>Purchased Item Quantity</td>
</tr>
<tr>
<td>I_COST</td>
<td>Purchased: Item Cost</td>
</tr>
</tbody>
</table>

Data Transformation and Cleaning

In this part of work involves forming the data warehouse and perform any transformations and cleaning needed in our collected data. The transformed and cleaned data can be useful to be mined using Apriori algorithm. Here, about 1,00,664 records were transformed and cleaned.

3 TEST CASES

Number of records = 1,00,000
Number of nodes = 4

Expected Performance

The expected results are tabulated in the following Table.

Table : 4

<table>
<thead>
<tr>
<th>No Of Records</th>
<th>Duration(In Secs)</th>
<th>Single System</th>
<th>N=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.06</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.08</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.11</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.19</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>0.27</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>0.58</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1.1</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

Actual Performance

The actual results are tabulated in the following table

Table 5

<table>
<thead>
<tr>
<th>No Of Records</th>
<th>Duration(In Secs)</th>
<th>Single System</th>
<th>N=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.06</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.08</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.11</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.19</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>0.27</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>0.58</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1.1</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

Deviation

There is no deviation between the expected result and the actual result. So the status of the test case is said to be success.

7. PERFORMANCE ANALYSIS

The Figure 3 shows the performance comparison of data mining in the single system versus distributed system with 2 workstations. The processing time required for mining 1,00,000 records with single system is 1.1 secs, where as, with the distributed system, took 0.64 secs for the same number of records. This clearly indicates the effectiveness of distributed data mining in terms of maximum utilization of resources.

FIGURE: 3 Performance analysis – Single Vs Distributed System. (N=2)
FIGURE: 4 Performance analysis – Single Vs Distributed System (N=4)

The Figure: 4 show the performance comparison of data mining in the single system versus distributed system with 4 workstations. The processing time required for mining 1,00,000 records with single system is 1.1 secs, whereas with the distributed system, took 0.58 secs for the same number of records. This clearly indicates the effectiveness of distributed data mining in terms of maximum utilization of resources.

FIGURE: 5 Performance analysis – Single Vs Distributed System (N=6)

The Figure 5 shows the performance comparison of data mining in the single system versus distributed system with 6 workstations. The processing time required for mining 1,00,000 records with single system is 1.1 secs, whereas with the distributed system, took 0.49 secs for the same number of records. This clearly indicates the effectiveness of distributed data mining in terms of maximum utilization of resources.

By seeing the performance analysis graph, conclusion made that distributed data mining saves the time as well as reduces the average idle time of the CPU in PCclusters.

8. CONCLUSION

The data mining is an essential activity that is required for any industry that is interested in forecasting the business trends and analysing the behavioural study of customers. The effective and maximal utilisation of the existing resources is the need of the hour for these activities. This paper achieved this objective of effective utilisation of computing resources in efficient manner. Association rules are generated to study and forecast trends in the academic environment by the implementation of Apriori algorithm. Finally the effective utilisation of the CPU is identified.

REFERENCES


Effective Utilization of Processor in a Distributed Environment
For PC Clusters Analysis


[7] Chris Clifton, Murat Kantarcioglu, Jaideep Vaidya, Xiaodong Lin, Michael Y. Zhu, "Purdue University,".Tools for privacy preserving Data mining.” In proceedings of the ACM SIGKDD Explorations, v.4 n.2, p.28-34, December 2002


[16] www.csc.liv.ac.uk/~frans/KDD, 2004


Author’s Biography

N. Kavitha Received UG Degree from Bharathiar University and PG Degree from mother Teresa Women’s University. I did my Phil Degree at Bharathidasan University. Currently Pursuing Ph.D in Karpagam University. Having 8yrs and 6 months in teaching Experience. My Research Area is Data Mining. I have presented 9 papers in various conferences. Other Areas include Computer Networks.

Karthikeyan S. received the Ph.D. Degree in Computer Science and Engineering from Alagappa University, Karaikudi in 2008. He was working as a Professor and Director in School of Computer Science and Applications, Karpagam University, Coimbatore. At present he is in deputation and working as Assistant Professor in Information Technology, College of Applied Sciences, Sohar, Sultanate of Oman. He has published more than 14 papers in National/International Journals. His research interests include Cryptography and Network Security
A Comparative Study of Texture Features for Image Segmentation

G. Madasamy Raja and V. Sadasivam

ABSTRACT

Image segmentation is one of the most significant tasks in image processing. The outcome of image segmentation is a group of regions that collectively cover the entire image, each of the pixels in a region are homogeneous with respect to some characteristic or computed property, such as color, intensity, or texture. Already there are many approaches proposed for texture feature extraction which can be useful for image segmentation. One of the important issues here is how well these methods work on differentiating various textures that are available in a single image. This paper considers two texture measures namely Texture Spectrum and Uniform Local Binary Pattern for texture segmentation and evaluates their performance based on the segmentation accuracy. Two different synthetic images are used in experiments. One image contains four different textures and another one contains two different textures. MATLAB has been used for the implementation purpose.

Keywords: Texture, Texture Spectrum, Local Binary Pattern, Uniform Local Binary Pattern, Texture Segmentation, Supervised Segmentation.

I. INTRODUCTION

MAGE segmentation refers to the process of partitioning a digital image into multiple regions (sets of pixels) [1]. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image.

Each of the pixels in a segmented region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Some of the practical applications of image segmentation are:

- Medical Imaging
  - Locate tumors and other pathologies
  - Measure tissue volumes
  - Computer-guided surgery
  - Diagnosis
  - Study of anatomical structure

- Locate objects in satellite images (roads, forests, etc.)

- Face recognition.

Segmentation merely based on the grey value alone is not efficient and features like color, texture, gradient magnitude or orientation, measure of a template match etc., can also be considered for the better output. Texture is an important characteristic for the analysis of many types of images [2].

Image texture is believed to be a rich source of visual information about the nature and three-dimensional shape...
of physical objects. Nowadays texture based finger print matching is an active research area because textures are complex visual patterns composed of entities or sub-patterns that have characteristic brightness, color, slope, size, etc., [3].

Texture measurements can also be used to segment an image and classify its segments [4]. Texture segmentation is to segment an image into regions according to the textures of the regions. Texture classification or segmentation is not an easy problem because there is not any precise definition of what a texture is.

This paper evaluates two important texture measures that are widely used in the present research for image segmentation in variety of applications. In this paper rest of the portion is organized as follows. Section II describes the texture measures that are used in this study and section III discusses the algorithm which is used for the segmentation. Experimental results are provided in section IV and finally the results with concluding remarks are discussed in section V.

II. TEXTURE MEASURES USED IN THIS STUDY

A. Texture Spectrum Method

The Texture Spectrum, one of a statistical method of texture analysis, focuses on texture characterization and discrimination [5]-[7]. The texture spectrum is based on the computation of the relative intensity relations between the pixels in a small neighborhood and not on their absolute intensity values. The importance of the texture spectrum method is determined by the extraction of local texture information for each pixel and of the characterization of textural aspect of a digital image in the form of a spectrum. The texture spectrum method results in a vector which characterizes the original image and the output spectrum maintains the texture characteristics of the input image.

Texture Spectrum method uses a basic concept called Texture Unit (TU). A Texture unit is characterized by eight pixels each of which has one of three possible values (0, 1, 2), obtained from a neighborhood of 3x3 pixels. Fig. 1. shows the method of forming the Texture Unit. If the intensity value of the central pixel is considered as X0 and the intensity value of each neighboring pixel as Xi, the set that is considered as the smallest complete unit of the under consideration image is: X = {X0, X1, X2, ..., X8}. This technique compares the greylevel of the central pixel (the one which is currently being processed), X0, with those of its neighbors, Xi (1<i<8), and records three logical relationships: smaller, equal and greater; noted by E, and coded as “0” “1” or “2”, respectively(1). By this way, each image pixel generates a Texture Feature Vector, called as Texture Unit which is defined as: TU = {E1… E8}. Equation (1) is used for deriving TU.

![Figure 1. Texture Unit.](image1)

![Figure 2. A basic LBP operator.](image2)
According to equation (1), each element can be assigned one of three possible values so that the total number of possible texture units for the eight elements can be estimated as \(3^8 = 6561\). As there is no unique method for labeling the texture units, equation (2) is followed for numbering the texture units.

\[
N_{TU} = \sum_{i=1}^{8} E_i \times 3^{i-1} \quad (2)
\]

In the equation (2) \(N_{TU}\) varies from 0 to 6560. Texture spectrum method is still a dominating texture measure in the research to reveal texture information in digital images and it has a promising discriminating performance for different textures [8].

**B. Uniform Local Binary Patterns Method (ULBP)**

Local Binary Patterns (LBP) operator introduced by Ojala et al [9] is a simple filter that labels an image by thresholding the neighborhood of each pixel with the value of the center pixel and gives the obtained result as binary values. Histogram of the labeled image is then used as a mean of texture description. The basic LBP operator is illustrated in Fig. 2. Many researches have been going on this texture descriptor method to enhance its uses for the various applications mainly for the reason that it is having very low computational complexity [10]-[13].

Equation (3) describes how each pattern in the image should be assigned a unique label.

\[
P_v = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (3)
\]

where \(s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \)

Figure 3: The 36 unique rotation invariant binary patterns that can occur in the circularly symmetric neighborhood. Black and white circles correspond to bit values of 0 and 1 in the 8-bit output of the operator. The first row contains nine “uniform” patterns.
In equation (3) \( g_c \) is the intensity of the center pixel, \( g_p \) is the intensity of the neighbor \( p \), \( s(x) \) is the step function, \( P \) represents angular resolution and \( R \) represents spatial resolution. The LBP \((P,R)\) operator produces \( 2^P \) different patterns. It is clear that LBP defined in the equation (3) is not rotation-invariant as the intensity value of \( g_p \) changes when the neighborhood circle is rotated by a specific angle. When the image is rotated, the pixel values will correspondingly move along the perimeter of the circle around. In order to remove the effect of rotation, Rotation Invariant Local Binary Pattern method was introduced and was defined as:

\[
\text{ROR}(x,i) = \text{ROR}(LBP_{P,R}^i) = \min \{x\}
\]

where \( \text{ROR}(x,i) \) performs a circular bit-wise right shift on \( x \) \( i \) times.

Two patterns should be treated as “Uniform”, if one can be obtained from the other through rotating by a certain angle. The extension of Local binary pattern method in which uniformity measure, \( U \), is defined as the number of spatial transitions between 1s and 0s in the pattern is called Uniform Local Binary Pattern method. Patterns that have uniformity values of at most 2 are designated as uniform patterns. The extended LBP which deals with uniform pattern is defined as per the equation (5).

\[
\text{LBP}_{P,R}^U = \begin{cases} 
1 & \text{if } \sum (g_p - g_c) s'(LBP_{P,R}) \leq 2 \\
0 & \text{otherwise}
\end{cases}
\]

\( \text{LBP}_{P,R}^U \) can have 36 unique rotation invariant binary patterns in the circularly symmetric neighbor set, which are shown in the Fig. 3. It was observed that more than 90 percent of all 3x3 neighborhood pixels present in any texture images are falling within this uniform patterns.

As Uniform Local Binary Pattern was not capable of effectively retrieving the textural information by merely considering the histogram of the uniform patterns, a new concept called DLBP was introduced. The dominant Local binary pattern (DLBP) concept, by S. Liao et al. [14], considers the most frequently occurred patterns to capture descriptive textural information. DLBP does not contain any information about the dominant pattern types but the occurrence frequencies only. Recently the DLBP method has been successfully used for detecting the bleeding regions in human digestive tract [15] also.

In the conventional Uniform Local Binary Pattern method, sometimes the availability of noise converts useful patterns into non-uniform patterns so that they are not considered as uniform patterns. This problem is avoided in DLBP method [14]. Though DLBP was very good in encoding the pixel-wise information in the texture images, it does not consider the pixel interaction that takes place outside the coverage of its circular neighborhood system, which plays an important role in feature extraction in texture models.

**Figure 4**: Block diagram for segmentation.
III. ALGORITHM FOR SEGMENTATION

By following supervised segmentation, texture sample distributions are obtained by scanning the texture samples with the corresponding texture descriptors and corresponding histograms are calculated. The Fig. 4. shows the Block diagram for the following algorithm which is used for segmentation.

The segmentation algorithm is described as follows:

1. A random sample sub-image with the size of 30*30 pixels from each texture image (one sample per texture) is retrieved.
2. Texture Model is calculated for all sample texture images.
3. The input image is scanned by a window of 30*30 pixels and again the Texture Model is calculated for each window.
4. Texture model for every window of size 30*30 of the input image, is compared with the texture model of the each sample and the absolute difference (D) between them is calculated.
5. The central pixel of the window considered will be assigned to class K such that D(K) is minimum among all the D(i), for i = 1,2,3,4,......,where i represents the sample texture image.

IV. EXPERIMENTAL RESULTS

For this supervised study, five different texture images are used and samples are taken from the image with 25*25 pixels in size. Two synthetic images, one with four different textures and another with two different textures, are used as input images.

Experiment Setup #1

The above segmentation algorithm is applied to the four texture synthetic image by using Texture Spectrum as the feature descriptor with varying scanning window size. The segmented results are shown in the Fig. 5., where four different textures are represented by four different grey levels.

In the second approach, the same four texture synthetic image is given as the input for the same segmentation algorithm by using Uniform Local Binary Pattern method as the texture feature descriptor with varying scanning window size and the results are shown in Fig. 6.

Experiment Setup #2

In these experiments, a two texture synthetic image is given as the input image for the segmentation algorithm. First Texture Spectrum method is used as the texture descriptor and the results are shown in the Fig. 7., where two different textures are represented by two different grey levels. Fig. 8. shows the segmentation results, when the Uniform Local Binary Pattern is the feature extraction method.
A Comparative Study of Texture Features for Image Segmentation

Fig. 5. (Texture Spectrum).
(a) Original Input Image
(b) Window Size is 25*25
(c) Window Size is 20*20
(d) Window Size is 15*15
(e) Window Size is 10*10

Fig. 7. (Texture Spectrum, LBP Pattern).
(a) Original Image
(b) Window Size is 25*25
(c) Window Size is 20*20
(d) Window Size is 15*15
(e) Window Size is 10*10

Fig. 8. (Uniform Local Binary Pattern).
(a) Original Image
(b) Window Size is 25*25
(c) Window Size is 20*20
(d) Window Size is 15*15
(e) Window Size is 10*10
By this study, it is noted that both Texture Spectrum and Uniform Local Binary Pattern methods were simple for implementation because only few mathematical operations were needed. These methods facilitate a very straightforward and efficient implementation, which may be necessary in time critical applications. Table I shows the segmentation accuracy for the varying window size with the four-texture image and Table II shows the segmentation accuracy for the varying window size with the two-texture image. Here, segmentation accuracy rates are calculated over all the pixels including the region near the boundaries of textures. If we remove these pixels from the counter, the segmentation accuracy will be higher.

Both Texture Spectrum and uniform Local Binary Pattern methods have been evaluated from the point of view of discriminating performance that includes the influence of the boundaries of different textures. When Texture Spectrum method was used, promising segmentation results have been obtained with the average segmentation rate of 92%, whereas Uniform Local Binary Pattern gives an average segmentation rate of 86%. The bar chart (Fig. 9.) shows the comparative result analysis between Texture Spectrum and Uniform Local Binary Pattern methods on the basis of segmentation error rate for the given four-texture image.

Further evaluation shows that Texture Spectrum is sensitive to the directional aspect of texture whereas uniform Local Binary Pattern is sensitive to uniformity of texture in nature. In the case of Uniform Local Binary Pattern, it can be noted that the top left texture (in the four-texture input image) was not segmented properly because of its non-uniformity in the textural aspect. Greater the uniformity in the texture, more quality we can get in the output, which was proved in the case of the bottom left texture. As far as Uniform Local Binary Pattern method is concerned, it considers only 9 patterns as “Uniform” and it considers only these 9 patterns for segmentation. This is the reason for not producing smooth boundaries in the process of texture segmentation. The comparative result analysis between
A Comparative Study of Texture Features for Image Segmentation

Texture Spectrum and Uniform Local Binary Pattern method, in segmenting the given two-texture image on the basis of segmentation error rate is shown in the Fig. 10.

![Figure 10](image)

Figure 10: Comparative result on the basis of segmentation error rate with varying window size for the two-texture image.

According to our results, it is noted that, Uniform Local Binary Pattern gives higher segmentation accuracy when the window size becomes high, irrelevant of number of textures available in a single image. Uniform Local Binary Pattern method is very robust in segmenting the images which are affected even with of grey scale variation due to poor lighting.

Though Texture Spectrum method has given better performance, it is still possible to change the Uniform Local Binary Pattern method even powerful, by joining with some filters or contrast measures. Since both methods use the texture features from the neighborhood window, they cannot produce the smooth boundaries. With smaller window size, the texture features are not extracted completely from the neighborhood and at the same time, the larger value of window size will result in inaccurate segmentation near the boundaries, especially when the input image contains more number of textures. Moreover, when the window size is increased it increases the computation time also. So, the window size must be chosen in such a way that it covers the whole smallest unit of a texture unit or pattern. As a future enhancement, this study can be improved by applying various similarity measures, various segmentation algorithms and of course various texture features can also be included.

REFERENCES


Energy Efficient by Removing Redundant and forming Dynamic Clustering of Nodes in Wireless Sensor Network

S. Manju Priya ¹, S. Karthikeyan ²

ABSTRACT

In the development of various large scale sensor networks, a challenging problem is how to increase the lifetime of the sensors and to save the energy. In this paper, we solve this problem by using EERRDC protocol (Energy Efficient by Removing Redundant and using Dynamic Clustering), first by removing the redundant node and then clustering the sensor nodes. Here we compare the energy of the random nodes and the clustered nodes. The simulation results shows that the clustered nodes after turning off the redundant nodes takes only less amount of energy when compared to random nodes.

Keywords - clustering, energy, network lifetime, redundant nodes, wireless sensor nodes.

1. INTRODUCTION

A wireless sensor network consists of low power, thousands and hundreds of small size sensor nodes. These sensor nodes are battery powered. The sensor nodes collect the data and transmit it to a base station, referred as sink node. The base station collects data from all the nodes and analyzes this data to draw conclusion about the activity in the area of interest [1]. These wireless sensor networks has innumerable applications including weather monitoring, security, military surveillance, environmental monitoring, forest fire detection, healthcare monitoring and so on.

Many studies on WSN have been carried out [2] [3] [4]. Incase of a network with high density of sensor nodes, some problems may arise such as intersection of sensing area, redundant data, communication interference, and energy waste. The main key issue in the wireless sensor network is the energy. Since the nodes are battery powered, main researchers are still going on to improve the lifetime of the sensor. This occurs when there is energy difference to some threshold between an individual sensor and its neighbors, either due to the introduction of new sensor or re-energization of sensor nodes, or by any changes in the network settings which may be needed for any applications. An inefficient use of the available energy will lead to poor performance and short lifetime of the network.

2. RELATED WORK

Wireless sensor networks have attracted much research in recent years. In order to minimize the energy consumption in WSN’s several energy efficient routing protocols and routing algorithms has been developed [5] [6]. A node scheduling scheme was developed and described by D.Tian and N.D. Georganaa [7]. In their approach, nodes take turn in saving the energy without affecting the service provided. The node scheduling scheme turns some nodes on or off but still some redundancy occurs. Sleep/wake scheduling has been proposed to reduce energy consumption in sensor networks [8][9]. The basic idea is to put the radio to sleep.
state during idle and wake it up before transmission. But it requires fine-grained synchronization between the sender and the receiver, so that they can wake up at the same time to communicate with each other.

The authors B. Chen, K. Jameson, H. Balakrishnan, R. Morris in article [10], proposes an algorithm to turn off the nodes based on the necessity for neighbor’s connectivity. It intends to reduce the system energy consumption without significantly diminishing the connectivity of the network. The research article [11], proposes a schema in which energy is conserved by turning off their nodes when they are not involved in sending, forwarding or receiving any data.

Lien et al [12] focused on increasing the total data capacity by only considering the energy spent on the data transmissions. A clustering protocol is Power Efficient Gathering in Sensor Information System [PEGASIS], in which each node can only talk to it neighbors and can talk directly to the base station. This protocol assumes that each node has the capability to link directly to the base station, which is not possible in most cases. PEGASIS also expects that every node has a database to store the data about the location of it neighboring nodes which leads to greater demands for memory and power. In the research article [14], the authors proposes a technique to distribute the role of cluster head among some of the wireless sensor nodes based on the weight value containing the remaining energy of the node.

3. PROPOSED WORK

The proposed work in this research paper is divided into two sections. In the first section, the redundant nodes are turned off. Later, the nodes are clustered.

A. Removing Redundant (RR)

In this section, we will present the system model, define a schema for node scheduling in our work. The figure 1 shows the system model that we use. The data from the nodes are sent to the base station. Some problems may arise if the network has a high density. Generally, the node scheduling problem is of two sub problems. First in what basis that each node should follow to determine whether it should turn off or on? Second, when the node should be turned off.

Figure 1 : Sensor Network

A large number of sensor nodes in a sensor area make it infeasible to collect redundant detailed state of information from each individual sensor node, given energy and communication constraints. Reducing redundant sensing information is an important task in a wireless sensor network which helps to save energy and can increase the lifetime of the nodes. Consider the figure 2. Suppose if both nodes N1 and N2 is sensing the same area and collecting the same information (redundant data arises), then any of the nodes can be turned off.

Figure 2 : Example of redundant node
The RR-SPIN protocol is used to identify the redundant node. The idea behind this protocol is, before transmitting the data, nodes negotiate with each other to overcome implosion and overlap. The useful information alone can be transferred. Among the nodes in the network one of the node act as a META node which collects the observed data from the entire region.

The RR-SPIN protocol works on two different messages namely ADV and REQ. During any process, the sink can send queries to certain regions and can wait for the data from the sensors located in the selected region. Before transmitting, the metadata send an ADV message to all the nodes about the information of its data. So when any of the nodes doesn’t have the data, it can retrieve it by sending an REQ message to the metadata. If it doesn’t give any REQ message, it shows that the corresponding node has collected the same information as metadata. This confirms that it has done a redundant collection and that node can be temporarily turned off.

Consider the figure 3, let N5 act as a metadata. When N5 advertises to all other nodes, if any of the nodes has the same data as N5, then that node acts as a subset of N5. Suppose if N4 has similar data as N5, then N4 acts as a subset of N5 and N4 can be turned off, so that collecting redundant information is avoided. There is more chance for a node to act as a redundant node when both the nodes are placed in the same range. By turning off the nodes temporarily, the energy used by those nodes are saved.

After turning off the redundant node, the clustering process starts. All the nodes in the network act as a sensor nodes collecting information from the environment, apart from that they can act as a cluster head, forwarding the information to the sink, interconnecting different clusters. The entire network area can be divided into various regions or clusters with each cluster having a cluster head (CH) and all other nodes act as its members. The nodes within the clusters can communicate through the CH to the sink in an effective way.

1) Selection of CH:

The desired characteristics to identify a cluster head are

- The nodes are randomly divided into clusters.
- In each clusters, which node has the highest energy can act as the cluster head (CH).
- If the current CH is loosing out too much of energy, then the next node which has the higher energy, can take up the role of CH, thereby providing reliable transmissions.

After the clusters are formed, the network starts its normal operation. Like LEACH (Low-energy adaptive clustering hierarchy), in order to reduce the probability of collision among joint REQ messages during the setup phase, CSMA (Carrier Sense Multiple Access) is utilized as the MAC layer protocol [15].
In this section we propose the model, in which, the clusters will be formed after removing the redundant node. Initially when the clusters are to be formed, each node checks its residual energy. Since the clusters are formed randomly, there can be a chance of more nodes in one cluster and less nodes in another cluster. So to avoid this, each CH from the clusters sends a message to other CH’s stating how much nodes are there in their region and about their energy. If any of the nodes has more energy than any of the CH, then that node is moved into that cluster group and elected as CH and the current CH becomes as a member.

The cluster table (CT) is formed at each cluster head, which have all the details of each cluster members in the corresponding cluster. During transmission of the data, to reduce the energy, the nodes which are not involved in sending any data can be turned off. In the CT, it has to maintain which nodes are active and which are sleep. After each round of sending the data from the source to the destination, each CH checks its energy level whether it has fallen below the threshold value. If so, it changes its role and hands its control to the next node which has highest energy. This change has to be updated in the CT and message it to other CH also.

4. SIMULATION RESULTS

To validate the performance of this, we simulated with a network of about 100 nodes. In the first part, to remove the redundant nodes, the MATLAB is used, which is a scalable simulation library for wireless network systems while in the second part for dynamic clustering of nodes the NS2 is used.

Initially sensor nodes around 100 are uniformly distributed over a terrain range of (800,800) and with the routing protocol as RR-SPIN, the simulation time is setup to 300s. The radio transmission range is 50m. The following table shows the performance of the energy consumption after using the RR-SPIN protocol.

The initial energy consumption for 100 nodes is about 322.2 Joules. By using the proposed system after turning off 4 nodes the energy is only 282.9 J. We can see that the proposed system reduces the nodes as well as the energy is also saved which is shown in the Table 1.

Table 1: Energy consumption for whole sensor n/w

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>Energy consumed (joules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>322.2</td>
</tr>
<tr>
<td>96</td>
<td>282.9</td>
</tr>
<tr>
<td>88</td>
<td>166.5</td>
</tr>
<tr>
<td>80</td>
<td>142.3</td>
</tr>
<tr>
<td>76</td>
<td>102.6</td>
</tr>
</tbody>
</table>

The following figure 4 clearly shows that after turning off the redundant node the energy consumed is reduced.

Figure 4: Energy Consumption

In the next stage, the nodes are clustered using DC technique. The nodes are clustered into five groups with one CH in each cluster. (as shown in figure 5).

Figure 5: Nodes are clustered
The data has been send from 14th node to the 38th node (as shown in Fig 6).

Figure 6: Data transmission from source to destination node

The results of the proposed system are compared with randomly distributed nodes. The proposed system proves that it uses only less amount of energy while sending the data. The results of the EERRDC protocol is also shown in figure 7.

Figure 7 shows that the total remaining energy of the clustered nodes is more than the random nodes. It shows that EERRDC utilizes only less amount of energy, so that the lifetimes of the nodes are also increased.

V. Conclusion

Here we introduced a novel way to utilize the nodes after turning off the redundant nodes, so that the energy has been saved. Other technique along with this is, using dynamic clustering of the nodes, which has also saved the energy. The energy efficiency and ease of deployment makes EERRDC a desirable protocol for wireless sensor networks. Both the simulation results show that the energy has been saved, so that the lifetime of the nodes is also increased. Our future work will be further investigating in clustering techniques.

References


**Author’s Biography**

Manju Priya S completed M.Sc, M.Phil in Computer Science from Bharathiar University, Coimbatore. Currently she is pursing Ph.D in Computer Science in Karpagam University, Coimbatore. She has presented more than 5 papers in the national and international conferences. Her research area includes Wireless sensor networks, network communications.

Karthikeyan S. received the Ph.D Degree in Computer Science and Engineering from Alagappa University, Karaikudi in 2008. He is working as a Professor and Director in School of Computer Science and Applications, Karpagam University, Coimbatore. At present he is in deputation and working as Assistant Professor in Information Technology, College of Applied Sciences, Sohar, Sultanate of Oman. He has published more than 14 papers in National/International Journals. His research interests include Cryptography and Network Security.
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>PAGE NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 70 MHz CMOS Gm-C Bandpass Filter</td>
<td>221</td>
</tr>
<tr>
<td>P Sampath, C M Preethi, K Gunavathi</td>
<td></td>
</tr>
<tr>
<td>2. Shunt Active Filter Based Voltage Regulation and Harmonic Damping</td>
<td>226</td>
</tr>
<tr>
<td>K. Suresh Kumar, P. Vijaya kumar, L. Ganesh</td>
<td></td>
</tr>
<tr>
<td>3. A Classification of Character Usage in Unique Addresses Employed for Accessing Yahoo! Groups Service</td>
<td>233</td>
</tr>
<tr>
<td>Jatinderkumar R. Saini, Apurva A. Desai</td>
<td></td>
</tr>
<tr>
<td>4. Effective Utilization of Processor in a Distributed Environment For PC Clusters Analysis</td>
<td>241</td>
</tr>
<tr>
<td>N. Kavitha, S. Karthikeyan</td>
<td></td>
</tr>
<tr>
<td>5. A Comparative Study of Texture Features for Image Segmentation</td>
<td>248</td>
</tr>
<tr>
<td>G. Madasamy Raja, V. Sadasivam</td>
<td></td>
</tr>
<tr>
<td>S. Manju Priya, S. Karthikeyan</td>
<td></td>
</tr>
</tbody>
</table>