AVOID SYNCHRONIZATION LATENCY USING VLSI IMPLEMENTATION

R.Keerthiga¹, T. M. Senthil Ganesan²

¹ PG Scholar, Department of VLSI Design, ² Assistant Professor
Theni Kammavar Sangam College of Technology, Tamilnadu, (India)

ABSTRACT
The phenomenon of metastability is inherent in clocked digital logic. Many techniques have been presented for minimizing metastability, both for crossing clock domains, and for handling asynchronous inputs. Flip-flops are among these systems and can take an unbounded amount of time to decide which logic state to settle to once they become metastable. This problematic behavior is often prevented by placing the setup and hold time conditions on the flip-flop’s input. However, in applications they induce catastrophic failures. These events are fundamentally impossible to prevent but their probability can be significantly reduced by employing synchronizer circuits. The latter grant flip-flop longer decision time at the expense of introducing latency in processing the synchronized input. The main contributions include two novel solutions for the problem of synchronization. Speculation technique is the main theme that helps reducing synchronization latency.

Keywords: Metastability, Setup Time, Hold Time, Synchronization, Latency

I. INTRODUCTION
Metastability is a phenomenon that can cause system failure in digital devices, including FPGAs, when a signal is transferred between circuitry in unrelated or asynchronous clock domains. This paper describes metastability in FPGAs, explains why the phenomenon occurs, and discusses how it can cause design failures. The calculated mean time between failures (MTBF) due to metastability indicates whether designers should take steps to reduce the chance of such failures. This paper explains how System reliability can be improved by reducing the chance of metastability failures with design techniques and optimizations.

II. METASTABILITY
The purpose of enforcing the setup and hold time conditions on combinational paths is to constrain the input of every flip-flop: to ensure that it is held stable for at least tsu seconds before the clock edge and that it remains stable for no less than th seconds afterwards. By doing so, flip-flop outputs are guaranteed to behave in a predetermined manner: they transition to the logic level of the input monotonically, with a nominal transition time and within a nominal clock-to-q delay. These properties are essential for the design of deterministic synchronous systems.
In some applications, however, the setup and hold times of a flip-flop’s input cannot be always satisfied. For example, when the flip-flop is used to sample a real-time signal, input transitions can occur at any time relative to the clock edge. For a clock edge occurring at tclk, if a transition occurs after tclk-tsu and before tclk + th (this interval is referred to as the setup-hold time window), the flip-flop may not behave in the predetermined manner.
described above. In other words, it may transition or not transition at all, it may transition after a long delay with a longer rise/fall time or it may produce multiple output transitions (behave non-monotonically).

Historically, flip-flops were not known to behave in this manner in the early days following their invention. It was believed that a flip-flop whose setup and hold time conditions were violated will either succeed or fail to capture the logic value of the input. The impact of these violations on the delay, transition time and monotonicity of the flip-flop output had not been foreseen. In consequence, multiple early synchronous computers which have included unconstrained flip-flops exhibited mysterious failures whose root cause was not identified until the first mathematical analysis of the problem was published in 1952. The anomalous behavior of unconstrained flip-flops was attributed to metastability: a pseudo-stable state in which a bistable element is neither logic high nor low but somewhere in between.

The duration of the metastable condition is a probabilistic phenomenon, and therefore there is no guaranteed maximum time. One can't build a bistable device such as a flip-flop that cannot go metastable. Metastability can appear as a flip-flop that switches late or doesn’t switch at all. It can present a brief pulse at a flip-flop output (called a runt pulse) or cause flip-flop output oscillations. Any of these conditions can cause system failures.

For a simple CMOS latch, valid data must be present on the input for a specified period of time before the clock signal arrives (setup time) and must remain valid for a specified period of time after the clock transition (hold time) to assure that the output functions predictably. This leaves a small window of time with respect to the clock (t<sub>0</sub>) during which the data is not allowed to change. If a data edge occurs within this aperture, the output may go to an intermediate level and remain there for an indefinite amount of time before resolving itself either high or low.

This metastable event can cause a failure only if the output has not resolved itself by the time that it must be valid for use (for example, as an input to another stage); therefore, the amount of resolve time allowed a device plays a large role in calculating its failure rate. Whenever there is any such violation, the output voltage is anywhere between a logic high and a logic zero. In such a condition the flip-flop takes additional time to settle to a stable output. And this stable output depends upon the process technology, manufacture and environment conditions that may force the output to go to a particular value. There is no way that the final state can be predicted.

This operation is analogous to a ball rolling over a hill. Each side of the hill represents a stable state, and the top of the hill represents the metastable state. Just as the slightest air current would eventually cause a ball on the
illustrated hill to roll down one side or the other, thermal and induced noise will jostle the state of the flip-flop causing it to move from the quasi-stable state into either the logic 0 or logic 1 state. Hence, the output is random.

In any case the CP-Q delay of the f/f is increased. The extra delay may be ten or twenty times longer than the normal Clk-Q delay. This extra time is called the metastable resolution time. However, metastability may not always result in unpredictable output. If provided sufficient time with proper excitation, the frequency to frequency can in fact settle to a stable state.

Metastable events are associated with data transitions occurring close to the active edge of the clock. Because a timing violation has occurred, the flip-flop will exhibit erratic behavior. The erratic behavior manifests itself in the form of an extended propagation delay with an unpredictable resolution of the Q output (of f/f). This metastable event can cause a failure only if the output has not resolved itself by the time that it must be valid for use therefore, the amount of resolve time allowed for a device plays a large role in calculating its failure rate.

2.1 Scenarios of Metastability Occurrence
Whenever setup and hold violation time occurs, metastability occurs, so it is to be seen when this signal violates this timing requirement.

- When the input signal is an asynchronous signal.
- When the clock skew is more (rise time and fall time is more than the tolerable Values).
- When interfacing two domains operating at two different frequency.
- When the combinational delay is such way that, it changes flip-flop’s input in the required window (setup + hold window)

2.2 Metastability Measurement
In order to define the metastability characteristics of a device, 3 things must be known

- Possible sequence that the device will enter a metastable state
- Period of device remain in that state
- Measured propagation delay of the device.

2.3 Methods to Avoid Metastability

- Synchronize any asynchronous input through one path that has at least one and preferably two flip-flops in series.
- The flip-flops should be running on the same edge of your system clock as the rest of the circuit. This will limit the area of potential problems to one path instead of several.
- Use buffered flip-flops, or un-buffered flip-flops with minimum load.
- Ensure that setup time of the destination flip-flop is met. This will avoid the creation of metastable conditions inside the circuit and minimize the propagation of any should they occur
- Use metastability hardened Flip-flops.

III. METHODS TO AVOID SYNCHRONIZATION LATENCY

3.1 SPECULATION
An alternative strategy to mitigate synchronization latency is to use redundant hardware to perform speculative computations during synchronization cycles. This “hides” synchronization latency by overlapping it with an
equivalent number of computation cycles. If computing an output based on an asynchronous input requires $n$ synchronization cycles and $m$ computation cycles, this method yields a processing time of $\max(m, n)$ cycles as opposed to $m + n$ for conventional synchronization. This reduces the total latency to $t_b + T \max(m, n)$ where $t_b$ is the bundling delay. Metastable states occur relatively rarely compared to handshake requests and that incurring two cycles to synchronize each individual handshake is thus unwarranted. Their scheme involves using a single flip-flop $k$ as a synchronizer and speculating that it does not become metastable. A detector circuit can then reliably identify, $n$ cycles later, whether $k$ has actually become metastable. If this was the case, each register in the synchronous block is restored to a backup copy which is kept in an $n$-level stack. Using this form of speculation, the latency of processing the asynchronous request is reduced to a single cycle only (plus $t_b$). The cost is that each register needs to be duplicated $n$ times.

### 3.1.1 Advantages

- Approach is entirely architectural and does not target the synchronization process itself.
- It does not rely on any assumptions about the relationship between the communicating clocks.
- It does not require fast metastability-resolving flip-flops.
- Trading reliability and low-latency with duplicated hardware will be an increasingly-affordable option in future technologies because of the continuous growth of available design area.

### 3.2 Datapath Unfolding

Speculation is the use of either time or resource redundancy to perform potentially useful work. Modern digital systems employ speculation at different abstraction levels. For example, memory management speculatively populates cache hierarchies with pre-fetched data to reduce the impact of slow memory access on processing speed. Also, processors that use branch prediction execute the instructions following branches speculatively to increase throughput. This is because restoring the state of a pipeline in the case of un-speculation is trivial. For example, when a branch condition in a pipelined processor is evaluated, invalid instructions in the fetch and decode stages can be discarded by flushing these stages. On the other hand, speculative computations cannot be “reversed” in a similarly straight-forward manner in non-pipelined systems. This is because non-pipelined systems have loop dependencies (i.e. feedback paths) such as the one represented by the expression $x + 1$. The existence of loop dependencies can corrupt the system state in the case of un-speculation (pipelined systems are free from such dependencies by definition). Nevertheless, arbitrary designs can be converted into functionally equivalent pipelines by unfolding. Although pipelining a design by unfolding is used primarily to increase throughput, it can also be used to perform speculative computations during synchronization cycles. To demonstrate how, consider the generic synchronous module. The module is represented by a Moore machine consisting of the state register $R$, the combinational block $C$ and the asynchronous port $[req, d, ack]$. To maintain reliability, two flip-flops are added to synchronize req. The latency introduced by this chain can be “hidden” by speculatively computing what the machine state would have been if req changed two cycles earlier.

### IV. PROPOSED METHOD

#### 4.1 Sequenced Latching

A novel technique to latch data reliably during synchronization cycles is the technique of Sequenced Latching. In short, a synchronizer is used as a state machine to sequence a series of latching operations. The synchronizer is constrained such that its state does not change when a latching operation fails. Therefore, any failed latching attempts are automatically retried in the subsequent cycles.
V. COMPARISON OF SPECULATIVE TECHNIQUES

In Speculative Synchronization, method uses a single flip-flop k to synchronize the asynchronous handshake and a detector to reliably identify. When a metastable state is identified, the machine is restored to a previous correct state. This approach can be summarized as “assume, execute, verify then correct if necessary”.

In Data path unfolding, additional instances of the entire machine are used to speculatively compute the machine states following the arrival of data. The assumption used here is that the value of the asynchronous data bus was valid n cycles earlier. This approach can be summarized as “assume, verify then execute”.

In Sequenced Latching, unlike the other two speculative methods, sequenced latching makes an individual assumption on each synchronization cycle. The assumption is that the transition of the synchronizer flip-flop Si is captured by its successor Si+1. The delays between the synchronizer flip-flops and the sequenced pipeline stages are constrained such that data moves through the pipeline safely when this assumption holds.

![Comparison Table]

VI. CONCLUSION

The growing number of asynchronously-clock core in modern systems means that the negative performance impact of clock domain crossing latency is likely to increase. Two novel architectural solutions (data path unfolding and sequenced latching) that are free from these limitation are proposed. The proposed methods leverage hardware duplication to speculatively compute the first few system states following a change in the asynchronous input. This allows a system to hide synchronization latency by overlapping it with the computation of the first few data-dependent states. The simulation results shows that metastability occurrence can be detected and based on the count the occurrence can be reduced which helps in reducing synchronization latency. The future work can be extended to work on the clock domain interface performing adaptive synchronization.

REFERENCES


**BIOGRAPHICAL NOTES**

MS.R.KEERTHIGA is presently pursuing M.E final year in Electronics and Communication Engineering Department (specialization in VLSI design) from Theni Kammavar Sangam College of Technology, Theni, Tamilnadu, India.

Mr.T.M.SENTHIL GANESAN is working as a Assistant Professor in Electronics and Communication Engineering Department, from Theni Kammavar Sangam College of Technology, Theni, Tamilnadu, India.
A REVIEW ON TECHNOLOGIES IN BUSINESS ANALYTICS

Pankaj Roy¹, Dr. Ruchi Agarwal²

¹ Dept. of Computer Science, School of Engineering and Technology, Sharda University, (India)
² Asst. Professor, Dept. of Computer Science, School of Engineering and Technology, Sharda University, (India)

ABSTRACT

Analytics are becoming more and more part of the decision making process of management and operational work. Business analytics is the combination of skills, technologies, applications and processes used by organizations to gain insight in to their business based on data and statistics to drive business planning. The definition of Analytics in simple can be defined as a set of practices, skills and technologies designed for investigating and analyzing business performance in order to achieve more strategic decision making and structuring in the future. Business analytics is used to evaluate organization-wide operations, and can be implemented in any department from sales to product development to customer service. Business analytics solutions typically use data, statistical and quantitative analysis and fact-based data to measure past performance to guide an organization's business planning. Big data analytics means the process of collecting, organizing and analyzing Big Data (large sets of data) to discover patterns and other useful information. Big data analytics help you to understand the information contained within the data, but it will also help identify the data that is most important to the business and future business decisions. Big data analysts basically want the knowledge that comes from analyzing the data. Recently, Big Data has attracted a lot of attention from academia, industry as well as government. Big Data is term defining collection of large and complex data sets that are difficult to process using conventional data processing tools. We require new tools and techniques to organize, manage, store, process and analyze Big Data. This paper reviews the trends, opportunities and underline technologies in business analytics & various research issues related to Big Data analytics.

Keywords: Analytics, Business Analytics, Big Data, Big Data Analytics, Descriptive, Predictive.

1. INTRODUCTION

Big Data Analytics has changed the way organizations make decisions, manage business processes, and create new products and services. Business analytics is the use of data, information technology, statistical analysis, and quantitative methods and models to support organizational decision making and problem solving. The main categories of business analytics are descriptive analytics, predictive analytics, and prescriptive analytics. Big Data is data that exceeds the processing capacity of conventional database systems and is typically defined by three dimensions known as the Three V's: Volume, Variety, and Velocity. Big Data brings big challenges. Big Data not only has influenced the analytics that are utilized but also has affected technologies and the people who use them. At the same time Big Data brings challenges, it presents opportunities. Those who embrace Big Data and effective Big Data Analytics as a business imperative can gain competitive advantage. [1]
Business Analytics is about using data and models to make better decisions, to improve policy making, to develop strategy and to improve day to day performance of organizations. In other words, it’s the application of management science. Business analytics focuses on developing new insights and understanding of business performance based on data and quantitative methods. It can answer questions like why is something happening, what if these trends continue, what will happen next, what is the best that can happen? Organizations cannot afford to base their decisions on hunches and intuition – increasingly they use sophisticated data analysis techniques to improve operational efficiency, spot emerging trends and understand what types of products and services customers will value. Business Analytics has come to stay as a growth driver for many new generation organizations. Gone are those days when managers will make decisions on the basis of their own guts or by extrapolating macro-economic indicators and their likely impact on individual businesses. Decisions made in the absence of information and data have proven to be disastrous for many organizations. With the advent of information technology and geometric rise in the information processing capability of computers, managers are using multiple criteria, algorithms and structured programs to envision future of business and improve profitability of the enterprise. Higher is the accuracy level in decision making, lower is the risk involved in business. This proposition leads to an idea that business analytics concepts and tools can be used for enhancing customer intimacy and organizational profitability. Application of descriptive and predictive analytics, customer relationship management tools and other process enhancement tools not only brings success but also increases the profitability of the organization. The whole business world is looking at Big Data as an opportunity and source of competitive advantage. The Three Categories of Business Analytics:

- Business analytics can be subdivided into of three fairly MECE (mutually exclusive and collectively exhaustive) categories: descriptive, predictive, and prescriptive analytics.
- Descriptive analytics deals with describing actual or known data.
- Predictive analytics deals with predicting the unknown.
- Prescriptive analytics deals with prescribing the best course of action, whether optimal or near-optimal. [3]

- What is going on? (Descriptive Analytics)
- What will happen? (Predictive Analytics)
- What should we do? (Prescriptive Analytics)

Figure 1: Three Types of Analytics Layers

The base layer is descriptive analytics, the middle layer is predictive analytics, and the apex is prescriptive analytics. The reason is that descriptive analytics often provides the basis for predictive analytics, and both
II. PROBLEM DEFINITION

Analytics can help make sense of big data, by helping companies identify the types of data and data sets that they should be examining to address specific business challenges. Still, there are significant challenges companies must overcome in order to exploit big data. Here are few of them:

1. **A comprehensive approach to using big data.** Most companies collect gobs of data but they don’t have comprehensive approaches for centralizing the information.

2. **Getting the right information into the hands of decision makers.** Companies should use analytics “to avoid getting buried under the humongous amount of information they generate through various outlets,”

3. **Effective ways of turning “big data” into “big insights.”** No matter how you slice it, data is just that – data. In and of itself, data doesn’t necessarily provide decision makers with the kind of insights they need to do their jobs effectively

4. **Big data skills are in short supply.** There’s already a shortage of data scientists in the market. This includes a scarcity of people who know how to work well with large volumes of data and big data sets.

III. DATA ANALYTICS

Data analytics (DA) is the science of examining raw data with the purpose of drawing conclusions about that information. Data analytics is used in many industries to allow companies and organization to make better business decisions and in the sciences to verify or disprove existing models or theories. Data analytics is distinguished from data mining by the scope, purpose and focus of the analysis. Data miners sort through huge data sets using sophisticated software to identify undiscovered patterns and establish hidden relationships. The science is generally divided into exploratory data analysis (EDA), where new features in the data are discovered, and confirmatory data analysis (CDA), where existing hypotheses are proven true or false. Qualitative data analysis (QDA) is used in the social sciences to draw conclusions from non-numerical data like words, photographs or video. The term “analytics” has been used by many business intelligence (BI) software vendors as a buzzword to describe quite different functions. Data analytics is used to describe everything from online analytical processing (OLAP) to CRM analytics in call centers. Banks and credit cards companies, for
instance, analyze withdrawal and spending patterns to prevent fraud or identity theft. Ecommerce companies
examine Web site traffic or navigation patterns to determine which customers are more or less likely to buy a
product or service based upon prior purchases or viewing trends. Modern data analytics often use
information dashboards supported by real-time data streams. [4]

Nathaniel Lin define “analytics” (or “business analytics”) this way [2]

- More than just numbers- Analytics is more than just working with numbers and data to find and report
observed correlations and/or statistical distributions.

- Knowledge and results centric- Business analytics is focused on the process of discovery of actionable
knowledge and the creation of new business opportunities from such knowledge

- Tools-agnostic- Analytics can use any computation or visualization tools from statistics, computer science,
machine learning and operation research to recognize patterns, seek and validate relationships and trends.

IV. DESCRIPTIVE ANALYTICS

Descriptive analytics deals with organizing, manipulating, visualizing, and describing actual business data.
Actual data is data that is strictly deterministic or known. It is often data about past performance, but can be
current state data or data about a certain future state. Examples would include past sales, customer locations, or
warehouse capacities. In my opinion, descriptive analytics is synonymous with business intelligence. It also
includes dashboards, reports, and advanced visualizations which often display KPI’s (key performance
indicators), which are business specific metrics that allow managers to better gauge the company’s performance.

Descriptive analytics is the most widely used and understood of the three types, and often provides the basis for
the other types of analytics. Most consulting firms that provide analytic services focus primarily on Descriptive
Analytics. Descriptive analytics does provide significant insight into business performance and enables users to
better monitor and manage their business processes. Additionally, descriptive analytics often serves as a first
step in the successful application of predictive or prescriptive analytics. Organizations that effectively use
descriptive analytics typically have a single view of the past and can focus their attention on the present, rather
than on reconciling different views of the past. Characterized by the use of key performance indicators,
descriptive analytics drills down into data to uncover details such as the frequency of events, the cost of
operations and the root cause of failures. The most common type of analytics used by organizations, it typically
displays information within a report or dashboard view. Solutions can be automated to issue alerts when
potential problems arise that fit data patterns the solution has discovered.

V. PREDICTIVE ANALYTICS

Predictive analytics deals with predicting unknown data based on actual data and other knowledge. The
unknown data is often data that pertains to the future i.e. forecasting, but it can be current and even past data that
is not explicitly known, sometimes referred to as “now casting”. The data that results from predictive analytics
is stochastic or subject to uncertainty. Predictive analytics is generally performed using statistical methods such
as regressions and simulations, probabilistic models, and increasingly data mining or machine learning
techniques. Most consulting firms that offer advanced analytics as a service are referring to predictive analytics.
VI. PRESCRIPTIVE ANALYTICS

Prescriptive analytics deals with prescribing optimal or near-optimal business actions based on actual and/or predicted data. This is the elusive answer to the “so what?” that skeptical management often poses to their analytics teams. I’ve seen this included in predictive analytics, but I believe it is distinct. It is not about predicting what will happen; it is about deciding what should be done. Prescriptive analytics is primarily the realm of optimization and mathematical programming (MP) which includes linear programming (LP), integer programming (IP), non-linear programming (NLP), constraint programming (CP) etc. It also includes heuristic algorithms for sub-optimal solutions, and simulation based and stochastic optimization. Additionally, decision trees, what-if analyses, and scenario planning are less sophisticated forms of prescriptive analytics. Prescriptive analytics is probably the least used and least understood type of business analytics, and I believe it presents an incredible opportunity for businesses to gain a competitive edge.

Once the past is understood and predictions can be made about what might happen in the future, one need to know what the best action will be, given the limited resources of the enterprise. This is the area of prescriptive analytics. Prescriptive analytics use data to propose the best course of action to increase the chance of realizing the finest outcome. Optimization and simulation techniques are being used for this kind of analytics. Prescriptive analytics are based on the concept of optimization, which can be divided into two areas:

- Optimization: How can we achieve the best results?
- Stochastic optimization: How can we achieve the best result and tackle improbability in the data to make better decisions?
A perfect example of prescriptive analytics is the use of product recommendation engines such as Amazon and Netflix. The analytic results, associated content, and delivery of said content are all controlled through an automated workflow. Big Data combined with prescriptive Analytics, is how we would make the Big Data trend productive. [6]

VII. OPERATIONS RESEARCH & BUSINESS ANALYTICS

Both predictive and prescriptive analytics are components of Operations Research (OR), though prescriptive analytics is the primary component. OR got its name from its genesis in the military operations of World War II and is used quite extensively in the manufacturing, logistics, distribution, and transportation industries, but there are extensive opportunities to apply it to almost every other industry including hospitality, retail, healthcare, finance, insurance, education, public policy, business services, and utilities. It is also known as decision science, management science, and in the UK as operational research. Due to its name, many people associate operations research with the operations management and supply chain disciplines, but there are many potential applications of OR in almost every other business discipline including: marketing, human resources, strategic planning, revenue management, financial planning, and pricing.

VIII. BIG DATA, TEXT ANALYTICS & CLOUD COMPUTING

Big data refers to extremely large data sets that are often difficult to store and process. Due to continuously improving technology, advancements in data networks, increasing use of the internet and social media, and the proliferation of “smart” devices there is an incredible amount of data that can potentially yield useful insights. Big data is part of the reason that business analytics has become such a hot topic, because the more data there is out there the more potential uses there are. Statistical and data mining techniques are used to convert these vast amounts of data in sophisticated and useful models and visualizations. Many of the challenges associated with data are more the realm of IT than business analytics. Cloud computing is a decentralized approach to storing big data, so that individual computers and devices do not need to store vast amounts of data on internal drives. Instead, they can access the data in the cloud via the internet. The easiest data to use in analytics is quantitative and structured in nice neat tables, because that lends itself nicely to statistics and mathematics. However there is an incredible amount of valuable information can be gleaned from social media posts, customer reviews, and countless other sources of information that is text based and unstructured. Within the umbrella of business analytics, text analytics (also referred to as text mining) uses machine learning, statistical, and other computational, techniques to yield useful business information from vast amounts of unstructured text. The use of big data in business analytics, whether structured or unstructured, quantitative or text based, mostly falls within the descriptive and predictive categories. My belief is that with the increased insights gleaned from big data, there are increasing opportunities to apply prescriptive analytics, the elusive “so what?” of business analytics to help companies increase revenues and decrease costs. Businesses that do so effectively will have a significant competitive advantage, and those that miss this opportunity risk falling significantly behind competition.
IX. TECHNOLOGIES

Big data analytics, until recently the exclusive domain of only the largest and best bankrolled organizations, has become more widely accessible with the advent of Hadoop, the open source implementation of the Map Reduce massively parallel programming framework pioneered by Google and Yahoo. Cloud-based Hadoop offerings have extended the big data analytics opportunity further, to include organizations that don't want to build and maintain their own Hadoop cluster. Many organizations looking to collect, process and analyze big data have turned to a newer class of technologies that includes Hadoop and related tools such as YARN, Map Reduce, Spark, Hive and Pig as well as NoSQL databases. Those technologies form the core of an open source software framework that supports the processing of large and diverse data sets across clustered systems. In some cases, Hadoop clusters and NoSQL systems are being used as landing pads and staging areas for data before it gets loaded into a data warehouse for analysis, often in a summarized form that is more conducive to relational structures. Increasingly though, big data vendors are pushing the concept of a Hadoop data lake that serves as the central repository for an organization's incoming streams of raw data. In such architectures, subsets of the data can then be filtered for analysis in data warehouses and analytical databases, or it can be analyzed directly in Hadoop using batch query tools, stream processing software and SQL on Hadoop technologies that run interactive, ad hoc queries written in SQL.

Potential pitfalls that can trip up organizations on big data analytics initiatives include a lack of internal analytics skills and the high cost of hiring experienced analytics professionals. The amount of information that's typically involved, and its variety, can also cause data management headaches, including data quality and consistency issues. In addition, integrating Hadoop systems and data warehouses can be a challenge, although various vendors now offer software connectors between Hadoop and relational databases, as well as other data integration tools with big data capabilities.

9.1 Understanding the Differences between Three Key Analytics Types

The goal of any analytics solution is to provide the organization with actionable insights for smarter decisions and better business outcomes. Different types of analytics, however, provide different types of insights. So it is important for managers to understand what each analytics type delivers and to match analytics functions to the organization’s operational capabilities across its real estate, facilities and asset management functions. The three types build on one another, with descriptive analytics being the most common and prescriptive analytics the most advanced. Yet they share goals for improving real estate, facilities and asset operations with capabilities that help provide an understanding for an event or action, uncover relationships in data, develop what-if scenarios and simplify business decisions.

X. CONCLUSION

The industry clusters of analytics ecosystem show that there are many opportunities for students, academics and industry professionals. With increases in data availability and the realization that more business insights could be generated from an otherwise junk volume of data, research in Big Data has increased. This trend in Big Data analytics is expected to continue to grow and complement traditional data analytics methods and existing data management strategies. The increase in the amount of attention analytics is receiving from the press also indicates that this is an expanding field which can provide a competitive edge to many types of businesses. As illustrated in this paper, descriptive, predictive and prescriptive analytics applications and research opportunities
are wide and growing. Management and managers need advice on how to select analytics technology, on the measures most useful for them, and on best practices and common mistakes in choosing business and operational measures, metrics and key indicators. They also need more reliable information than is currently available about integrating historical and predictive analytics into business systems and processes so they can make better use of existing investments and plan new ones that provide deeper insight. And at last I conclude that with proper planning and execution, companies that deploy and use analytics effectively can achieve a competitive advantage.

REFERENCES

[1] Business Analytics and Big Data: Driving organizational change Dennis T. Kennedy (La Salle University, USA), Dennis M. Crossen (La Salle University, USA) and Kathryn A. Szabat (La Salle University, USA) http://www.igi global.com/ chapter/ business-analytics-and-big-data/122745# chapter-preview


REVIEW OF INTELLIGENT TRANSPORT SYSTEM ARCHITECTURE AND ITS APPLICABILITY IN TRAFFIC MONITORING AND CONTROL, USING SENSOR NETWORKS

Garima Mathur
Research Scholar, Department of ECE, School of Engineering & Technology, Poornima University Ramchandrapura, Jaipur, (India)

ABSTRACT
As traffic and population is increasing, its monitoring and automation is finding a scope for development and it is emerging as a lucrative field for research and development. Intelligent transportation systems (ITS) are advanced applications which aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated, and 'smarter' use of transport networks. Realizing its enormous potential in automation, this review work was built around ITS and its applications. As ITS majorly involves the Wireless Sensor Networks, it was also explored to some extent. Various architectures were explored and compared to deduce the strengths and weaknesses of proposed solution.

Keywords: ITS, WSN

I. INTRODUCTION
Intelligent Transport Systems (ITS), a combination of Information Technology and telecommunications, allows the provision of on-line information in all areas of public and private administration. ITS can be applied to road transport to improve efficiency and safety through the provision of on-line information to drivers in their vehicles and by equipping the vehicle with computerized systems which assist the driver (e.g. following and lane keeping). It also improves the efficiency of transport by use of electronic systems to improve traffic control and enforcement of traffic rules and regulations. ITS will have beneficial effects on the environment by reducing air and noise pollution on highways and by helping to create traffic free zones in cities. This paper is categorized in six sections. The first section being introduction, second section details the review process adopted for this literature survey, third and fourth section gives the brief about the various issues which came up during review and their common findings along with comparative tables with associated strengths, weaknesses and unique features of each. Fifth section concludes with the scope and sixth sections lists down the references used.

II. REVIEW PROCESS
The review process is divided into five stages in order to make the process simple and adaptable by every researcher. As it reflects from the literature that while beginning the finding of research objectives, it is necessary to start with a broader domain of any area / sub area of interest and narrow down to specific issue, the process described in the diagram includes the narrowing down. I have followed one of the typical processes to
make a literature review and frame the objectives of research. The process diagram is shown in Fig. 1, which includes all five stages defined as under:

![Fig 1. Literature Review Process](image)

**Stage 0: Get the Feel**
This stage is the beginning of literature review process wherein one has to broadly select his / her area of interest and start searching the scientific research papers from valid sources. I have gone through large number of research papers and selected 19 research papers published in various research journals using this stage.

**Stage 1: Get the “big picture”**: In order to understand the paper broadly and get an idea whether the paper exactly belong to the research area / sub area selected or it deviates, if deviates how much, these concepts are made clear this stage, known as Get Big Picture. It helps in classifying the research papers according to key issues and narrowing down the survey. Using this stage 19 research papers were clustered according to regions and the type analysis presented by the researchers.

**Stage 2: Get the “details”**: Stage 2 deals with going in depth of each research paper and understand the details of methodology used to justify the problem. Total 19 paper’s data has been extracted under this process pertaining to my research focus.

**Stage 3: “Evaluate the details”**: This stage evaluates the details in relation to significance of the problem, Novelty of the problem, significance of the solution, novelty in approach, validity of claims etc. All the selected papers have been processed through this stage.

**Stage 3+: “Synthesize the detail”**: This stage deals with synthesis of the data, concept & the results presented by the authors. Here imagination of situations different from those presented and expected results has to be predicted. In short it uses interpolation and extrapolation to find out the gaps in the published research. I could exercise this process for all the 19 research papers, of which the final findings and the outcomes have been discussed in the upcoming sections. This process helped me to easily classify the literature, evaluate, synthesize, summarize and build the problem statement of my research proposal. The outcome of the review was in the
form of various findings, and comparative study of similar approaches. The findings included Technologies and methodologies used to solve particular research problem, along with their strengths and weaknesses and the scope for the future work in the area

III. ISSUES AND SOLUTION APPROACHES

After an exhaustive review process, the 19 research papers could be categorized in following three issues:
1. ITS Architecture
2. Traffic Monitoring and Control
3. Vehicle collision avoidance

3.1. Findings Of Its Architecture (Issue 1)

- Generally all proposed architectures are based on Wireless Sensor Networks. Three kinds of nodes viz., vehicle unit, road side unit, and street or intersection unit are used for inter communication of vehicle with networks.
- Either mobile nodes on vehicle (Mobile Adhoc network: VANET) or stationary nodes on road sides to sense vehicle have been used in all the techniques to collect traffic information.
- WSN based ITS\(^4\), named as WITS gave a setup of ITS using the basic methodology of road side units and vehicular units. This technology took the advantage of peculiar features of the WSN being low energy consumption, small size, non-cable and full-road coverage which helped in achieving efficient ITS networking. However it was limited by the requirement of every vehicle to be equipped with the sensors.
- UTOSPF\(^{10}\), Urban Traffic Open Shortest Path First technique which is a novel distributed system based on WSN and OSPF protocol collected real-time traffic information from roads. While using all merits of WSN, the use of OSPF protocol concepts strengthened UTOSPF by making it autonomous, distributed, simple, inexpensive, highly scalable and totally independent of vehicles movement. This technology was only implemented to find optimal paths, but, this structure could also be used to detect and send other traffic information.
- DRGS\(^{11}\), Dynamic Route Guidance System, emerged to be another set up for capturing real time traffic information for route guidance, also used the features of wireless sensor networks, as it strengths. The major limitation of DRGS was broadcasting if data to all drivers despite of its irrelevance to some.
- It has also been proposed to combine the WSN and VANET\(^3\) technology to complement each other’s limitations. This approach combined the strength of both networks and emerged to be a wise solution for ITS. However, it was limited by the need of new protocol and equipment design.
- IGCS\(^{16}\), Intelligent Guiding and Control System based on Wireless Sensor Network Technology showed its merit over few, as it used mobile wireless sensors inside the vehicle, instead of fixed WSN nodes and was inexpensive set up. But, the response time of the system depended on the speed of network transmission and the operating database performance

3.1.1. Strengths

- All technologies were based on WSN (WITS, UTOSPF, DRGS, IGCS etc.), utilizing its features like low energy consumption, small size, non-cable and full-road coverage which helped in achieving efficient ITS networking.
• UTOSPF, Urban Traffic Open Shortest Path First technique used OSPF protocol concepts to make it autonomous, distributed, simple, inexpensive, highly scalable and totally independent of vehicles movement.

• DRGS, Dynamic Route Guidance System, emerged to be another set up for capturing real time traffic information for route guidance using inexpensive sensors.

• IGCS, Intelligent Guiding and Control System superseded few techniques, as it used mobile wireless sensors inside the vehicle, instead of fixed WSN nodes and was inexpensive set up.

3.1.2. Weaknesses

• WITS were limited by the requirement of every vehicle to be equipped with the sensors.

• UTOSPF was only implemented to find optimal paths, but, this structure could also be used to detect and send various other traffic information.

• DRGS, was broadcasting if data to all drivers despite of its irrelevance to some.

• Proposed to combine the WSN and VANET technology to complement each other’s limitations was constrained by the need of new protocol and equipment design.

• In IGCS, response time of the system depended on the speed of network transmission and the operating database performance.

3.1.3. DISCUSSION

• Generally, a Wireless Sensor Network set up for Intelligent Transport System, is composed of following subsystems:
  o sensing subsystem – road side units, on-board sensors (magnetic, ultrasound, accelerometer, Bluetooth)
  o distribution subsystem – intersection units, access units on roads, Traffic information Centers
  o Decision making and execution subsystem – data traffic monitoring and control by Master Node and Computer system, TMC.

• Many architectures were proposed like WITS, UTOSPF, DRGS, IGCS, WSN based ITS along with its prototypes in some, and simulations in others.

• [Fernando Losilla, et al,2012] stated the key requirement of combining WSN and VANET, as WSN had their main weakness in the constrained use of scarce available energy which was a minor issue in VANET. On the contrary, achieving high technology penetration rates in VANETs in order to boost performance was not straightforward, but the installation of WSN nodes on selected roads was a simpler task. A combination of both could enable that power consuming tasks such as dissemination of data rely on VANET nodes meanwhile WSN nodes offer a permanent monitoring of a given location, allowing the system to work properly even if the penetration rate of VANET was low

• In this light, [Djamel Djenouri, 2013] have proposed WVSN, wireless vehicular sensor network which seems to be a self-consistent technology, although he proposed this for

• [Hemjit Sawant, et al, 2004] have in turn proposed a system using Bluetooth for inter vehicular communication, for exchange of information like collision detection, congestion detection etc. which was sensed by an inexpensive on-board millimeter wave radar. Though this approach reduced the infrastructure cost directly, it did not consider the challenges of integrating this with the WSN set up.
Architectures like WITS and DRGS, were limited by requiring all vehicles to be equipped by the respective sensor units, specific to that requirement, which added the cost of implementation.

Scope still lies for better utilization, and improvement lies in redefining protocols to operate at high vehicle speeds, improving signal transmission, studying WSN application in network security issues, improving power efficiency etc. to make WSN work as an efficient standalone technology for ITS applications.

Comparative study have been given in following Table 1.

### Table 1 ITS Architecture: Solution Approaches

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Technology &amp; Year</th>
<th>Author &amp; Year</th>
<th>Solution Approach</th>
<th>Features</th>
<th>Results / Conclusion</th>
<th>Limitation and Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WITS: WSN for Intelligent Transportatio n System</td>
<td>Wenjie Chen, et al, 2006 [4]</td>
<td>Set up with 3 nodes Vehicle unit Road side unit Intersection unit-calculate location and velocity of vehicle</td>
<td>- Unique ID used for each WSN node for specific destination - Use of solar cell</td>
<td>Data aggregation and transfer successful by prototype.</td>
<td>Vehicles must be equipped with sensor and transceiver - Challenges in large scale network</td>
</tr>
<tr>
<td>2</td>
<td>UTOSPF: Urban traffic Open shortest Path-distributed system</td>
<td>Karim Faez, Moham mad Khanjar y, 2008 [10]</td>
<td>3 nodes or units Street Unit, Intersection Unit, Repeater Unit Estimated Street Travel Time of every street used to calculate optimal routes by Dijkstra's algorithm</td>
<td>- Based on WSN and Open shortest path first protocol. - Infoms via Variable Message Signs (VMS) or special transceivers - fast low cost deployment, high scalability, autonomy distributed processing (features of OSPF)</td>
<td>Simulation results showed that Average speed of vehicles could improve up to 40% as compared to random routes in the simulated scenario</td>
<td>Detecting average speed of vehicles from streets was only focused; could be used to detect and send various other traffic information and could play the role as an infrastructure for further applications</td>
</tr>
<tr>
<td>3</td>
<td>DRGS: Dynamic Route Guidance System</td>
<td>Chao Long &amp; MengShuai, 2010[11]</td>
<td>On-board magnetic sensors Traffic message Centre (TMC) TMC receiver</td>
<td>- On-board sensors used. - Use WORA (Write ones, run anywhere) features of JAVA</td>
<td>Could obtain real time and “real space” traffic information.</td>
<td>Information was delivered to drivers in spite of their relevance TMC receiver on board</td>
</tr>
<tr>
<td>4</td>
<td>IGCS</td>
<td>Yi-Ping et al 2012 [16]</td>
<td>Using Cyber Physical system with on board mobile sensor Data acquisition system, Control system, Monitoring system</td>
<td></td>
<td>Average response time 1010 ms run successfully on smart phones</td>
<td>Response time depends on the speed of network Transmission and operating atabase</td>
</tr>
<tr>
<td>5</td>
<td>WVSN</td>
<td>Djamel Djenouri, 2013[3]</td>
<td>2 sensors Magnetic Sensor-provide distance from and the velocity of the front vehicle, embedded within the front bumper Accelerometer sensor including GPS receiver -</td>
<td>Combines the features of both WSN (sensors inside vehicles) and Ad-hoc network (road side units). Tiny low-cost sensors replace infrastructure-based systems in rural</td>
<td>Evaluation and practical implementation by simulation to be done</td>
<td>Dealing with other kind of collisions, such as rear-end collisions, and use in other applications for traffic monitoring</td>
</tr>
<tr>
<td>6</td>
<td>Bluetooth based detection</td>
<td>Hemjit Sawant et al 2004 [2]</td>
<td>Formation of vehicular mobile ad-hoc networks (piconets and scatternets) using Bluetooth, which could exchange data in real time situations sensed by the on-board millimeter-wave radar sensor.</td>
<td>Bluetooth Characteristics, viz. 1 Mbps B.W., 3 level security, strong short range communication Being cost effective solution as data sensed by one sensor could be used by other vehicle easily, and low sensor cost</td>
<td>The set up was simulated on bluetooth simulator which revealed consistent connection set up when nodes were increased from 10 to 50.</td>
<td>Integration of the vehicle-based ad hoc WVSN with road side infrastructures to involve the vehicle-to-roadside sensor comm.</td>
</tr>
<tr>
<td>7</td>
<td>WSN based ITS</td>
<td>Fernand o Losilla, et al 2012 [15]</td>
<td>4 subsystems defined: Sensing subsystem: nodes in holes on roads. Distribution subsystem, Decision Making subsystem and Execution subsystem which define a reference architecture</td>
<td>General WSN set up</td>
<td>No results as it was a proposed prototype.</td>
<td>Though WSN is effective, it may still work better with other network as VANET, complementing the weaknesses</td>
</tr>
</tbody>
</table>

### 3.2 Findings Of Traffic Monitoring & Control, And Vehicle Collision Avoidance

Four papers were reviewed under this category, which is related to the previous, but focuses on traffic monitoring and control. The common findings could be listed as below:
In all approaches, sensors were used on vehicles: Bluetooth for inter-vehicle communication of traffic information, speed sensors and mobile bug to sense mobile phone usage; or tracking sensor and vehicular sensor to sense the traffic congestion.

Passenger Management System Based on Face Recognition for Intelligent Transport was proposed to make registration of passengers in crowded vehicles convenient and 20 times faster than existing technologies. It was enhanced using wireless network giving it remote control features. But, remote monitoring was practically irrelevant in dynamic and overcrowded vehicles.

Intelligent Vehicle Monitoring System using Wireless Communication was proposed to impose traffic rule to ban messaging while driving, using GSM and WSN set up. Though this technique was a novel approach, but required lot of training to stakeholders, and supporting infrastructure for this application. It could have been more preferred if local mobile jammers were used.

Intelligent Transport Management System using wireless sensor networks and dedicated traffic servers was conceptualized with the aim to provide the shortest time path for the driver in the variable traffic conditions. This technique could improve and reduce response time for a server to consider the user request and give reply, but was limited by security and reliability of the network performance.

Bluetooth and Sensor Networks for Intelligent Transportation Systems were proposed to increase the safety of road travel inter-vehicle communication among vehicles equipped with Bluetooth device and other on-board sensors. This was practically possible implementation, but, this paper did not consider several issues, such as, communication efficiency, integration of the vehicle-based ad hoc WSN with road side infrastructure.

WSN also found its application in traffic safety issues like collision avoidance. Two papers were reviewed which depicted the use of WSN in this application. Following are the common findings for this issue:

- Both approaches specifically pertained to collision avoidance scheme, still required entire WSN or WVSN set up.
- Wireless Vehicular Sensor Network technology was proposed to prevent vehicle crashes and avoid frontal collisions. The strength of this solution was that it was based on WVSN which was a combination of WSN and VANET technology. This could complement each other’s weaknesses. This application was limited in its scope of only dealing with frontal collision, and not rear-end or sideways.
- Several algorithms were presented to calculate collision time based on current speed and estimated time intervals and implemented using WSN in ITS. Though these algorithms returned faster simulation results, but were not practically validated.

### 3.2.1. STRENGTHS

- Passenger Management System Based on Face Recognition for Intelligent Transport could give 20 times faster than existing technologies. It was enhanced using wireless network giving it remote control features.
- Intelligent Vehicle Monitoring System using Wireless Communication was a novel approach to ban messaging while driving.
- Intelligent Transport Management System using wireless sensor networks and dedicated traffic servers could improve and reduce response time for a server to consider the user request and give reply.
- Bluetooth and Sensor Networks for Intelligent Transportation Systems was practically implementable and user friendly. It did not attract more infrastructure cost.
Wireless Vehicular Sensor Network technology\(^3\) was proposed to prevent vehicle crashes and avoid frontal collisions. The strength of this solution was that it was based on WVSN which was a combination of WSN and VANET technology. This could complement each other’s weaknesses.

- It could give information of probable collision before 2 seconds when entering in that zone.
- Several algorithms \(^5\) presented to calculate collision time based on current speed and estimated time intervals and implemented using WSN in ITS, returned faster simulation results.

Table 2 Shows Comparative Study of Different Traffic Monitoring and Control Schemes

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Technology</th>
<th>Author &amp; Year</th>
<th>Solution Approach</th>
<th>Features</th>
<th>Results / Conclusion</th>
<th>Limitation and Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bluetooth and sensor networks</td>
<td>Hong, Y., et al. 2004</td>
<td>Bluetooth communication and WSN set up, with mobile ad hoc network</td>
<td>Bluetooth strong short range communication</td>
<td>10-50 Bluetooth nodes could communicate easily and consistently</td>
<td>Integration of the vehicle-based ad hoc WSN with road side infrastructures to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>On board Radar senses traffic info, communicates to Bluetooth device</td>
<td></td>
<td></td>
<td>involve the vehicle-to-roadside sensor comm. communications overhead</td>
</tr>
<tr>
<td>2</td>
<td>WSN-ITS collision avoidance</td>
<td>Verna, V.K., et al. 2008</td>
<td>Modified best first path algo implemented using masking sensor on road and vehicle sensor to calculate shortest path</td>
<td>Hierarchical model to reduce response time of servers compatible with conventional protocols</td>
<td>Information to driver about shortest path time.</td>
<td>Security and trust</td>
</tr>
<tr>
<td>3</td>
<td>Passenger management by face detection</td>
<td>Byung-Gil Han, et al. 2013</td>
<td>Use AdaBoost method utilizing Local Binary Pattern (LBP) histogram</td>
<td>Could be integrated to remote units for monitoring</td>
<td>Detection speed is about 20 times faster</td>
<td>Only registered passengers could use this technology</td>
</tr>
<tr>
<td>4</td>
<td>Stop Messaging while driving</td>
<td>Ananthapadmanabhan, N., 2013</td>
<td>Traffic safety by detecting mobile phone usage while driving using GPS and GSM comm.</td>
<td>Detecting mobile phone usage while driving using GPS and GSM comm. Sending alert message to monitoring bodies.</td>
<td>Restricts use of mobile phones while driving.</td>
<td>Implementation calls for major infrastructural changes.</td>
</tr>
<tr>
<td>5</td>
<td>Collision Prediction at Intersection in Sensor Network Environment</td>
<td>Oje Kwon, et al. 2006</td>
<td>Several algorithms based on calculation of collision time interval, current speed, distance between vehicles etc.</td>
<td>All algorithms designed and tested. Practically not implemented</td>
<td>Algorithm based on current speed calculations was found to be most effective.</td>
<td>Challenges in large scale network, when the intersection units in the same city itself forms a huge network</td>
</tr>
<tr>
<td>6</td>
<td>Wireless Vehicular Sensor Network</td>
<td>Dyson, D., 2008</td>
<td>- Unique ID of the segment in which vehicle travels given.</td>
<td>- WVSN; using WSN + VANET.</td>
<td>Sensor received the overtaking information within seconds which could reduce accidents.</td>
<td>More investigations into the solution and its parameters, - Dealing with other kind of collisions,</td>
</tr>
</tbody>
</table>

3.2.2. WEAKNESSES

- In passenger management using face recognition, remote monitoring was practically irrelevant in dynamic and overcrowded vehicle
Using GSM and GPS for traffic rule imposition, like stopping driver to message, was limited by the fact that it required lot of training to stakeholders, and supporting infrastructure for this application. It could have been more preferred if local mobile jammers were used.

ITS using WSN was limited by security and reliability of the network performance.

The Bluetooth technique did not consider several issues, such as, communication efficiency, integration of the vehicle-based ad hoc WSN with road side infrastructure.

WVSN This application was limited in its scope of only dealing with frontal collision, and not rear-end or sideways.

Design of vehicle collision detection system was specifically designed only for frontal collisions, others were ignored.

Proposed algorithms were not practically implemented.

Proposed system included entire ITS system for specific application. Other features could be implemented.

3.2.3. DISCUSSION

Many approaches were given to control and monitor the traffic which were managing passengers in crowded vehicles using faster face detection, enforcement of traffic rule to stop messaging while driving, use of Bluetooth sensor to detect shortest path etc.

Various algorithms were defined using best first search algorithm to define shortest path time, and guide the driver to take up congestion free path.

The IGCS [16] system set up discussed in above category also proposed to guide the driver in case of congestion and heavy traffic.

All the systems and approaches discussed lead to the conclusion that various opportunities could be explored to inbuilt intelligence in transport system.

The passenger management system [18] was designed and tested and results could fetch 20 times faster face detection. This could manage public in crowded transport vehicle efficiently. On other hand, using GSM and GPS to detect mobile phone usage was a good approach to improve traffic safety. Other papers which emerged out to be techniques to detect congestion free path found their application in increasing traffic scenario, and the one based on Bluetooth was feasible and practically implementable.

For collision avoidance small sensors (magnetic sensor and accelerometer sensor) [3] were used which could replace the infrastructure-based systems in rural and suburban areas, where the deployment of such infrastructure was constrained. It gave an alternative solution for unsophisticated vehicles that were not equipped with aboard computers. It used the WVSN infrastructure for communication.

Second approach of designing several algorithms based on overlapping ratio between two collision time intervals, current speed of vehicles, and hybrid of two led to development of a system which could report collision warning to drivers at least two seconds before entering intersection, thus preventing collision. Out of thoses proposed algorithms, first algorithm based on overlapping ratio between two collision time intervals was suggested to be most robust, giving minimum number of false warnings.

IV. CONCLUSION

It emerged from the review that WSN was a cost effective and compatible technology to in-build intelligence in transport system as it could make ITS autonomous, adapting, efficient and reliant. Wireless Vehicular Sensor
Network Technology (WVSN) emerged out as a solution to mitigate weakness of standalone WSN and VANET. Various ITS architectures came up while the different papers were reviewed viz. WITS, UTOSPF, IGCS, DRGC, Bluetooth Technology, etc. All were similar differing in the type of sensors, location of sensor, and communicating channel. The review of few papers pertaining to the traffic management in ITS, highlighted the enormous scope of WSN inn ITS, may it be passenger management, route guidance, detection of shortest path, detection of traffic jams, imposition of traffic rules, remote access to dynamic vehicles and passengers, collision avoidance and many more. The scope which emerged out during this literature review was the acute requirement of stronger, energy efficient and secure protocol design for WSN. A protocol for the WVSN infrastructure also requires to be looked into. The limitations of each technology and solution approaches defines a new scope of study, yet to be explored.

REFERENCES


[18] Byung-Gil Han, Kil-Taek Lim, Yun-Su Chung, Soo-In Lee, “Passenger Management System Based on Face Recognition for Intelligent Transport Vehicles”, Ubiquitous and Future Networks (ICUFN), Fifth International Conference, IEEE, 2013

APPLICATION OF ANALYTIC HIERARCHY PROCESS (AHP) TO PRIORITIZE THE FACTORS INDIAN CONSUMERS CONSIDER WHILE BUYING SMART PHONES IN INDIA

Soham Chakraborty\(^1\), Shashwat Srivastava\(^2\), Karishma.Das\(^3\)

\(^1\), 2, 3 School Of Mechanical Engineering, KIIT University, Orissa, (India)

ABSTRACT
Smart phones have almost become a necessity for Indian mobile phone users. The study clearly identifies and prioritizes the features Indian consumers consider while purchasing smart phones in India. Analytic Hierarchy Process is a decomposition multiple-attribute decision making (MADM) method, in this study it has been used for ranking of the features present in smart phones. A mathematical model based on pair-wise comparison values has been developed by applying AHP. This paper provides relevant data to help the mobile phone manufacturing companies to focus on the features that Indian customers demand and at the same time make efforts to improve them in order to increase the sales of their smart phones in the country.

**Keywords:** Analytic Hierarchy Process, Features, Indian Consumers, Multi Attribute Decision Making (MADM)

I. INTRODUCTION

It is the era of information and technology and we are surrounded by many such gadgets and equipments which make our lives something more than easy. Smart phones are one such class of items which has totally revolutionized the way people used mobile phones even a decade ago. A smart phone is a cellular telephone with built-in applications and Internet access in addition to digital voice service, modern smart phones provide text messaging, e-mail, Web browsing, still and video cameras, MP3 player and video playback and calling \[1\]. They include functions similar to personal computers. Initially PDAs (personal digital assistant) were smart phones possessing calling features. But, nowadays the presence of added media- players, high resolution cameras, GPS, Wi-Fi are a must with any smart phone. Additionally, every smart phone these days have high-density screen resolution allowing the handsets to display various websites in their standard formats as they appear on our computer screens.

1.1 The Indian Smart Phone Scenario
Smart phones are no longer a gadget of luxury and sophistication they have now become a sensation in the Indian mobile phone market. India is ranked third among the top countries with smart phone users only behind China and USA with an estimated 118 million subscribers. A study by telecom equipment maker Ericsson shows that smart phone users in the country have among the highest rates of usage of smart phones daily globally, spending over three hours on an average on their devices. According to the study, Indian users spend 3 hours 18 minutes on average everyday with their smart phones while in the US, where the average is 132 minutes (2 hrs 12 mnts \[2\).
Analytic Hierarchy Process is a decomposition multiple-attribute decision making (MADM) method. It is a method that can represent human decision making process and help in the achievement of better judgments based on hierarchy, pair-wise comparisons, judgment scales, allocation of criteria weights and selection of the best alternative from a finite number of variants by calculation of their utility functions. Subsequently, there has been a growth of applications and mathematical development to this methodology. The developments were focused on different parts of the method.

II. AIM OF THE RESEARCH

In this study we try to prioritize the factors the Indian consumers consider while purchasing smart phones. The Analytic Hierarchy Process (AHP) methodology has been used to determine the factors which play the most influential role for Indian customers while buying smart phones. AHP is a mathematical and psychological decision making technique developed by T.L.Satty in the 1970s and has been widely studied and refined since then. The BPMSG AHP priority calculator online tool was used for obtaining preferences on criteria and alternatives for the determination of the relative priority of factors for smart phones in the Indian scenario.

III. REVIEW OF EXISTING LITERATURE

3.1 Smart phones

Hsiao and Chen did an empirical study on Smart phone demand and on the relationships between phone handset, Internet access, and mobile services. Their study explored the smart phone demand by emphasizing the differences between the three demand dimensions: (1) mobile or smart phone handset, (2) subscription to the 2G/3G network, and (3) mobile services, and then examined the relationship between them, and the effect of users’ demographic characteristics on these three dimensions as well, by an empirical study in Taiwan [3]. Tseng and Lo in their study mentioned that the characteristics of the mobile/smart phone handset industry is multi-faceted; e.g., rapidly evolving nature with short product life-cycles [4]. Economides and Grousopoulou found that students tend to consider the following features important: battery life, mp3 player, video camera, photo camera, storage memory, Bluetooth, design and elegance, clock, calendar, organizer and reminder, while most of the respondents in their study do not consider the following important: touch screen, voice commands, chat, teleconference, encryption and cryptography, common use of files, printing [5]. Malviya et al. evaluated the factors influencing consumer’s purchase decision towards smart phones in Indore. Using the confirmatory factor analysis model they concluded that people in Indore were buying Smart phones irrespective of its prices [6]. Lay-Yee et al. studied the factors affecting smart phone purchase decision among Malaysian generation Y [7].

3.2 Analytic Hierarchy Process (AHP)

Dožić and Kalić applied AHP for aircraft selection process. By consideration of the selected criteria (aircraft seat capacity, aircraft price, total baggage, MTOW, payment conditions and CASM), various aspects of aircraft purchasing were covered, allowing airline’s planner to choose the right aircraft from the set of alternatives. Their study showed that the AHP can be successfully used as a support tool in the decision making process related to aircraft selection problem, regarding criteria defined in their research [8]. Franek and Kresta in their research investigated the application and characteristics of different judgment scales developed by scholars for use in AHP. Results and their comparison showed that judgement scales played a significant role in AHP.
decision making [9]. Zuo et al. researched on the current situation of peasant-workers in construction industry based on AHP. They concluded that the whole level of the peasant-workers’ living was on the poor level. The indexes of the physiological need, the security need and the social need weighed high but obtained low scores. The indexes of the respected need and the self-realization need weighed low, but the comprehensive scores of the five indexes had no obvious difference, that means the peasant-workers in construction industry not only had physiological, security and social needs, but also had respected and self-actualization needs [10].

Wang and Pan researched on the influence factors of Wuhan Housing Industry based on the AHP. They concluded that Government, enterprise, consumer and market were the key four factors of housing industrialization development in Wuhan through analysis and investigation and additionally they established an AHP structure model of influence factors to decide which factor was the most important one [11]. Tyagi et al. analysed the e-SCM performance by a hybrid approach using AHP-TOPSIS. Their research showed that the criteria ‘improvement in production efficiency’ and ‘on time delivery’ achieved higher priority weights [12].

Lijuan and Shinnan made use of the approach of AHP for human factors analysis in the Aircraft Icing Accident.[13]. Khanmohammadi and Rezaeiahari did an AHP based classification of algorithm selection for clinical decision support system. In their study, a meta-learning algorithm was proposed to choose a machine learning classification algorithm that could be used for the development of CDSS [14]. Podgórski in his study demonstrated an AHP based study for selection of leading key performance indicators measuring operational performance of OSH management system. His paper presented a concept of making use of operationally focused minimum set of key performance indicators assigned to individual OSH MS components [15].

IV. PROBLEM DESCRIPTION

Indian customers are very much choosy when it comes to purchasing smart phones. They are driven by a variety of factors starting from pricing, promotions and advertisements, durability, configuration, battery life, storage, camera resolution, connectivity options which are considered as the criteria in this study. The alternatives considered for this study are the affordability, design, brand, operating system (like Android, IOs, Windows etc.), functionality and user experience. Considering the selected criteria and the alternatives the analytical hierarchy process (AHP) has been applied.

V. METHODOLOGY OF RESEARCH

The study initiated by the identification of criteria and alternatives for the features Indian customers consider while purchasing Smart phones and selection of appropriate MADM methods. Depending upon the decision maker’s inputs, the criteria weights were computed using the Analytic Hierarchy Process and additionally further computations were done to prioritize the factors while purchasing smart phones in the country. AHP is a Multi Attribute Decision Making technique and it is designed to incorporate tangible as well as intangible factors especially where the judgments are subjective for different individuals constitute an important part of decision making. A five step process is used in AHP to solve decision problems.

Analytic Hierarchy Process is a mathematical and psychological tool for the systematic analysis of expert opinions. Consultation with more experts helps avoid bias which may be found while considering the judgments of a single expert. The decision making for this study included people from all age groups who possess smart phones as well as from industry experts of the mobile phone sector. For the survey, Importance scale from 9 being Extremely Important to 1 being Equally Important was used to obtain the judgments. After building the
matrix it becomes possible to compute the priority vector. Comparison between the elements based upon a single entity for building the pairwise comparison matrices for the criteria with the pairwise comparison matrix for the alternatives helps in the computation of global and local priorities as well as ranking of the alternatives. The calculation of priorities from pairwise comparison matrices can be done in different ways which include:
(a) eigenvector method; (b) geometric mean method; (c) arithmetic average method

In this research we have used the eigenvector method proposed by Saaty(1980), according to whom the priority estimation of elements can be done by finding the principal eigen vector \( w \) of any matrix \( M \), \( Mw=\lambda_{\text{max}}w \), where \( \lambda_{\text{max}} \) is the maximum eigenvalue of matrix \( M \). After normalising vector \( w \), it becomes the vector of priorities of elements. During building of each of pairwise comparison matrices involved in the decision making process, calculation of Consistency Ratio is necessary to check for consistency which is the ratio of Consistency Index (CI) and Random Index (RI).

\[
\text{Consistency Index} = \frac{\lambda_{\text{max}} - r}{n-1} \tag{1}
\]

\[
\text{Consistency Ratio} = \frac{CI}{RI} \tag{2}
\]

The BPMSG AHP priority calculator online tool helped in obtaining the preferences on criteria and alternatives from the data collected for this study as well as the principal eigen value.

5.1 Questionnaire Design and Survey

In case of purchase of smart phones, pricing, durability, battery life, promotions and advertisements, camera resolution, configuration, storage and connectivity options are the main criteria which influence the features which customers consider while buying any smart phone in India. Thus the research objective is to prioritize the features related to smart phones while purchasing them from the Indian scenario. A country wide survey was conducted which included people from all age groups as well as industry experts from various mobile phone manufacturers. The medium of data collection was from online surveys and email, telephonic conversations, interview with company officials. The questionnaires was distributed among 2000 people nationwide from which 1981 positive responses were received which converted to percentage value comes to 99.05%. A successful survey is only possible when the questions are kept simple and to the point. This helps in the achieving valid, relevant and reliable data. Some of the sample questions have been tabulated below.
<table>
<thead>
<tr>
<th>S.NO.</th>
<th>SAMPLE QUESTIONS</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Smart phone should be Durable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Battery-life has significance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Pricing is considered while purchasing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Camera resolution should be high</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Sample Questionnaires

1: Equal Importance ; 3: Moderate Importance ; 5: Strong Importance ; 7: Very Strong Importance ; 9: Extreme Importance ; (2,4,6,8 are the values in between)

5.2 Establishment of the AHP Structure System

![Hierarchial Structure](image)

NOMENCLATURE OF CRITERIA & ALTERNATIVES

<table>
<thead>
<tr>
<th>CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 DURABILITY</td>
</tr>
<tr>
<td>C2 PRICING</td>
</tr>
<tr>
<td>C3 ADVERTISEMENTS AND PROMOTIONS</td>
</tr>
<tr>
<td>C4 BATTERY LIFE</td>
</tr>
<tr>
<td>C5 STORAGE</td>
</tr>
<tr>
<td>C6 CAMERA RESOLUTION</td>
</tr>
<tr>
<td>C7 CONNECTIVITY OPTIONS</td>
</tr>
<tr>
<td>C8 CONFIGURATION</td>
</tr>
<tr>
<td>A1 AFFORDABILITY</td>
</tr>
<tr>
<td>A2 DESIGN</td>
</tr>
<tr>
<td>A3 BRAND</td>
</tr>
<tr>
<td>A4 OPERATING SYSTEM</td>
</tr>
<tr>
<td>A5 FUNCTIONALITY</td>
</tr>
<tr>
<td>A6 USER EXPERIENCE</td>
</tr>
</tbody>
</table>
VI. RESULTS

6.1 Pairwise Comparison of Criteria

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>RI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.58</td>
<td>0.90</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
<td>1.45</td>
<td>1.49</td>
</tr>
</tbody>
</table>

Table 3. Random Inconsistency Index (RI) (Saaty 1980, Saaty 1990)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>Priority Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1</td>
<td>5.00</td>
<td>8.00</td>
<td>1.00</td>
<td>3.00</td>
<td>0.20</td>
<td>5.00</td>
<td>0.41</td>
<td>0.124</td>
</tr>
<tr>
<td>C2</td>
<td>0.20</td>
<td>1</td>
<td>4.00</td>
<td>0.33</td>
<td>2.00</td>
<td>0.20</td>
<td>2.00</td>
<td>0.71</td>
<td>0.054</td>
</tr>
<tr>
<td>C3</td>
<td>0.12</td>
<td>0.25</td>
<td>1</td>
<td>0.25</td>
<td>0.50</td>
<td>0.14</td>
<td>1.00</td>
<td>0.12</td>
<td>0.024</td>
</tr>
<tr>
<td>C4</td>
<td>1.00</td>
<td>3.00</td>
<td>4.00</td>
<td>1</td>
<td>2.00</td>
<td>0.25</td>
<td>6.00</td>
<td>0.20</td>
<td>0.099</td>
</tr>
<tr>
<td>C5</td>
<td>0.33</td>
<td>0.50</td>
<td>2.00</td>
<td>0.50</td>
<td>1</td>
<td>0.33</td>
<td>4.00</td>
<td>0.17</td>
<td>0.053</td>
</tr>
<tr>
<td>C6</td>
<td>5.00</td>
<td>5.00</td>
<td>7.00</td>
<td>4.00</td>
<td>3.00</td>
<td>1</td>
<td>6.00</td>
<td>0.33</td>
<td>0.238</td>
</tr>
<tr>
<td>C7</td>
<td>0.20</td>
<td>0.50</td>
<td>1.00</td>
<td>0.17</td>
<td>0.25</td>
<td>0.17</td>
<td>1</td>
<td>0.17</td>
<td>0.026</td>
</tr>
<tr>
<td>C8</td>
<td>7.00</td>
<td>6.00</td>
<td>8.00</td>
<td>5.00</td>
<td>6.00</td>
<td>3.00</td>
<td>6.00</td>
<td>1</td>
<td>0.386</td>
</tr>
</tbody>
</table>

Table 4. Decision Matrix for Criteria

$\lambda_{max} = 8.878 \quad CR=0.09 \quad CI=0.125$

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>0.33</td>
<td>0.17</td>
<td>4.00</td>
<td>1.00</td>
<td>4.00</td>
<td>0.101</td>
</tr>
<tr>
<td>A2</td>
<td>3.00</td>
<td>1</td>
<td>1.00</td>
<td>8.00</td>
<td>3.00</td>
<td>9.00</td>
<td>0.301</td>
</tr>
<tr>
<td>A3</td>
<td>6.00</td>
<td>1.00</td>
<td>1</td>
<td>8.00</td>
<td>6.00</td>
<td>9.00</td>
<td>0.401</td>
</tr>
<tr>
<td>A4</td>
<td>0.25</td>
<td>0.12</td>
<td>0.12</td>
<td>1</td>
<td>0.17</td>
<td>2.00</td>
<td>0.035</td>
</tr>
<tr>
<td>A5</td>
<td>1.00</td>
<td>0.33</td>
<td>0.17</td>
<td>6.00</td>
<td>1</td>
<td>8.00</td>
<td>0.128</td>
</tr>
<tr>
<td>A6</td>
<td>0.25</td>
<td>0.11</td>
<td>0.11</td>
<td>0.50</td>
<td>0.12</td>
<td>1</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Table 5. Decision Matrix for Alternatives with respect to Durability

$\lambda_{max} = 6.345 \quad CR=0.055 \quad CI=0.069$

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>6.00</td>
<td>0.33</td>
<td>3.00</td>
<td>4.00</td>
<td>7.00</td>
<td>0.0245</td>
</tr>
<tr>
<td>A2</td>
<td>0.17</td>
<td>1</td>
<td>0.17</td>
<td>0.50</td>
<td>0.14</td>
<td>1.00</td>
<td>0.038</td>
</tr>
<tr>
<td>A3</td>
<td>3.00</td>
<td>6.00</td>
<td>1</td>
<td>8.00</td>
<td>6.00</td>
<td>9.00</td>
<td>0.479</td>
</tr>
<tr>
<td>A4</td>
<td>0.33</td>
<td>2.00</td>
<td>0.12</td>
<td>1</td>
<td>1.00</td>
<td>3.00</td>
<td>0.079</td>
</tr>
<tr>
<td>A5</td>
<td>0.25</td>
<td>7.00</td>
<td>0.17</td>
<td>1.00</td>
<td>1</td>
<td>7.00</td>
<td>0.127</td>
</tr>
<tr>
<td>A6</td>
<td>0.14</td>
<td>1.00</td>
<td>0.11</td>
<td>0.33</td>
<td>0.14</td>
<td>1</td>
<td>0.031</td>
</tr>
</tbody>
</table>
### Table 6. Decision Matrix For Alternatives With Respect To Pricing

\[ \lambda_{max} = 6.517 \quad CR=0.082 \quad CI=0.1034 \]

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>0.17</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.33</td>
<td>0.023</td>
</tr>
<tr>
<td>A2</td>
<td>6.00</td>
<td>1</td>
<td>0.33</td>
<td>2.00</td>
<td>0.12</td>
<td>2.00</td>
<td>0.106</td>
</tr>
<tr>
<td>A3</td>
<td>9.00</td>
<td>3.00</td>
<td>1</td>
<td>4.00</td>
<td>0.50</td>
<td>7.00</td>
<td>0.272</td>
</tr>
<tr>
<td>A4</td>
<td>8.00</td>
<td>0.50</td>
<td>0.25</td>
<td>1</td>
<td>0.17</td>
<td>3.00</td>
<td>0.097</td>
</tr>
<tr>
<td>A5</td>
<td>9.00</td>
<td>8.00</td>
<td>2.00</td>
<td>6.00</td>
<td>1</td>
<td>6.00</td>
<td>0.453</td>
</tr>
<tr>
<td>A6</td>
<td>3.00</td>
<td>0.50</td>
<td>0.14</td>
<td>0.33</td>
<td>0.17</td>
<td>1</td>
<td>0.049</td>
</tr>
</tbody>
</table>

### Table 7. Decision Matrix for Alternatives With Respect To Advertisement and Promotions

\[ \lambda_{max} = 6.465 \quad CR=0.074 \quad CI=0.093 \]

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>3.00</td>
<td>1.00</td>
<td>0.33</td>
<td>0.11</td>
<td>0.12</td>
<td>0.056</td>
</tr>
<tr>
<td>A2</td>
<td>0.33</td>
<td>1</td>
<td>0.25</td>
<td>0.17</td>
<td>0.12</td>
<td>0.17</td>
<td>0.030</td>
</tr>
<tr>
<td>A3</td>
<td>1.00</td>
<td>4.00</td>
<td>1</td>
<td>1.00</td>
<td>0.14</td>
<td>0.20</td>
<td>0.081</td>
</tr>
<tr>
<td>A4</td>
<td>3.00</td>
<td>6.00</td>
<td>1.00</td>
<td>1</td>
<td>0.33</td>
<td>0.50</td>
<td>0.135</td>
</tr>
<tr>
<td>A5</td>
<td>9.00</td>
<td>8.00</td>
<td>7.00</td>
<td>3.00</td>
<td>1</td>
<td>1.00</td>
<td>0.382</td>
</tr>
<tr>
<td>A6</td>
<td>8.00</td>
<td>6.00</td>
<td>5.00</td>
<td>2.00</td>
<td>1.00</td>
<td>1</td>
<td>0.316</td>
</tr>
</tbody>
</table>

### Table 8. Decision Matrix For Alternatives With Respect To Battery Life

\[ \lambda_{max} = 6.326 \quad CR=0.052 \quad CI=0.0652 \]

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>8.00</td>
<td>0.500</td>
<td>4.00</td>
<td>0.25</td>
<td>3.00</td>
<td>0.158</td>
</tr>
<tr>
<td>A2</td>
<td>0.12</td>
<td>1</td>
<td>0.12</td>
<td>0.33</td>
<td>0.11</td>
<td>0.25</td>
<td>0.025</td>
</tr>
<tr>
<td>A3</td>
<td>2.00</td>
<td>8.00</td>
<td>1</td>
<td>8.00</td>
<td>0.25</td>
<td>2.00</td>
<td>0.221</td>
</tr>
<tr>
<td>A4</td>
<td>0.25</td>
<td>3.00</td>
<td>0.12</td>
<td>1</td>
<td>0.14</td>
<td>1.00</td>
<td>0.052</td>
</tr>
<tr>
<td>A5</td>
<td>4.00</td>
<td>9.00</td>
<td>4.00</td>
<td>7.00</td>
<td>1</td>
<td>6.00</td>
<td>0.471</td>
</tr>
<tr>
<td>A6</td>
<td>0.33</td>
<td>4.00</td>
<td>0.50</td>
<td>1.00</td>
<td>0.17</td>
<td>1</td>
<td>0.073</td>
</tr>
</tbody>
</table>

### Table 9. Decision Matrix for Alternatives with respect to Storage.

\[ \lambda_{max} = 6.410 \quad CR=0.065 \quad CI=0.082 \]
<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>4.00</td>
<td>0.50</td>
<td>0.50</td>
<td>8.00</td>
<td>7.00</td>
<td>0.232</td>
</tr>
<tr>
<td>A2</td>
<td>0.25</td>
<td>1</td>
<td>0.12</td>
<td>0.14</td>
<td>2.00</td>
<td>1.00</td>
<td>0.052</td>
</tr>
<tr>
<td>A3</td>
<td>2.00</td>
<td>8.00</td>
<td>1</td>
<td>0.50</td>
<td>5.00</td>
<td>2.00</td>
<td>0.256</td>
</tr>
<tr>
<td>A4</td>
<td>2.00</td>
<td>7.00</td>
<td>2.00</td>
<td>1</td>
<td>6.00</td>
<td>6.00</td>
<td>0.359</td>
</tr>
<tr>
<td>A5</td>
<td>0.12</td>
<td>0.50</td>
<td>0.20</td>
<td>0.17</td>
<td>1</td>
<td>1.00</td>
<td>0.042</td>
</tr>
<tr>
<td>A6</td>
<td>0.14</td>
<td>1.00</td>
<td>0.50</td>
<td>0.17</td>
<td>1.00</td>
<td>1</td>
<td>0.058</td>
</tr>
</tbody>
</table>

Table 10. Decision Matrix for Alternatives with respect to Camera resolution

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>1.00</td>
<td>0.11</td>
<td>0.12</td>
<td>0.50</td>
<td>0.12</td>
<td>0.035</td>
</tr>
<tr>
<td>A2</td>
<td>1.00</td>
<td>1</td>
<td>0.12</td>
<td>0.12</td>
<td>0.50</td>
<td>0.25</td>
<td>0.043</td>
</tr>
<tr>
<td>A3</td>
<td>9.00</td>
<td>8.00</td>
<td>1</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00</td>
<td>0.274</td>
</tr>
<tr>
<td>A4</td>
<td>8.00</td>
<td>8.00</td>
<td>1.00</td>
<td>1</td>
<td>3.00</td>
<td>0.50</td>
<td>0.262</td>
</tr>
<tr>
<td>A5</td>
<td>2.00</td>
<td>2.00</td>
<td>0.50</td>
<td>0.33</td>
<td>1</td>
<td>0.50</td>
<td>0.101</td>
</tr>
<tr>
<td>A6</td>
<td>8.00</td>
<td>4.00</td>
<td>1.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1</td>
<td>0.283</td>
</tr>
</tbody>
</table>

$\lambda_{max} = 6.408 \quad CR=0.065 \quad CI=0.081$

Table 11. Decision Matrix for Alternatives with respect to Configuration

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>7.00</td>
<td>1.00</td>
<td>2.00</td>
<td>0.11</td>
<td>2.00</td>
<td>0.122</td>
</tr>
<tr>
<td>A2</td>
<td>0.14</td>
<td>1</td>
<td>0.25</td>
<td>0.12</td>
<td>0.11</td>
<td>1.00</td>
<td>0.031</td>
</tr>
<tr>
<td>A3</td>
<td>1.00</td>
<td>4.00</td>
<td>1</td>
<td>1.00</td>
<td>0.11</td>
<td>2.00</td>
<td>0.093</td>
</tr>
<tr>
<td>A4</td>
<td>0.50</td>
<td>8.00</td>
<td>1.00</td>
<td>1</td>
<td>0.12</td>
<td>1.00</td>
<td>0.096</td>
</tr>
<tr>
<td>A5</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>8.00</td>
<td>1</td>
<td>7.00</td>
<td>0.601</td>
</tr>
<tr>
<td>A6</td>
<td>0.50</td>
<td>1.00</td>
<td>0.50</td>
<td>1.00</td>
<td>0.14</td>
<td>1</td>
<td>0.057</td>
</tr>
</tbody>
</table>

$\lambda_{max} = 6.201 \quad CR=0.0323 \quad CI=0.0402$

Table 12. Decision Matrix for Alternatives with respect to Connectivity options

<table>
<thead>
<tr>
<th>Alternative</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Priority vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>7.00</td>
<td>1.00</td>
<td>2.00</td>
<td>0.11</td>
<td>2.00</td>
<td>0.122</td>
</tr>
<tr>
<td>A2</td>
<td>0.14</td>
<td>1</td>
<td>0.25</td>
<td>0.12</td>
<td>0.11</td>
<td>1.00</td>
<td>0.031</td>
</tr>
<tr>
<td>A3</td>
<td>1.00</td>
<td>4.00</td>
<td>1</td>
<td>1.00</td>
<td>0.11</td>
<td>2.00</td>
<td>0.093</td>
</tr>
<tr>
<td>A4</td>
<td>0.50</td>
<td>8.00</td>
<td>1.00</td>
<td>1</td>
<td>0.12</td>
<td>1.00</td>
<td>0.096</td>
</tr>
<tr>
<td>A5</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>8.00</td>
<td>1</td>
<td>7.00</td>
<td>0.601</td>
</tr>
<tr>
<td>A6</td>
<td>0.50</td>
<td>1.00</td>
<td>0.50</td>
<td>1.00</td>
<td>0.14</td>
<td>1</td>
<td>0.057</td>
</tr>
</tbody>
</table>

$\lambda_{max} = 6.622 \quad CR=0.099 \quad CI=0.1244$
Table 13 Global Priority Weights for Alternatives With Respect To the Chosen Criteria

<table>
<thead>
<tr>
<th>ALTERNATIVES</th>
<th>GLOBAL PRIORITY WEIGHTS</th>
<th>RANKING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affordability</td>
<td>0.1002</td>
<td>5</td>
</tr>
<tr>
<td>Design</td>
<td>0.0759</td>
<td>6</td>
</tr>
<tr>
<td>Brand</td>
<td>0.2709</td>
<td>1</td>
</tr>
<tr>
<td>Operating system</td>
<td>0.2161</td>
<td>2</td>
</tr>
<tr>
<td>Functionality</td>
<td>0.1609</td>
<td>4</td>
</tr>
<tr>
<td>User Experience</td>
<td>0.1657</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 14 Ranking of Alternatives With Respect To Global Priority Weights Calculated

VII. CONCLUSIONS AND SCOPE FOR FUTURE WORK

Selection and prioritizing of features related to Smart phones in the Indian mobile phone market is a very important development which will help the companies to identify the specific features to cater to the needs of the Indian customers. India being third in the Smart phone users list globally certainly acts as a lucrative trading zone for the leading Smart phone producers in the world. Smart phones today have more than expected features
to offer to the customers but what Indian customers want has been the main concern in this research. Thus it can be concluded that the feature which Indian customers consider while purchasing smart phones is the Brand of the Smart phones. It is followed by the Operating system such as Android, Windows, IOs, etc., and then comes User experience functionality and so on. The feature which least affects the buying decision of Indian Smart phone buyers is the Design of Smart phones compared to all these given features. In the present advancement of Smart phone technology Affordability which earlier used to be a great determinant of purchasing decision is now on the lower side of the rankings. Analytic Hierarchy Process has been put into use for the calculation of the global priority weights of the Alternatives with respect to the Criteria considered for this study. The results seem pretty much satisfactory matching Smart phone buying trend among Indian customers belonging to different age groups. Hence by the application of AHP for the analysis of features affecting the Smart phone buying decision of Indian customers, it has been found that Brand is the best alternative. With such rapid development of new technology in the area of Smart phones in near future the present trend may change and considering these features along with many more added features, other statistical decision-making methods apart from AHP may be implemented for this study to obtain results showing the future trends.

VIII. ACKNOWLEDGEMENTS

The author would like to thank six anonymous referees for their excellent suggestions and valuable advice in this paper.

REFERENCES

[1] (http://www.pcmag.com)
[5] A.A. Economides, A. Grousopoulou, Students’ thoughts about the importance and costs of their mobile devices’ features and services, Telematics Information, 26, 57-84.


**BIBLIOGRAPHY**

Soham Chakraborty is a research scholar in School of Mechanical Engg in KIIT University. He has currently published two papers in various conferences and journals. His interest area of research is TQM, Sustainability MGMT etc.

Shashwat Srivastava is a research scholar in School of Mechanical Engg in KIIT University. He has currently published two papers in various conferences and journals. His interest area of research is Service operation MGMT, Supply chain MGMT etc.

Karishma Das is a research scholar in School of Mechanical Engg in KIIT University. She has currently published two papers in various conferences and journals. Her interest area of research is TQM, Sustainability MGMT, Service operation MGMT etc
“IMPORTANCE OF HADOOP IN THE MODERN ERA” - PERFORMANCE AND ITS PORTABILITY

Meetali

Computer Science & Engineering ITM University, Gurgaon, (India).

ABSTRACT

Hadoop is a flexible and open source implementation for analyzing large datasets using Map Reduce. The file system is developed in java that encourages portability around multiple heterogeneous software and hardware platforms. This paper analyzes the bottlenecks in the architecture of Hadoop that causes the time delay in the tasks and lead to the prolonged execution time of the jobs. Secondly, the limitations in the portability and other issues and challenges of Hadoop are discussed and for them what can be the solutions. We see Hadoop features, extensions, and tools as well as significant opportunities for optimization. We discussed remarkable miscellany in the styles, including various respective workloads that helps in the motivation of the new tools in the ecosystem. We have also found some standard techniques that helps in improving the potential and interpretation effectively but also suggested some of their substitutes. Altogether, some outstanding ways of opportunity for the simplification of the usage and the optimization of the Hadoop technology are discussed that also helps in making future recommendations for further research work in this area.

Keywords - Hadoop, Metadata, Map Reduce, Nodes, Security.

I. INTRODUCTION

In the daily lives, the data is generated at such unprecedented rates that the data intensive computations are performed in every second and hence so truly called “The digital universe” (EMC Corporation) [1]). The mindset of researchers are changed due to the phenomenal raised usage of the internet to store and analyze voluminous data. Nowadays, companies need to acquire some reliable solutions that can process multi PetaByte set of data efficiently and hence various groups of engineers in the companies like Nokia and some other telecommunications had already began experimenting with Hadoop. Clustering is an example of such one of the technology out of many existing methodologies in data mining that captures the visiting behavior of the web users accordingly their interest and thus can improve the infrastructure of the website to make it more user friendly. Sectors that employ these techniques includes retailers, banking institutions and some insurance companies as well as other health related fields (Huang et al. [2]; Lavrac et al. [3]). Hadoop also provides some methods to enhance the availability of applications by making the data replications but it lacks for the high availability for itself which is called SPOF. It means single point of failure or failure due to failure of single node i.e. mainly the critical node.
II. HADOOP ECOSYSTEM

![Fig 1: hadoop architecture](image)

a) NAME NODE

The NameNode of the system manages the namespace tree and the physical location of file data. An HDFS user that wants to access (reading) the file will first reach the NameNode to get the location of data blocks holding the file and then reads contents from the DataNode nearest to the client. When writing the data, the client sends requests to the NameNode to nominate a suite of three DataNodes for hosting the block replicas and then client writes the data to the DataNodes in a pipeline manner. Everytime the record of the image is stored in the local host’s native file system persistently and it is called a checkpoint. The NameNode stores the log of modification of the image that is referred as the journal. To have everlasting durability, redundant copies of the checkpoint and journal can be obtained at other physical and virtual servers.

b) DATA NODES

When the System starts every DataNode connects to the NameNode and performs handshaking that helps in to identify the namespace ID and the DataNode’s version. If either does not match, the NameNode and the DataNode automatically stop working and the system shuts down. After handshaking the DataNode registers with the NameNode for further processing. Under normal operation of the System DataNodes sends ping packets to the NameNode after a definite interval of time. Ping packet basically contain the all concern information of the data node and its existing application files. If the NameNode does not receive a ping packet from a DataNode after a certain time period NameNode considers that DataNode temporarily to be out of service and in turn NameNode schedules creation of new replicas of such blocks on other active DataNodes.

III. MAP REDUCE FRAMEWORK

MapReduce is proposed by Google (Apache Hadoop [4]. It is a programming model and implemented for large-scale data processing in distributed cluster HDFS.

Two Stage Performance-
Stage I
Map function is applied in parallel to each partition of the input raw data, to carry out the grouping and clustering operations.

Stage II

Reduce function is applied in parallel to each group produced in the first stage, to achieve the final aggregation.

The MapReduce model enables the users to develop the programs for data analysis that can be ascend upto thousands of nodes, without Concerning about the details of implementing parallelism. Being popular and open source implementation, Hadoop, has been used extensively by companies like Yahoo and Facebook to process Multi Petabyte Datasets efficiently in cloud computing. Therefore it is vitally important to oversee and track the status of distributed cluster through MapReduce-based data analysis using Hadoop (Shvachko et al. [5])

Fig 2: Map Reduce Framework

Hadoop MapReduce framework is used for executing applications holding enormously immense amounts of data (terabytes of data) in parallel on largely built clusters with numerous no of nodes in a reliable and fault-tolerant way. Though it can run on a single machine, its real power reside in its potential to scale up to several thousands of systems each with several processors. Hadoop is contructed in a way that it distributes the data effectively and efficiently across multiple nodes in the cluster. It involes the distributed file system that is responsible for the distribution of the massive amount of data sets efficiently across the nodes in the cluster. MapReduce framework splits the task into various numbers of chunks and the Map tasks process all them in parallel. The yield from the map tasks are sorted by the framework and provided as input to Reduce tasks. Both the input and output of the following tasks are registerd in a file system. The scheduling of the tasks while monitoring the failed ones and reexecuting them is done by the framework itself.

IV ROLE OF TRACKERS
a) Job tracker
Every cluster carry only one JobTracker sometimes called a “daemon service” for submitting and tracking MapReduce jobs in Hadoop and hence also responsible for the occurrence of single point of failure so if it goes out of the service all the currently running jobs will also be halted.

b) Task tracker
The slaves configured perform tasks as directed by the JobTracker. Each slave node has only one TaskTracker which maintains the track of task instances and each time it gives notification to the JobTracker about the status of the level of task implemented. Map and Reduce functions overall depend on implementing the appropriate interfaces and abstract classes by the applications. The client submits the job and configuration to the JobTracker which further redistributes the configuration to the slaves and then schedules tasks and oversee them. After this the job report is submitted to the client. The report consists of the status and diagnostic statistics of the tasks.

IV. BACKGROUND AND RELATED WORK
This paper proposed a metadata replication method to empower the Hadoop’s high availability by removing single point of failure (SPOF) in the existing Hadoop. Hadoop Distributed File System (HDFS), is a key component of Hadoop used used to store the input and output data for applications. Earlier phases in the implementation-

a) Initialization phase
Each slave node is registered as Active or primary node and its initial metadata or information like version file and file system image are lined up with existing active/primary node.

b) Replication phase
The runtime metadata like outstanding operations and lease states for failover in future are replicated.

c) Failover phase

Newly elected primary node overtakes all the communications. For this Hadoop make attempts of some methods to upgrade the availability of applications running on it. Various techniques were developed like keeping and retaining multiple replicas of application data and redeploying application tasks depending upon the failure, but it doesn’t provide high availability for itself. In the architecture of Hadoop, there exists SPOF (Single Point of Failure), due to which the whole system break down and become out service due to the failure of critical node. Therefore the SPOF is a big hazard to the opportunity of Hadoop.

V. PROPOSED WORK

In this new execution scenario of high two architecture for the topology of nodes in execution environment are discussed.

i) active-standby topology made up of one active critical node and one standby node.

ii) primary-slaves topology made up of primary critical node and several slave nodes. This new solution to generate high availability Hadoop involves major phases: Before going through the phases the meaning of metadata in Hadoop System must be cleared.

Metadata: Metadata is the most essential management statistical data or knowledge replicated in case of the Name node failover. Metadata carries two types of files: version file which contains the version information of running HDFS and file system image (fsimage) file which is a persistent checkpoint of the file system.

a) Initialization phase

Every standby/slave node is registered as Active or primary node and its initial metadata like version file and file system image are written into the system same as that of Earlier process. Initialization helps in synchronizing the metadata consistency between primary node and the slave nodes.

b) Replication phase

This is the core phase of the solution proposed here, the runtime metadata like outstanding operations, lease states that are the basic reasons for failover are replicated but also to save penalty it only replicates the metadata that are the most valued information instead of the whole data. While processing for high availability Hadoop, there are some challenges like-

a) Spotting the SPOF: the two critical components in Hadoop are Namenode and jobtracker, they are SPOF in Hadoop. Identifying them and determining the state information in not a easy going task.

b) Optimized overhead: to obtain high availability Hadoop, different nodes in the system must be synchronized which creates additional runtime cost for the system. That is why it is necessary to obtain a low costing overhead system.
c) **Adaptable configuration and layout**: for implementing high availability Hadoop, adjustable workloads are required so that they can work in various execution states without lacking in the performance at any stage, that is network bandwidth and the system latency arrangements should be versatile and in will to go with the changes in their configuration.

Some other challenges are like non sharing nature of Hadoop, the related information about various failures is neither shared between the same tasks in the job noer among different tasks. Actually HDFS is designed in such a way that it can be scalable as much as it can be so keeping this in mind and to avoid this additional load on the Name Node, much of the functionality like failure detection and recovery, is relegated to the compute nodes. One other upcoming issue of Hadoop is Heterogenity in clusters which is solved by Implementing MapReduce. Its framework is already discussed.

A vital boon of MapReduce is that it handles the failures undoubtedly and in a significant manner while keeping the complexity of fault tolerance out of sight of the programmer the. Whenever a node crashes, the task of MapReduce is to put its tasks on a different machine and also in case when a node is available but with the failing performance, situation is called straggler, MapReduce took this task (back up task) on another machine and completes the computation on the task.

Hadoop support FIFO mechanism for handling the jobs which is wholly unfair in case of multiple jobs in a cloud. Thus, an upright scheduling scheme is required so that all the running jobs can utilize the resource in the system. In case of the heterogeneous cloud, as each node has unconnected discrete capability and task in hands or the workload, the nodes having high capability or low load should wait for the nodes with low capability or high load before assimilating the intermediary results given by these nodes. This way, the total execution time is lengthened or extended. Therefore some other intelligent schedulers, are required to obtain achieve better resource utility which will provisionally improve the system overhead and will lower the running time of all the submitted jobs especially in the case of the heterogeneous cloud having many constraining resources. For example, the capacity scheduler in Hadoop, supports multiple queues and job priority, is more flexible and suitable for heterogeneous clouds with various job types. Best effort resource allocation should be employed. Hadoop is enclosed through many technologies that includes various networks, databases, operating systems, resource scheduling, load balancing, concurrency control and memory management and that ultimately leads to the growing security concerns. For example, the network which interconnects the systems in a Hadoop cluster needs to be fixed and reliable. Data security means the encryption of the data while ensuring for the suitable tactics for the data sharing. In addition to the optimal resource Sharing and memory management algorithms also needs to be secure and for this data mining Techniques can be applied.

Hadoop is very useful and there are many security concerns with it. In Hadoop data is stored in HDFS. This file system have no control on read and write operation. User can use the input files and analyse the results. All the jobs on the system are run as Hadoop user, which can execute the files without agreement or approval. For example, the user is having the controlled accessibility to the number of jobs that it can run, on any data set of the cluster. At client level, control list of accessibility checks should be performed at the start of any read or write when it should
VI. HADOOP BENCHMARKING

For Big Data, performance is an intrinsic part of storage and data retrieval within Hadoop. In addition to this, IBSG predicts that there will be 25 billion devices that will connect to the Internet by 2015 and approximately 50 billion by 2020 (Cisco Internet Business Solutions Group (IBSG) [6]). Thus, the information stored in such solutions will be statistically representative making the data analytics performed on it reliable. There are the principle guidelines that are used to examine the performance in contrast between the systems to distinguish between likely alternatives. When setting up a Hadoop cluster we always want to confirm if the cluster is precisely and accurately configured and this can be concluded by running the tasks linearly and then analyzing the results. For measuring the performance of the cluster, the preferred choice is to isolate the cluster and execute benchmarks so that we can evaluate the resources utilized and the cluster speed of the whole processing.

This section examine the various benchmarks that are used to test the performance of Hadoop clusters in real world scenarios. Hadoop enclosed many benchmark packages that encapsulated in Java Archive JAR files -

- **Test DFSIO**: helps in Testing the Distributed File System input and output performance evaluation of Hadoop Distributed File System.
- **Map Reduce sort**: This actually test the Map Reduce performance in the Hadoop by creating Map Reduce tasks to accomplish the initial partial sorting of input and then transfer the input by the shuffle. This test is performed in three steps: generation of the random data, sorting of the data, and validation of the result.
- **Grid mix suit**: It utilizes the mixture of the synthetic jobs, and assembles the realistic cluster workloads by resembling the all patterns obtained on real time systems in which the data is accessed.

<table>
<thead>
<tr>
<th>Internal factors</th>
<th>External factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Performance optimization parameters)</td>
<td>Environment</td>
</tr>
<tr>
<td>Number of maps</td>
<td>Number of cores</td>
</tr>
<tr>
<td>Number of reducers</td>
<td>Number of cores</td>
</tr>
</tbody>
</table>
Table 1: factors that affects the performance of Hadoop’s cluster

<table>
<thead>
<tr>
<th>Combiner</th>
<th>Memory size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Custom serialization,</td>
<td>The Network</td>
</tr>
<tr>
<td>Intermediate compression</td>
<td></td>
</tr>
</tbody>
</table>

Experimentally it is found that Hadoop cluster that runs in virtual scenario significantly perform low than the one that runs over the physical environment (physical machine). Other factors also affect the performance but in this study we focused on the environment effect which is 68% affecting the performance of the Hadoop cluster. A comparison chart is prepared between the execution of the Hadoop cluster in both the environments. Figure represents the difference in the performance between the clusters running in different environments.

![Physical vs. Virtual Performance](image)

From the above observation it can be concluded that virtual Hadoop cluster performance is notably lower than the cluster running on physical machine because of the overhead of the virtualization on the CPU of the physical host. The various factors like (RAM size, network bandwidth), were considered in our experiment. However, our aim was directed towards the effect of environmental aspects on the performance.

Whenever a Hadoop System is modeled and checked for various factors affecting the whole processing it is observed that more than 50% the external factors effect the system mainly the environment. Errors are also evaluated and always sum of the modeled errors should be equal to zero, as they are the difference between measured response and estimated response.
Errors are constantly independent and are normally distributed over the system. It is very difficult to find a specific distribution pattern for the errors against response which simply indicates that they are independent of them.

VII. CONCLUSION

In this paper it is tried to obtain issues that mainly comes while discussing the performance of HDFS on heterogeneous clusters. Performance degradation happens because of the heterogeneity. The problems concerned with the access control and ownership are discussed in the terms of security. Our conclusion is that the use of Hadoop for academic research is in the stage of adolescence. Making the use of Hadoop easy by enhancing the design are important research directions for future. It is evidenced by IBM’s 2011 Global study that the business intelligence and the data analytics are the main focus areas for the next five years (IBM [7]). Optimization techniques for HDFS discussed in this paper will definitely help in boosting up the efficiencies of the running algorithms in the background and hence motivate the use of this parallel computing paradigm. Hadoop is ultimately a “mission critical” at this platform and a productive environment where it is widely used and adopted.

REFERENCES

Book references

[1] Feng Wang, Bo Dong, Jie Qiu, Xinhu Li, Jie Yang, Ying Li “Hadoop High Availability through Metadata Replication” 2009 ACM 978-1-60558-802-5/09 Pg 37-44


Journal references


ON SEMI-INVARIANT SUBMANIFOLDS OF A NEARLY HYPERBOLIC COSYMPLECTIC MANIFOLD WITH SEMI-SYMMETRIC NON-METRIC CONNECTION

Toukeer Khan¹, Shadab Ahmad Khan², Mobin Ahmad³

¹Department of Mathematics, Integral University, Lucknow (India)
²Assistant Professor, Department of Mathematics, Integral University, Lucknow (India)
³Associate Professor, Department of Mathematics, Jazan University, Jazan (Saudi Arabia)

ABSTRACT
We consider a nearly hyperbolic cosymplectic manifold and study semi-invariant submanifolds of a nearly hyperbolic cosymplectic manifold admitting semi-symmetric non-metric connection. We also find the integrability conditions of some distributions on nearly hyperbolic cosymplectic manifold with semi-symmetric non-metric connection and study parallel distributions on them.

Keywords: Integrability Condition, Nearly Hyperbolic Cosymplectic Manifold, Parallel Distribution, Semi-Invariant Sub Manifolds, Semi-Symmetric Non-Metric Connection.

2000 AMS Mathematics Subject Classification: 53D05, 53D25, 53D12.

I. INTRODUCTION

In 1978, A. Bejancu [1] initiated the concept of CR-sub manifolds of a Kaehler manifold as generalization of invariant and anti-invariant sub manifolds. The extension of the concept of a CR-sub manifold of Kaehler manifold is a semi-invariant sub manifold to sub manifolds of an almost contact manifold. A semi-invariant sub manifold of a Sasakian manifold was initially studied by Bejancu - Papaghuic [2]. In 1983, K. Matsumoto [3] and Yano-Kon [4] studied the same concept under the name of contact CR-sub manifold. The study of semi-invariant sub manifolds in almost contact manifold was enriched by several geometers (see, [5], [6], [7], [8], [9], [10]). On the otherhand, Golab [11] introduced the idea of semi-symmetric and quarter-symmetric connection. Upadhyay and Dube [12] studied and define the almost hyperbolic (f, g, η, ξ)-structure. A semi-invariant sub manifolds of an almost r-contact hyperbolic metric manifolds was studied by Joshi and Dube [13]. Ahmad M. and Ali K., studied semi-invariant sub manifolds of a nearly hyperbolic cosymplectic manifold in [14].

Let be a linear connection in an n-dimensional differentiable manifold . The torsion tensor and curvature tensor of are given respectively by

\[ T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \]

\[ R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z \]

The connection \( \nabla \) is symmetric if its torsion tensor \( T \) vanishes, otherwise it is non-symmetric. The connection \( \nabla \) is metric connection if there is a Riemannian metric \( g \) in \( \tilde{M} \) such that \( \nabla g = 0 \), otherwise it is non-metric. It is well known that a linear connection is symmetric and metric if and only if it is the Levi-Civita connection.
Friedmann and J. A. Schouten [15] introduced the idea of a semi-symmetric connection. A linear connection is said to be a semi-symmetric connection if its torsion tensor $T$ is of the form:

$$T(X, Y) = \eta(Y)X - \eta(X)Y$$

Many geometers (see, [16], [17]) have studied properties of semi-symmetric non-metric connection.

In this paper, we study semi-invariant submanifolds of a nearly hyperbolic cosymplectic manifold with semi-symmetric non-metric connection.

This paper is organized as follows. In section 2, we give a brief description of nearly hyperbolic cosymplectic manifold. In section 3, we study some properties of semi invariant submanifolds of a nearly hyperbolic cosymplectic manifold with semi-symmetric non-metric connection. We also study parallel horizontal distribution on nearly hyperbolic Kenmotsu manifold with semi-symmetric non-metric connection. In section 4, we discuss the integrability conditions of some distributions on nearly hyperbolic cosymplectic manifold with semi-symmetric non-metric connection.

II. PRELIMINARIES

Let $\tilde{M}$ be an $n$-dimensional almost hyperbolic Contact metric manifold with the almost hyperbolic contact metric structure $(\mathcal{O}, \xi, \eta, g)$, where $\mathcal{O}$ is a tensor of type $(1,1)$, $\xi$ is a vector field called structure vector field and $\eta$ is the dual 1-form of $\xi$ and the associated Riemannian metric $g$ satisfying the following

$$\mathcal{O}^2X = X + \eta(X)\xi \quad \quad (2.1)$$
$$\eta(\xi) = -1, \quad g(X, \xi) = \eta(X) \quad \quad (2.2)$$
$$\mathcal{O}(\xi) = 0, \quad \eta\mathcal{O} = 0 \quad \quad (2.3)$$
$$g(\mathcal{O}X, \mathcal{O}Y) = -g(X, Y) - \eta(X)\eta(Y) \quad \quad (2.4)$$

de for any $X, Y$ tangent to $\tilde{M}$ [8]. In this case

$$g(\mathcal{O}X, Y) = -g(\mathcal{O}Y, X) \quad \quad (2.5)$$

An almost hyperbolic contact metric structure $(\mathcal{O}, \xi, \eta, g)$ on $\tilde{M}$ is called nearly hyperbolic cosymplectic manifold [8] if and only if

$$\nabla_X\mathcal{O}Y + \nabla_Y\mathcal{O}X = 0 \quad \quad (2.6)$$
$$\nabla_X\xi = 0 \quad \quad (2.7)$$

de for all $X, Y$ tangent to $\tilde{M}$, where $\nabla$ is Riemannian connection $\tilde{M}$.

Now, we define a semi-symmetric non-metric connection

$$\nabla_XY = \nabla_XY + \eta(Y)X \quad \quad (2.8)$$

such that

$$\langle \nabla_Xg(Y, Z) \rangle = -\eta(Y)g(X, Z) - \eta(Z)g(X, Y) \quad \quad (2.9)$$

From (2.6) & (2.8), replacing $Y$ by $\mathcal{O}Y$, we have

$$\nabla_X\mathcal{O}Y + (\nabla_Y\mathcal{O})X = -\eta(X)\mathcal{O}Y - \eta(Y)\mathcal{O}X \quad \quad (2.9)$$
$$\nabla_X\xi = -X \quad \quad (2.10)$$

An almost hyperbolic contact metric manifold with almost hyperbolic contact structure $(\mathcal{O}, \xi, \eta, g)$ is called nearly hyperbolic Cosymplectic manifold with semi-symmetric non-metric connection if it is satisfied (2.9) & (2.10).
III. SEMI-ININVARIANT SUBMANIFOLDS AND SOME BASIC RESULTS

Let $M$ be submanifold immersed in $\overline{M}$, we assume that the vector $\xi$ is tangent to $M$, denoted by $\{\xi\}$ the 1-dimensional distribution spanned by $\xi$ on $M$, then $M$ is called a semi-invariant submanifold [7] of $\overline{M}$ if there exist two differentiable distribution $D \& D^\perp$ on $M$ satisfying

(i) $TM = D \oplus D^\perp \oplus \xi$, where $D, D^\perp \& \xi$ are mutually orthogonal to each other.

(ii) The distribution $D$ is invariant under $\mathcal{O}$, i.e. $\mathcal{O}D_x = D_x$ for each $X \in M$.

(iii) The distribution $D^\perp$ is anti-invariant under $\mathcal{O}$, i.e. $\mathcal{O}D^\perp_x \subset T^\perp M$ for each $X \in M$, where $TM \& T^\perp M$ be the Lie algebra of vector fields tangential & normal to $M$ respectively.

Let Riemannian metric $g$ and $\nabla$ be induced Levi-Civita connection on $M$ then the Gauss formula & Weingarten formula are given by

\[ g(A_y X, Y) = g(h(X, Y), N) \]  \hspace{1cm} (3.3)

Any vector $X$ tangent to $M$ is given as

\[ X = PX + QX + \eta(X) \xi \]  \hspace{1cm} (3.4)

where $PX \in D \& QX \in D^\perp$.

Similarly, for $N$ normal to $M$, we have

\[ \mathcal{O}N = BN + CN \]  \hspace{1cm} (3.5)

where $BN$ (resp. $CN$) is tangential component (resp. normal component) of $\mathcal{O}N$.

Using the semi-symmetric non-metric connection the Nijenhuis tensor is expressed as

\[ N(X, Y) = (\nabla_X \mathcal{O}Y - (\nabla_Y \mathcal{O})X - \mathcal{O}(\nabla_X \mathcal{O})Y) + \mathcal{O}(\nabla_Y \mathcal{O})X \]  \hspace{1cm} (3.6)

Now from (2.9) replacing $X$ by $\mathcal{O}X$, we have

\[ (\nabla_X \mathcal{O}Y = -\eta(Y)X + \eta(X)\eta(Y)\xi - (\nabla_Y \mathcal{O})X \]  \hspace{1cm} (3.7)

Differentiating (2.1) conveniently along the vector and using (2.10), we have

\[ (\nabla_Y \mathcal{O})X = (\nabla_Y \eta)(X) \xi - \eta(X)Y - \mathcal{O}(\nabla_Y \mathcal{O})X \]  \hspace{1cm} (3.8)

From (3.7) \& (3.8), we have

\[ (\nabla_X \mathcal{O}Y = \eta(X)Y - \eta(Y)X - (\nabla_Y \eta)(X)\xi - \eta(X)\eta(Y)\xi + \mathcal{O}(\nabla_Y \mathcal{O})X \]  \hspace{1cm} (3.9)

Interchanging $X \& Y$, we have

\[ (\nabla_X \mathcal{O}Y = \eta(Y)X - \eta(X)Y - (\nabla_Y \eta)(Y)\xi - \eta(X)\eta(Y)\xi + \mathcal{O}(\nabla_Y \mathcal{O})Y \]  \hspace{1cm} (3.10)

Using equation (3.9), (3.10) and (2.9) in (3.6), we have

\[ N(X, Y) = 4\eta(X)Y + 2g(\mathcal{O}X, Y)\xi + 4\eta(X)\eta(Y)\xi + 4\mathcal{O}(\nabla_Y \mathcal{O})X \]  \hspace{1cm} (3.11)

As we know, \[ (\nabla_Y \mathcal{O})X = \nabla_Y \mathcal{O}X - \mathcal{O}(\nabla_Y X) \]

Using Guass formula (3.1), we have

\[ \mathcal{O}(\nabla_Y \mathcal{O})X = \mathcal{O}(\nabla_Y \mathcal{O}X) + \mathcal{O}h(Y, \mathcal{O}X) - \nabla_Y X - \eta(\nabla_Y X)\xi - h(Y, X) \]  \hspace{1cm} (3.12)

Using equation (3.12) in (3.11), we have
Lemma 3.1. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then

$$2(\tilde{\nabla}_Y \phi)Y = \nabla_X \phi Y - \nabla_Y \phi X + h(Y, \phi Y) - h(Y, \phi X) - \phi [X, Y]$$

for all $X, Y \in D$.

Proof. By Gauss formula (3.1), we have

$$\nabla_X \phi Y - \nabla_Y \phi X = \nabla_X \phi Y - \nabla_Y \phi X + h(Y, \phi Y) - h(Y, \phi X)$$

(3.14)

Also, by covariant differentiation, we know that

$$\nabla_X \phi Y - \nabla_Y \phi X = (\tilde{\nabla}_X \phi)Y - (\tilde{\nabla}_Y \phi)X + \phi [X, Y]$$

(3.15)

From (3.14) and (3.15), we have

$$(\nabla_X \phi Y)' - (\nabla_Y \phi)X = \nabla_X \phi Y - \nabla_Y \phi X + h(Y, \phi Y) - h(Y, \phi X) - \phi [X, Y]$$

(3.16)

Adding (2.9) and (3.16), we obtain

$$2(\tilde{\nabla}_X \phi)Y = \nabla_X \phi Y - \nabla_Y \phi X + h(Y, \phi Y) - h(Y, \phi X) - \phi [X, Y]$$

for all $X, Y \in D$.

Hence lemma is proved.

Lemma 3.2. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then

$$2(\tilde{\nabla}_Y \phi)X = \nabla_Y \phi X - \nabla_X \phi Y + h(Y, \phi Y) + h(X, \phi X) + \phi [X, Y]$$

for all $X, Y \in D$.

Proof. Using Weingarten formula (3.2), we have

$$\nabla_X \phi Y - \nabla_Y \phi X = A_{\phi X}Y - A_{\phi Y}X + \nabla^X_X \phi Y - \nabla^X_Y \phi X$$

(3.17)

Comparing equation (3.15) & (3.17), we have

$$\nabla_X \phi Y - (\nabla_Y \phi)X = A_{\phi X}Y - A_{\phi Y}X + \nabla^X_X \phi Y - \nabla^X_Y \phi X - \phi [X, Y]$$

(3.18)

Adding (2.9) & (3.18), we have

$$2(\tilde{\nabla}_X \phi)Y = A_{\phi X}Y - A_{\phi Y}X + \nabla^X_X \phi Y - \nabla^X_Y \phi X - \phi [X, Y]$$

for all $X, Y \in D^\perp$.

Hence lemma is proved.

Lemma 3.3. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then

$$2(\tilde{\nabla}_X \phi)Y = A_{\phi X}Y - A_{\phi Y}X + \nabla^X_X \phi Y - \nabla^X_Y \phi X - \phi [X, Y]$$

for all $X, Y \in D^\perp$.
Lemma 3.5. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then

$$2(\overline{\nabla}_X\phi)Y = -A_{gY}X + \nabla^X_H\phi Y - \nabla_Y\phi X - h(Y, \phi X) - \phi[X, Y]$$

for all $X \in D$ and $Y \in D^\perp$.

Proof. By Gauss formulas (3.1) and Weingarten formula (3.2), we have

$$\nabla^X_H\phi Y - \nabla_Y\phi X = -A_{gY}X + \nabla^X_H\phi Y - \nabla_Y\phi X - h(Y, \phi X) \quad (3.19)$$

Comparing equation (3.15) and (3.19), we have

$$(\nabla^X_H\phi)Y - \nabla^X_H\phi X = -A_{gY}X + \nabla^X_H\phi Y - \nabla_Y\phi X - h(Y, \phi X) - \phi[X, Y] \quad (3.20)$$

Adding equation (2.9) & (3.20), we get

$$2(\overline{\nabla}_X\phi)Y = -A_{gY}X + \nabla^X_H\phi Y - \nabla_Y\phi X - h(Y, \phi X) - \phi[X, Y]$$

for all $X \in D$ and $Y \in D^\perp$.

Hence lemma is proved.

Lemma 3.6. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then

$$2(\overline{\nabla}_Y\phi)X = A_{gY}X - \nabla^X_H\phi Y + \nabla_Y\phi X + h(Y, \phi X) + \phi[X, Y]$$

for all $X \in D$ and $Y \in D^\perp$.

Lemma 3.7. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric semi-metric connection, then

$$(3.21)$$

$$(3.22)$$

$$(3.23)$$

$$(3.24)$$

for all $X, Y \in TM$.

Proof. Differentiating covariantly equation (3.4) and using equation (3.1) and (3.2), we have

$$(\overline{\nabla}_X\phi)Y + \phi(\overline{\nabla}_X Y) + \phi(h(X, Y)) = \nabla^X_H\phi Y + h(X, \phi Y) - A_{qY}X + \nabla^X_H\phi QY \quad (3.25)$$

Interchanging $X \& Y$, we have

$$(\overline{\nabla}_Y\phi)X + \phi(\overline{\nabla}_Y X) + \phi(h(Y, X)) = \nabla^X_H\phi X + h(Y, \phi X) - A_{qX}Y + \nabla^X_H\phi QX \quad (3.26)$$

Adding equations (3.25) & (3.26), we have

$$(\overline{\nabla}_X\phi)Y + (\overline{\nabla}_Y\phi)X + \phi(\overline{\nabla}_X Y) + \phi(\overline{\nabla}_Y X) + 2\phi(h(X, Y)) = \nabla^X_H\phi Y + \nabla^Y_H\phi X + h(X, \phi Y) + h(Y, \phi X) - A_{qX}Y + \nabla^X_H\phi QX + \nabla^Y_H\phi QX \quad (3.27)$$

By Virtue of (2.9) & (3.27), we have

$$-\phi(h(X, Y)) - \phi(h(Y, X)) + \phi(\overline{\nabla}_X Y) + \phi(\overline{\nabla}_Y X) + 2\phi(h(X, Y)) = \nabla^X_H\phi Y + \nabla^Y_H\phi X + h(X, \phi Y) + h(Y, \phi X) - A_{qX}Y + \nabla^X_H\phi QX + \nabla^Y_H\phi QX$$
Using equations (3.4), (3.5) & (2.3), we have
\[-\eta(X)\partial PY - \eta(X)\partial QY - \eta(Y)\partial PX - \eta(Y)\partial QX + \partial P(\nabla_X Y) + \partial Q(\nabla_X Y) + \partial P(\nabla_Y X)
+ \partial Q(\nabla_Y X) + 2Bh(X, Y) + 2Ch(X, Y) = P(\nabla_X \partial PY) + Q(\nabla_X \partial QY) + \eta(\nabla_X \partial QY) \xi
+ P(\nabla_Y \partial PX) + Q(\nabla_Y \partial QX) + \eta(\nabla_Y \partial PX) \xi + h(X, \partial PX) + h(Y, \partial PY) - PA_{\partial QY} X
- QA_{\partial QY} X - \eta(A_{\partial QY}) \xi - PA_{\partial QX} Y - QA_{\partial QX} Y - \eta(A_{\partial QX}) \xi + \nabla_X \partial QY + \nabla_Y \partial QX\]

Comparing horizontal, vertical and normal components we get desired results.

Hence lemma is proved.

**Definition 3.8.** The horizontal distribution \( D \) is said to be parallel \([2]\) on \( M \) if \( \nabla_X Y \in D \), for all \( X, Y \in D \).

**Theorem 3.9.** Let \( M \) be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold \( \bar{M} \) with semi-symmetric non-metric connection. If horizontal distribution \( D \) is parallel, then
\[h(X, \partial Y) = h(Y, \partial X)\]
for all \( X, Y \in D \).

**Proof.** Let \( X, Y \in D \), as \( D \) is parallel distribution, then
\[\nabla_X \partial Y \in D \cap \nabla_Y \partial X \in D\]

Then, from (3.22) and (3.23), we have
\[Q(\nabla_X \partial PY) + Q(\nabla_Y \partial PX) - QA_{\partial QY} X - QA_{\partial QX} Y + h(X, \partial PY) + h(Y, \partial PX) + \nabla_Y \partial QY
+ \nabla_X \partial QX = -\eta(X)\partial QY - \eta(Y)\partial QX + \partial Q(\nabla_X Y) + \partial Q(\nabla_Y X) + 2Bh(X, Y) + 2Ch(X, Y)\]

As \( Q \) being a projection operator on \( D^+ \) then we have
\[h(X, \partial Y) + h(Y, \partial X) = 2\partial h(X, Y)\] \hspace{1cm} (3.28)
Replacing \( X \) by \( \partial X \) in (3.28) & using (2.1), we have
\[h(\partial X, \partial Y) + h(Y, X) = 2\partial h(\partial X, Y)\] \hspace{1cm} (3.29)
Replacing \( Y \) by \( \partial Y \) & using (2.1) in (3.28), we have
\[h(X, Y) + h(\partial Y, \partial X) = 2\partial h(X, \partial Y)\] \hspace{1cm} (3.30)
By virtue of (3.29) and (3.30), we have
\[h(X, \partial Y) = h(Y, \partial X)\]
for all \( X, Y \in D \).

Hence theorem is proved.

**Definition 3.10.** A semi-invariant submanifold is said to be mixed totally geodesic \([2]\) if \( h(X, Y) = 0 \), for all \( X \in D \) and \( Y \in D^+ \).

**Theorem 3.11.** Let \( M \) be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold \( \bar{M} \) with semi-symmetric non-metric connection. Then \( M \) is a mixed totally geodesic if and only if
\[A_{\partial Y} X \in D \quad \text{for all} \quad X \in D.\]

**Proof.** Let \( A_{\partial Y} X \in D \) for all \( X \in D \).

Now,
\[g(h(X, Y), N) = g(A_{\partial Y} X, Y) = 0, \quad \text{for} \quad Y \in D^+.\]

Which is equivalent to \( h(X, Y) = 0 \).
Hence $M$ is totally mixed geodesic.

Conversely, let $M$ is totally mixed geodesic.

That is $h(X, Y) = 0$ for $X \in D$ and $Y \in D^\perp$.

Now, $g(h(X, Y), N) = g(A_\eta X, Y)$.

This implies that $g(A_\eta X, Y) = 0$.

Consequently, we have $A_\eta X \in D$, for all $Y \in D^\perp$.

Hence theorem is proved.

IV. INTEGRABILITY OF DISTRIBUTION

Theorem 4.1. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then the distribution $D \oplus \{\xi\}$ is integrable if

$$h(X, \partial Z) = h(\partial X, Z)$$

for each $X, Y, Z \in (D \oplus \{\xi\})$.

Proof. The torsion tensor $S(X, Y)$ of an almost hyperbolic contact manifold is given by

$$S(X, Y) = N(X, Y) + 2\phi_\eta(X, Y)\xi$$

Where $N(X, Y)$ is Neijenhuis tensor.

If $(D \oplus \{\xi\})$ is integrable,

then $N(X, Y) = 0$, for any $X, Y \in (D \oplus \{\xi\})$.

Hence from (3.13), we have

$$4\phi_\eta(X, Y) + 4\phi_\eta(X)\eta(Y)\xi + 4\phi_\partial(X, \partial Y) + 4\phi_\partial(\partial Y, \partial X) - 4(\nabla Y, X) - 4(\nabla Y, X)\xi - 4\phi(Y, X)$$

Comparing normal part both side of (4.2), we have

$$\phi_\partial(\partial Z, \partial X) - h(\partial Z, X) + C\phi(Y, \partial X) = 0$$

for $X, Y \in (D \oplus \{\xi\})$.

Replacing $Y$ by $\partial Z$, where $Z \in D$ in (4.3), we have

$$\phi_\partial(\partial Z, \partial X) - h(\partial Z, X) + C\phi(\partial Z, \partial X) = 0$$

Interchanging $X$ and $Z$, we have

$$\phi_\partial(\partial Z, \partial X) - h(\partial Z, X) + C\phi(\partial Z, \partial X) = 0$$

Subtracting (4.4) from (4.5), we obtain

$$\phi_\partial(\partial Z, \partial X) - h(\partial Z, X) + h(\partial Z, X) = 0$$

Since $(D \oplus \{\xi\})$ is integrable,

So that $[\partial X, \partial Z] \in (D \oplus \{\xi\})$, for $X, Z \in D$.

Consequently, (4.6) gives

$$h(\partial Z, X) = h(\partial Z, X)$$

for each $X, Y, Z \in (D \oplus \{\xi\})$.

Hence theorem is proved.
Theorem 4.2. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then
\[ A_{\tilde{g}Y}Z - A_{\tilde{g}Z}Y = \frac{1}{3} \partial P [Y, Z] \]
for each $Y, Z \in D^\perp$.

Proof. Let $Y, Z \in D^\perp$ and $X \in TM$, from (3.3), we have
\[ 2g(A_{\tilde{g}Z}Y, X) = g(h(Y, X), \partial Z) + g(h(X, Y), \partial Z) \quad (4.7) \]
Using (2.9) & (3.1) in (4.7), we have
\[ 2g(A_{\tilde{g}Z}Y, X) = -g(\nabla_Y \partial X, Z) - g(\nabla_X \partial Y, Z) - \eta(X)g(\partial Y, Z) - \eta(Y)g(\partial X, Z) \quad (4.8) \]
From (3.2), we have
\[ \nabla_X N = -A_N X + \nabla_Z N \]
Replacing $N$ by $\partial Y$
\[ \nabla_X \partial Y = -A_{\tilde{g}Y} X + \nabla_Z \partial Y \]
As $\nabla$ is a Levi-Civita connection, using above, then from (4.8), we have
\[ 2g(A_{\tilde{g}Z}Y, X) = -g(\nabla_Y \partial Z, X) + g(A_{\tilde{g}Y} Z, X) \quad (4.9) \]
Transvecting $X$ from both sides from (4.9), we obtain
\[ 2A_{\tilde{g}Z} Y = \nabla_Y \partial Z + A_{\tilde{g}Y} Z \quad (4.10) \]
Interchanging $Y$ & $Z$, we have
\[ 2A_{\tilde{g}Z} Z = \nabla_Z \partial Y + A_{\tilde{g}Z} Y \quad (4.11) \]
Subtracting (4.10) from (4.11), we have
\[ (A_{\tilde{g}Y}Z - A_{\tilde{g}Z}Y) = \frac{1}{3} \partial P [Y, Z] \]
Comparing the tangential part both side in above equation, we have
\[ (A_{\tilde{g}Y}Z - A_{\tilde{g}Z}Y) = \frac{1}{3} \partial P [Y, Z] \]
where $[Y, Z]$ is Lie Bracket.

Hence theorem is proved.

Theorem 4.3. Let $M$ be a semi-invariant submanifold of a nearly hyperbolic cosymplectic manifold $\tilde{M}$ with semi-symmetric non-metric connection, then the distribution is Integrable if and only if
\[ A_{\tilde{g}Y}Z - A_{\tilde{g}Z}Y = 0 \quad (4.12) \]
for all $Y, Z \in D^\perp$.

Proof. Suppose that the distribution $D^\perp$ is integrable, that is $[Y, Z] \in D^\perp$
For any $Y, Z \in D^\perp$, therefore $P[Y, Z] = 0$.
Consequently, from (4.11) we have
\[ A_{\tilde{g}Y}Z - A_{\tilde{g}Z}Y = 0 \]
Conversely, let (4.12) holds. Then by virtue of (4.11), we have
\[ \partial P [Y, Z] = 0 \]
For all $Y, Z \in D^\perp$. Since $\text{rank} \; \partial = 2n$
Therefore, either $P[Y,Z] = 0$ or $P[Y,Z] = k\xi$.

But $P[Y,Z] = k\xi$ is not possible as $P$ being a projection operator on $D$.

So

$$P[Y,Z] = 0,$$

This implies that $[Y,Z] \in D^\perp$, for all $Y, Z \in D^\perp$.

Hence $D^\perp$ is integrable.

Hence theorem is proved.

REFERENCES

[16] AHMAD, M. and JUN, J.B., ‘On Semi-invariant submanifold of nearly Kenmotsu manifold with semi-
257-266.

manifold endowed with a semi-symmetric non-metric connection’, *Journal of Applied Analysis, Vol. 17,
No. 1*, (2011), *pp 119-130.*

**Biographical Notes**

Mr. Toukeer Khan is working as a Lecturer in Department of Mathematics, Integral University, Lucknow and
presently pursuing in Ph.D from Integral University, Lucknow, India.

Dr. Shadab Ahmad khan is working as Assistant Professor in Department of Mathematics, Integral University,
Lucknow, India.

Dr. Mobin Ahmad is working as Associate Professor in Department of Mathematics, Faculty of Science, Jazan
University, Jazan-2069, Saudi Arabia.
SOME SUFFICIENT CONDITIONS FOR POISSON DISTRIBUTION SERIES ASSOCIATED WITH CONIC REGIONS

Divesh Srivastava$^1$, Saurabh Porwal$^2$

$^1$Department of Mathematics, Integral University, Lucknow, (India)
$^2$Department of Mathematics, UIET, CSJM University, Kanpur, (India)

ABSTRACT
The purpose of the present paper is to investigate some sufficient conditions for the convolution operator $I(m)f(z)$ belonging to the classes $k−UCV(\alpha)$, $k−Sp(\alpha)$, $S^*(\lambda)$ and $C(\lambda)$.

Keywords: Convex Functions, Harmonic Functions, Poisson Distribution Series, Starlike Function, Univalent Functions.

2010 Mathematics Subject Classification: 30C45.

I. INTRODUCTION

Let $A$ denote the class of functions $f(z)$ of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n \quad (1.1)$$

which are analytic in the open unit disc $U = \{z : z \in \mathbb{C} \text{ and } |z| < 1\}$ and satisfy the normalization condition $f(0) = f'(0) − 1 = 0$. Further, we denote by $S$ the subclass of $A$ consisting of functions of the form (1.1) which are also univalent in $U$. A function $f$ of the form (1.1) is said to be starlike of order $\alpha$ if it satisfies the following condition

$$\text{Re} \left\{ \frac{zf'(z)}{f(z)} \right\} > \alpha, \quad z \in U.$$ 

and is said to be convex of order $\alpha$ if it satisfies the following condition

$$\text{Re} \left\{ 1 + \frac{zf''(z)}{f'(z)} \right\} > \alpha, \quad z \in U.$$ 

The classes of all starlike and convex functions of order $\alpha$ are denoted by $S^*(\alpha)$ and $C(\alpha)$, respectively, studied by Robertson [16], (see also [19]).

Bharti et al. [1] introduced the subclasses of $k$-uniformly convex functions of order $\alpha$ and corresponding class of starlike functions as follows

If $f \in A$, $0 \leq k < \infty$ and $0 \leq \alpha < 1$ then $f \in k−UCV(\alpha)$, if and only if

$$\text{Re} \left\{ 1 + \frac{zf''(z)}{f'(z)} \right\} \geq k \left| \frac{zf''(z)}{f'(z)} \right| + \alpha \quad (1.2)$$

For $\alpha = 0$ the class $k−UCV(\alpha)$ reduce to the class $k−UCV$ introduced and studied by Kanas and Wisniowska [7] and for $k = 1, \alpha = 0$ it reduce to the class uniformly convex functions $UCV$ studied by Goodman [4]. Using the Alexander transform we can obtain the class $k−Sp(\alpha)$ in the following way $f \in k−UCV(\alpha) \Leftrightarrow zf^\alpha \in k−Sp(\alpha)$. 


For more results on these directions we refer the reader to ([2], [5], [6], [8], [9], [17], [20]) and references therein. A function $f \in A$ is said to be in the class $P^{\gamma}_{\beta}(\beta)$ if it satisfies the following inequality
\[
\left| \frac{(1 - \gamma) f(z)}{z} + \gamma f'(z) - 1 \right| < 1
\]
where $0 \leq \gamma < 1$, $\beta < 1$, $\tau \in \mathbb{C}/\{0\}$ and $z \in U$. The class $P^{\gamma}_{\beta}(\beta)$ was introduced by Swaminathan [21].

Next, we introduce the classes $S^{\lambda}_{\ast}$ and $C^{\lambda}$ as follows
\[
S^{\lambda}_{\ast} = \left\{ f \in A : \left| \frac{zf'(z)}{f(z)} - 1 \right| < \lambda, \ (z \in U, \lambda > 0) \right\}
\]
\[
C^{\lambda} = \left\{ f \in A : \left| \frac{zf''(z)}{f'(z)} \right| < \lambda, \ (z \in U, \lambda > 0) \right\}.
\]
From (1.3) and (1.4) it is easy to see that
\[
f(z) \in C^{\lambda} \iff zf'(z) \in S^{\lambda}_{\ast}, \ (\lambda > 0)
\]
The classes $S^{\lambda}_{\ast}$ and $C^{\lambda}$ were introduced by Ponnusamy and Rønning [10].

Very recently, Porwal [12] introduce a power series whose coefficients are probabilities of Poisson distribution
\[
K(m, z) = z + \sum_{n=0}^{\infty} \frac{m^n}{(n-1)!} e^{-m} z^n
\]
By ratio test the radius of convergence of above series is infinity. Using the above series they obtain some interesting results on certain classes of analytic univalent functions.

The convolution (or Hadamard product) of two series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{n=0}^{\infty} b_n z^n$ is defined as the power series
\[
(f \ast g)(z) = \sum_{n=0}^{\infty} a_n b_n z^n.
\]
Now, we consider the linear operator $I(m) : A \rightarrow A$ defined by
\[
I(m)f = K(m, z) \ast f(z) = z + \sum_{n=0}^{\infty} \frac{m^n}{(n-1)!} e^{-m} a_n z^n
\]
In the present paper, motivated by results of [12] and on connections between various subclasses of analytic univalent functions by using hypergeometric functions (see [3], [10], [14] , [18]) and by work of ([11], [13], [15]) we establish some sufficient conditions for convolution operator $I(m)f(z)$ belonging to the classes $k-UCV^{\alpha}(a)$, $k-S^{\mu}_{\lambda}(a)$, $C^{\lambda}$ and $S^{\lambda}_{\ast}$.

II. MAIN RESULTS

To establish our main results, we shall require the following lemmas.

**Lemma 2.1.** ([1]) A function $f \in A$ is in $k-UCV^{\alpha}(a)$ if it satisfies the following condition
\[
\sum_{n=3}^{\infty} n(n + k) - (k + \alpha)|a_n| \leq 1 - \alpha \quad (2.1)
\]
**Remark 1.** It was also found that the condition (2.1) is necessary if $f \in A$ is of the form
\[
f(z) = z - \sum_{n=2}^{\infty} a_n z^n, \ a_n \geq 0 \quad (2.2)
\]
Lemma 2.2. ([1]) A function \( f \in A \) is in \( k - S_p(\alpha) \) if it satisfies the following inequality
\[
\sum_{n=0}^{\infty} |n(1 + k) - (k + \alpha)| |a_n| \leq 1 - \alpha. \quad (2.3)
\]
The condition (2.3) is necessary for functions of the form (2.2).

Another sufficient condition is also given for the class \( k - UCV(\alpha) \) in [8] which is given by the following way.

Lemma 2.3. ([7]) Let \( f \in S \) and have the form (1.1). If for some \( k, 0 \leq k < \infty \), the inequality
\[
\sum_{n=0}^{\infty} n(n - 1)|a_n| \leq \frac{1}{(k + 2)}
\]
holds, then \( f \in k - UCV \). The number \( 1/k + 2 \) can not be increased.

Lemma 2.4. Let \( f \in A \) be of the form (1.1). If
\[
\sum_{n=0}^{\infty} (\lambda + n - 1)|a_n| \leq \lambda, \quad (\lambda > 0), \quad (2.4)
\]
then \( f \in S_{\lambda}^* \).

We further note that when \( f(z) \) is of the form (2.2), the condition (2.4) is both necessary and sufficient for \( f \in S_{\lambda}^* \).

Lemma 2.5. Let \( f \in A \) be of the form (1.1). If
\[
\sum_{n=0}^{\infty} n(\lambda + n - 1)|a_n| \leq \lambda, \quad (\lambda > 0), \quad (2.5)
\]
then \( f \in C_{\lambda} \).

Lemma 2.6. If \( f \in P_{\gamma}(\beta) \) is of the form (1.1) then
\[
|a_n| \leq \frac{2|\tau|(1 - \beta)}{1 + \gamma(n - 1)}
\]

Theorem 2.1. Let \( f \in A \) be defined as in (1.1) suppose that \( m > 0, k \geq 0, 0 \leq \alpha < 1 \) and the inequality
\[
(k + 1)m + (1 - \alpha) - e^{-m} \leq \frac{\gamma(1 - \alpha)}{2|\tau|(1 - \beta)}
\]
is satisfied then for \( f \in P_{\gamma}(\beta), 0 < \gamma \leq 1 \) and \( 0 \leq \beta < 1 \), \( I(m)f(z) \in k - UCV(\alpha) \).

Proof. Since
\[
I(m)f(z) = z + \sum_{n=0}^{\infty} \frac{m^{n-1}}{(n-1)!} e^{-m} a_n z^n.
\]
To prove that \( I(m)f(z) \in k - UCV(\alpha) \), from Lemma 2.1, it is sufficient to show that
\[
\sum_{n=0}^{\infty} n[n(1 + k) - (k + \alpha)]|A_n| \leq 1 - \alpha \quad (2.6)
\]
where
\[
A_n = \frac{m^{n-1}}{(n-1)!} e^{-m} a_n, \quad n \geq 2
\]
Now
\[
\sum_{n=2}^{\infty} n[n(1 + k) - (k + \alpha)] \frac{m^{n-1}}{(n-1)!} e^{-m} |a_n| \leq 2|\tau|(1 - \beta) \sum_{n=2}^{\infty} n[n(1 + k) - (k + \alpha)] \frac{m^{n-1}}{(n-1)!} e^{-m} \frac{1}{1 + \gamma(n - 1)} \quad (\text{using Lemma } 2.6)
\]
\[ \leq 2|\tau|(1-\beta)e^{-m}\sum_{n=2}^{\infty} (n-1)^{-1} \alpha \frac{m^{n-1}}{(n-1)!}. \]  
\begin{align*}
\text{(Since } 1 + \gamma (n-1) \geq \gamma n) \\
= \frac{2|\tau|(1-\beta)e^{-m}}{\gamma} \sum_{n=2}^{\infty} \left( n-1 \right) \frac{m^{n-1}}{(n-1)!} \\
= \frac{2|\tau|(1-\beta)e^{-m}}{\gamma} \left[ (k+1) \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-2)!} + (1-\alpha) \sum_{n=2}^{\infty} \frac{m^{n-1}}{(n-1)!} \right] \\
= \frac{2|\tau|(1-\beta)e^{-m}}{\gamma} \left[ (k+1)m + (1-\alpha)(e^m - 1) \right] \\
\leq 1 - \alpha \\
\text{by the given hypothesis.}
\end{align*}

This completes the proof of Theorem 2.1.

**Theorem 2.2.** Let \( f \in A \) be defined as in (1.1) suppose that \( m > 0, k \geq 0, \alpha < 1 \) and the inequality
\[ (k+1)(1-e^{-m}) - \frac{(k+\alpha)}{\alpha} (1-e^{-m} - me^{-m}) \leq \frac{\gamma(1-\alpha)}{2|\tau|(1-\beta)} \]
is satisfied then for \( \beta < 1 \) and \( \gamma < 1 \), \( I(m)f(z) \in k + S_\beta(\alpha) \).

**Proof.** The proof of this theorem is much akin to that of Theorem 2.1 so we omit the details involved.

**Theorem 2.3.** Let \( m > 0 \) be such that
\[ \frac{2|\tau|(1-\beta)e^{-m}}{\gamma} \left[ 1 - e^{-m} + \frac{1}{\gamma} (1-\lambda) - me^{-m} \right] \leq \lambda \]
is satisfied then for \( f \in P_{\gamma}^\beta(\beta) \), \( 0 < \gamma, \beta \leq 1 \) and \( \lambda > 0 \), \( I(m)f(z) \in S_\lambda^* \).

**Proof.** To prove that \( I(m)f(z) \in S_\lambda^* \), from Lemma 2.4 it is sufficient to prove that
\[ \sum_{n=3}^{\infty} (n-1)A_n \leq \lambda \]
where
\[ A_n = \frac{m^{n-1}}{(n-1)!} e^{-m}a_n, \quad n \geq 2 \]
Since \( f \in P_{\gamma}^\beta(\beta) \) using Lemma 2.6 and \( 1 + \gamma(n-1) \geq \gamma n \) we need only to show that
\[ \sum_{n=3}^{\infty} (n-1)A_n \leq \lambda \]
Now
\[
\sum_{n=2}^{\infty} \frac{(n + \lambda - 1)}{(n - 1)!} \frac{m^{n-1}}{e^{m}} |a_n|
\leq \sum_{n=2}^{\infty} \frac{(n + \lambda - 1)}{(n - 1)!} \frac{m^{n-1}}{e^{m}} \frac{2|\tau|(1 - \beta)}{1 + \gamma(n - 1)}
\leq \frac{2|\tau|(1 - \beta)}{\gamma} \sum_{n=2}^{\infty} \frac{(n + \lambda - 1)}{(n - 1)!} \frac{m^{n-1}}{e^{m}}
= \frac{2|\tau|(1 - \beta)}{\gamma} e^{-m} \left[ \sum_{n=3}^{\infty} \frac{m^{n-1}}{(n - 1)!} + \frac{(\lambda - 1)}{m} \sum_{n=2}^{\infty} \frac{m^n}{n!} \right]
= \frac{2|\tau|(1 - \beta)}{\gamma} \left[ (1 - e^{-m}) + \frac{(\lambda - 1)}{m} (1 - e^{-m} - me^{-m}) \right] < \lambda
\]
by the given hypothesis.

Thus the proof of Theorem 2.3 is complete.

**Theorem 2.4.** Let \( m > 0 \) and the inequality
\[
\frac{2|\tau|(1 - \beta)}{\gamma} [m + \lambda - 1 - e^{-m}] \leq \lambda
\]
is satisfied then for \( f \in P_{\gamma}(\beta); \ 0 < \gamma \leq 1, \ \beta < 1 \) and \( \lambda > 0, I(m)f(z) \in C_\lambda. \)

**Proof.** The proof is similar to that of Theorem 2.3 therefore we omit the details.

**Theorem 2.5.** Let \( m > 0 \). If for \( k \geq 0, \ 0 \leq \alpha < 1 \) and the inequality
\[
e^{m} \left[ (1 + k)m^3 + (6 + 5k - \alpha)m^2 + (7 + 4k - 3\alpha)m \right] \leq 1 - \alpha
\]
is satisfied then \( I(m)f(z) \) maps \( f(z) \in S \) of the form (1.1) into \( k - UC\mathcal{V} \).\( \alpha \).\( \)

**Proof.** Let \( f(z) \in S \) be of the form (1.1). In view of Lemma 2.1 it is enough to show that
\[
T = \sum_{n=3}^{\infty} \frac{n(n(k + 1) - (k + \alpha))m^{n-1}}{(n - 1)!} e^{-m} |a_n| \leq 1 - \alpha.
\]

Now
\[
T = \sum_{n=2}^{\infty} \frac{n(n(k + 1) - (k + \alpha))m^{n-1}}{(n - 1)!} e^{-m} |a_n|
\leq \sum_{n=2}^{\infty} n^2 |a(k + 1) - (k + \alpha)| \frac{m^{n-1}}{(n - 1)!} e^{-m}
= \sum_{n=2}^{\infty} [(k + 1)(n - 1)(n - 3) + (5k + 6 - \alpha)(n - 1)(n - 2)
+ (4k + 7 - 3\alpha)(n - 1) + (1 - \alpha)] \frac{m^{n-1}}{(n - 1)!} e^{-m}
= (1 + k)n^3 + (6 + 5k - \alpha)m^2 + (7 + 4k - 3\alpha)m + (1 - \alpha)(1 - e^{-m}) \leq 1 - \alpha
\]
by the given hypothesis.

Thus the proof of Theorem 2.5 is established.

**Theorem 2.6.** Let \( m > 0 \) if for \( k \geq 0, \ 0 \leq \alpha < 1 \) and the inequality
\[
e^{m} \left[ (1 + k)m^3 + (3 + 2k - \alpha)m \right] \leq 1 - \alpha. \ (2.7)
\]
then \( I(m)f(z) \) maps \( f(z) \in S \) of the form (1.1) into \( k - S_\alpha(\alpha) \).

**Proof.** The proof of this theorem is much akin to that of Theorem 2.5. Therefore we omit
Theorem 2.7. Let $m > 0$, $\lambda > 0$ and the inequality
\[ e^n m^5 + (\lambda + 5)m^2 + (3\lambda + 4)n \leq \lambda \]
is satisfied then $I(m)f(z)$ maps $f(z) \in S$ of the form (1.1) into $C$. 

**Proof.** The proof of this theorem is similar to that of Theorem 2.3. Therefore we omit the details involved.

III. AN INTEGRAL OPERATOR

In the following theorem, we obtain analogues results in connection with a particular integral operator $G(m,z)$ which is defined as follows

$$G(m, z) = \int_0^z \frac{I(m)f(t)}{t} \, dt \quad (3.1)$$

**Theorem 3.1.** Let $f$ be defined by (1.1) is in the class $P^r_{\alpha}(\beta)$ with $m > 0$ and the inequality
\[ (k + 1)(1 - e^{-m}) - \frac{k + \alpha}{n} (1 - e^{-m} - m e^{-m}) \leq \frac{\gamma(1 - \alpha)}{2|\tau| (1 - \beta)} \]
is satisfied then $G(m,z)$ defined by (3.1) is in the class $k - UCV(\alpha)$. 

**Proof.** Since
\[ G(m, z) = z + \sum_{n=0}^{\infty} \frac{m^{n-1}}{n!} e^{-m} a_n z^n \quad (3.2) \]
To prove it in $k - UCV(\alpha)$, we have to show that
\[ T = \sum_{n=2}^{\infty} n[n(k + 1) - (k + \alpha)] |A_n| \leq 2|\tau| (1 - \beta) \sum_{n=2}^{\infty} \frac{m^{n-1}}{n!} e^{-m} \frac{1 + \gamma(n-1)}{1 + \gamma(n-1)} \]
\[ \leq 2|\tau| (1 - \beta) \sum_{n=2}^{\infty} \frac{m^{n-1}}{n!} e^{-m} \frac{1 + \gamma(n-1)}{1 + \gamma(n-1)} \]
\[ = 2|\tau| (1 - \beta) \sum_{n=2}^{\infty} \frac{m^{n-1}}{n!} e^{-m} \frac{1 + \gamma(z+1)}{(1 + \gamma(z+1))(n-1)!} \]
\[ \leq \frac{1 - \alpha}{\gamma} \frac{[(1 + k)(1 - e^{-m}) + (1 - \alpha)(1 - e^{-m} - m e^{-m})]} \]
\[ \leq 1 - \alpha \]
by the given hypothesis.

This completes the proof of Theorem 3.1. 

**Theorem 3.2.** Let $f$ be defined by (1.1) in the class $S$ with $m > 0$ and the inequality (2.7) is satisfied then $G(m,z)$ defined by (3.1) is in $k - UCV(\alpha)$. 

**Proof.** Since
\[ G(m, z) = z + \sum_{n=0}^{\infty} \frac{m^{n-1}}{n!} e^{-m} a_n z^n \quad (3.3) \]
To prove that $G(m,z) \in k - UCV(\alpha)$, we have to show that
\[ \sum_{n=2}^{\infty} n[n(k + 1) - (k + \alpha)] |A_n| \]
\[ \leq 1 - \alpha \]
Using the well-known inequality $|a_n| \leq n$ and proceeding the previous theorem we obtain the required condition.
The proof of following Theorems 3.3-3.5 are similar to Theorem 3.1 therefore we only state these theorems.

**Theorem 3.3.** Let $f$ be defined by (1.1) in the class $S$ with $m > 0$ and the inequality 
$$(k + 1)me^m \leq 1 - \alpha$$
is satisfied then $G(m,z)$ is in the class $k - S_p(\alpha)$.

**Theorem 3.4.** Let $f$ be defined by (1.1) in the class $S$ with $m > 0$ and the inequality 
$$me^m \leq \lambda$$
is satisfied then $G(m,z)$ is in the class $S^*(\lambda)$.

**Theorem 3.5.** Let $f$ be defined by (1.1) in the class $S$ with $m > 0$ and the inequality 
$$e^m \left[ m^2 + (\lambda + 2)m \right] \leq \lambda$$
is satisfied then $G(m,z)$ is in the class $C(\lambda)$.

**REFERENCES**


STUDY ON THE EFFECTS OF SiC PARTICLES ON TENSILE PROPERTIES FOR Al/SiC COMPOSITES

T R Mohan¹, C M Sharanaprabhu², Shashidhar K Kudari³

¹,² Department of Mechanical Engineering, Bapuji Institute of Engineering & Technology, Davanagere 577004, (India)
³Principal, Bapatla Engineering College, Bapatla-522101 (India)

ABSTRACT
Owing to the strength and stiffness, Aluminium matrix composites (AMCs) are employed in aerospace and automobile industries. The various weight proportions of silicon carbide (SiC) particle reinforced with the pure aluminum (Al) matrix are prepared to manufacture AMCs using stir casting technique. The tensile properties and microstructure analyses were carried out for various weight proportions of SiC particles. The observations of distributed SiC particles before and after tensile test specimens were carried out using Scanning Electron Microscope (SEM). The result signifies that the modulus of elasticity and yield strength increases with increase in weight proportions of SiC particles in Al/SiC composites. The SEM fractographs for fractured surface of tensile test specimens illustrates that the addition of SiC particles to the Al/SiC composites will change the properties of the composites from ductile to brittle.

Keywords: Metal matrixes composite; silicon carbide; stir casting, microstructure.

I INTRODUCTION
Now a day's composite materials emerged as a distinct technology which provides a motivation for development of materials. A composite material is a combination of two or more materials that results in better properties than those of the individual components used alone. The two constituents are reinforcement and a matrix. Composite materials are usually classified on the basis of the physical or chemical nature of the matrix phase, e.g., polymer matrix, metal-matrix and ceramic composites. Among these composite material, a metal matrix composite (MMCs) provides improved performance of an advanced military systems for a primary motivation [1]. Aluminum based metal matrix composites have been utilized considerable due to their high performance and class of light weight [2-3]. Consequently, Al based matrix composites depends on: The base matrix of Al composition; the nature of reinforcing material; the methods implemented for manufacturing the composite [4-6].

Many investigators [7-9] have evaluated that the volume fraction of reinforced particle is less than 30% will be utilized for structural and wear applications. Miller [10] and Pavan et.al. [11] examined the addition of the SiC particles in Al composite will increase the elastic modulus their by decrease in the ductility and toughness. The SiC particle reinforcement in Al matrix composites is an important parameter for optimizing the mechanical properties. Consequently, the work includes both experimental studies and mechanical modeling. Withers et.al.[12] and
Nardone [13] uses Eshelby-type model to estimate the yield stress and elastic modulus of the Al/SiC composites. Researchers [14-15] showed the relationship between tensile ductility and fracture toughness of the SiC reinforced Al composites. The mechanical properties related to ductility and strength of Al composites depend on the volume fraction and granularity of the SiC particles [16]. In this paper, an attempt is made to fabricate Al matrix composites with various weight proportions of reinforcing phase, SiC by stir casting method. The effects of various weight proportions of the SiC particles on tensile properties and the microstructure for Al/SiC composites were studied.

II PROCEDURE FOR STIR CASTING PROCESSES OF Al/SiC COMPOSITE

The SiC particles of 20 μm mean size were used in the present work. The various weight proportions of SiC particles such as: 3%, 5%, 7%, 9%, 12% and 15% were added to the commercially available Al with purity of 99.5%. A circular furnace maintaining temperature between 900°C-950°C with silicon carbide crucible was utilized for manufacturing Al/SiC composites. The pure Al metal is taken in a crucible and placed in a furnace then heated above its liquids temperature. The reinforced particles are preheated at 800°C to remove the volatile substances, to ensure homogeneous mixing and intimate dispersion. The degas tablet was dropped into the mixture to prevent the defects in the cast. The mechanical stirrer is introduced into the melt and stirred at with a speed of about 700 rpm to homogenize the dispersion of the particles and then cast. The stirrer was immersed 1/3rd heights of the molten melt from the bottom of the crucible. A 15 minute stirring time was permitted to improve the wetting property of the Al/SiC composites. The procedure of manufacturing the Al/SiC composites in this work is similar to the earlier works of Aniban et al.[17] and Vikram and Prasad [18].

III EXPERIMENTAL METHOD FOR TENSILE TEST

The rectangular slabs of 55mmX100mmX120mm cast composites of Al/SiC were machined to obtained dog-bone tensile test specimen as per ASTM standards [19]. The flat tensile specimen along with specifications of the gauge thickness 6±0.1 mm and gauge length of 25mm as shown in Fig.1(a). The Fig.1(b) will show the photograph of tensile test specimens for various weight proportions of SiC particles reinforced with pure. Three identical specimens were prepared for each composition of Al/SiC composite to estimate the modulus of elasticity, yield strength, and ductility. The testing was carried out using a 100kN servo hydraulic computer interfaced universal testing machine. The test setup is shown in Fig.2. Extensometers are attached to measure the elongation.

IV RESULTS AND DISCUSSION

Scanning Electron Microscopy (SEM) was conducted at magnification of 500X for various weight proportions of the SiC particles reinforced with Al. Figure 3 (a-f) illustrates the different weight proportions of SiC particles reinforce with Al matrix composite. It is clearly seen from Fig.3 that the SiC particles reinforcements in the respective matrix are fairly uniform. The unity between the SiC particles and matrix is relatively weak. Since, pores co-exist with the SiC particles. Figure 3 reveals that the SiC particles are less in number for 3wt% composite and
situated relatively farther when compared with other SiC volume fractions. The SiC particles are compactly located as the volume fraction increases and more in number for 15wt%.

Figure 1: (a) Tensile specimen as per ASTM B557 M-94, G is gage length, W is width, T is thickness, R is radius of fillet, L is overall length, A is length of reduced section, B is length of grip section and C is width of grip section [19], (b) photograph of tensile test specimens for various weight proportions of SiC particles of Al/SiC composites.

Figure 2. Tensile test up to evaluate tensile properties for various Al/SiC composite
4.1 Tensile properties of A/SiC Composite

The tensile properties such as: elastic modulus, yield strength, percentage of elongation was evaluated for various Al/SiC composites. The average values of three tensile test specimen results were taken for the analyses. The tensile properties for various Al/SiC composites as a function of the weight percentage of SiC particles are shown in Table.1. The modulus of elasticity, yield strength values are increases as the weight percentage of SiC particles increases in composites. However, the percentage of elongation decreases with the increase in the weight percentage of SiC particles. Hence, the results from Table.1 clearly demonstrate that the SiC reinforced Al can improve the strength but degrade the ductility of the composites compare to the monolith.

<table>
<thead>
<tr>
<th>Material</th>
<th>Modulus of Elasticity /GPa</th>
<th>Yield strength/MPa</th>
<th>% Elongation (in terms of strain)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al+0wt% SiC</td>
<td>64</td>
<td>54.3</td>
<td>25.6</td>
</tr>
<tr>
<td>Al+3wt% SiC</td>
<td>68</td>
<td>56.8</td>
<td>20.4</td>
</tr>
<tr>
<td>Al+5wt% SiC</td>
<td>71</td>
<td>57.4</td>
<td>16.4</td>
</tr>
<tr>
<td>Al+7wt% SiC</td>
<td>73</td>
<td>59.4</td>
<td>13.7</td>
</tr>
</tbody>
</table>

Table 1: Tensile properties of Al/SiC
4.2 Fracture surfaces analyses of tensile specimen for Al/SiC composite

The fracture surfaces after tensile testing with the pure Al and various weight proportions SiC particles of Al/SiC composite were investigated for recognizing the micro mechanisms of failure. The fractured surface of the specimens cut to a specified dimensions and ultrasonically cleaned in acetone then examined in SEM. Figure 4 shows the SEM fractographs for pure Al and various weight proportions SiC particles of Al/SiC composites. The numerous dimples were observed over the fracture surface for pure Al, indicating a ductile fracture behavior. The size of the dimple is about 5 μm. This type of ductile fracture behavior was observed in the work of Min [16]. As the SiC reinforced increases for Al composites the dimples should be a result of the void nucleation and subsequent coalescence which results brittle fracture. The Fig. 4(b-f) clearly analyses for weight proportions of 3% to 12% SiC particle of Al composite undergoes the transition state of ductile and brittle fracture. For weight proportion of 15% of SiC, the presence of void nucleation and subsequent coalescence are more which leads to brittle fracture.

V CONCLUSION

The significant conclusions of the studies are: a) the conventional low cost stir casting method was successfully adopted in the preparation of Al/SiC composite. Since, the microstructural studies revealed the uniform distribution of the particles in the matrix system, b) the SiC particle additions to Al matrix will increases the modulus of elasticity and yield strength of the composite. But the elongation decreases with the increase in the volume fraction of the SiC particles to Al metal matrix composite, c) The analyses of microstructured fracture surface for the tensile specimen clearly demonstrate that the wt 15% of SiC particles to the Al metal matrix composite will undergo brittle fracture indicating the composite is of brittle material.
VI ACKNOWLEDGEMENTS
Authors gratefully acknowledge the Research Centre, Bapuji Institute of Engineering & Technology, Hubli-577 004, BiSS, Bangalore and TUV Rheinland, Bangalore for their support in using the testing facilities.

REFERENCES


SYNTHESIS AND STUDIES OF MONOAZO DISPERSE DYES BASED ON 2, 4-DIHYDROXY-6-CHLORO QUINOLINE

Dina Yogesh Champaneri¹, A.G. Mehta²

¹ I/C Principal, B.P. Baria Science Institute Navsari, Gujarat, (India)
² Ex. Principal, P.T. Sarvajanik College of Science, Surat, Gujarat, (India)

ABSTRACT

2, 4 – dihydroxy – 6 – chloroquinoline has been coupled with various diazotized coupling components to prepare various disperse dyes. All the dyes were characterized by their percentage yield and elemental analysis. IR, H¹ NMR, CMR of some selected synthesized compounds have been carried out. Their dying performance were assessed on polyester fabric. The percentage dye bath exhaustion were found to be good and acceptable. The dyed fibres showed poor to fair good light fastness and good to excellent wash fastness.

Keywords: 2, 4 Dihydroxy – 6 – Chloro Quinolone, Dyeing, Exhaustion, Light Fastness, Wash Fastness.

I. INTRODUCTION

Disperse Dyes are relatively small molecules, with very low water solubility, which possess substantivity for hydrophobic fibres such as cellulose acetate and polyester. Most of the new disperse dyes introduced during the last decade have been intended for application to polyester and the majority represent advances in application and fastness properties over dyes previously available [1,2]. Useful reviews on structural factors affecting the fastness to light of dyed fibres[3] and on relations between the molecular structures of dyes and their technical properties have appeared [4].

The development of disperse dyes is due to significant increase in the world production of polyester compared to other fibres. The disperse dyes have the most satisfactory results due to their simple application properties and abilities to cover regularities in yarn. Over 90% of disperse dyes usage is for the colouration of polyester and its blends.

Some of the dyes based on heterocyclic ring system are known to possess high tinctorial power and excellent fastness properties. The compactness extension of conjugation in the structures of heterocyclic compounds are important for the disperse dyes derived from them. Heterocyclic coupling components give heterocyclic azo disperse dyes with colour ranging from yellow to deep red. The synthesis and application of monoazo disperse dyes derived from 4-hydroxy-1-methyl-2-quinolone [5,6], 4-hydroxy-1-phenyl-2-quinolone [7], 5-hydroxyquinoline[2,1-b]-quinazolin-12[H]-one [8] have been earlier. Malankar and Desai [9] have prepared heterocyclic dyes using quinoline moiety. Desai and Desai [10] and Vashi and Mehta [11,12] have earlier reported quinoline based disperse dyes using quinoline as a coupling component. The synthesis and application of monoazo disperse dyes derived from 3-[(2,6-dimethyl-4-quinolinyl)amino] phenol [13] and 4-hydroxy-1-
methylquinololin-2-[1H]-one \[14\] and 4-hydroxy-2-phenyl-6/7 substituted quinolone \[15\], 2,4-dihydroxy quinoline \[16\] and 2,4-dihydroxy-6-methyl quinolone \[17\] systems have also been reported to show excellent shades on polyester. These compounds when applied on polyester fabric gave brown, yellow and grey shades with excellent exhaustion and fastness properties. In view of the encouraging reports about the technical applications of the heterocyclic dyes, it was thought interesting to undertake the synthesis and study of dyeing properties of the monoazo disperse dyes based on 2,4-dihydroxy-6-substituted quinoline system. Here, 2,4-dihydroxy-6-chloro quinoline was synthesized and series of monoazo disperse dyes were synthesized. The characterization of the dyes, evaluations of their technical properties were also performed. The aim of the present work was to find out the possibilities of increasing the shade range of these dyes.

II. EXPERIMENTAL

All the dyes synthesized were purified by crystallization. Melting points were determined by open capillary method and are uncorrected. IR spectra of the selected synthesized compounds were recorded in KBr pellets on a Perkin-Elmer Model – 377 spectrophotometer for structural elucidation, particularly for presence of functional groups. 1H NMR and CMR of some selected synthesized compounds has been carried out for the structural elucidation, particularly for the presence of protons and carbons. “High Temperature High Pressure” method has been employed for the application of disperse dyes on polyester fabrics\[18,19,20\]. Wash fastness was assessed with Grey Scale \[21\]. Light fastness was also assessed for synthesized dyes \[22\]. The exhaustion study of the synthetic dyes was also carried out.

2.1 Ethyl-3-[(4-Chloro Phenyl)Amino-3-Oxopropanoate

A mixture of p-chloroaniline (5.1 g , 0.04 mole) and diethylmalonate (6.5 g , 0.04 mole) was refluxed in boiling xylene (30 ml) with an air condenser for six hours on a sand bath. The solvent was evaporated leaving a residue.

2.2 2, 4 – Dihydroxy-6-Chloro Quinoline (DHCQ)

The residue obtained as above was mixed with freshly prepared poly phosphoric acid (P\(_2\)O\(_5\) 40 g , orthophosphoric acid 24 ml ), stirred well for some time and the temperature was slowly raised to 120° ; it was kept in a desiccator overnight. Next day the temperature was slowly raised and lowered by 10°, until it reached 140° and the heating was continued at this temperature for two hours. This treatment helps in getting clean product in good yield. The reaction mass was poured slowly in to ice-cold water and neutralized with ammonium hydroxide on the acidic side. The product was filtered, washed with water, dried and crystallized from ethanol. Yield 60 %, mp>360° , [Found : N, 7.06 %; C\(_9\)H\(_6\)O\(_2\)NCl required N, 7.20 % ]; IR spectra (KBr) : 3200-2600 (-OH), 1251 (C -O), 1589,1502 (C=C), 817 (C -Cl) ; NMR spectra (DMSO-d\(_6\) ) , δ 5.76 (1H,4-OH) , 11.36-11.56 ( 1H,2-OH); CMR Spectra (Chemical shift),163 (C\(_2\),C-OH), 136 (C\(_3\) ), 161 (C\(_4\),C-OH), 122 (C\(_5\) ), 117 (C\(_6\) ), 89 (C\(_7\) ), 116 (C\(_8\) ), 125 (C\(_9\) ), 131 (C\(_{10}\) ).
Table-1 Results for IR of 2, 4-Dihydroxy-6-Chloro Quinolone (DHCQ)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O-H stretching vibrations</td>
<td>3200-2600 cm⁻¹</td>
</tr>
<tr>
<td>C-H stretching vibrations</td>
<td></td>
</tr>
<tr>
<td>aromatic /aliphatic</td>
<td></td>
</tr>
<tr>
<td>C=(\text{C}) stretching vibrations</td>
<td>1589,1502 cm⁻¹</td>
</tr>
<tr>
<td>C-O stretching vibrations</td>
<td>1251 cm⁻¹</td>
</tr>
<tr>
<td>C-Cl stretching vibrations</td>
<td>817 cm⁻¹</td>
</tr>
</tbody>
</table>

Table-2 Results for NMR of 2, 4-Dihydroxy-6-Chloro Quinolone (DHCQ)

<table>
<thead>
<tr>
<th>No.</th>
<th>Chemical Shift (δ)</th>
<th>Number of Relative Protons</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.76</td>
<td>1</td>
<td>4-OH</td>
</tr>
<tr>
<td>2</td>
<td>11.36-11.56</td>
<td>1</td>
<td>2-OH</td>
</tr>
</tbody>
</table>

Table-3 Results for CMR of 2, 4-Dihydroxy-6-Chloro Quinolone (DHCQ)

<table>
<thead>
<tr>
<th>No.</th>
<th>Chemical Shift</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>163</td>
<td>(\text{C}_2) C-OH</td>
</tr>
<tr>
<td>2</td>
<td>136</td>
<td>(\text{C}_3)</td>
</tr>
<tr>
<td>3</td>
<td>161</td>
<td>(\text{C}_4) C-OH</td>
</tr>
<tr>
<td>4</td>
<td>122</td>
<td>(\text{C}_5)</td>
</tr>
<tr>
<td>5</td>
<td>117</td>
<td>(\text{C}_6)</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>(\text{C}_7)</td>
</tr>
<tr>
<td>7</td>
<td>116</td>
<td>(\text{C}_8)</td>
</tr>
<tr>
<td>8</td>
<td>125</td>
<td>(\text{C}_9)</td>
</tr>
<tr>
<td>9</td>
<td>131</td>
<td>(\text{C}_{10})</td>
</tr>
</tbody>
</table>

2.3 Diazotisation of Amino Compounds

Seventeen amino compounds were diazotized in the usual manner.
2.4 Coupling of Diazo Solution with DHCQ

A clear solution of DHCQ (0.976 g, 0.005 mole) in rectified spirit (15 ml) and sodium hydroxide (15 ml, 10%) was cooled below 50, stirred well and the diazo solution was added drop-wise over a period of 15 minutes maintaining the pH at 8-9. The stirring was continued for 2 hours at 0 – 50. The reaction mixture was diluted with water and the pH was adjusted at 7.0 with acetic acid. The product was washed with water and crystallized from DMF.

![Chemical Reaction](attachment:chemical_equation.png)

**Table No. – 4 Characterisation Data of Disperse Dyes (D-1 to D-17)**

<table>
<thead>
<tr>
<th>Dye No.</th>
<th>Coupling Component</th>
<th>Molecular Formula</th>
<th>Mol. Wt. (g)</th>
<th>Yield (%)</th>
<th>M.P.° C (d)</th>
<th>% Nitrogen Found</th>
<th>% Nitrogen Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-1</td>
<td>2-Amino-5,6-dichlorobenothiazole</td>
<td>C_{16}H_{12}N_{4}Cl_{2}S</td>
<td>425.5</td>
<td>61</td>
<td>188</td>
<td>12.90</td>
<td>13.16</td>
</tr>
<tr>
<td>D-2</td>
<td>2-Amino-6-nitrobenzothiazole</td>
<td>C_{16}H_{12}N_{4}SCl</td>
<td>401.5</td>
<td>61</td>
<td>210</td>
<td>17.18</td>
<td>17.43</td>
</tr>
<tr>
<td>D-3</td>
<td>3-Chloro-4-fluoroaniline</td>
<td>C_{16}H_{12}O_{2}N_{4}Cl_{2}F</td>
<td>352.0</td>
<td>64</td>
<td>165</td>
<td>11.72</td>
<td>11.93</td>
</tr>
<tr>
<td>D-4</td>
<td>4-Aminoacetanilide</td>
<td>C_{17}H_{13}O_{2}N_{4}Cl</td>
<td>356.5</td>
<td>64</td>
<td>218</td>
<td>15.47</td>
<td>15.71</td>
</tr>
<tr>
<td>D-5</td>
<td>4-Nitroaniline</td>
<td>C_{15}H_{12}O_{2}N_{4}Cl</td>
<td>344.5</td>
<td>82</td>
<td>219</td>
<td>16.03</td>
<td>16.26</td>
</tr>
<tr>
<td>D-6</td>
<td>p-Toluidine</td>
<td>C_{16}H_{12}O_{2}N_{4}Cl</td>
<td>313.5</td>
<td>75</td>
<td>168</td>
<td>13.12</td>
<td>13.39</td>
</tr>
</tbody>
</table>
Table - 5  Percentage exhaustion, Light fastness and wash fastness on Polyester for Disperse dyes (D -1 to D-17)

<table>
<thead>
<tr>
<th>Dye No.</th>
<th>Wavelength for absorbance measurement $\lambda_{\text{max}}$(nm)</th>
<th>% Exhaustion = (Y/ 10) 100</th>
<th>Light Fastness</th>
<th>Wash Fastness</th>
</tr>
</thead>
<tbody>
<tr>
<td>D – 1</td>
<td>287</td>
<td>47.5</td>
<td>2</td>
<td>4-5</td>
</tr>
<tr>
<td>D – 2</td>
<td>372</td>
<td>50.0</td>
<td>4</td>
<td>4-5</td>
</tr>
<tr>
<td>D – 3</td>
<td>399</td>
<td>42.5</td>
<td>3-4</td>
<td>4-5</td>
</tr>
<tr>
<td>D – 4</td>
<td>460</td>
<td>5.75</td>
<td>2</td>
<td>4-5</td>
</tr>
<tr>
<td>D – 5</td>
<td>424</td>
<td>71.0</td>
<td>2-3</td>
<td>4</td>
</tr>
<tr>
<td>D – 6</td>
<td>436</td>
<td>77.5</td>
<td>3</td>
<td>4-5</td>
</tr>
<tr>
<td>D – 7</td>
<td>424</td>
<td>77.5</td>
<td>3</td>
<td>4-5</td>
</tr>
</tbody>
</table>
III. RESULT AND DISCUSSION

Quinoline based monoazo disperse dyes have been prepared and characterized by elemental and spectral analysis. The % yields of the synthesized dyes were 51 to 90 %. Their dyeing performance on polyester fabric was assessed. All the synthesized dyes almost produce variety of different yellow shades with poor to fairly good light fastness, good to excellent wash fastness and the percentage dye bath exhaustion 77.5 % to 37.5 %.m In general, brilliance and beauty of the shade and excellent wash fastness reveal that some of the disperse dyes would prove to be useful dyes for dyeing polyester fabrics.

IV. ACKNOWLEDGEMENT

The authors express their gratitude to the principal, Sir P. T. Sarvajanik College of Science, Surat for providing research facilities and to Colour Text., Surat for dyeing facilities.

REFERENCES


[22] Colour Fastness Tests, SDC Publications
A FAULT NODE RECOVERY ALGORITHM TO REINFORCES THE LIFETIME AND ROUTING IN WIRELESS SENSOR NETWORK

Divyashree S1, Sanjay Kumar N.V2, Dr. Y. P. Gowramma3

1PG Student, Dept. of CS&E, 2Assistant Professor, Dept. of CS&E, 3Professor and HOD of CS&E
Kalpataru Institute of Technology, Tiptur, Karnataka, (India)

ABSTRACT

Wireless sensor networks have seen tremendous advances and utilization in the past two decades. Ranging from fossil oil exploration, mining, weather and even battle operations, all of these need detector applications. In the WSN, reduction of energy consumption is incredibly necessary for every each sensor node as a result of it will extend WSN period. The Wireless sensor network may be an assortment of sensors that are meet giant geographical region. Since sensors are wide unfold and enormous in range, the occurrences of faults within the network are additional. Therefore to detect the fault node and to replace the fault node an efficient algorithm is proposed. In this paper proposes a fault node recovery (FNR) to reinforce the period of a wireless sensor network (WSN) once a number of the sensor nodes stop working, either as a result of they now not have battery energy or they need reached their operational threshold. Using the FNR algorithm may end up in replacements of sensor nodes and more reused routing paths. Thus, the algorithm not solely enhances the WSN period however additionally reduces the cost of replacing the sensor nodes.

Keywords: Genetic Algorithm, Grade Diffusion (GD) Algorithm, Gradient Diffusion Algorithm, and Wireless Sensor Networks (WSN)

I. INTRODUCTION

RECENT advances in small process, wireless and battery technology and good sensors have enhanced information processing [5], wireless communication, and detection capability. In sensor networks, every sensor node has restricted wireless computational power to process and transfer the live data to the base station or data collection center. Therefore, to extend the sensor area and also the transmission area the wireless sensor network typically contains several sensor nodes. Generally, every sensor node features a low level of battery power that can't be replenished. Once the energy of a sensor node is exhausted, wireless sensor network leaks can seem, and also the unsuccessful nodes won't relay information to the opposite nodes throughout transmission process. Thus, the other sensor nodes are going to be burdened with exaggerated transmission process. In the WSN, reduction of energy consumption is incredibly necessary for every each sensor node as a result of it will extend WSN period. The Wireless sensor network may be an assortment of sensors that are meet giant geographical region. Since sensors are wide unfold and enormous in range, the occurrences of faults within the network are additional. Therefore to detect the fault node and to replace the fault node an efficient algorithm is proposed. This work proposes a fault node recovery (FNR) to reinforce the period of a wireless sensor network (WSN) once a number of the sensor nodes stop working, either as a result of they now not have battery energy
or they need reached their operational threshold. Using the FNR algorithm may end up in replacements of sensor nodes and more reused routing paths. Thus, the algorithm not solely enhances the WSN period however additionally reduces the cost of replacing the sensor nodes.

1.1 Wireless Sensor Networks

Wireless sensor networks have seen tremendous advances and utilization in the past two decades. Ranging from fossil oil exploration, mining, weather and even battle operations, all of those need detector applications. One reason behind the growing quality of wireless sensors is that they'll add remote areas while not manual intervention. All the user needs to do is to collect the data sent by the sensors, and with bound analysis extract important information from them. Typically sensor applications involve several sensors deployed along. These sensors form a network and collaborate with one another to collect data and send it to the base station. These nodes combine with routers and gateways to form a WSN system. The WSN is formed of nodes from a couple of too many hundred, wherever every node is connected to at least one or many sensors as shown in figure 1.

II. RELATED WORK

Failures are unavoidable in Wireless Sensor Networks due to the lack of observation and unattended deployment. There is a unit several problems associated with energy, memory and process ability of a sensor node. The occurrences of faults are a unit largely due the presence of faulty sensor nodes [1]. To identify a fault node and to interchange it, several techniques are proposed.

C. Intanagonwiwat et al. [1] explored the directed-diffusion paradigm for such coordination. Directed diffusion is data-centric there in all communication is for named data. All nodes in an exceedingly directed-diffusion-based network are a unit application aware. This permits diffusion to realize energy savings by choosing through empirical observation smart ways and by caching and process knowledge in-network (e.g., data aggregation). They explore and evaluated the employment of directed diffusion for a straightforward remote-surveillance device network analytically and by experimentation.

Fault management for WSNs [2] is totally different from traditional networks. Recent analysis has developed many schemes and techniques that affect differing kinds of faults at totally different layers of the network. All these techniques consider only few issues of WSNs. An efficient Fault Management system or model should take into account most problems or forms of faults.
J. A. Carballido et al. [3] mentioned the foundations and implementation of a genetic algorithm (GA) for instrumentation functions. The GA constitutes a format module of a call network for sensor network style. The tactic development entailed the definition of the individual’s illustration additionally because the design of a graph-based fitness operates, alongside the formulation of many alternative impromptu enforced options. The performance and effectiveness of the GA were assessed by initializing the instrumentation design of an ammonia synthesis plant. The format provided by the GA succeeded in fast the sensor network design procedures. It conjointly accomplished a great improvement within the overall quality of the ensuing instrument configuration. The GA will constitute a valuable tool for the treatment of real industrial issues.

H. C. Shih et al. had given the Grade Diffusion (GD) [4] algorithm in 2012 to enhance the ladder diffusion algorithm using ant colony optimization (LD-ACO) for wireless sensor networks. The GD algorithm not solely creates the routing for every sensor node however additionally identifies a collection of neighbor nodes to scale back the transmission loading. Every sensor node can choose a sensor node from the set of neighbor nodes once its grade table lacks a node able to perform the relay.

Hong-Chi Shih et al. proposed a fault node recovery (FNR) algorithm [6] to boost the life of a wireless sensor network (WSN) once a number of the sensor nodes pack up, either as a result of they not have battery energy or they need reached their operational threshold. Mistreatment the FNR algorithmic rule may end up in fewer replacements of device nodes and additional reused routing paths. Thus, the algorithm not solely enhances the WSN lifespan but also reduces the cost of exchange the sensor nodes. The normal approaches to sensor network routing embrace the directed diffusion (DD) [1] algorithm and the grade diffusion (GD) algorithm [4]. The algorithm proposed in this paper relies on the GD algorithm, with the goal of exchange fewer sensor nodes that are a unit inoperative or have depleted batteries, and of reusing the most variety of routing paths. These optimizations can ultimately enhance the WSN lifespan and reduce sensor node replacement cost.

III. EXISTING SYSTEM

3.1 Directed Diffusion Algorithm

A series of routing algorithms for wireless sensor networks have been proposed in recent years. C. Intanagonwiwat et al. given the Directed Diffusion (DD) algorithm [1] in 2003. The goal of the DD algorithm is to scale back the data relay transmission counts for power management. The DD algorithm may be a query-driven transmission protocol. The collected data is transmitted providing it matches the query from the sink node. Within the DD algorithm, the sink node provides the queries within the type of attribute-value pairs to the opposite sensor nodes by broadcasting the query packets to the total network. After, the sensor nodes send the data back to the sink node only when it fits the queries.

3.2 Grade Diffusion Algorithm

H. C. Shih et al. had given the Grade Diffusion (GD) [4] algorithm in 2012 to enhance the ladder diffusion algorithm using ant colony optimization (LD-ACO) for wireless sensor networks. The GD algorithm not solely creates the routing for every sensor node however additionally identifies a collection of neighbor nodes to scale back the transmission loading. Every sensor node can choose a sensor node from the set of neighbor nodes once its grade table lacks a node able to perform the relay.

The GD algorithm may also record some information relating to the data relay. Then, a sensor node will choose a node with a lighter loading or a lot of offered energy than the opposite nodes to perform the additional relay.
operation. That is, the GD formula updates the routing path in real time, and also the event data is so sent to the sink node quickly and properly and shown in Figure 3.

![Figure 3: Grade Diffusion Flowchart](image)

Whether the DD or the GD algorithmic rule is applied, the grade making packages or interested query packets should first be broadcast. Then, the sensing element nodes transfer the event data to the sink node, in step with the algorithmic rule, once appropriate events occur. The sensor routing paths are shown in Figure 4.

![Figure 4: Wireless Sensor Node Routing](image)

The WSN could fail due to a spread of causes, together with the following: the routing path would possibly expertise a break; the WSN sensing space would possibly expertise a leak; the batteries of some sensor nodes could be depleted, requiring additional relay nodes; or the nodes wear out when the WSN has been in use an extended amount of your time. In Figure 5, the situation in which the outside nodes transfer event data to the sink node via the inside nodes (the sensor nodes near the sink node) in a WSN illustrate the accommodation measures for non-working nodes. The inside nodes therefore have most important data transmission loading, intense energy at a quicker rate. If all the inside nodes spend their energy or otherwise stop to operate, the event data will now not be sent to the sink node, and therefore the WSN can now not operate.

![Figure 5: Wireless Sensor Node Routing Path When Some Nodes Are Not Working](image)

The power consumption of the sensor nodes in WSNs is inevitable. However, proposes an algorithm to search for and replace fewer sensor nodes and to utilize the foremost routing ways. Typical search techniques are often incapable of optimizing nonlinear functions with multiple variables. One scheme, the genetic algorithm (GA), may be a directed random search technique developed in 1975, supported the conception of natural biology. The present work proposes a fault node recovery (FNR) algorithmic rule supported the GD algorithm combined with the GA. The FNR algorithm creates a routing table using the GD algorithm and replaces sensor nodes using the GA once the amount of sensor nodes that are not functioning exceeds the brink. This algorithmic rule not solely reuses the foremost routing ways to reinforce the WSN life however additionally reduces the cost.
IV. PROPOSED SYSTEM

This work proposes a fault node recovery (FNR) [6] algorithm to boost the period of time of a wireless sensor network (WSN) once a number of the sensor nodes pack up, either as a result of they not have battery energy or they need reached their operational threshold. Victimization the FNR algorithm may end up in fewer replacements of sensor nodes and additional reused routing paths. Thus, the rule not solely enhances the WSN period of time however additionally reduces the value of replacement the sensor nodes. The standard approaches to sensor network routing embrace the directed diffusion (DD) [1] algorithm and also the grade diffusion (GD) [4] algorithm. The rule projected during this work relies on the GD algorithm, with the goal of replacement fewer sensor nodes that square measure inoperative or have depleted batteries, and of reusing the utmost variety of routing paths. These optimizations can ultimately enhance the WSN lifetime and reduce sensor node replacement cost.

This work proposes a fault node recovery (FNR) algorithm for WSNs supported the grade diffusion algorithm [6] combined with the genetic rule. The Block diagram and flow chart is shown in Figure 6. The FNR algorithm creates the routing table, grade value, neighbor nodes, and payload value for every sensor node using the grade diffusion algorithm. In the FNR algorithm, the amount of nonfunctioning sensor nodes is calculated throughout the wireless sensor network operation, and also the parameter $B^h$ is calculated consistent with (1).

In Figure 6, the FNR algorithm creates the grade value, routing table, a group of neighbor nodes, and payload value for every sensor node, victimization the grade diffusion rule. The sensor nodes transfer the event data to the sink node consistent with the GD rule once events seem. Then, $B^h$ is calculated consistent with (1) within the FNR algorithm [2]. If $B^h$ is larger than zero, the rules are going to be invoked and replace nonfunctioning device nodes by purposeful nodes selected by the genetic rule. Then the wireless device network will still work as long because the operators square measure willing to switch sensors.

In (1), Grade is the sensor node’s grade value. The variable $N_i^{original}$ is the number of sensor nodes with the grade value $i$. The variable $N_i^{now}$ is the number of sensor nodes still functioning at the current time with grade value $i$. The parameter $\beta$ is about by the user and should have a worth between 0 and 1. If the amount of sensor nodes that operate for every grade is a smaller amount than $\beta$, $T_i$ can become 1, and $B^h$ are going to be larger than zero. Then, the algorithm can calculate the sensor nodes to replace using the genetic algorithm. The parameters square measure encoded in binary string and function the chromosomes for the GA. The elements (or bits), i.e., the genes, within the binary strings square measure adjusted to reduce or maximize the fitness value. The fitness operate generates its fitness value that consists of multiple variables to be optimized by the GA. Within every, iterations of the GA a pre-determined number of individuals will produce fitness values related to the chromosomes. There are five steps within the genetic algorithm: Initialization of data format, Evaluation, Selection, Crossover, and Mutation of corresponding data.

In initialization chromosome are generated. Every Chromosome is associate degree expected. The quantity of chromosomes depends on number of sensors to be replaced. Next is the evaluation method the quantity of routing path on the market if some non functioning sensors square measure replaced is evaluated supported fitness value. This fitness value is calculated with variety of sensor nodes grade values, number of reusable routing methods, total number of sensor nodes in original WSN, and total number of routing methods in original WSN. And in selection step chromosomes with lowest fitness values are eliminated. Two individual
chromosomes are chosen and compared and a part of it is replaced with the other to produce new offspring are obtained by crossover process. At last a single gene is replaced after comparison happened in mutation process.

\[ B^i = \sum_{i=1}^{\text{max (Grade)}} T_i \]

\[ T_i = \begin{cases} 
1, & \frac{N_{\text{new}}}{N_i} < \beta \\
0, & \text{otherwise} 
\end{cases} \]

V. CONCLUSION

In real wireless sensor networks, the sensor nodes use battery power provides and so have restricted energy resources. Additionally to the routing, it's vital to analysis the improvement of sensor node replacement, reducing the replacement cost, and reducing the cost, and reusing the foremost routing ways once some sensor nodes are a unit nonfunctional. This work proposes a fault node recovery algorithm for WSN supported the grade diffusion algorithm combined with a genetic algorithm. The FNR algorithm needs replacement fewer sensor nodes and reuses the foremost routing paths, increasing the WSN lifespan and reducing the cost.

REFERENCES


