ISSN: 2005-4262

International Journal of Grid and Distributed Computing

Volume No. 16

Issue No. 3

September - December 2023

ENRICHED PUBLICATIONS PVT. LTD

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

ISSN: 2005-4262

International Journal of Grid and Distributed Computing

Aims and Scope

The topics covered by IJGDC include the following:-

Architectures and Fabrics
Autonomic and Adaptive Systems
Cluster and Grid Integration
Creation and Management of Virtual Enterprises and Organizations
Dependable and Survivable Distributed Systems
Distributed and Large-Scale Data Access and Management
Distributed Multimedia Systems
Distributed Trust Management
eScience and eBusiness Applications
Fuzzy Algorithm
Grid Economy and Business Models
Histogram Methodology
Image or Speech Filtering
Image or Speech Recognition
Information Services
Large-Scale Group Communication
Metadata, Ontologies, and Provenance
Middleware and Toolkits
Monitoring, Management and Organization Tools
Networking and Security
Novel Distributed Applications
Performance Measurement and Modeling
Pervasive Computing
Problem Solving Environments
Programming Models, Tools and Environments
QoS and resource management
Real-time and Embedded Systems
Security and Trust in Grid and Distributed Systems
Sensor Networks
Utility Computing on Global Grids
Web Services and Service-Oriented Architecture
Wireless and Mobile Ad Hoc Networks
Workflow and Multi-agent Systems

ISSN: 2005-4262

International Journal of Grid and Distributed Computing

ISSN: 2005-4262

International Journal of Grid and Distributed Computing

(Volume No. 16, Issue No. 3, September - December 2023)

Contents

Blockchain Based Secure and Energy Efficient Routing in Wireless Sensor Network

Umesh Kumbhalkar[1], Dr. Sharad Gangele[*]

[1]Research Scholar, Professor, Department of Computer Science & Application R.K.D.F. University, Bhopal (MP), India E-mail: sharadgangele@gmail.com

[*] Professor, Department of Computer Science & Application R.K.D.F. University, Bhopal (MP), India E-mail: sharadgangele@gmail.com

ABSTRACT

Embedded systems and wireless sensor networks (WSN) are found today in increasingly critical areas of applications. They have become integrated and embedded in nearly all aspects of everyday life, including manufacturing, healthcare, education, critical infrastructure, and entertainment. The number of connected devices continues to grow, and due to the insecure nature of these devices, the amount of risk continues to grow as well. These risks, however, can be mitigated with the creation and adoption of WSN security standards developed to create an environment of safety, security, and confidence in the technology. Designing the security policy for WSNs requires asking somepreliminary questions. These questions are particularly important in the case of WSNs because their use is highly decentralized. Blockchain's ability on governing decentralized networks makes it especially suitable for designing a self-managing system on WSN devices. This article proposes a routing protocol that uses Blockchain technology to offer a shared memory between the network's nodes. The simulation results have shown that this solutioncanbeapplicableandcanresolvetheissuescitedabove.

Keywords: Wireless Sensor Network, Blockchain, Peer-to-Peer, Node Security

1. INTRODUCTION

The modern microelectronic mechanical systems (MEMS) based advanced technology provided low cost, tiny for the development of a large number of tiny, low cost, and low power wireless sensors [1]. Sensor network tends to be very application sensitive; they are deployed and utilized in various civilian and military settings [2]. Regardless of these advantages, wireless sensor networks (WSNs) introduce various security vulnerabilities such as different types of attacks and intruders due to the open nature of sensor nodes and unreliable wireless links [3–5].

The sensors transmit mission-critical information over the network, so there is a need for security services such as authentication, encryption, key management, and so forth. Because of the low cost of sensors, lack of tamper-resistant hardware, and meager resources, sensor network's security presents many formidable and unique challenges. Security needs should be taken into account to ensure data protection (also called data survivability) in these sensors at the design time. Distributed security schemes are preferable over centralized techniques because centralized procedures are prone to single point failure. The security of the transaction is supposed to be granted by the Blockchain's signatures [6].Furthermore, a lightweight Blockchain mechanism can be implemented to validate this approach, and it is used tosimulate our proposition solution through a homogeneous static network. Unlike some existing articles and to the best of the knowledge, this work considers the Blockchain as support to share network status in real-time to enhance the routing process [7]. The main contributions of this article are as follows:

• Considering nodes as coins and transfer their ownership between eachother;
• Use Blockchain as shared memory to broadcast the status of the network'snodes;
• Usethepastnode'sactivitiestodeterminethetrafficload.

Similar contribution was given by Gangele et al.(2018) in which threshold-based energy-estimated distributed routing algorithm was presented in prospect of wireless sensor network[22]. Multi-path and multi-hop energy efficient routing was also estimated by Gangele et al.(2019) [23] and heterogeneous multi-clustered energy efficient routing protocol [24] in wireless sensor network environment.

2. SECURITY IN WIRELESS SENSOR NETWORKS

WSN is based on several enabling technologies suchasradio frequency identification (RFID), near field communication, machine-to-machine communications (M2M), ultra- wideband, or IPv6 low power wireless personal area networks (6LowPAN). WSNs experience various security threats due to their essential characteristics and vulnerability to numerous security threats due to their open communication environment compared with wired networks [8–10]. Figure 1 shows major security issues in wireless sensor network.

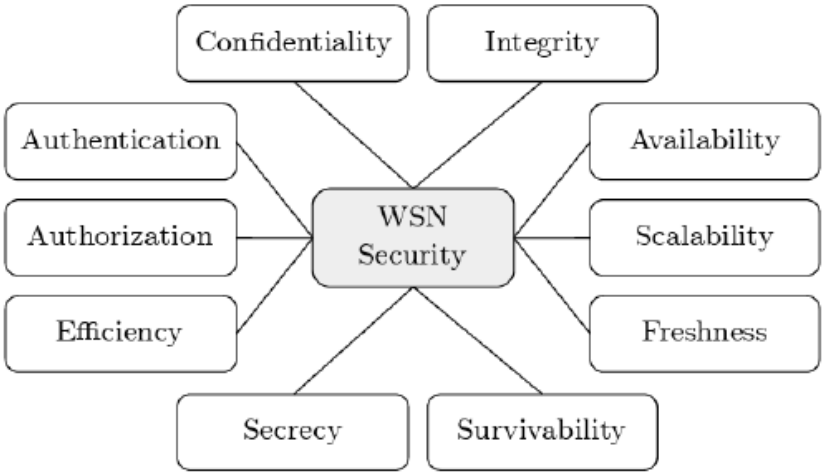

Fig. 1-Security Issues in Wireless Sensor Network

WSN and IoT cover a wide range of applications and will touch on almost all areas that we face daily.

Smart Cities: the IoT allows better management of the networks that supply our cities (water, electricity, gas, and so forth) by enabling continuous [11].Monitoring inreal-time and precise. Sensors can be used to improve the management of car parks and urban traffic and reduce congestion and CO2 emissions.

• **Energy:** the management of electric grids will be improved thanks to telemetry, allowing real-time management of the energy distribution infrastructure. This large- scale interconnection facilitates maintenance and control of consumption and detection offraud.

• **Transport:** in this area, the WSN will support intelligent vehicles current efforts for road safety and driver assistance.

This concerns communication between vehicles and between vehicles and road infrastructure. The WSN be a natural extension of "intelligent transport systems" and their contributions to road safety, comfort, efficient traffic management, and saving time andenergy.

• **Health:** in the health field, the IoT allows the deployment of personal networks (PAN) for the control and monitoring of clinical signs, especially for the elderly. This makes it possible to facilitate remote monitoring of patients at home and provide solutions for people's autonomy with reduced mobility.

• **Industry:** in industry, the IoT will allow full tracking of products, from the production chain to the logistics and distribution chain, by overseeing supply conditions. This end-to-end traceability facilitated the fight against counterfeiting, fraud, and cross-border economic crimes.

The IoT sees its roots back to the M2M to control the remote production process. European Telecommunications Standards Institute advocates an evolution of the M2M paradigm toward the IoT. This architecture allows the coexistence of the various current and future technologies that enter the IoT development landscape. We find the different technologies of the interconnection of objects (M2M, RFID, IEEE802.15.4, 6LowPAN), and gateways to core transport networks [12, 13].

In the area of M2M applications and client applications, we find M2M platforms, middleware, and application programming interface (API) APIs for M2M applications, business processes using the IoT, and soforth. An immediate needis the development of effective security mechanisms for new technologies. Current results in sensor networks, actuators, RFID technology, and mobile computing show the limits of the devices that will constitute the IoT in terms of resources and capabilities [14]. It is needed to verify the applicability of modern cryptography in the context of IoT. Indeed, the limitations of resources and capacities of objects make it difficult to use current cryptographic algorithms because of theircomputationandmemoryconsumption[15,16].The emergence of robust and inexpensive cryptography in terms of resources combined with technological advances would make it possible to overcome these difficulties in the medium term. Indeed, several research studies have shownthat cryptography based on elliptic curves offers a level of security robustness similar to conventional asymmetric cryptography with the advantage of being inexpensive in terms of resources. Besides, one of the technologies recommended at the Internet Engineering Task Force (IETF) for the interconnection of IoT networks is IPv6 [17].

One of the significant advantages is the exploitation ofthe immense addressing capacity of 128 bits of IPv6, which would meet the very large-scale addressing needs of an IoT, potentially containing tens of billions of objects. However, the objects' limited resources, particularly in terms of energy and the environment of intermittent connectivity (also called LLN: Low power lossy networks), make this technology difficult [18]. One of the avenues explored today for communication in low energy and low connectivity LLN environments is adapting IPv6 to these environments through a series of protocols such as 6LowPAN and RPL. 6LowPAN is a working group at the IETF responsible for adapting IPv6 technologyto PAN with low energy LLN connectivity. Among the results of this working group, 6LowPAN technology (IPv6 for PAN) and its adaptation to the IEEE802.15.4 standard widely used in PAN and recommended to be an integral part, among other technologies, of IoT[19]."Routing Over Low power Lossy Networks" (ROLL) is another working group at the IETF whose role is to address the problem of routing in LLN networks. An extensive study reveals various well-known ad hoc routing protocols, suchas ad hoc on-demand distance vector, optimized link state routing protocol, and ad hoc

on-demand distance vector, are not suitable for the LLN networks. As a result, ROLL offers a new protocol for these specific environments called RPL: IPv6 routing protocol for low power and lossy networks. The resource constraints of the LLN environments in which IoT objects will require new solutions adapted for the securityof exchanges and potential threats. The two groups at the IETF 6LowPAN and ROLL are taking a close interest in these security problems and developing a specific "framework" for security in LLN environments[20].

3. RELATED WORK

Cui et al.[21] proposed hybrid Blockchain based multi sensor network authentication method in distributed system of IoT network. This this structure, a private block chain is constructed between cluster heads in a single WSN, and base stations of all WSN are added to the public block chain. A hybrid block chain model is constructed between the whole networks. In this model, the identity information registration between cluster head nodes and ordinary nodes and communication authentication between nodes are completed.

To detect malicious node in WSN, She et al.[22] proposed a Block chain based trust model. Through 3D space it is realized by using block chain intelligent contract and WSNs quadrilateral measurement for localization of the detection of malicious nodes in, and the consensus results of voting are recorded in the block chain distributed. The simulation results show that the model can effectively detect malicious nodes in WSNs and ensure the trace ability of the detection process.

Goyat et al. [23] implemented a trust-based range-free secure localization algorithm is successfully for WSNs. Block chain technology is used for sharing the evaluated trust values of beacon nodes with neighboring beacon nodes. The highly trusted beacon nodes are selected as a miner for mining process of blocks so that unknown nodes can get trustworthy information from highly honest beacon nodes to perform the localization process.

Nguyen et al.[24] developed energy efficient and secure clustering based data transmission in pervasive wireless net- works using RDAC-BC technique. The RDAC-BC technique is initialized primarily and then clusters are produced using the derived fitness function. The RDAC model includes a fitness function, which is solely based on the energy, power density as well as node density, distance to nearby nodes, and distance to sink. When the CH election process is completed, blockchain based secure communication process takes place. The experimental validation of the RDAC-BC technique is assessed under several aspects and the results are compared with existingmethods.

4. PROPOSED METHOD

Blockchain system is based on a ledger that keeps track of all the transactions circulating in a network. Thus, as we need some way to figure out which nodes are transmitting and through which path, we will store the active paths in real-time as transactions in the Blockchain. To achieve this, we treat the network's nodes as coins. More precisely, when some nodes are carrying a message from a source node to the sink, their ownership will be affected by the source node. At the first stage, all the nodes are owned by the sink.

Each node that is owned by the sink is considered to be inactive. Otherwise, all the nodes which are not owned by the sink are considered active. When a node senses some event, it looks up in the Blockchain and defines a list of all inactive nodes; then, it finds among them which ones optimize its path to the sink. The route selection process will be selected on the basis of the Blockchain.

Next, it asks the sink to transfer the path's nodes ownership to it. Once the transaction is registered to the Blockchain, the node starts transmitting over the chosen route. When the data is carried successfully to the sink, the transmitting node transfers back the ownership of the path's nodes, including itself, to the sink to inform the network's peers that the transmission was finished and these nodes were released. A source node is assumed to be u nodes in the network while u < n. The nodes are assumed transmit data into two channels. The first one is dedicated to the paths claiming and to the Blockchain transactions transferring. The second one is designated to carry the sensed data. We are interested, primarily, in the second channel, which is used to transmit the message. We suppose that each intermediate node could be owned only by one source node, and a source node is owned only by itself. The node notifies the sink, through the first channel, in order to be added to a waiting queue. The sink mainly manages the waiting queue, and it is necessary to apply a kind of priority to the waiting nodes.

As we have seen, this technique allows for a good knowledge of the source nodes and the paths they transmit on at agiven moment. It is also necessary to mention that the nodes are represented in the Blockchain by their Ids. Hence, the traffic load could be easily determined through the Blockchain. Actually, it suffices to determine, directly from the chain, how many times a node's status has changed to be active. This change number is the number of messages carried by a node since a node status changes only when it is in the path on which a message is transmitted. After we defined the traffic load at each node, we have to explain our model's routing determination process.

As each node knows the network's map and can access the Blockchain and find which nodes are transmitting and which nodes are not, it becomes simpler to define the shortest path to the sink through a set of inactive nodes. However, as said previously, our primary goal is to balance the traffic load and to reduce the interferences in the routing phase. So, we have to define a cost function that optimizes the path. First of all, let us explain the signal and interference to noise ratio (SINR) as,

$$SINR_{(i,j)} = \frac{\frac{Pi}{d^{a}_{i,j}}}{N_0 + \sum_{k=1,k\neq i}^{n} \frac{P_k}{d^{a}_{k,j}}} \quad ...(1)$$

N_0= Signal Noise,
P_i= Signal Power of i^{th} Node
$d^{a}_{i,j}$ = Distance between i^{th} Node to j^{th} Node

5.RESULT ANALYSIS

To validate our work, we simulated a network of 1500nodes that contains one sink. The system is connected components in distributed graphs, as shown in Figure 1. The simulation was done using Python3, NumPy, and Matplotlib. First of all, we created a primary Blockchain mechanism, which response to our work's requirements regarding transaction structure and cryptographic tools.

We assumed that the nodes are broadcasting at a transmission power of 1 mW and we considered that the path loss and the noise are 2 and 3. Next, to verify the proposed mechanism's ability to balance the traffic load, we assumed that only one node is transmitting several messages over the network. To evaluate whether there is an enhancement or not, we applied the shortest path routing protocol to the same scenario, and we analyzed the returned data.

Figure 2 shows a comparison of the traffic load when only one node is transmitting, using the shortest path routing protocol against our proposed approach.

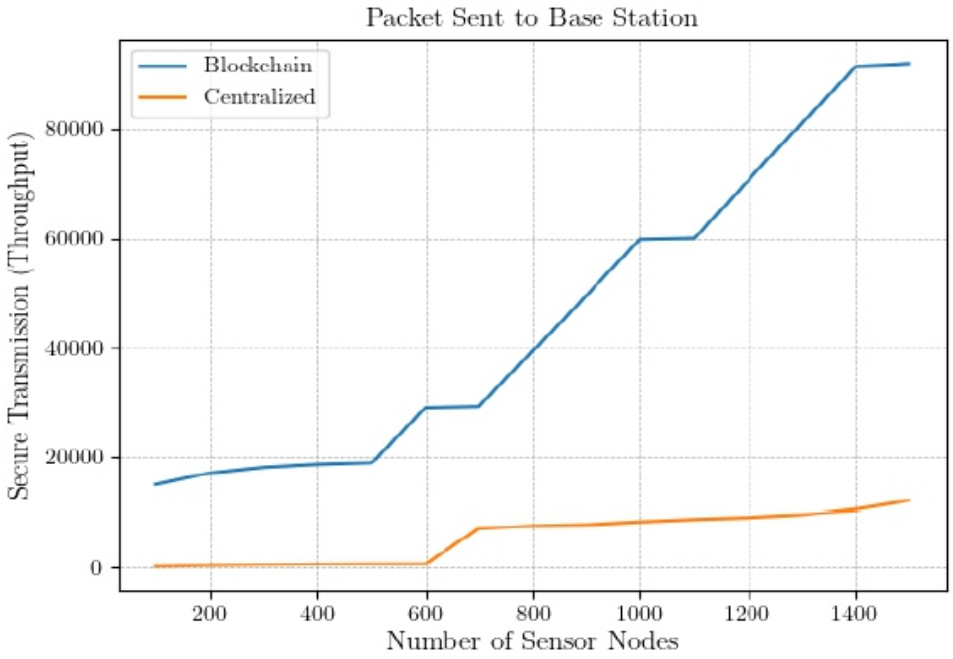

Fig.2- Packet sent to Base Station

6.CONCLUSION

In this article, optimal routing protocol in WSNs based on Blockchain has been discussed. The technology is relatively new. It was first known in the Blockchain domain, and lately, it has been employed in many fields where it has proven its efficiency. To achieve a better solution to the traffic load unbalance, the high interference levels, and the security issues, we proposed a protocol that takes advantage of the Blockchain technology benefits. The approach consists of using Blockchain as a shared memory between nodes and storing all the network's activities on it. The nodes are considered coins owned by the sink when they are inactive or owned by the source node if they are carrying some message. Moreover, a cost function is proposed to optimize the chosenpath. Regarding the protocol's security, it is granted by the Blockchain. Simulation results show that the proposed protocol can improve the traffic load balance, decrease the interference levels, and strong guaranty security during the routing phase.

REFERENCES

1. *V. Rajaram, Z. Qian, S. Kang, N. E. McGruer, and M. Rinaldi, "Mems-based near-zero power infrared wireless sensor node," in 2018 IEEE Micro Electro Mechanical Systems (MEMS), pp. 17–20, 2018*
2. *C. He, M. E. Kiziroglou, D. C. Yates, and E. M. Yeatman, "A mems self-powered sensor and rf transmission platform for wsn nodes," IEEE Sensors Journal, Vol. 11, No. 12, pp. 3437–3445, 2011. doi: https://doi.org/10.1109/JSEN.2011.2160535*
3. *S. C. Polavarapu and S. K. Panda, "A survey on industrial applications using mems and wsn," in 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 982–986. doi: 10.1109/I-SMAC49090.2020.9243514*
4. *H. Wang, L. Dong, W. Wei, W. Zhao, K. Xu, and G. Wang, "The wsn monitoring system for large outdoor advertising boards based on zigbee and mems sensor," IEEE Sensors Journal, Vol. 18, No. 3, pp. 1314–1323, 2018. doi:https://doi.org/10.1109/JSEN.2017.2770324*
5. *A. Heragu, D. Ruffieux, and C. Enz, The Design of Ultralow- Power MEMS-Based Radio for WSN and WBAN. Cham: Springer International Publishing, 2014, pp. 265–280. doi:https://doi.org/10.1007/978-3-319-01080-9 15*

6. L. Feng, H. Zhang, L. Lou, and Y. Chen, "A blockchain- based collocation storage architecture for data security process platform of wsn," in 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD)), 2018, pp. 75–80. doi: https://doi.org/10.1109/CSCWD.2018.8465319
7. Y. Zeng, X. Zhang, R. Akhtar, and C. Wang, "A blockchain- based scheme for secure data provenance in wireless sensor networks," in 2018 14th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN), 2018, pp. 13–18. doi: https://doi.org/10.1109/MSN.2018.00009
8. S. Ali, T. Al Balushi, Z. Nadir, and O. K. Hussain, Wireless Sensor Network Security for Cyber-Physical Systems. Cham: Springer International Publishing, 2018, pp. 35–63. doi: https://doi.org/10.1007/978-3-319-75880-0 3
9. M. Almasri, K. Elleithy, A. Bushang, and R. Alshinina, "Terp: A trusted and energy efficient routing protocol for wireless sensor networks (wsns)," in 2013 IEEE/ACM 17th International Symposium on Distributed Simulation and Real Time Applications, 2013, pp. 207–214. doi:https://doi.org/10.1109/DS-RT.2013.30
10. A. Araujo, J. Blesa, E. Romero, and D. Villanueva, "Security in cognitive wireless sensor networks. challenges and open problems," EURASIP Journal on Wireless Communications and Networking, Vol. 2012, No. 1, p. 48, Feb 2012. doi:https://doi.org/10.1186/1687-1499-2012-48
11. S. Kushch and F. Prieto-Castrillo, "Blockchain for dynamic nodes in a smart city," in 2019 IEEE 5th World Forum on InternetofThings(WF-IoT),2019,pp.29–34.
12. K. Sohraby, D. Minoli, B. Occhiogrosso, and W. Wang, "A review of wireless and satellite-based m2m/iot services in support of smart grids," Mobile Networks and Applications,Vol.23, No.4,pp.881–895,Aug2018.doi:https://doi.org/10.1007/s11036-017-0955-1
13. R. Ahmed and M. A. Matin, Cognitive M2M Communications: Enablers for IoT. Cham: Springer International Publishing, 2020, pp. 67–79. doi:https://doi.org/10.1007/978-3-030-42573-9 6
14. S. Iraji, P. Mogensen, and R. Ratasuk, "Recent advances in m2m communications and internet of things (IoT)," International Journal of Wireless Information Networks, vol. 24, no. 3, pp. 240–242, Sep 2017. doi:https://doi.org/10.1007/s10776-017-0362-3
15. R. Prasad and V. Rohokale, Internet of Things (IoT) and Machine to Machine (M2M) Communication. Cham: Springer International Publishing, 2020, pp. 125–141. doi:https://doi.org/10.1007/978-3-030-31703-4 9
16. D. Seo, Y.-B. Jeon, S.-H. Lee, and K.-H. Lee, "Cloud computing for ubiquitous computing on m2m and iot environment mobile application," Cluster Computing, vol. 19, no. 2, pp. 1001–1013,Jun 2016. doi: https://doi.org/10.1007/s10586-016-0573-x
17. C. Cath and L. Floridi, "The design of the internet's architecture bytheinternetengineeringtask force(ietf) and humanrights," ScienceandEngineeringEthics,vol.23,no.2,pp.449–468, Apr 2017. doi: https://doi.org/10.1007/s11948-016-9793-y
18. S. Shorgin, K. Samouylov, Y. Gaidamaka, A. Chukarin, I. Buturlin, and V. Begishev, "Modeling radio resource allocationschemewithfixedtransmissionzonesformultiservice m2m communications in wireless iot infrastructure," inIntelligentInformationandDatabaseSystems,N.T.Nguyen,B.Trawin´ski,andR.Kosala,Eds.Cham:SpringerInternational Publishing, 2015, pp. 473–483. doi: https://doi.org/10.1007/978-3-319-15705-4 46
19. K.L. Chen and F. J. Lin, "Ensuring iot/m2m system security under the limitation of constrained gateways," in Security, Privacy, and Anonymity in Computation, Communication, and Storage, G. Wang, M. Atiquzzaman, Z. Yan, and K.-K. R. Choo, Eds. Cham: Springer International Publishing, 2017, pp. 302–311. doi: https://doi.org/10.1007/978-3-319-72395-2 28
20. H. Lazrag, A. Chehri, R. Saadane, and M. D. Rahmani, "Efficient and secure routing protocol based on blockchain approach for wireless sensor networks," Concurrency and Computation: Practice and Experience, p.e6144. doi:https://doi.org/10.1002/cpe.6144
21. Z. Cui, F. XUE, S. Zhang, X. Cai, Y. Cao, W. Zhang, andJ. Chen, "A hybrid blockchain-based identity authentication scheme for multi-wsn," IEEE Transactions on Services Computing, vol. 13, no. 2, pp. 241–251, 2020
22. U. Kumbhalkar & S. Gangele ,"Threshold-Based Energy-Estimated Distributed Routing Algorithm in Wireless Sensor Network", International Journal of Scientific Research in Science, Engineering and Technology, Vol. 4, No.9, pp.614-622, 2018
23. U. Kumbhalkar& S. Gangele,"Multi-path and Multi-hop Energy Efficient Routing in Wireless Sensor Network", International Journal of Computer Applications, Vol.178, No.8, pp. 1-6, 2019
24. U. Kumbhalkar& S. Gangele, "Heterogeneous Multi-Clustered Energy Efficient Routing protocol in Wire sensor Network", International Journal of Future Generation Communication and Networking , Vol. 13 No.4, pp. 4007-4014, 2020

Predictive Analysis of Student Performance using Educational Data Mining

Kirti Soni[1], Dr. Sharad Gangele[*]

[1]Research Scholar, Department of Computer Science & Application R.K.D.F. University, Bhopal (MP), India

[*] Professor, Department of Computer Science & Application R.K.D.F. University, Bhopal (MP), India E-mail: sharadgangele@gmail.com

ABSTRACT

This paper examines the factors of course study in various subjects which affect the persistence and graduation outcomes of over 1000 students in different villages of the Madhya Pradesh, and their control group counterparts. This work addressed four questions: (1) how did the timing of student's Mathematics, Biology, Commerce, Arts and Agriculture courses affect their performance, persistence, and graduation outcomes; (2) whether students who progressed farther through the prescribed foundation course sequences of the program exhibited higher persistence and graduation rates; (3) what were the most frequently taken sequences of courses, and whether students who progressed farther through those sequences exhibited higher persistence and graduation rates; and (4) whether greater progress was more important than other demographic and academic factors for predicting persistence and graduation. We found that students who took their Math course in the second year showed higher fifth-term and seventh-term persistence than students who took it in the first year. Also, students who progressed farther through course sequences consistently exhibited higher persistence and graduation rates. Furthermore , a student's persistence was a more reliable predictor of graduation than other features. Overall, these findings can potentially inform an institution's strategies for maximizing persistence and graduation by emphasizing a student's progress through the curriculum.

Keywords: Educational Data Mining, Predictive Analysis, Higher Education.

1. INTRODUCTION

Higher education institutions face ongoing pressure to maximize their retention and graduation rates despite constrained resources. The graduation rate from the first institution at- tended by first-time, full-time, bachelor's degree-seeking students at 4-year institutions was 40.6% nationally for the 2010 starting cohort [1]. Retention of first-time bachelor's degree- seeking undergraduates at degree-granting 4-year institutions has also remained a challenge, with the percentage of first time undergraduates retained for their second year at 75.3% nationally for the 2015 starting cohort [2].

The applications of educational data mining and machine learning techniques to student data have proliferated rapidlyin recent years as the availability of such data has increased [3–5]. Among many other questions, increasing attention has been devoted to predicting student performance, specifically persistence and graduation outcomes of students based on various demographic and academic factors. Data mining [5–7]andmachinelearning[4,8]havebeenemployedtofindthe strongest predictors of a student's outcome, as well as to help identify intervention methods that have a positive impact on that outcome. However, the impacts of certain factors that operate at the curriculum level rather than at the student level have been less extensively researched. Specifically, the impact of the ordered sequence of courses taken by a student on the likelihood of that student to persist and graduate has not been studied in depth [9]. Furthermore, it has been reported that successful completion of Mathematics courses during

the first two years has a significant impact on degree completion, with more than 70% of students who complete their degree having completed the required Mathematics courses during the first two years [9]. These situations are not only for Mathematics as it applies in the courses of the University.

Institutional efforts to improve retention and graduation have recently been directed at the curriculum level, rather than at the student-specific level. Learning communities represent one such effort, and have been implemented at many institutions with the objective of improving student's integration into university life [10,11]. An emerging trend is the evolution of the learning community into the long-duration learning community, which offers academic momentum, integration into university life, and access to timely support services over several semesters at the start of a student's college education [12].

The challenges of retention and graduation are especially acute for under-represented, under-privileged, and/or first- generation students. For example, graduation rates for black and Hispanic students from the first institution attended for first-time, full-time, bachelor's degree-seeking students at 4- year institutions was 21.4% and 31.7% nationally for the 2010 starting cohort, compared to 45.2% for white students [1]. One long-duration learning community program, the Metro College Success Program ("Metro") at various Universities[13],is endeavoring to narrow this gap by concentrating its efforts on these students. The Metro program groups its participants into different academies based on the student's anticipated major or area of interest (e.g., science, ethnic studies, engineering, among others). A novel feature of the Metro program is that it incorporates a carefully-designed sequence of scaffolded foundation courses, to be taken during the first two yearsofcollege[13].

A survey conducted in some rural area sin Madhya Pradesh, India and subject-wise data of boys and girls were collected as per their choice, preference and interest for higher studies, and Table I represents the subject-wise survey report for higher studies collected from boys and girls in rural area. 2. Related Work Educational data mining (EDM) has proliferated rapidly in recent years [3,14]. The users of EDM applications include learners, educators, course developers, institutions, and administrators. In which some of the educational tasks for which EDM has been applied, and lists some of the techniques that have been applied to each task type [4].It should be noted that the Table does not do justice to the tremendous variety of techniques that have been used in tackling EDM problems, and as such it should not be interpreted as an exhaustive treatment of the subject.Some other similar contribution was given by Gangele et al.(2020 a, b)in which data mining based study was elaborated and derived some interested results on it [25-26].

Gangele et al.(2018) also suggest some survey result for student assessment integrated with higher education [24]. In a comprehensive meta-analysis, Romero and Ventura[4] provide an excellent introductory look at EDM and its applications. The authors describe the numerous unanswered questions and areas where EDM can improve and grow. The authors give very brief descriptions of over two hundred papers in this survey, broken down both by data types, by tasks, and by data mining techniques used. It is clear that EDM, despite being in its adolescence according to the authors, has seen a great many publications and applications as of 2010. As such, this paper provides a very useful high-level view of the EDM landscape.

A more recent meta-analysis of EDM research [14] provides some insights on the current trends and future applications of EDM for improving education. The modeling of student behavior and performance has experienced tremendous growth in the last few years, as researchers strive to identify

criteria that place certain students at risk of poor class performance (e.g., reduced attendance or a failing grade) [14, Sections 3.1-3.2]. Likewise, scholarship of the best algorithms and modeling techniques has increase rapidly since 2010 [14,Section 4]. Thus, the ongoing proliferation of EDM studies include work both to improve education and to improve the tools and approaches to analysis that are employed by EDM scholars. As described in more detail in the following paragraphs, a subset of the broad landscape of EDM research described in[4] and [14] involves prediction of student performance and long-term outcomes, such as persistence and graduation. Many of the features considered in generating predictive models include demographic and personal factors, such as age, race, gender, and family educational history (e.g., highest education level attained by the mother and father of a student) [7].Another category of student features relates to the activities of the student, such as whether and how much students participate in extracurricular activities and the extent of their social networks. Additional factors such as student interests, study behavior, and family support have been included in some analyses [7, Section 3.1]. Lastly, student performance in their courses, asme assured by either cumulative GPA or the grade points earned in particular courses, is another frequently-used factor in predicting long-term student outcomes.

The work reported by Asif [5] exemplifies the use of EDM to improve a curriculum through the use of student performance data. The problems addressed in the work were to identify courses as indicators of a student's overall, curriculum- long performance, and to identify progressions of student performance up to and including the indicator course, in an effort to be stunderst and student out comes. The study analyzed the performance of students seeking a four-year degree in the Information Technology major of a Pakistani university. The available dataset included the grades of 210 Information Technology-majoring students in the 2007-2008 and 2008- 2009 entering cohorts in a few selected courses (specifically, five required courses) taken by those students [5,Table 1]. Thus, performance in each of those courses by a student comprised the input feature set for that student. The response variable was coded as a single value, computed as 10% of the first-year average examination mark, 20% of the second-year averagemark,30%ofthethird-yearaveragemark,and40%of the fourth-year average mark. This value was then transformed into an ordinal categorical variable with five levels. One disadvantage of the dataset was its unbalanced nature: only one student in the entire dataset achieved the highest category out of the five ordinal categories of the response variable; over 80% of the students in the dataset had the second or third out of five response variable values (i.e., over 80% of the students in the dataset were in level two or level three out of five levels of the response variable). Naive Bayes, random forest, neural network, and 1-nearest neighbors classifiers were used to predict the aforementioned response variable for each student, with naive Bayes representing the most accurate model. The random forest model revealed evidence that low performance in Applied Physics in the first year, or in Logic Design in the second year, were the courses most strongly predictive of poor long-term performance. However, the paper did not report any feature elimination strategies for any of the models, which would have been useful for determining which features were most important in reaching the highest classification accuracy scores reported.

More recently, Saa [6] investigated the problem of determining the effect of social and personal factors of students, as well as the effect of their academic performance in a certain semester, on the student's performance in the following semester.

TABLE-1 SUBJECT-WISE SURVEY REPORT FOR HIGHER STUDIES

S.No.	Village Name	Arts		Math		Biology		Agriculture		Commerce		Total
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	
1	THUNIYA	2	2	0	0	1	2	1	0	0	0	8
2	GOURIYA	2	4	0	0	1	3	0	0	1	1	12
3	MOHKHEDMAU	0	2	4	3	2	3	0	1	0	0	15
4	MOHKHED	1	9	4	6	1	1	4	3	0	0	29
5	BISAPUR	3	8	3	3	1	4	9	9	1	0	41
6	UMRANALA	3	6	2	4	10	3	6	5	0	7	46
7	SHIKARPUR	1	2	1	1	0	5	0	1	1	2	14
8	IKALBHIRI	0	1	0	0	0	0	0	0	0	0	1
9	KHRWADA	0	2	0	0	0	0	1	0	0	0	3
10	KHUNAJHI	2	8	5	0	0	0	1	0	0	0	16
11	R KALA											
12	LINGA	4	5	3	0	1	3	0	0	2	5	23
13	TIWDA KAMATH	0	2	0	0	0	0	0	0	0	0	02
14	MURMARI	0	1	0	0	0	0	2	0	0	0	3
15	PALAMAU	0	1	0	0	0	0	0	0	0	0	1
16	SHIVLALDANA	1	0	1	0	0	0	2	0	0	0	4
17	JAITPUR	0	1	0	0	0	0	0	0	0	0	1
18	KUKDA CHIMAN	0	0	1	0	0	0	2	0	0	0	3
	IMLIKHEDA	0	0	1	0	0	0	0	0	0	0	1
	Total	19	54	25	17	17	24	28	19	5	15	223
		73		42		41		47		20		

As such, the study sought to develop predictive models for student performance in the short-term, i.e., the next semester. The dataset consisted of voluntary surveys completed by 270 students, across different departments of a four-year university in the United Arab Emirates. The survey gathered information about demographic (gender, race, national origin, living location, commute length and transport mode), economic (financial aid, family income) and personal (language spoken, family size, parents' marital status and careers) factors, and combined these with the student's high school GPA and the GPA in the semester prior totaking the survey ([6,Table 1]). All of these factors were coded as categorical variables and used as input features for decision trees that employed different feature importance algorithms: C4.5, CART, CHAID, and ID3. No hyper-parameter tuning wasreportedforanyofthemodels,whichwerethencompared with a naive Bayes classification model. For the response variable (GPA in the next semester), four ordinal categorical values were created and used (Excellent, Very Good, Good, Pass) as the outputclasses. The classification models achieved accuracy between 33% and 40%, with the CART implementation of the decision tree model yielding the highest accuracy. However, the work does not report any feature elimination strategies or any attempt to identify the most important features used by the models. The authors interpreted the conditional probability distribution matrix as indicating that gender, high school grades, mother's occupation status in a service profession, and whether a student is a scholarship recipient are the factors with the highest prior probability for those students with very good or excellent grades in the next semester (the response variable). The authors did not share any insight as to how these findings can lead to an improved curriculum, or how they should inform whether and when the university conducts interventions with certain students. Moreover, it would have been interesting to see the results of feature selection or elimination efforts, so as to determine whether the accuracy improved by leaving out features with minimal or no importance to the classifications. The next section focuses on a technique, sequential patternmining, that has emerged relatively recently in the EDM landscape but is seeing increasing use.

3. APPLICATIONS OF SEQUENTIAL PATTERN MINING

Sequential pattern mining is well-suited for addressing questions wherein the data can be characterized as an ordered collection of events or actions [15]. In the majority of applications of sequential pattern mining for educational data, the actions under study involve learning behavior, such as when students participate in discussion forums or view class materials [16,17] or when students perform various actions in collaborative problem-solving environments [18–20].The next few paragraphs describe examples of these applications in more detail.

As an interesting example of the use of sequential pattern mining to understand learner behavior, Martinez [18] tackled the problem of determining frequent patterns in the actions of primary-school learners in collaborative settings. The dataset included the log traces of a problem-solving tabletop application for use by elementary school students between the ages of eleven and fourteen. The table top application presented each of the six groups of three students each with a question on any subject (e.g., math, history or physics) and the information necessary to solve the question. The tabletop application’s log trace recorded all of the transactions between the students and the system. The dataset collected over 17,000 distinct actions performed by the six groups of students over twelve different logged sessions. There were seven distinct possible actions: (1) move; (2) enlarge; (3) reduce to normal size; (4) shrink; (5) add to a group; (6) remove from a group; and (7) combine contents of two screens into one, so that both are visible at the same time. Each of these actions was performed on digital screens that contained specific contents identified by a unique label; hence a certain action was a tuple consisting of (student-identifier, action, content-identifier, time- of-action). Within each session, these actions took place in an

ordered sequence, making them amenable to sequential pattern mining. Students were categorized into high-achieving and low-achieving groups based on their success in solving the questions.

Using the foregoing dataset, the authors used sequential pattern mining to find the most frequent sequential patterns of actions. The authors also associated those sequential patterns with the students' achievement level. The results of clustering revealed interesting insights about the sequences of steps undertaken by more successful students and by less successful students [18,Table 1]. For example, the students who performed fewer of the "combine" actions (i.e., students who combined multiple information screens with less frequency) performed better than students who performed many "combine" actions. The authors interpreted this finding as suggesting that students who let their attention be drawn to too much information at a time have greater difficulty in solving the problem. Thus, this study provided greater transparency to the learning process of the student participants and revealed interesting insights about how groups of students solve problems.

Perera [20] utilized a similar problem-solving approach with university seniors in a software engineering course. The authors addressed the problem of identifying which actions and resources in the software development process contributed most to the students' successful realization of the course goal: to build an effective software application by the end of the course. More specifically, the students, working in groups of five to seven, were tasked with developing a software application for a client. The students were required to usea development tracking system that included the SVN version control system, a ticketing system, and a group wiki for shared web pages [20, Section 3.2]. The available dataset thus consisted of all events in the activity log of the tracking system. Those events included: commits to the SVN repository; the creation, modification, or removal of the wiki pages; and the creation and resolution of bug tickets. The dataset included between 1,400 and 2,500 distinct tracked events for each of seven student groups. Lastly, the authors categorized each group as successful or unsuccessful.

Using the foregoing dataset, the authors performed clustering and sequential pattern mining. In the clustering phase, the authors determined that the most successful groups made the most extensive use of the wiki pages while resolving tickets. The authors subsequently incorporated this finding into future teaching of the course. For sequential pattern mining, the authors used a modified form of the a priori algorithm for frequent item sets [20, Section 6.1]. The authors translated the raw sequence data from the tracking system into sequences of character strings usable by the Weka data mining tool [20, Table 9]. By associating frequent sequential patterns with the course outcomes of the groups who performed those patterns, the authors found that the best-performing groups had the highest frequency of alternating SVN-wiki events (in other words, these groups had the highest frequency of accessing the wiki pages before and after making commits to the SVN repository). Meanwhile, the lowest-performing groups had a high frequency of wiki access but a low frequency of SVN commits. The authors concluded that for the low-performing groups, the students' efforts at understanding the wiki pages were not being used to support software development. Thus, through the use of sequential pattern mining, the authors gained actionable insights into how to better teach the course and how to better advise students on the best ways to complete their development projects.

Massive online open courses (MOOC) have represented another context for sequential pattern mining research, since the learning actions of the student participants in those courses are saved in activity logs. The study reported by Jiang [16] is illustrative. The problem that the authors address is to create a way to

better manage forum contents by matching video clips viewed at certain times with forum threads whose contents are edited at similar times. Thus, this work investigates the association between forum threads and subtitles of video clips that are available as learning resources to the students. The available dataset included the subtitles of video clips, discussion forum contents, and learners' click-stream logs associated with a Coursera course and a course of China University (a leading MOOC platform in China). The courses had over 3,000 and over 10,000 student participants respectively. The authors indicate that only a limited set of labeled data was available, and therefore they consider the work an unsupervised learning exercise and do not seek to generate any classification or regression models. The labeled data was used to evaluate several different approaches for clustering related forum contents together based on the video clips and click- stream events. The authors' proposed approach utilizes the idea that the order of click-stream events associated with the learners' viewing of videos, reading threads, or posting threads can reflect document-level latent similarity [16, Figure 1]. The authors reveal that for the Coursera course, their proposed approach yielded higher P@1, P@3, and P@5 precision scores than bag-ofwords, Word2Vec, or Para2Vec models against which their approach was compared. (Recall and F1 scores were not reported.) For the larger dataset (10,000 students), their approach offered lower precision than the other models at P@3,andP@5.Overall,thisworksuggeststhepotentialutility of sequential pattern mining as another option for revealing insights from the activities of learners. Previous works have employed sequential pattern mining in analyzing educational data in various contexts. As explained in more detail in the preceding chapter, a growing body of work has applied sequential pattern mining to the activities of learners in massive online open courses (MOOC) [16]as well as to how groups of students at the elementary school [18] and university [20] levels collaborate to solve problems and complete complex projects. These studies demonstrated that sequential pattern mining can lend transparency to the learning process by uncovering the frequent patterns in which learning actions occur, and those patterns can then treated as features when analyzing how successful the students were in completing their tasks.

However, the application of sequential pattern mining to sequences of courses taken by students during their academic careers is still very much an emerging research direction. Campagni [21] applied sequential pattern mining to address the problem of how the timing of computer science students' completion of course examinations affected students' graduation outcomes. In the higher education institution under study, students were allowed to take the examination in a different semester from which the course is taken, so that the order of examinations is not necessarily the same as the order of courses.

The dataset consisted of the courses taken by 141 graduate students in the Computer Science department of an Italian university. Each student's academic career was thus represented as a sequence of examinations ordered by semester, and examinations taken during the same semester comprised a non-ordered set.

For example, a student's academic career might be represented by the ordered sequence of examinations<$\{2,3,4\} \rightarrow \{1,4\} \rightarrow \{5,6\}$. This sequence represents that the student took the examinations for courses 2, 3 and 7 in semester 1; examinations for courses 1 and 4 in the second semester; and examinations for courses 5 and 6 in the third semester [21, Section 3]. This work assumed the existence of an ideal sequence, in which each student takes the examination at the end of the semester in which the associated course was taken. The study used sequential pattern mining to generate each student's sequence, and to compute each student's deviation from the ideal sequence using bubble sort distance. Using this approach, the study revealed that greater deviation from the ideal sequence of

examinations (i.e., taking more examinations in semesters other than the ones in which the corresponding courses were taken) led to decreasing likelihood of graduation. The authors also used the sequential patterns to generate new features. For example, students who took the examinations for courses 1 and 5 at the same time were given a positive value for a new feature reflecting that those students exhibited the {1, 5} sequential pattern [21,Table 8]. The authors used these features to identify particular examinations that the students delayed taking the longest, leading to the largest separation in time between the underlying course and the examination. These findings provided insights for the authors for advising students not to delay the examination in that particular course.

As exemplified in [21], sequential pattern mining offers the ability not only to extract sequential patterns from course sequences, but also to generate new features that can be added to existing feature sets to provide a new dimension to the analysis of student data for predicting persistence and graduation. Recently, Gopala krishnan [8] addressed the problem of identifying useful features for the prediction of fifth-term persistence, seventh-term persistence, and graduation for the student participants in a long-duration learning community program, the Metro College Success Program ("Metro") at San Francisco State University [13]. The Metro program, while open to any incoming first-time, full-time freshmen, targets students from under-represented minorities and disadvantaged socioeconomic and personal backgrounds, as reflected by Pell eligibility and first-generation student status. Students of color exhibit lower persistence [2] and graduation [1] rates. Metro's mission is to bolster these rates by providing enhanced services to these students during their first two years of college. The enhanced services include: reminders and support during course registrations; academic advising; access to a tutoring center; and frequent coordinating and counseling interactions. Metro also emphasizes learning communities, to instill a sense of belonging among students [10,17].

The dataset in [8, 22] consisted of twelve different features for 651 students in 2009-2013 entering cohorts of the Metro program. The available features included Pell eligibility; first- generation status; EOP status; ELM and EPT entrance scores; household income; education level of each parent; race; start term; department; and gender, among others. The study applied different feature selection algorithms to the dataset to identify those features most strongly predictive of the third-term, fifth- term, and seventh-term persistence and graduation outcomes of the students [8,Table 3]. ELM score or the mother's education level were found to be the most important features for predicting graduation by six out of the seven feature selection algorithms. This work also developed and tested classification models based on Adaboost, extratrees,K-nearest neighbors, linear SVC, and naive Bayes classifiers [22].The study revealed the difficulty of predicting graduation based solely on personal and demographic factors: the algorithm that achieved the highest accuracy was naive Bayes, with an accuracy score of roughly 66% [22, Figure 16, top]. The algorithm with the lowest classification accuracy was extra trees, with an accuracy of 54%.

Overall, these findings demonstrated that a student's university career is a complicated phenomenon, and the student's long-term outcome is not easily predicted based on demo- graphic or personal factors about a student. Consequently, this work seeks to leverage sequential pattern mining in order to find curriculum-level factors (i.e., factors that operate over multiple-semester course sequences) that can be added to the personal and demographic student attributes studied in [8] to yield not only improved classification models for predicting persistence and graduation, but also new insights about the importance of multiple-semester course sequences and how they affect a student's long-term outcome.

This study extends the aforementioned applications of sequence pattern mining to the courses taken across multiple semesters by the students at San Francisco State University to determine whether the most commonly-observed sequences offer better persistence or graduation for the students who take those sequences. The study also extends the study of Gopala krishnan [8] by making use of a larger student dataset that now includes the years up to and including 2016 as well, and as such, this research work covers many more students over a longer period. Furthermore, this the study considers features that capture the extent by which students progress through their chosen course sequences, extracted using sequential pattern mining techniques as described next.

4. PROPOSED METHOD

This section describes the problem-solving methodologies that were employed to address the aforementioned four problems, which are:

a) The timing of students' completion of certain Math courses, and its effect on persistence and graduation.
b) The effect of increased progress through the Metro program's sequence of three foundation courses on persistence and graduation.
c) The extraction of the most frequently-taken course sequences and the effect of increased progress through those sequences on persistence and graduation.
d) Determination of the relative importance of sequence progress as a feature, compared to demographic and personal features, in predicting graduation outcomes.

Figure 1 depicts the software modules used to address these problems, each of which is discussed in turn in the remainder of this section.

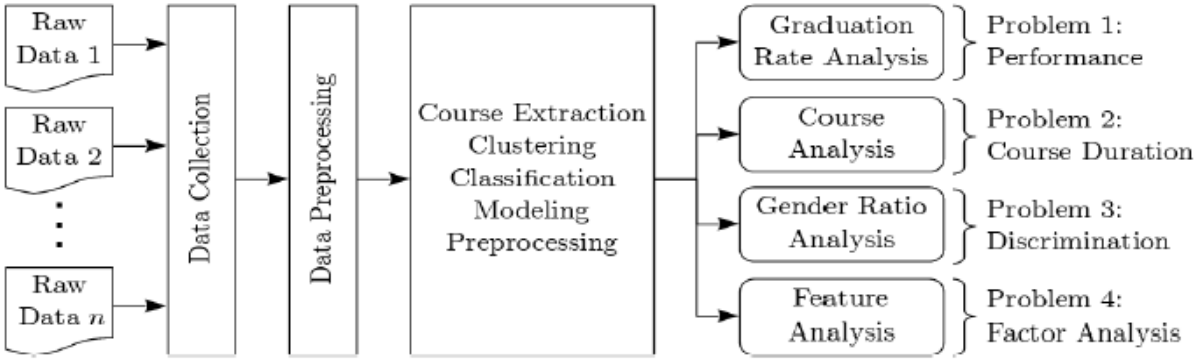

Fig. 1- Proposed Method

5. SEQUENTIAL PATTERN MINING

A brief description of the concept of sequential pattern mining is presented next; for interested readers, Pei [23]presents a more comprehensive treatment of the subject and of the Prefix Span algorithm employed in the study.

Let $S=\{s_1,s_2,s_3.....s_N\}$ be the collection of N student records, where in each student recordsi consists of the entire under graduate academic career of studenti. Specifically, each recordsi consists of an ordered collection of semesters, each of which in turn consists of the non-ordered set of kj course (s)

$c_1, c_2 ... c_{k_j}$ taken in semester j: $Si = < \{c_1, c_2 ... c_{k_1}\} \rightarrow \{c_1, c_2 ... c_{k_2}\} .. >$ Moreover, semesters in which no courses were taken arerepresented as an empty set, as in the following example in which a student took no courses in the second semester:

$Si = < \{c_1, c_2 ... c_{k_1}\} \rightarrow \{\} \{c_1, c_2 ... c_{k_3}\} .. >$

Let the minimum support m_{sup}represent the proportion ofN student records that contain a given sub-sequence. Sequential pattern mining algorithms extract all sub-sequences that occur in at least N*msupof the student records. In a simple three-student example, consider the following set of student records:

Student 1: < {MATH100,ENG30}→{MATH101,ENG35}→ {MATH102}>

Student 2: < {MATH100,PHIL50} →{PHY140}>

Student 3: < {MATH100} →{MATH102}>

If msup= 0.5, then the sub-sequence < {MATH100} →{MATH102}>

Would be the only sequential pattern of length 2 returned. The pattern <MATH100> and <MATH102>, each of length 1, would also be returned.

The elements in the sub-sequence are not required to be in adjacent sets. Hence, in the above example, <MATH100 MATH102>is a valid sequential pattern even though the pattern spans three semesters for Student 1 rather than two semesters for Student 3. Thus, in this study, given the set of student records and a certain msupthreshold, a sequential pattern mining algorithm identifies the complete set of sequential patterns in the collection of studentrecords.Given a set S of the course registrations per semester for all students, msupthreshold, and minimum length lmin, then the pseudo-code associated with mining and processing the sequential patterns from the collection of student records in thispaper.

6. CONCLUSION

In this paper the study has explored how certain curriculum-level factors, including progress through a student's sequence of courses as well as the timing of the first Mathematics course, affect the student's likelihood of persistence and graduation. In the study also assessed the utility, challenges, and potential future directions of sequential pattern mining in the analysis of student course sequences. The technique proved to be a useful tool for exploring sequences of courses and generating new student features based on the extracted sequential patterns. Specifically, these findings can potentially inform an institution's strategies for maximizing persistence and graduation by emphasizing a student's progress through the curriculum. Intervention and support services should be targeted at keeping students on track and progressing towards graduation, so as to maintain their academic momentum. In this paper sequential pattern mining technique demon strated with its utility for creating new features of interest about students within particular subject selection.

REFERENCES

1. *N. C. for Education Statistics, "Graduation rate from first institution attended for first-time, full-time, bachelor's degree- seeking students at 4-year postsecondary institutions, by race/ethnicity, time to completion, sex, control of institution, and acceptance rate: Selected cohort entry years, 1996*
2. *N. C. for Education Statistics, "Retention first-time degree- seeking undergraduates at degree-granting postsecondary institutions, by attendance status, level and control of institution, and percentage of applications accepted: Selected cohort entry years, 1996 through 2010*
3. *C. Romero and S. Ventura, "Educational data mining: A survey from 1995 to 2005," Expert Systems with Applications, vol. 33, pp. 135–146, 2007*
4. *C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," Transactions on Systems, Man, and Cybernetics,vol.40,no.6,pp.601–618,2010*
5. *R. Asif, A. Merceron, S. Ali, and N. Haider, "Analyzing undergraduate students' performance using educational datamining,"Computers&Education,vol.113,p.2017,177–194*
6. *A. A. Saa, "Education Data Mining & Students' Performance Prediction," International Journal of Advanced Computer ScienceandApplications,vol.7,pp.212–220,2016*

7. A. Shahiri, W. Husain, and N. Rashid, "A Review onPredictingStudents' Performance using Data Mining Techniques," Procedia Computer Science, vol. 72, pp. 414–422, 2015
8. A. Gopalakrishnan, H. Yang, C. Graterol, R. Kased, M. Love, and A. Shada, "A Multifaceted Data Mining Approach to Understanding What Factors Lead Students to Persist andGraduate," Computing Conference, pp. 372–381,2017
9. S. Bhaskaran, K. Lu, and M. A. Aali, A data mining approach for investigating students' completion rates. Taylor & Francis Group,London,UK,2018,pp.105–116
10. V. Tinto, "Classroom as Communities: Exploring the Educa- tional Character of Student Persistence," Journal of Higher Education, vol. 68, pp. 599–623, 1997
11. C. Zhao and G. Kuh, "Adding Value: Learning Communities and Student Engagement," Research in Higher Education, vol.45,no.2,p.2004,2004
12. Z. Kolenovic, D. Linderman, and M. Karp, "Improving Stu- dent Outcomes via Comprehensive Supports: Three-Year Out- comes From CUNY's Accelerated Study in Associate Programs (ASAP)," Community College Review, vol. 41, no. 4, pp. 271– 291,2013
13. M. C. S. Program, "About the metro college success program," https://metro.sfsu.edu/about-metro-academies,accessed02.20.2019.
14. A. Pena-Ayala, "Educational data mining: a survey and a data mining-based analysis of recent works," Expert Systems with Applications, vol. 41, pp. 1432–1462, 2014
15. R. Agrawal and R. Srikant, "Mining Sequential Patterns,"Eleventh Intl. Conf. on Data Eng'g, pp. 3–14, 1995
16. Z. Jiang, S. Feng, G. Cong, C. Miao, and X. Li, "A Novel Cascade Model for Learning Latent Similarity from Heterogeneous Sequential Data of MOOC," in Proceedings of the 2017 Conference on Empirical Methods in Natural LanguageProcessing, 2017, pp.2768–2773.
17. W. Chen, C. Joe-Wong, C. Brinton, L. Zheng, and D. Cao, "Principles for Assessing AdaptiveOnlineCourses,"http://www.andrew.cmu.edu/user/cjoewong/VelocityChess EDM.pdf, accessed 02.20.2019
18. R. Martinez, K. Yacef, J. Kay, A. Al-Qaraghuli, and A. Khar-rufa "Analysis frequent sequential pattern of collaborative learning activity around an interactive tabletop", in Proceeding of fourth international conference on Education data mining,2011,pp111-120
19. M. Munk, M. Drlik, L. Benko, and J. Reichel, "Quantitative and Qualitative Evaluation of Sequence Patterns Found by Application of Different Educational Data Preprocessing Techniques," IEEE Access, vol. 5, pp. 8989–9004, 2017
20. D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. Zaiane, "Clustering and Sequential Pattern Mining of Online Collaborative Learning Data," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 6, pp. 759–772, 2009
21. R. Campagni, D. Merlini, R. Sprugnoli, and M. Verri,"Data mining models for student careers," Expert Systems with Applications, vol. 42, pp. 5508–5521, 2015
22. A. Gopalakrishnan, "A Multifaceted Data Mining Approach to Analyzing College Students' Persistence and Graduation Rates," Master's thesis, San Francisco State University, 2016
23. J. Pei, "Mining sequential patterns by pattern growth:The prefix span approach," IEEE Transactions on Knowledge andDataEngineering,vol.16,no.10,pp.1424–1440,2004 24. K. Soni, S. Gangele & S. Patil, "Data Mining Approach towards Students Behavior Assessment Methods for Higher Studies", International Journal of Computer Applications, Vol. 181, No.30, pp.11-14, Nov. 2018
25. K. Soni & S. Gangele "Predictive Analysis for Higher Studies in Rural Region using Linear Regression Model", International Journal of Advanced Science And Technology Vol. 29 No.8(S), pp. 3829-3840,2020
26. K. Soni & S. Gangele "Gender Discrimination-Aware Classification Model for Course Selection of Higher Studies using Real-World Data-Set",International Journal of Future Generation Communication and Networking , Vol. 13 No.4, pp. 3997-4006, 2020

An Efficient Transition Region based Feature Extraction and Combining Classification Model for Semantic Image Segmentation

N. Shanmugapriya[1], Dr. S. Pannirselvam [2]

[1] Research Scholar, Department Of Computer Science , Erode Arts And Science College Erode

[2]M.Sc., M.Phil.,Ph.D., Associate Professor Department Of Computer Science Erode Arts And Science College Erode

ABSTRACT

SIS (Semantic Image Segmentation) is the process of assigning image pixels to object classes based on previously defined nomenclatures. SIS classifies pixels and is a very demanding task. Existing SIS studies are effective and simple, but they are dependent on robust extractions of required areas. SIS techniques show reduced performances while working with textured backgrounds as gray levels of background and foregrounds overlap. This work proposes improvements in DWTs (Discrete Wavelet Transforms). The scheme called IDWT (Improved DWT) is proposed in this paper to extract features for transition regions and overcomes the issue of background/foreground gray level overlaps. IDWT uses image local variances for getting feature variances where transitional features are separated. Otsu thresholding identifies transitional regions from generated transitional feature images. A morphological operation extracts image edges while filling operations get its object regions. Images are recognized using MLTs (Machine Learning Techniques) while SIS is performed using DLTs (Deep Learning Techniques). EXGB (Enhanced XGBoost) detects super pixels while CNNs (Convolution Neural Networks) trained on monographic images classify objects into SLs (Semantic Labels). The proposed SIS+ -Q2`method significantly improves semantic segmentation compare to the other image semantic segmentation.

Keywords: Semantic Image Segmentation,Transition region based feature extraction, Improved Discrete Wavelet Transform, Otsu thresholding, Morphological operations, Enhanced Extreme Gradient Boost classifier, Convolutional Neural Network.

1. INTRODUCTION

Segmentations have been one of the most challenging tasks in computer [1]. Segmentation of images are not like classifications which do not have knowledge about objects and classify only when the data is labelled like horse or house. Effective segmentation techniques segment unknown or new objects [2]. Segmentations have found use in many applications including cultural heritage image preservations to satellite imagery analysis.

Segmentations problems can also be approached semantically [3]. For example, content based image retrievals each image is added to a database on segmentations and can be used to match user queries matching them or when user's query is segmented. When video frames are segmented it could be used by users in HCIs (Human Computer Interactions) at a micro level. For example, security teams at Airports would be more interested in unattended baggage images when looked from the angle of security as it could have dangerous materials like explosives [4] and queries on object left behind by passengers can give them clues.

SIS is a classification at the pixel-level and can be clustered together to form objects by aligning them into a class. Thus, image segmentation and classification are two significant parts of image processing

where classification is categorizing identical images [5]. Objects are localized for recognitions where segmentations ease this process by identifying pixel level demarcations and thus help in improving classifications. Further, instance segmentations combine detections in their segmentations. Traditional segmentation techniques use clustering with extra information on edges and contours [6]. Example of such segmentations can be satellite image segmentation where pixel wavelengths are used to group pixels based on similarity of neighbouring pixels. Clustering has been evolving with several enhancements.

Markov process is a significant approach used in modelling segmentation processes [7]. Contour based detections, region growing techniques using unsupervised learning have been used in image segmentations. DLTs have also been used to segment colours and edges from medical images for analysis [8]. Current evolution of algorithm has made older segmentation techniques obsolete. Hence, current studies turn towards these evolutions and train themselves on standard and well-known international datasets. SIS methods can be divided into three types [9] namely region-based, FCN-based and feebly weakly supervised segmentations.

Region based methods segment based on recognition where free-form regions are first extracted from images, described and then classified based on region's characteristics. These predicted regions are transformed into pixels and labelled according to their scores. RCNNs (Recurrent CNNs) have been used in region-based image segmentations [10]. These networks can perform semantic segmentations. RCNNs first search the feature space to obtain multiple and then compute their CNN features. Each object regions are then classified using other techniques like SVMs (Support Vector Machines). RCNNs can be used to address even complicated tasks like object detections and is used in these areas. The disadvantages found in the use of RCNNs motivated significant research in this area [11]. These disadvantages can be listed as performance degradations when applied on image textures or while processing overlapped gray levels of image's foregrounds and backgrounds. This work proposes to improve SIS with IDWT filter based feature extractions for transition regions as it can overcomes the issue of gray level overlaps. This introductory section is followed by literature review in section two. The third section details on the proposed methodology followed simulation results and discussions in section four. This paper in concluded in the fifth and final section.

2. LITERATURE REVIEW

This section details on studies and methodologies related to SIS.

Chen et al [12] probed scaling of CNN features at multiple levels using an Atrous Spatial Pyramid Pooling module. Their scheme encoded image-level features in a global context afor enhancing performances. The study also elaborated on the implementations of their 'DeepLabv3' which performed significantly better than previous versions of DeepLab and without using DenseCRF post-processing. The scheme was benchmarked for image segmentations on PASCAL VOC 2012 dataset.

Automatic annotations for images were proposed in the study [13]. The framework segmented images and labelled objects simultaneously. The scheme was a SIS segmenting framework that operated at a semantic level with possible semantic labels. Fuzzy sets were used to manage image regions in place of visual features. Image segmentations were primary to the study and were executed using watershed and recursive shortest spanning tree region growing algorithms. The study also used visual context representations and analysis in interpreting local object with global knowledge. A new method using

fuzzy algebra and taxonomic ontology processed contextual information. This learning was then used to re-adjust semantic region growing labels and fine-tuned by degrees of membership. Their performance evaluations on image datasets from dual domains were satisfactory.

MRFs (Markov Random Fields) were introduced in the study [14] for labelling mixed label contexts. The study's novelty lay in using CNNs for optimizing MRFs instead of iterative algorithm usage. The study used DPNs (Deep Parsing Networks) in a singular forward pass for deterministic end-to-end computations. DPNs modelled unary terms by extending CNN architecture and by using additional layers to approximate MF (Mean Field) algorithm in identification of pair wise terms. Their scheme was novel in many ways: the use of CNN for MRF/MF iterations which require training of images in back-propagations; use of DPNs approximated one iteration of MF; DPNs obtained many pairwise terms making it special and DPNs made MF parallel operations hastened GPU (Graphical Processing Unit) operations. The proposed DPNs when evaluated on PASCAL VOC 2012 dataset scored 77.5% in accuracy.

EMs (Expectation-Maximizations) were used by [15] in their study who developed SIS methods for feeble supervised/semi-supervised configurations. Their proposed models were evaluated on PASCAL VOC 2012 dataset extensively where their models learnt better under configuration constraints delivering competitive results in image segmentations with minimized effort in annotations.

Transition regions were segmented by Li et al [16]. The proposal was a robust hybrid scheme for segmenting single- objects from images. Their system specifically used local complexities and variances to identify image's transition regions. The technique chose transition regions with the highest number of pixels as the main transition region. The study also used gray level intervals obtained from image's transition regions where one of the determined gray level was used as the segmentation threshold for identifying significant transition region. This framed image thresholding results were refined for final segmentation outputs where salient transition regions eliminated duplicate object regions. The study's proposed scheme was evaluated extensively with 170 real world single-object images where their technique achieved higher segmentation accuracy and was robust in identifying transition areas when compared with other similar techniques. Further, it was simple to implement Images were represented as a partition tree by Salembier et al [17]. Their scheme first estimated similarity of regions in images which were then used for partitioning. Their proposed scheme was a compromise between systematic processing of images and application specific image processing. Their partition tree demonstrated that their tree partitioning could be used in various kinds of image segmentation applications and feature spaces like spatial, motional, region-based coding/retrieval and semantic object extractions.

Ontology bases SIS was proposed by Zand et al [18]. The scheme OBSIS (ontology-based SIS) modelled image segmentations with object detections. Their Dirichlet process mixed model transformed low-level visual features to an intermediate transitional semantic space and applied dimensionality reduction on features which were then weighed individually and learnt with multiple CRFs for contexts. Their image segmentations into object's parts minimized load on classifiers as object inferences were passed to them. The model mimicked human visual understanding of images by its combination of cues, contexts and ontology learning based on rules. When experimentally evaluated on MSRC-21 and PASCAL VOC'2010 datasets, the scheme showed satisfactory results.Mackowiak et al [19] proposed an active learning-based strategy, called CEREALS (Cost-Effective REgion-based

Active Learning strategy) where a manually annotated labels were used to select regions from images. The scheme used these annotated semantics for further image segmentations thus minimizing annotation efforts while maximizing SIS. Their automatic selections learned costs from human annotations and exploited image spatial coherencies. The scheme when tested on Cityscapes reduced 17% annotation efforts while maintaining 95% of the model's mIoU (mean Intersection over Union) trained on annotated Cityscapes training set.

Local label descriptor figured in the study of Yang et al [20] where cell's label histograms were concatenated. Their use of these descriptors eliminated label patch misalignments by structuring label predictions for SIS labelling. The study took an input image's label maps and approximated label descriptors with sparse convex combinations during training, where weak regions were regularized by input image's local feature descriptor similarities in the training set. They incorporated a low-level over-segmentation in images to enhance efficiency. Their scheme was tested on CamVid and Barcelona datasets and displayed good results.

Bertasius et al [21] addressed boundary localizations in their study. The study used RWNs (Random Walk Networks) to overcome boundary localization issues and fragmented predictions of spatial data. Their RWNs optimized pixel affinities and SIS using novel random walk layer which enforced spatial grouping consistency in the network's deep layers. RWNs implemented convolution and matrix multiplications for integration with FCN frameworks and also enabled an end-to-end training of the network with back-propagations. The study used only 131 additional parameters when compared to traditional FCN parameters and showed consistency in labelling of scenes and SIS over FCNs.

FickleNet was used to explore multiple combinations of feature map locations generated by DNNs in [22]. The scheme randomly selected hidden units and then obtained activation scores classification of images. The study's FickleNet learnt generated feature maps location coherences for localizing the maps and thus identifying normal and discriminative object parts. The study's selection of random hidden unit pairs helped them achieve ensemble effects from a single network. This effect was then used on a single image to create multiple localized maps. Their scheme did not use or employ additional training steps and only added layers to the CNN. The proposed scheme outperformed many other techniques when benchmarked on the Pascal VOC 2012 dataset with weak and semi-supervised configurations.

It can be inferred from above discussions that image gradients help in detecting object boundaries in edge-based active contour methods and in the case of region-based active contours models using object and their background region energy optimums fit the image to its best. LEMs (Local Entropy Methods) take into account neighborhood's entropy for extracting transition regions. These methods have limitations when events change frequently as local entropy values rise, making models assume the neighborhood is transition. Existing methods have gaps in their proposed schemes and their performances degrade specifically when gray level intensities of backgrounds and foregrounds overlap. Hence, this research work attempts to overcome this stated issue and focuses on improving the image segmentation techniques with the proposed method.

3. PROPOSED METHODOLOGY

Image segmentations attempt to distinguish foreground and background regions in images. Segmentations can be based on local or global features which are utilized to separate objects from their surroundings. Localized segmentations involve division of images into constituent regions. Images

generically have transition between their foregrounds and backgrounds and effective extraction of these transitions lead to better results in image segmentations. This study improves the effectiveness is segmenting transition region in gray scale images by its use of IDWT based feature extractions from image transition regions. These feature extraction are then followed by image recognitions using MLTs and subsequent SIS by DNNs. The main aim of this work is overcoming hurdles in image segmentations while handling gray level overlaps of backgrounds and foregrounds. This paper's contributions are listed below:

- Local variances between image pixels are used in for extracting features from gray scale images. The extracted features are split based on their intensity values.
- DWT's improved version is used for identifying transition areas in images where pixel variances help separation of transition areas.
- Otsu thresholding extracts transitional features in images and for image's edge extractions a morphological thinning operation is used.
- Objects are obtained by region filling and morphological cleaning of image edges.
- SIS is executed by using MLT EXGB (detect super pixels) and CNN (training monographic images and classify object with semantic labelling.

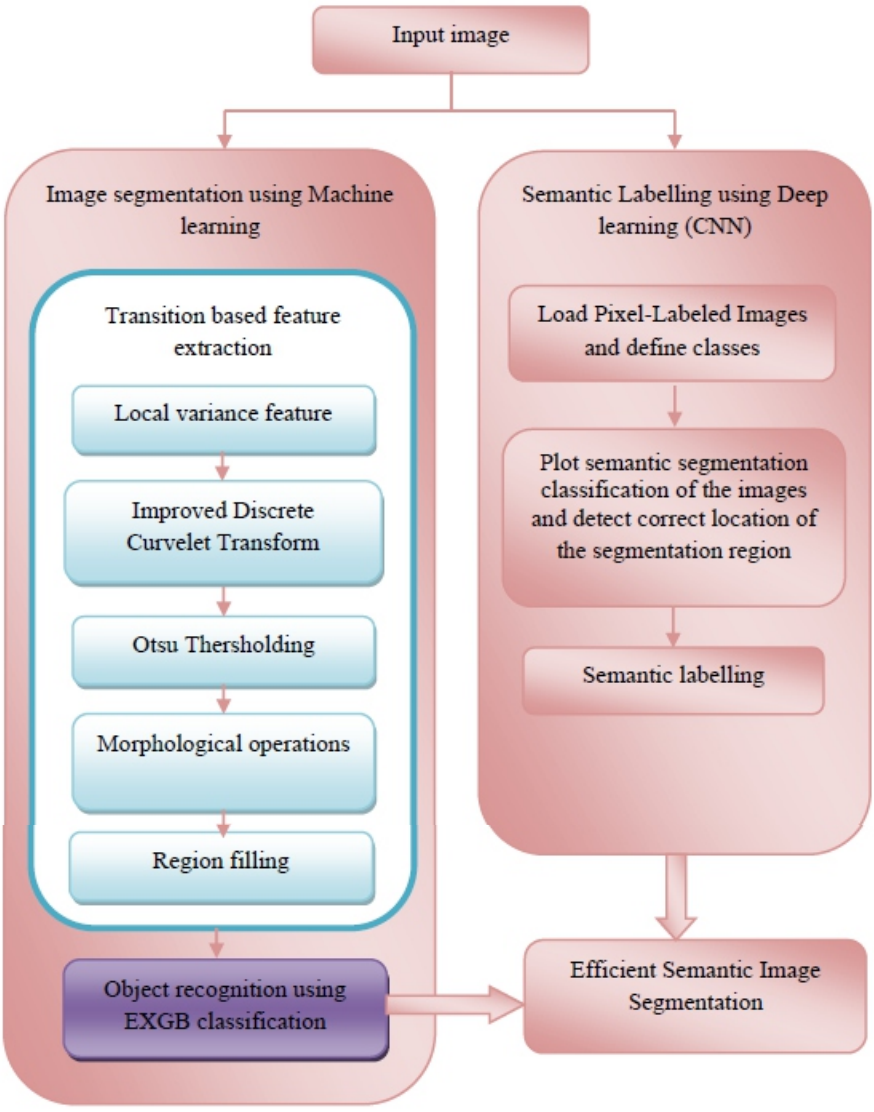

Figure .1. The overall flow of the research work

3.1. Feature extraction based on Local variances

Local variances of pixels in images are extracted first. Pixel values having higher variance are regions behind feature pixels and typically contain edges of objects. Harmonized regions show lesser variances when compared to image edges. Assuming a square region (n x n) of neighbourhood pixels where its central pixel is denoted by x(k,l) variances of pixels in the region can be extracted using Equation (1)

$$Lv(k,l) = \sigma^2 = \frac{1}{n^2-1}\sum_{k=1}^{n}\sum_{l=1}^{n}(f(k,l) - \bar{f})^2, \qquad for \ \begin{matrix} k=1,2,\ldots\ldots,M \\ l=1,2,\ldots\ldots,N \end{matrix} \qquad (1)$$

Where, (k,l) - $n \times n$ local neighbourhood coordinates, f - gray level mean of the neighbourhood, M,N – Height, Width of the image. The entire image is passed through a sliding window $n \times n$ and from left top to right bottom for extracting feature variances. The value of n used in this work is 3 implying feature extractions are done using a 3×3 window. These local variance extractions obtain dominating image features like object's edges and backgrounds which are clustered.

3.2. Improved discrete wavelet transform

The use of wavelet forms is predominant in much multi resolution analysis where studies have proposed tools to represent edges and singularities along curves better using wavelet forms. , Curvelets are much better in extraction when compared to wavelets which are limited in their directional features whilecurvelets can be used to describe signals with a group of matrices that can be scaled or be made multi-directional. Moreover, in scaling the count of direction also become finer [23]. In terms of accuracy achievements, the technique requires manual detections of core images like fingerprint images. Curvelet features are extracted after complex image enhancements including estimations of local ridge-orientations, local ridge frequencies across fingerprints, image filtering conversion of a gray-scale image to binary image form

• Improved Discrete Curvelet Transform (IDCT)

Curvelets represent a framework corresponding to singularities and edges of curves. In concept, Curvelet transforms represent a pyramid which can be scales with multiple orientations or positions at each scale which be intricately scaled with needle-shaped elements. Curvelets can be distinguished from their counterpart wavelet forms due to their geometric features. For example, Curvelets behaviours are highly anisotropic with variable widths and lengths. At a micro level, anisotropy increases when scales decrease. Considering these advantages this work uses Curvelets by warping, the fastest currently available Curvelet transform.

On Curvelet transform executions, multiple Curvelet coefficients groups are created at varying angles and scales. These obtained coefficients can be represented as a matrix $C_{j,l}$ where is can represent the finest to coarsest scale and l is the angle starting at the top-left corner and incremented clockwise.

Assuming an image's size is N_1, N_2, the image is represented as $f(t_1, t_2), 1 \leq t_1 \leq N_1, 1 \leq t_2 \leq N_2$ then its 2D discrete fourier transform is $\hat{f}[n_1, n_2]$.

Step 1: A two dimensional FFT (Fast Fourier Transform) is applied the image to obtain its 2D discrete fourier transforms.

Step 2: The 2D discrete fourier transforms is re-sampled in scale and direction in a frequency domain resulting in new samples depicted in Equation (2)

$$\hat{f}[n_1, n_2 - n_1 tan\theta_l], \quad (n_1, n_2) \in P_j, \qquad (2)$$

Where, $P_j = \{(n_1, n_2), n_{1,0} \leq n_1 < n_{1,0} + L_{1,j}, n_{2,0} \leq n_2 < n_{2,0} + L_{2,j}\}$ and$n_{1,0}n_{2,0}$ are two positions of the window function $\tilde{u}_{j,l}[n_1, n_2]$, $L_{1,j}$, $L_{2,j}$ - lengths in window function support interval and 2^j , $2^{j/2}$, - width of window function support interval.

Step 3: Multiplying new samples with the result is Equation (3)

$$\tilde{f}_{j,l}[n_1, n_2] = \hat{f}[n_1, n_2 - n_1 tan\theta_l]\, \tilde{u}_{j,l}[n_1, n_2] \qquad (3)$$

Where

$$\tilde{u}_{j,l}[n_1, n_2] = w_j(w_1, w_2) v_j \left(s_{\theta_l} \frac{\left(2^{\lfloor j/2 \rfloor} w_2\right)}{w_1} \right) \qquad (4)$$

$$w_j(w_1, w_2) = \sqrt{\phi_{j+1}^2(w^2) - \phi_j^2(w^2)} \qquad (5)$$

$$\phi_j(w_1, w_2) = \phi(2^{-j} w_1) \phi(2^{-j} w_2) \qquad (6)$$

$$s_{\theta_l} = \begin{bmatrix} 1 & 0 \\ -tan\theta_l & 1 \end{bmatrix} \qquad (7)$$

$$tan\theta_l = l \times 2^{\lfloor -j/2 \rfloor}, l = -2^{\lfloor -j/2 \rfloor}, \ldots, 2^{\lfloor -j/2 \rfloor} - 1. \qquad (8)$$

Step 4: The application of inverse 2DFFT to each $\tilde{f}j.$, results in obtaining the targeted discrete coefficients $Ci.l.$

Curvelet transform descriptions of signals contain scales, angles, and positions. They also use variable lengths and widths and thus have improved directional capabilities and can represent curve edges and curve singularities much better than other multi-scale transforms like wavelet transforms.

Disadvantages of curvelet's can be summarized as :its functions do not have the same average; curvelet coefficients are prearranged into multiple sized sub-bands; redundancy ratio changes even within ranges resulting is memory allocation issues. Hence this work uses discretecurvelet transforms with a SAFs (Symmetrical Angle Functions), a novel approach in implementing curvelet transform and is mainly used for its advantages in terms of reduced redundancy ratio and hierarchical data structures.

• Symmetrical Angle Functions

This section describes SAFs used in this work. SAFs are the same as angle functions in continuous states V(t) with their squares summed up to 1. Since, the new functions are defined for discrete frequencies shearing relationships are used in place of rotation relationships.

Assuming the defined angle functions need to be in the range$(-\pi/4, 3\pi/4)$, N number of intermediary functions denoted by $\tilde{v}_l(t), l = 1, \ldots, N,$ are defined as in Equation (9)

$$\tilde{v}_l(t) = \beta \left(\frac{\frac{2}{N} 1 - t}{\frac{2}{N} \eta_b} \right) \beta \left(\frac{t+1}{\frac{2}{N} \eta_b} \right) \quad with\ \eta_b \leq \frac{1}{2} \qquad (9)$$

$$\tilde{v}_l(t) = \tilde{v_1} \left(t - \frac{2(l-1)}{N} \right) \quad with\ l = 2, \ldots, N. \qquad (10)$$

In the above definition of $\tilde{v}_l(t)$ implies it is a smooth window function with support values between -1 to $-1+\frac{2}{N}.\eta_b$ controls the transition area's width. The function's values gradually change from 0 to 1/1 to 0 when the value of set t changes from $-1-\eta_b 2/N$ and $-1+\eta_b 2/N$ or from $-1+(1-\eta_b)2/N$ to $-1+(1+\eta_b)2/N$). Set t satisfies $\tilde{v}_l(t)$ by $2(l-1)/N$when $\eta_b \leq (1/2)$. $\beta(t)$is defined for verification of sum of squares of $\tilde{v}_l(t), l=1,\ldots..,N$ and thus the smooth window support is $-1-2/N\eta_b$to $1+2/N\eta_b$.

This work defines a function $T(\theta)$ to convert defined $\tilde{v}_l(t)$ function to Symmetrical functions and maps θ values in the range $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$to (-2,2) as defined in Equation (11)

$$T(\theta)=\begin{cases}\tan(\theta) \quad when -\pi/4 \leq \theta \leq \pi/4\\ 2-\tan(\pi/2-\theta) \quad when\ \pi/4<\theta<\pi/2\\ -2-\tan(\pi/2-\theta)\ when\ -\pi/2<\theta<-\pi/4\end{cases} \tag{11}$$

Thus, functions $\tilde{v}_l(t)$, define N Symmetrical angle functions of θ as in Equation (12)

$$v_l(\theta)=\tilde{v}_l(T(\theta)) \tag{12}$$

The functional mapping of $T(\theta)$ to one dimensional $\tilde{v}_l(t)$ and symmetrical angle functions $v_l(\theta)$, support 3-D views.

It is evident that functions $T(\theta \pm \pi/4) \pm 1$ are anti-symmetric. Remaining N Symmetrical angle functions are constructed by flipping $v_l(\theta), l=2,\ldots..,N-1$ with $\pi/4$

$$v_l(\theta)=v_{2N+1-l}\left(\frac{\pi}{2}-\theta\right), l=N+1,\ldots..,2N. \tag{13}$$

The anti-symmetric property of $T(\theta-\pi/4)-1$ can also illustrate that sum of square of $v_l(\theta)=1$ in overlapped regions of $v_N(\theta)$ and $v_{N+1}(\theta)$. Further, sum of squares of all angle Symmetrical functions $v_l(\pm\theta)$ is also equal to 1.

$$\sum_{l=1}^{2N} v_l^2(\theta)+v_l^2(-\theta)=1 \tag{14}$$

The end result of these processes obtains 3 classes of features in which class-3 features are transition features and are subsequently processed in this work.

3.3. Extraction of transition regions from transitional features

The disjoint sets of features obtained have been categorized as class-1 (Only Background), class-2(Overlapped Object inner and Background regions) and class-3 (transitional features) and depicted in figure 2. Otsu thresholding extracts transition regions from class-3 features where a value of 1 indicates the pixel belongs to the transition region are separated eventually.

Figure .2. Extracted Class features

3.4. Morphological operations thinning and cleaning

Transition regions surrounding objects are several pixels wide. For achieving an object pixel's closed contour width, transition regions are thinned (diluted) morphologically. This dilution operation results in the object's edge image including a few isolated pixels near the edge. Morphological cleaning removes this isolated pixels resulting from dilations. Thus, clear object contours are obtained.

3.5. Object region extractions using Morphological region filling

The vivid object contours obtained from morphological cleaning and dilation are fully connected as they are continuous in nature. The obtained contour regions inside the edge image are morphologically filled with a value of 1 (segmentation mask) while remaining portions are filled with 0. Figure .3 depicts the proposed feature extraction process of transition regions.

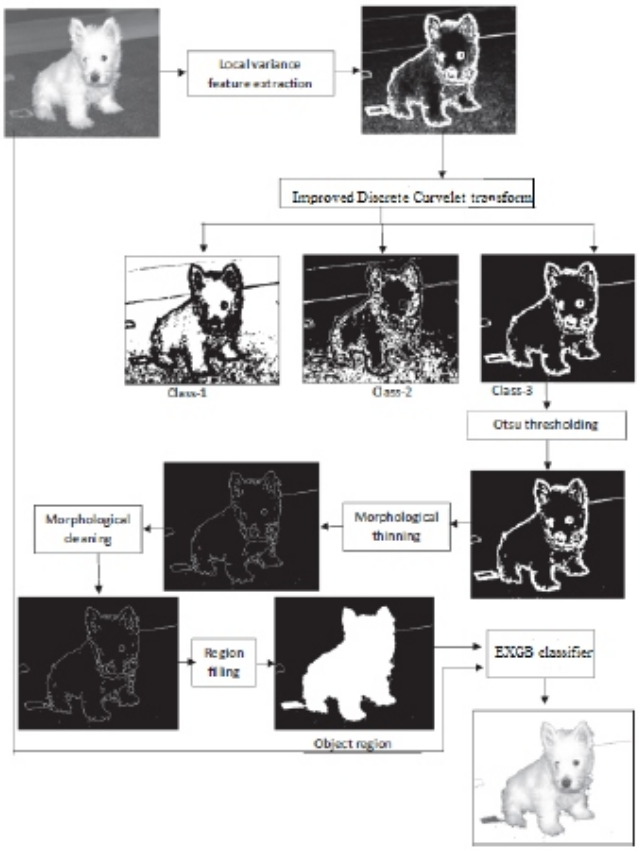

Figure .3. The proposed feature extraction process of transition regions

3.6. Object region extraction with EXGB classifier

The extracted object regions are in the form of binary image where 1 implies an object mask and 0 the background. All 1's are replaced with their original gray intensity values to segmented objects while 0's are assigned 255 to convert background colur as white.

XGBoost is a MLT used in regression and classification problems and is based on GBDT (Gradient Boosting Decision Tree) [24]. Its regression tree's nodes represent attribute's test values while their scores become leaf nodes. XGBoost learns additively with 2^{nd} order approximations, 1^{st} / 2^{nd} order loss functions for predictive model fits.

Assume the data count is m, n the count of features, z_i, is the 'raw prediction' before application of sigmoid function $\sigma(.)$ and $\hat{y}_i = \sigma(z_i)$ the probabilistic prediction, discrepancies in notations are denoted by z, true label are denoted by y_i, α, γ and are loss function parameters. The gradients obtained are in merged format and independent of y_i, for simplifying implementation and vectorization for other related programs.

The additive learning objective is depicted in Equation (15)

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, z_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (15)$$

Where, t – training's t^{th} iteration. Notations have been replaced in the equation and 2^{nd} order Taylor expansion application on Equation (15) results in Equation (16)

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n}[l\left(y_i, z_i^{(t-1)} + g_i f_t(x_i)\right) + \frac{1}{2}h_i(f_t(x_i))^2 + \Omega(f_t) \quad (16)$$

The term $l(y_i, z_i^{(t-1)})$ is unnecessary in Equation (16) as does not correspond to the model in t^{th} and hence the modified as Equation (17)

$$\propto \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i(f_t(x_i))^2] + \Omega(f_t) \quad (17)$$

XGBoost does differentiate automatically and hence the use of hand-derived derivatives are essential and can also be implemented MLTs. Sigmoid activation function is used for loss functions. The sigmoid property used in derivatives is depicted bekow:

$$\frac{\partial \hat{y}}{\partial z} = \frac{\partial \sigma(z)}{\partial z} \quad (18)$$

$$= \sigma(z)(1 - \sigma(z)) \quad (19)$$

$$= \hat{y}(1 - \hat{y}) \quad (20)$$

In addition to defining regularized objectives this study preventoverfitting of the model. Shrinkage scales are added weights W_j^* after each tree boosting which minimize each individual tree's influence, allowing for the growth of more trees to enhance the model.

If q(x) is a fixed structure the optimal weight W_j^* a leaf j can be computed using Equation (21)

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (21)$$

Where, $\in$ - factor of approximation. Intuitively this implies the presence of roughly selected points. Each data point is weighed by h_i and as per Equation (22)

$$\sum_{i=1}^{n} \frac{1}{2} h_i (f_t(x_i) - g_i/h_i)^2 + \Omega(f_t) \quad (22)$$

If $I_j = \{i|q(x_i) = j\}$ is defined as an instance set of leaf j. Equation (20) with the expansion Ω can be changed to the following

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i {f_t}^2(x_i)] + \Omega(f_t) \quad (23)$$

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2}h_i {f_t}^2(x_i)] + \frac{1}{2}\lambda \sum_{j=1}^{T} {w_j}^2 \quad (24)$$

which amounts to the exact weighted squared loss of labels gi/hi and weights hi. In large datasets, it is non-trivial to find candidate splits that satisfy this criteria.

3.4. Semantic Labeling with CNN

Semantic segmentations imply meaningful classification of image pixels of objects. Deep learning in machine learning learns multiple representations. CNNs are a kind of DNNs where forward processes and back propagations are used. They are used for their maximal speed and efficiency of operations.

a) Role of Convolutional Neural Networks

CNN is a MLP (MultiLayer Perceptron) specially designed for identifying 2D image information. CNNs use input, convolution, sample and output layers. CNN input neurons are connected to previous layers, for extracting local feature [25]. On extraction of local features, positional relationships with other features are determined in the feature map layer; CNN's computing layers are composed of feature maps where every feature map is a plane and their neurons have equal weights. Feature map use a sigmoid activation for feature map's invariance. Neurons in the same mapping plane share their weights thus reducing network's parameters. Each convolution layer is followed by a computing layer in CNNs where local averages are computed reduce feature extraction structure resolutions.

CNNs can identify displacements or other forms of distortions in two-dimensional feature spaces. CNN implicit learning from train samples avoids explicit feature extractions. A major advantage of CNNs is the inter-connection of neurons. Its tight layouts mimic natural biological neural networks. CNN's sharing of weights minimize complexities of the network. A multi-dimensional input vector image need not be reconstructed for feature extractions and classifications in CNNs.

• Back Propagation Algorithm

CNNs use convolution and sampling. Key properties of CNNs include local receptive fields, weight sharing, sub sampling in terms of time or space, extracting features and dimensionality reduction. Convolutions use a trainable filter Fx, de-convolution of input image or feature image of each layer (Feature Map), bias bx and a convolution layer Cx. In sampling, pixels of each neighbourhood are added scalar weights Wx + 1 and bias bx + 1 and processed using an activation function, for generating feature map Sx + 1.

$$O_{x,y}^{(l,k)} = tanh \sum_{t=0}^{f-1} \sum_{r=0}^{K_h} \sum_{c=0}^{K_w} W_{(r,c)}^{(k,t)} O_{(x+r,x+c)}^{(l-1,t)} + Bias^{(l,k)} \quad (25)$$

Where, f – count of convolution core's feature pattern output of neuron of row x , column y in the l th sub sample layer and kth feature pattern :

$$O_{x,y}^{(l,k)} = \tanh(W^k \sum_{r=0}^{S_h} \sum_{c=0}^{S_w} O_{(x \times S_h + r, y \times S_w + c)}^{(l-1,t)} + Bias^{(l,k)} \quad (26)$$

the output of the j th neuron in l th hide layer H :

$$O_{(i,j)} = \tanh\ (W^k \sum_{k=0}^{S-1} \sum_{x=0}^{S_h} \sum_{y=0}^{S_w} W_{(x,y)}^{(j,k)} O_{(x,y)}^{(l-1,t)} + Bias^{(l,k)} \quad (27)$$

among them, s is the number of feature patterns in sample layer. output of the ith neuron l th output layer F

$$O_{(i,j)} = \tanh\ (\sum_{j=0}^{H} O_{(l-1,j)} W_{(i,j)}^{l} + Bias^{(l,i)} \quad (28)$$

• **FP (Forward Propagation)**

FP is computation and storage of intermediate variables/outputs of NNs within models in the order input-o output layers. Assuming input is $x \in \mathbb{R}^d$ without bias, then the intermediate variable is as follows:

$$z = W^{(1)}x \tag{29}$$

$W^{(1)} \in \mathbb{R}^{h \times d}$ is hidden layer's parameter weight. The intermediate variable $z \in \mathbb{R}^h$ in the activation function ϕ is operated by basic elements obtains a hidden layer with the vector length of h,

$$h = \phi(z) \tag{30}$$

The hidden variable h is also an intermediate variable. Assuming the parameters of the output layer only possess a weight of $W^{(2)} \in R^{q \times h}$, obtain an output layer variable with a vector length of q:

$$o = W^{(2)}h \tag{31}$$

Assuming the loss function is l and the example label is y, then calculate the loss term for a single data example,

$$L = l(o, y) \tag{32}$$

According to the definition of ℓ_2 norm regularization, given the hyper-parameter λ, the regularization term is

$$s = \frac{\lambda}{2}\left(\left\|W^{(1)}\right\|_F^2 + \left\|W^{(2)}\right\|_F^2\right), \tag{33}$$

where the Frobenius norm of the matrix is equivalent to the calculation of the L2 norm after flattening the matrix to a vector. Finally, the model's regularized loss on a given data example is

$$J = L + s \tag{34}$$

Where J is the objective function of a given data example.

4. RESULTS AND DISCUSSION

This section displays results of the proposed method which was evaluated using the BSDS500 dataset. This dataset is an extension of BSDS300 with 200 test images. It also has 2,000 hand labelled segmented parts of thousand Corel images collected from 30 human subjects. Fifty percent of the samples are color image while the balace being in grayscale. Benchmarks were based on a mixture of grayscale and color segmentations of 300 images. The images are divided into a training set of 200 images, and a test set of 100 images.

The performance metrics are measured using of TPs (True Positives), FPs (False Positives), TNs (True Negatives) and FNs (False Negatives). Precision is proportion of retrieved instances that were relevant. The second performance metric was recall, defined as the proportion of relevant instances that were retrieved. Though conflicting, the measures, precision and recall are important in evaluating performance of prediction approaches. Hence, their combined metric F-measure is also used. Accuracy is defined as the proportion of correctly predicted instances relative to all predicted instances.

Precision is defined as the ratio of correctly found positive observations to all of the expected positive observations.

$$\text{Precision} = TP/TP+FP \qquad (35)$$

Sensitivity or Recall is defined the ratio of correctly identified positive observations to the over-all observations.

$$\text{Recall} = TP/TP+FN \qquad (36)$$

F - measure is defined as the weighted average of Precision as well as Recall. As a result, it takes false positives and false negatives.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \qquad (37)$$

Accuracy is calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = (TP+FP)/(TP+TN+FP+FN) \qquad (38)$$

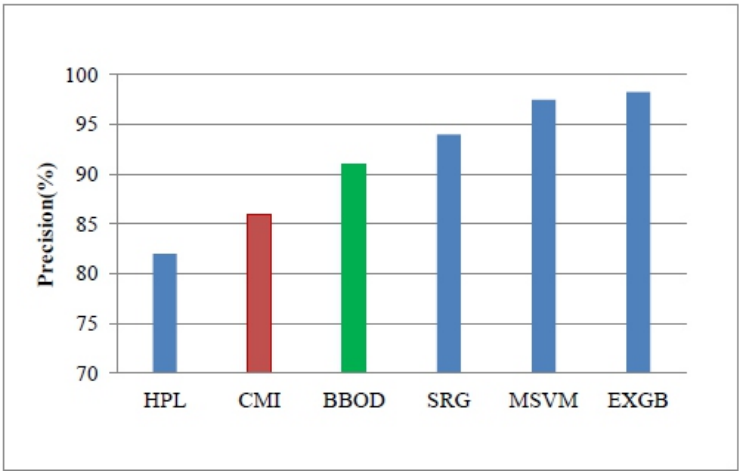

Fig.2. Precision comparison results between the proposed and existing techniques for Semantic Image Segmentation

The fig.2. illustrates that the precision comparison results between the proposed and existing techniques for Semantic Image Segmentation using the proposed EXGB and CNN based technique. From the results it concludes that the proposed EXGB and CNN technique has high precision results compare to the existing segmentation techniques.

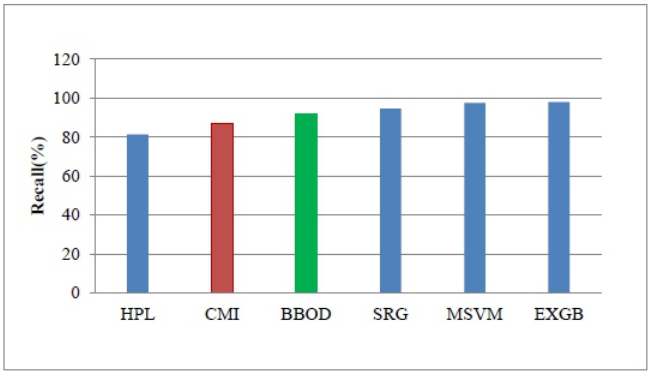

Fig.3. Recall comparison results between the proposed and existing techniques for Semantic Image Segmentation

The fig.3. illustrates that the recall comparison results between the proposed and existing techniques for Semantic Image Segmentation using the proposed EXGB and CNN based technique. From the results it concludes that the proposed EXGB and CNN based technique has high recall results compare to the existing Semantic Image Segmentation techniques.

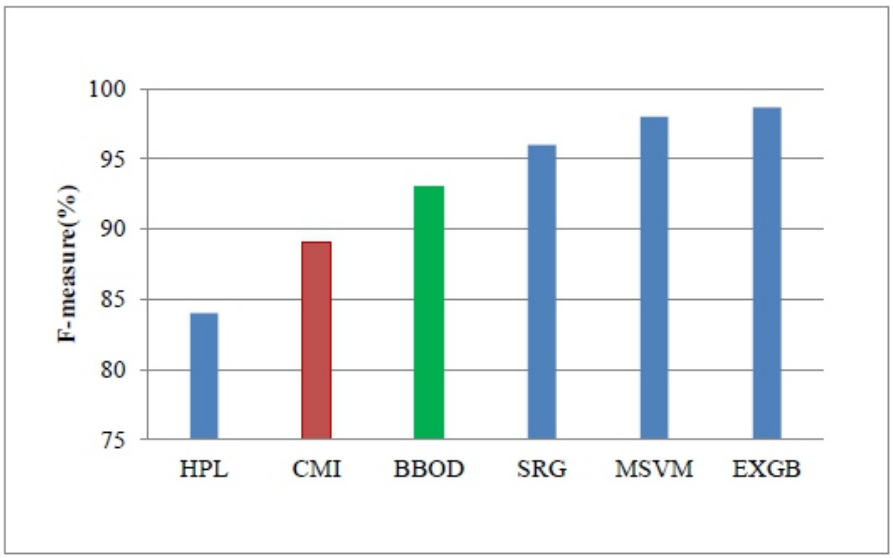

Fig.4. F-measure comparison results between the proposed and existing techniques for Semantic Image Segmentation

The fig.4. illustrates that the F-measure comparison results between the proposed and existing techniques for Semantic Image Segmentation using the proposed EXGB and CNN based technique. From the results it concludes that the proposed EXGB and CNN based technique has high F-measure results compare to the existing segmentation techniques.

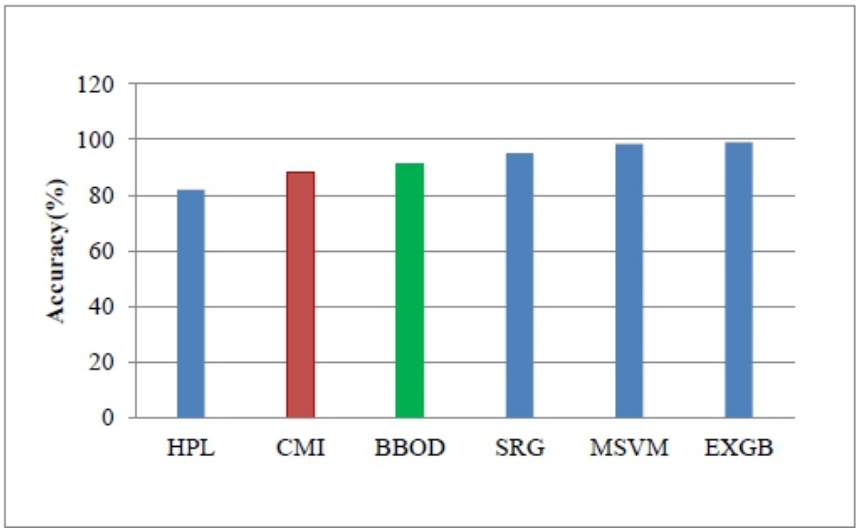

Fig.5. Accuracy comparison results between the proposed and existing techniques for Semantic Image Segmentation

The fig.5. illustrates that the accuracy comparison results between the proposed and existing techniques for Semantic Image Segmentation using the proposed EXGB and CNN based technique. From the results it concludes that the proposed EXGB and CNN based technique has high accuracy results compare to the existing segmentation techniques.

5. CONCLUSION

This work has proposed and demonstrated a novel method of feature extraction of transition areas from images using IDWT executes semantic image segmentations. The proposed technique of transition region extractions are robust in identifying and extracting transition regions accurately from all types of images including foreground, background and textured images. The approximations of image's decomposed wavelet coefficients suppress background and foreground texture features bringing out the image's objects. This suppression enhances transition region extractions in turn leading to efficient segmentations. These effective segmentations are used for better image recognitions by XGBoost's machine learning which detects the image's super pixels. CNN in this work is trained on monographic images categories for classifying each object based on semantic labels. The evaluation and experimental results of the proposed scheme depict that it outperforms other SIS methods.

REFERENCES

1. *Liu, X., Deng, Z., & Yang, Y. (2019). Recent progress in semantic image segmentation. Artificial Intelligence Review, 52(2), 1089-1106.*
2. *Jingyu, L. S. Y. (2007). Survey on Extraction Methods of Transition Region [J]. Engineering Sciences, 9.*
3. *Xiao-lin, Z. O. U. (2014). Anthresholding image segmentation algorithm based on local descriptors fusion. Journal of Hunan University of Science & Technology (Natural Science Edition), 02.*
4. *Ni, B., & Huang, J. (2011). Transition Region Extraction Algorithm of Infrared Image Based on Wavelet Transform. Video Engineering, 2011, 19.*
5. *Cao, Z., Zhang, K., &Lii, Y. (2008). Transition Region Extraction Algorithm based on Local Fuzzy Complexity. Fire Control and Command Control, 1.*
6. *Tianxu, Y. C. S. N. Z. (2004). Wavelet Transform based Image Transition Region Extraction and Seg mentation [J]. Computer Engineering and Applications, 18.*
7. *Wang, Y. C., Liang, D. Q., Wang, Y., & Wang, G. L. (2008). Image transition region extraction and segmentation based on neighborhoodunhomogeneity. Journal of Optoelectronics Laser, 19(3), 404.*
8. *GAN, B., ZHANG, X. W., & GAN, Z. M. (2001). Multi-scale image edge processing based on wavelet transform [J]. Tv Engineering, 8.*
9. *Yujin, Z. (2015). A Survey on Transition Region-Based Techniques for Image Segmentation. Journal of Computer-Aided Design & Computer Graphics, 3, 379-387.*
10. *KANG, W. W., KANG, W. Y., & KANG, X. T. (2013). Image transition region extraction and segmentation based on information measure. Journal of Jilin University (Engineering and Technology Edition), S1.*
11. *Liwei, G. U. O. (2013). Exploration of image segmentation algorithm. Journal of Shenyang Normal University (Natural Science Edition), 03.*
12. *Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.*
13. *Athanasiadis, T., Mylonas, P., Avrithis, Y., &Kollias, S. (2007). Semantic image segmentation and object labeling. IEEE transactions on circuits and systems for video technology, 17(3), 298-312.*
14. *Liu, Z., Li, X., Luo, P., Loy, C. C., & Tang, X. (2015). Semantic image segmentation via deep parsing network. In Proceedings of the IEEE international conference on computer vision (pp. 1377-1385).*
15. *Papandreou, G., Chen, L. C., Murphy, K. P., &Yuille, A. L. (2015). Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. In Proceedings of the IEEE international conference on computer vision (pp. 1742-1750).*
16. *Li, Z., Liu, G., Zhang, D., & Xu, Y. (2016). Robust single-object image segmentation based on salient transition region. Pattern recognition, 52, 317-331.*
17. *Salembier, P., &Marqués, F. (1999). Region-based representations of image and video: segmentation tools for multimedia services. IEEE Transactions on circuits and systems for video technology, 9(8), 1147-1169.*
18. *Zand, M., Doraisamy, S., Halin, A. A., &Mustaffa, M. R. (2016). Ontology-based semantic image segmentation using mixture models and multiple CRFs. IEEE Transactions on Image Processing, 25(7), 3233-3248.*
19. *Mackowiak, R., Lenz, P., Ghori, O., Diego, F., Lange, O., & Rother, C. (2018). Cereals-cost-effective region-based active learning for semantic segmentation. arXiv preprint arXiv:1810.09726.*
20. *Yang, Y., Li, Z., Zhang, L., Murphy, C., VerHoeve, J., & Jiang, H. (2012, October). Local label descriptor for example based semantic image labeling. In European Conference on Computer Vision (pp. 361-375). Springer, Berlin, Heidelberg.*

21. *Bertasius, G., Torresani, L., Yu, S. X., & Shi, J. (2017). Convolutional random walk networks for semantic image segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 858-866).*
22. *Lee, J., Kim, E., Lee, S., Lee, J., & Yoon, S. (2019). Ficklenet: Weakly and semi-supervised semantic image segmentation using stochastic inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5267-5276).*
23. *Ying, L., Demanet, L., &Candes, E. (2005, September). 3D discrete curvelet transform. In Wavelets XI (Vol. 5914, p. 591413). International Society for Optics and Photonics.*
24. *Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., & Cho, H. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4).*
25. *Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5, 23.*

Relational Coordination Parameters and Correlation Analysis for enhancing Work practices in Health Care

***Mr. M. Selwin[1], Dr. N. Rajini[2]**
[1] Assistant Professor,
[2]Professor,
Department of Mechanical Engineering, Kalasalingam Academy of Research and Education, Krishnankoil.

ABSTRACT

Organization is the place where quality and efficiency are really matters a lot. The most important factor which helps the organization to achieve both the outcomes is Relational Coordination (RC). The Communication and Relationship ties of RC correlates the work practices model with great impact. Here quality and efficiency outcome of health care sector is improvised by finding the flaws in the work practices model. This can be achieved by improving the relationship between various dependent and independent variables in both RC and Work practices mediation model. 2 sets of Questionnaires have been framed and tested for internal consistency. The reliability between individual variables have been found by Correlation analysis. The Correlation analysis suggests that the task interdependency between RC dimensions and work practices model variables are interlinked with each other. Proper improvement in RC dimensions will exhibit positive outcome in the work practices model which in turn improve the quality outcome.

Keywords: Relational Coordination (RC), Work Practices Model, Cronbach alpha, Correlation Analysis

SUPPLY CHAIN MANAGEMENT IN HEALTH CARE

In health care sectors SCM is a new way of conceptualizing Medical handled management. SCM is defined as "A imaginary functional web that manipulates the entire flow of of materials, raw communication data, description about the products, procurement and delivering of goods"

The group of people working for the attainment of a specific goal with mutual understanding is meant to be Coordination. In the alarming world of business race, it is impossible for a firm to stand tall without having a plan to reduce the overall expenses. Even the health care providers are sailing in the same boat which always have the danger of sinking in the situation of high cost rise and shrinking in terms of quality aspects.This uphill rise of cost always makes the health care providers to take a step back in the business environment. Thus, the SCM is becoming a hero in many situations which effortlessly handle both the price rise and quality. This grabs the lime light of all the stake holders in various service sectors. The role of SCM in medical care providers bind several materials and make various collaborators to involve effectively. The main function involves the timely assistance of care to the patients without any compromises. Information flow is always a two-way process which always stimulates a great relationship between the service providers and the patients.

RELATIONAL COORDINATION (RC)

It is the combined effort to monitor, organize, enhance and develop the overall performance of work force. The role of RC creates a greater impact in developing the above mentioned SCM base in the hospitals. For setting those targets various parameters have to be taken into account.

PARAMETERS TO MEASURE RC

The figure shows the various questions to be addressed for the proper implementation of RC in health sector.

Frequent Communication	How **frequently** do you communicate with people in each of these groups about the patients?
Timely Communication	Do people in these groups communicate with you in a **timely way** about the patients?
Accurate Communication	Do people in these groups communicate with you **accurately** about the patients?
Problem Solving Communication	When a problem occurs with care of the patients do the people in these groups work with you to **solve the problem**?
Shared Goals	How much do people in these groups **share your goals** regarding care of the patients?
Shared Knowledge	How much do people in each of these groups **know** about the work you do with the patients?
Mutual Respect	How much do people in these groups **respect** the work you do with the patients?

Fig 1 RC Parameters

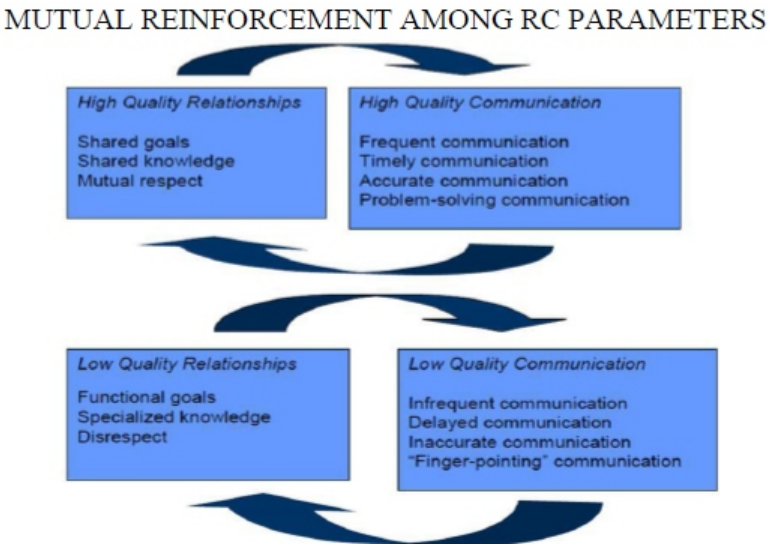

Fig 2 RC Reinforcement

On keeping the above parameters in mind and in order to achieve between outcome in terms of quality and care a new model for better work practice is framed.

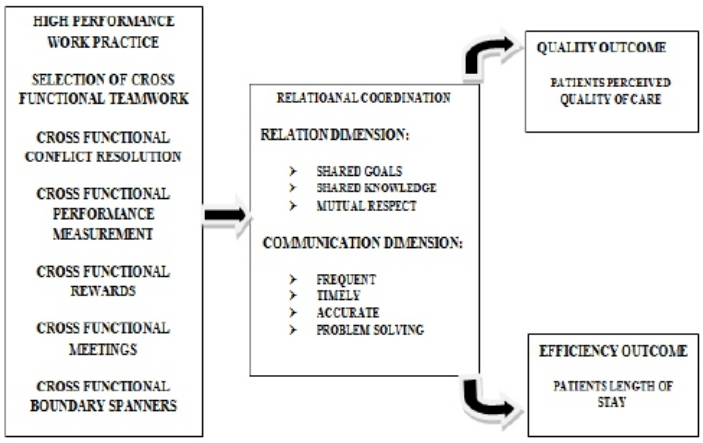

Fig 3 Work Practice RC model

TABLE 1 WORK PRACTICE MODEL

WORK PRACTICE MODEL	PURPOSE OF USING
Cross functional teamwork (CFTW)	Multi-disciplinary work involvement for solving issues to the patients.
Cross functional rewards (CFR)	Periodic appraisal and encouragement
Cross-functional conflict resolution (CFCR)	Framing a standard problem-solving methodology
Cross-functional performance measurement (CFPM)	Monitor the creditability of employers
Cross –functional team meetings (CFTM)	Enhance the Multi-disciplinary team for active participation

I. LITERATURE SURVEY

TABLE 2 LITERATURE ANALYSIS

AUTHOR	COORDINATION PROBLEM	COORDINATION MECHANISM	PERFORMANCE MEASURED
Jody Hoffer Gittell, Kathleen M Fairfield, Robert Jackson, William head (2000)	Length of Stay	Mutual respect	Understanding and equality for other care providers
Jody Hoffer Gittel (2002)	Low communication between primary physicians	Coordination and communication	Good communication and information transfer
Jody Hoffer Gittell ,Dana Beth Weinberg ,Cori Kautz (2007)	Coordination in complex work	Knowledge transfer	Effective knowledge transfer
Jody Hoffer Gittell(2008)	During the quality Maintaince hospitals feels difficult in cost reduction	Relational coordination in high module	Relational work system performance
Zhen Xiong Chen,Yizheng Shi,Da-Hai Dong(2008)	Reduced the customer relationship	Relationship between providers and consumers of	Outcomes of the relationship between providers and consumers

Jody Hoffer Gittell, Dana Weinberg, Susan Pfefferle (2008)	Difficult to knowledge sharing	Knowledge sharing	Effective knowledge sharing networks in professional complex systems
Abraham carameli ,Jody Hoffer Gittel (2008)	Way of communication and knowledge sharing	Collaborative communications, Knowledge transfer	Competencies within different types of healthcare organizations
Rob Seidner et.al (2010)	Organizational problems	Communication and integration	Enhancing improved integration to facilitate multi-disciplinary collaboration
Donna S. Havens et.al (2010))	Difficult in relationship, and customer satisfaction	Doctor-patient relationship, Customer satisfaction	Efficiency of quality of care
Jody Hoffer Gittell(2011)	Low performance	Relational coordination	Investigated performance effects of relational coordination
Chantal Sylvain and Lise Lamothe (2012)	Difficulties in integration service	Integrate professional services	Work on the dynamics of sense making in organizational life

II. PROCESS FLOW

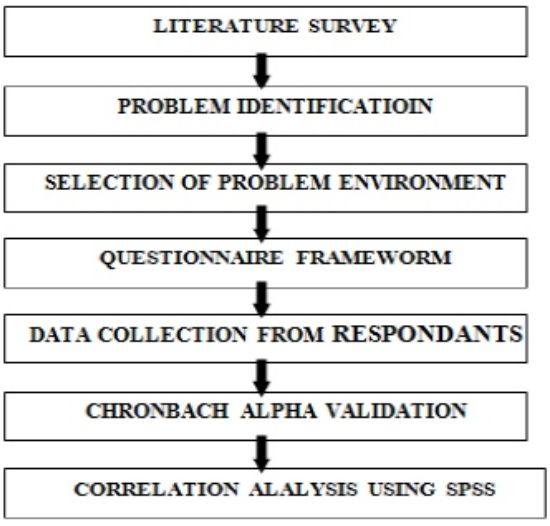

Fig 4 Process flow RC implementation

III. SUITABLE STUDY ENVIRONMENT SELECTION

The health care providers lie in the same area where our research travels. These work places rely more on quality as well as care. The perfect work environment which suits our study is Primary health centers or Medical Care proving centers. To achieve best quality a perfect environment, have to be created where all the gaps on the flow of information between the care providers should be filled unconditionally. The next step is to choose the set of respondents from various departments who are having multi-disciplinary links. Starting from the hierarchy of control the selected respondents are,

- Doctors
- Nurses Nurse Assistants
- Technicians
- Pharmacist
- Receptionist

SURVEY QUESTION:

CFTW

1. Do you believe working in team formation improves quality of care?
2. Is the previous health records of an individual very important for giving best treatment?
3. Is it fair to say lack of experience in team does not help to solve complicated problems?
4. Do you think working in a cross functional team improves outcomes in terms of quality?
5. Equality in work load allocation helps to share the load of individual worker

CFTR

6. Do you have the practice of getting appreciated for your good work?
7. Will your team gives you a better space to get reward for your work?
8. Do you think the appreciation you received your work matches the effort of your work?
9. Do you think you are in a wrong team which affects your performance?
10. Is there a scope for creativity and innovation in your team?

CFTM

11. Effective participation in team meetings helps to resolve problems?
12. Do you think all the problems you are facing are addressed in team meetings?
13. Are you getting proper invitation for attending meetings in different departments?
14. Do you think commencing and completing of meeting are planned perfectly?
15. Does the multi-disciplinary team meetings helps you to rectify your flaws?

CFCRP:

16. When a problem arises do you have a common strategy to address the problem?
17. Do you get frequent support from your colleague for facing a difficulty in work?
18. Do you face any misunderstandings with your co-worker during work?
19. Do you think the reason for Conflict of interest between the workers is due to improper information sharing from hierarchy of control?
20. The conflicts within the organization always creates a negative impact in the care given to the patients?

CFPM:

21. Do you think the cross functional treatment expands the patient's retention in the hospital?
22. Immediate response to the problem always give a positive feedback in patients duration of stay?
23. Do multi-disciplinary team meetings helps to sort out the problems of patients in several aspects?
24. Cross functional appreciation to the workers increases their morality to work?
25. The length of stay of a particular individual can be enhanced by properly resolving the conflict of interest?
26. Most of the patients are looking for health care providers who have multi-disciplinary skills?
27. The method of linkage between Cross functional approach and service quality is?
28. Steps taken for solving the patient's problem will have direct relation with the quality of service given to them?
29. Multi-disciplinary team meetings helps to improve the quality of care given to the patient?
30. Appreciating the workers in multi departmental level helps to improve the level of care given?
31. Resolving the conflict of interest between the workers results in better quality outcome?
32. Does the impact of cross functional team work made impact in service quality.

RC

33. Do you discuss the details of the patient with your colleague frequently?
34. Will you have the habit of sharing the case history of the patients to other people?
35. Do you communicate the criticality of the patients in a measured quantum of time?
36. Will you maintain timely cross functional approach to discuss about the patient?
37. Does the interrogation you are having about the patient is exactly matches the actual problem?
38. Does accurate communication is carried out with other department people about the patients?
39. Can you get help from your co-workers during critical situations of your patient?
40. Can you get help from the workers of other units during critical situations of your patient?
41. Does all the workers maintain mutual respect towards each other?
42. Do you get recognition for your effort outside your department?
43. Do you share the objective of your work to others in the department?
44. Will the persons outside your department understand your idea in patients care easily?
45. Does all the members in your department knows others roles in quality care?
46. Does the vision of your department is open to others for better understanding?

CHECKING THE RELIABILITY

It is necessary to fix a optimum range in Cronbach alpha for maintaining the consistency of the survey. The primary thing is to ensure the reliability of the questionnaire with respect to the selected variables should meet a certain requirement. If it fails we need to change the variables accordingly. The acceptable consistency range should be fixed at 0.70and above for empirical research. 0.80 and above should be fixed for non empirical research study.

CRONBACH ALPHA

Feasibility between the variables and their behaviour with respect to each other will be determined by calculating the Cronbach alpha using Kuder Richardson formula.

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} \sigma_{Y_i}^2}{\sigma_X^2}\right)$$

K- Questions taken for survey (in numbers).

$\sigma^2 X_i$ -Variance of choices given/ Total number of variables.

$\sigma^2 Y_i$ -Variance of number of samples.

ALPHA VALUE	INTERNAL CONSISTENCY
$\alpha \geq 0.9$	EXCELLENT
$0.7 \leq \alpha < 0.9$	GOOD
$0.6 \leq \alpha < 0.7$	ACCEPTABLE
$0.5 < \alpha < 0.6$	POOR

Fig 5 Cronbach alpha range

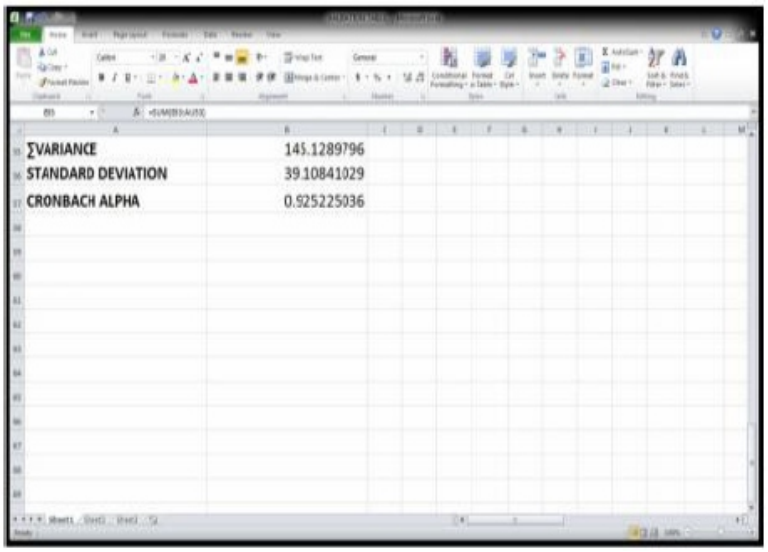

Fig 6 Cronbach alpha for Questionnaire set 1

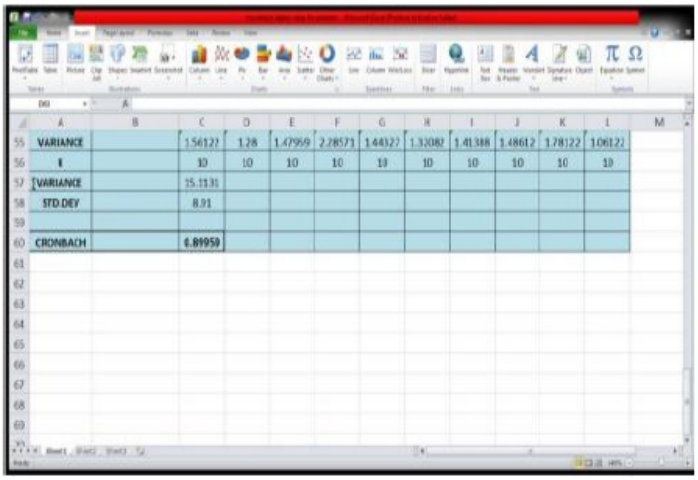

Fig 7 Cronbach alpha for Questionnaire set 2

The Cronbach alpha value have been calculated for 103 participants for 46 questions from set 1 and for 50 participants from set 2. The value for evaluating the question have been taken from 7 point likert scale. Based on the choice of the respondents the variance value got changed. The calculated Cronbach alpha value suggest that the feasibility of questionnaire is in excellent and good range.

CORRELATION ANALYSIS

The relationship capability of dependable variable with respect to in-dependable can be formulated by Correlation analysis.

Coefficient of Correlation = Co-variance of two variables/product of their STD Deviation.

From the above-mentioned stipulated study, it is analyzed that each variable has a relatively impactable relationship with one or many other variables at a certain level of either 0.05 or 0.01. The optimum level of significance we fix for this kind of study would be 0.05.

Correlation analysis

CORRELATIONS

		CFTW	CFTR	CFTM	CFCRP	CFPM	FC	TC	AC	PSC	MR	SG	SK
CFTW	Pearson Correlation	1	.017	.239*	.300**	.384**	.206	-.040	.229*	.160	-.264*	.270*	.090
	Sig.(2-tailed)		.886	.039	.009	.001	.077	.731	.048	.170	.022	.019	.441
	N	100	100	100	100	100	100	100	100	100	100	100	100
CFTR	Pearson Correlation	.017	1	.233	.035	.189	.024	.063	-.140	.163	-.186	-.056	-.107
	Sig.(2-tailed)	.886		.044	.769	.105	.835	.593	.231	.162	.11	.663	.360
	N	100	100	100	100	100	100	100	100	100	100	100	100
CFTM	Pearson Correlation	.239*	.233*	1	.403**	.303**	.245*	-.103	.097	-.056	-.208	.020	.234
	Sig.(2-tailed)	.039	.044		.000	.008	.034	.378	.410	.636	.073	.862	.043
	N	100	100	100	100	100	100	100	100	100	100	100	100
CFCRP	Pearson Correlation	.300**	.035	.403**	1	.412**	.237*	-.105	-.014	.247*	-.010	.007	.087
	Sig.(2-tailed)	.009	.769	.000		.000	.040	.68	.905	.033	.935	.954	.459
	N	100	100	100	100	100	100	100	100	100	100	100	100
CFPM	Pearson Correlation	.384**	.189	.303**	.412**	1	.229*	-.090	-.020	-.314**	-.256*	.046	.110
	Sig.(2-tailed)	.001	.105	.008	.000		.048	.445	.864	.006	.026	.696	.348
	N	100	100	100	100	100	100	100	100	100	100	100	100
FC	Pearson Correlation	.206	.024	.245*	.237*	.229*	1	-.362**	-.071	.213	-.207	-.147	-.030
	Sig.(2-tailed)	.077	.835	.034	.040	.048		.001	.547	.067	.0100	.208	.795
	N	100	100	100	100	100	100	100	100	100	100	100	100
TC	Pearson Correlation	-.040	.063	-.103	-.105	-.090	-.362**	1	.265*	-.096	.077	.038	-.093
	Sig.(2-tailed)	.731	.593	.378	.368	.445	.001		.022	.411	.511	.744	.425
	N	100	100	100	100	100	100	100	100	100	100	100	100
AC	Pearson Correlation	.229*	-.140	.097	-.014	-.020	-.071	.265*	1	-.119	-.088	.479**	.037
	Sig.(2-tailed)	.048	.231	.410	.905	.864	.547	.022		.309	.451	.000	.1003
	N	100	100	100	100	100	100	100	100	100	100	100	100
PSC	Pearson Correlation	.160	.163	-.056	.247*	.314**	.213	-.096	-.119	1	-.024	.016	-.063
	Sig.(2-tailed)	.170	.162	.636	.033	.006	.167	.411	.309		.835	.891	.589
	N	100	100	100	100	100	100	100	100	100	100	100	100
MR	Pearson Correlation	-.264*	-.186	-.208	-.010	-.256*	-.207	.077	-.088	-.024	1	-.063	-.145
	Sig.(2-tailed)	.022	.111	.073	.935	.026	.0100	.511	.451	.835		.589	.215
	N	100	100	100	100	100	100	100	100	100	100	100	100
SG	Pearson Correlation	.270*	-.056	.020	.007	.046	-.147	.038	.479**	.016	-.063	1	.049
	Sig.(2-tailed)	.019	.633	.862	.954	.696	.208	.744	.000	.891	.589		.677
	N	100	100	100	100	100	100	100	100	100	100	100	100
SK	Pearson Correlation	.090	-.107	.234*	.087	.110	-.030	-.093	.037	-.063	-.145	.049	1
	Sig.(2-tailed)	.441	.360	.043	.459	.348	.795	.425	.1003	.589	.215	.677	
	N	100	100	100	100	100	100	100	100	100	100	100	100

Fig 8 Correlation analysis using SPSS

*Correlation is Significant at the 0.05 level (2-tailed)

**Correlation is Significant at the 0.01 level (2-tailed)

RESULT AND DISCUSSION

CRONBACH ALPHA:

The internal consistency range taken for the employees in hospital from the first set Questionnaire was found to be 0.925. As per the Cronbach alpha value (Fig 5) internal consistency range of first set

Questionnaire is in excellent level. Similarly, the (Cronbach alpha) value of questionnaire collected from the patients is 0.89. This shows that the internal consistency is in Good level.

CORRELATION ANALYSIS:

The output thus obtained indicates that there exists a correlation (either positive or negative) for each of the variables with one or more other variables. For a significant level of 0.05 when the p value falls below 0.05 i.e., $p<0.05$ then the null hypothesis will be rejected and subsequently alternate hypothesis gets accepted.

It is found that the task interdependency between RC dimensions and Mediation model variables are interconnected with each other. If one increases the other also increases correspondingly and vice versa. So, RC parameters not only have its impact in the work practices model but also act as a deciding factor on the Quality and Efficiency outcomes. With this we come to conclude that the Relational Coordination dimensions have all the capability to control the entire setup of the work practice model. If a proper structure is framed there might be definitely a greater impact in both quality and efficiency outcome.

REFERENCES

1. *Jody Hoffer Gittell, Kathleen M Fairfield, Robert Jackson, William head (2000) "Impact of Relational Coordination on Quality of Care, Postoperative pain and Functioning, and Length of Stay A Nine-Hospital Study of Surgical Patients"-Journal of Medical Care, Vol.38 No.8, 807-819*
2. *Jody Hoffer Gittel (2002) "Coordinating Mechanisms in Care Provider Groups: Relational Coordination as a Mediator and Input Uncertainty as a Moderator of Performance Effects' Management Science Vol. 48, No. 11, November 2002. 1408–1426*
3. *Jody Hoffer Gittell, Dana Beth Weinberg, Cori Kautz (2007) "Modularity and the coordination of complex work" Health Services Research February 200742(1 Pt. 1):7–2*
4. *Jody Hoffer Gittell (2008), "Relationships and Resilience: Care Provider Responses to Pressures from Managed Care" Journal of Applied Behavioral Science 2008; 44; 25*
5. *Jody Hoffer Gittell, Dana Weinberg, Susan Pfefferle (2008) "Impact of relational coordination on job satisfaction and quality outcomes: a study of nursing homes" Human resource management journal, Vol 18 NO 2, 2008*
6. *Abraham carameli, Jody Hoffer Gittel (2008) "High-quality relationships, psychological safety, and learning from failures in work organizations" Journal of Organizational Behavior J. Organiz. Behav. 30, 709–729 (2009)*
7. *Rob Seidner, Jody Hoffer Gittell, Julian Wimbust (2010) "A relational model of how high-performance work system works" Organization Science Vol. 21, No. 2, March–April 2010, pp. 490–506*
8. *Donna S. Havens, Joseph Vasey, Jody Hoffer Gittell (2010) "Relational coordination among nurses and other providers: impact on the quality of patient care" Journal of Nursing Management, 2010, 18, 926–937*
9. *Jody Hoffer Gittell (2011)," The Southwest Airlines Way: Using the Power of Relationships to Achieve High Performance" Journal of Relational Coordination Research Collaborative*

A Literature Survey on Communication and Relationship Ties of Supply Chain Management in Health Care

[1]M. Selwin, [2] N. Rajini
[1] Assistant Professor,
[2] Professor
[1, 2] Department of Mechanical Engineering, Kalasalingam Academy of Research and Education
Email: selwin123@gmail.com

ABSTRACT

Hospitals are complex organizations in which multiple HealthCare professionals work interdependently to deliver care. Such conditions raise the potential for confusion, errors and delays. Consequently, quality has emerged as the prime focus in the delivery of hospital services. While RC research is growing to date, we have no knowledge about RC. Relational coordination (RC), is a mutually reinforcing process of interaction between communication and relationships carried out for purpose of task integration (Gittell, 2002a: 301). Relational coordination theory makes argues that coordination encompasses not only the management of interdependence between tasks but also between the people who perform those tasks. RC differs from other approaches to coordination by proposing three specific relationship dimensions that are needed for effective coordination: Shared Goals, Shared Knowledge and Mutual Respect; and four specific communication dimensions: Frequent, Timely Accurate and Problem Solving. The aim of this paper is to study coordination between RC parameter and High-performance work practices parameters. Questionnaire is prepared based on the seven dimensions of RC, High performance work practice and Quality then Efficiency outcome. Hospitals with highly structured and more inter departments is to be selected and survey will be conducted among them. The obtained results will be used to enhance the quality and efficiency of those hospitals.

Keyword: Relational coordination, Health care, High performance work practice and communication

INTRODUCTION

Globalization and growth of competitions causes the companies search for management concepts which enable the efficiency of operation to be improved and consequently, they lead to improvement in competitive position in the market or in holding the current position. Under current conditions, supply chain management seems to be a concept with fresh perspectives. It is caused by the fact that this concept does not means negation of current relationship between supplier and customer which are frequently contradictory and focused on taking advantage of the market situation at the expense of supply chain partner. In return supply chain management focus on close cooperation of organizations which are the link in supply chain in order to achieve stronger competitive position through effective management.

1. GENERAL

ANALYSIS OF THE CONCEPT OF SUPPLY CHAIN MANAGEMENT

Supply chain management is a relatively new concept, which was first meant by consultants in the early eighties and since then the concept has gained on popularity in scientific papers on logistics and management. However, the scientists working on supply chain management frequently make reference even to scientific output of the fifties or sixties. Practical use of the concept seems to be proved by the

fact that the results of investigations carried out in 1993 in the USA among 325 logistics experts, members of American Council of Logistics Management indicated growth in importance of cooperation between companies and inter institutional aspects of supply chain[1] This mainly concerns specific approach to a group of cooperating companies, which eliminates many currently existing barriers and which is performed in order to manage and coordinate flow of goods, from raw materials to finished goods purchased by customers and other users. This cooperation is focused on achieving high efficiency of each company and their network as a whole, due to integration and coordination, as well as optimization of the added value through all the links in the chain until the product expected by the customer. Success of supply chain management depends on integration and coordination of three types of flow[2] Information Goods Cash

1.2 HEALTH CARE

The past three decades has witnessed fundamental changes in the way health care is financed and delivered in many countries. Faced with mounting economic, social, and political pressures, health services institutions have struggled to adopt methods of quality improvement to enhance both patient outcomes and cost-effectiveness. One consequence of the increased emphasis on quality has been a greater reliance on multidisciplinary collaboration and teamwork in patient care. Ideally, the principal members of the health care team – doctors, nurses, technicians, pharmacist and receptionist – should work together in a true partnership, marked by mutual understanding of roles and responsibilities, and shared mutually-derived clinical goals. Achieving such collaborative relationships appears to offer tangible benefits in terms of improved patient outcomes [1-3]. In health care organizations, supply chain is a new way of conceptualizing Medical supply management. A supply chain is defined as ―A virtual network that Facilitates the movement of product or information from its production, distribution, and consumption‖ (McFadden and Leahy, 2000). In the age of competition, no industry can survive without pondering much about reducing expenditures wherever possible. The same is true for health care industry, which is witnessing sharp rise in price and decrease in Quality in almost all its products and services. The alarmingly high pace of upward movement of cost is making the product of the industry beyond the reach of the mass. Supply chain in this industry being a significant driver of cost and quality of care is therefore grabbing all the attention from industry stakeholders.

Virtual centralization of the supply chain on the other hand helps improving level of cooperation in hospitals thereby helping them controlling costs and improving services. SCM brings together geographically dispersed healthcare units together and allows them to work together to attain maximum efficiencies in procurement, contracting and customer service. The healthcare supply chain involves the flow of many different product types and the participation of several stakeholders. The main purpose of the healthcare supply chain is to improvise communication flow, in order to fulfill the needs of providers.

The healthcare supply chain is frequently described as highly fragmented and relatively inefficient (Schneller and Smeltzer 2006). A major problem with the traditional healthcare supply chain is that each stage of the supply chain operates independently, leading to misaligned incentives and conflicting goals that prevent the supply chain from operating as a system. In hospitals, SCM covers both medical and non-medical items as well as both operational and capital expenditures. The medical category includes hospital administrators who strive to improve operations in their facilities because of the far-reaching impact it has on patient experience, staff morale, and overall costs, as well as quality of care. To have the best quality of service provided to the customers, the coordination between various departments of the

hospital should be effective. interrelationship of each care provider with that of the others is depicted clearly in fig

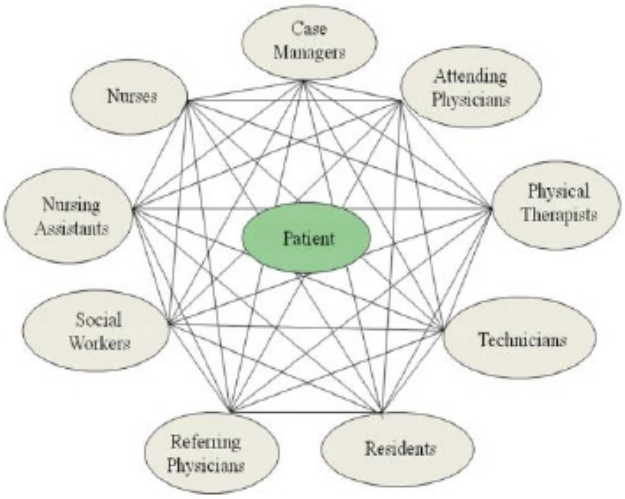

Figure 1.: Inter-relationship of various providers in health care(Foreign),Gittell,2012

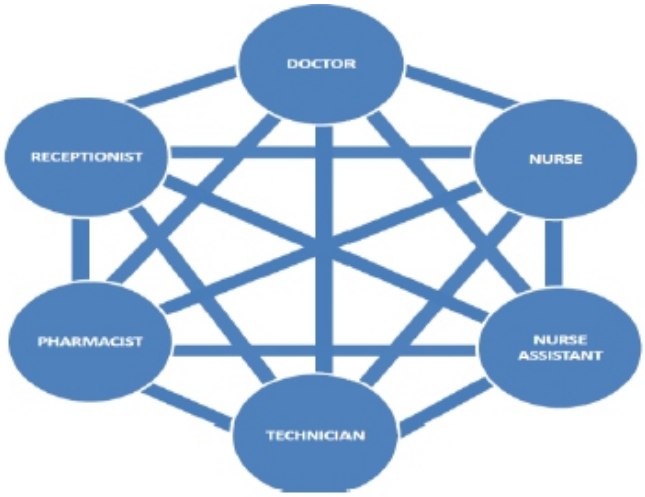

Figure 2 : Inter-relationship of various providers in health care system (INDIA)-modified

1.3 COORDINATION

Coordination is the act of organizing, making different people or things work together for a goal or effect to fulfill desired goals in an organization. Coordination, the management of interdependencies among tasks is believed to be critical for organizational performance. Well-coordinated work processes are expected to produce higher-quality outcomes. Care coordination is one of six priorities identified for Quality Improvement in Health Care. These priorities were chosen because they have the potential to rapidly improve health outcomes and increase the effectiveness of care in health care. For many patients today, health care is delivered in multiple settings and by multiple professionals and organizations. Collaboration and effective communication among healthcare professionals, health plans, patients, and caregivers in all care settings are necessary to provide high-quality care.

1.4 RELATIONAL COORDINATION (RC)

Relational coordination is an emerging theory for understanding the relational dynamics of coordinating work. Other theorists have argued for the importance of relationships for coordinating work, based on the argument that coordination is the management of task interdependence and is therefore a

fundamentally relational process (Crowston and Kammerer, 1997; Bechky, 2006; Faraj and Sproull, 2000; Gittell, 2006; Weick and Roberts, 1994). According to the theory of relational coordination, coordination that occurs through frequent, high quality communication supported by relationships of shared goals, shared knowledge and mutual respect enables organizations to better achieve their desired outcomes. Specifically, relational coordination is a mutually reinforcing process of interaction between communication and relationships carried out for the purpose of task integration(Gittell,2002a:301). According to this theory, three dimensions of relationships are integral to the process of coordination: shared knowledge, shared goals and mutual respect.

In many of the contexts where it has been explored, relational coordination appears to have a significant positive impact on key measures of performance, including both quality and efficiency. Relational coordination is an example of a fundamental process improvement that enables a work group, department or organization to shift out its production possibilities frontier to a more favorable position, achieving higher levels of quality while simultaneously achieving greater efficiencies. More specifically, relational coordination improves a work process by improving the quality of work relationships between people who perform different functions in that work process, thus leading to higher quality communication. Task interdependencies are therefore managed in a more seamless way, with fewer redundancies, lapses, errors and delays. Relational coordination enables employees to more effectively coordinate their work with each other, thus pushing out the production possibilities frontier to achieve higher quality outcomes while using resources more efficiently for example, enabling hospital workers to achieve higher patient perceived quality of care along with shorter patient lengths of stay. Relational coordination is therefore particularly relevant in industries that must maintain or improve quality.

1.5 DIMENSIONS OF RELATIONAL COORDINATION

The seven tenants of RC fig 3; frequent, timely, accurate, problem solving communication in an environment of shared goals, shared knowledge and mutual respect, can be tied directly to two of the three dimensions found in social capital theory (Nahapiet and Ghoshal 1998). Shared goals and the related HRM practice of shared rewards are structural practices underpinning the relational dimension by encouraging the subordination of individual or functional goals for the greater cross-functional good. Setting shared goals and rewards as a management strategy offers a practical method for management to affect coordination. The effective communication and implementation of these goals between management and workers can be expected to build mutual respect between workers (Ahmed et al. 2010).

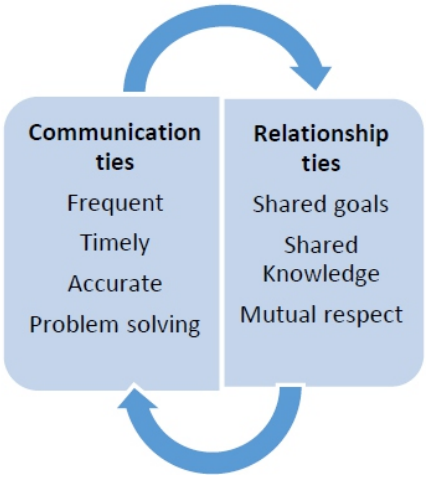

Figure3: Dimensions of Relational Coordination (Gittell 2002)

1.6 Frequent Communication:

Communication is the process of exchanging information, both verbal and nonverbal, within an organization. An organization may consist of employees from different parts of society. In order to unite the activities of all employees, communication is crucial. Communicating necessary information to the entire workforce becomes necessary. While formal communication that is done too rarely or too often is not good for an organization. Information must be communicated as and when required rather than holding unnecessary meetings frequently. At the same time crucial information must not be held till the last hour or day, instead they have to be communicated as early as possible to get the employees in tune with the objectives of the organization. (Picardi, Richard p. (2001). Skills of workplace communication: a handbook for T & D specialists and their organization. West port, conn.: Quorum Books) By contrast, some argue that high quality connections can exist independent of the frequency of communication (Dutton and Heaphy, E.D, 2003). While recognizing the importance of frequent communication for coordinating highly interdependent work, relational coordination encompasses far more than simply the frequency of communication.

Timely communication:

Communication can be frequent and still be of poor quality. For one thing, it can lack timeliness. Knowledge of each participant's contribution to the overall work process enables everyone to communicate in a timely way across functions, grounded in an understanding of who needs to know what, why, and with what degree of urgency. In coordinating highly interdependent work, timing can be critical (2013 Jody Hoffer Gittell and Anthony L. Suchman). Delayed communication may result in errors or delays, with negative implications for organizational outcomes. Research supports the importance of timely communication for successful task performance (e.g. Waller, 1999), though timely communication has not been widely recognized as essential to the coordination of highly interdependent work.

Accurate communication:

The effective coordination of work depends not only on frequent and timely communication, but also on accurate communication. If updates are received frequently and in a timely way, but the information is inaccurate, the consequences for error or delay are apparent. (O.Reilly and Roberts 1977) argued and showed that accurate communication plays an important role in task group effectiveness. The accuracy of communication can also be conceptualized in terms of its trustworthiness, as in the knowledge seeking literature, and specifically in recent work by Levin and Cross (2004).

Problem Solving Communication:

Task interdependencies often result in problems that require joint problem solving. Hence, effective coordination requires that participants engage in problem solving communication. However, the more common response to interdependence is conflict (Gladstein, 1984) as well as blaming and the avoidance of blame (Donnellon, 1994). The resort to blaming rather than problem solving reduces opportunities to solve problems, with negative consequences for performance. Others have recognized the role that problem solving plays in the coordination of highly interdependent work (Stevenson and Gilly, 1993; Rubinstein, 2000).

2. RELATIONSHIP DIMENSIONS OF RC

Participants ability to effectively coordinate their work is also influenced by the quality of their relationships, particularly the extent of shared goals, shared knowledge and mutual respect.

2.1 Shared Goals:

Shared goals increase participants' motivation to engage in high quality communication and predispose them towards problem-solving rather than blaming when things go wrong (2013 Jody Hoffer Gittell and Anthony L. Suchman). Effective coordination depends upon participants having a high level of shared goals for the work process in which they are engaged. With a set of shared goals for the work process, participants have a powerful bond and can more easily come to compatible conclusions about how to respond as new information becomes available. However, shared goals are often lacking among those from different functional areas who are involved in the same work process. In their classic work on organizations, (March and Simon 1958) described the negative outcomes that occur when participants pursue their own functional goals without reference to the super ordinate goals of the work process in which they are engaged. More recently, group theorists have identified shared goals as playing an important role in the coordination of highly interdependent work (Saavedra, Earley and Van Dyne, 1993; Wageman,1995).

2.2 Shared Knowledge:

Shared knowledge enables participants to communicate with each other with greater accuracy, based on an understanding of how their own tasks relate to the tasks of others functions (2013 Jody Hoffer Gittell and Anthony L. Suchman). Furthermore, effective coordination depends upon participants having a high degree of shared knowledge regarding each other's tasks. When participants know how their tasks fit together with the tasks of others in the same work process, they have a context for knowing who will be impacted by any given change and therefore for knowing who needs to know what, and with what urgency. But shared knowledge is often lacking. Consistent with sociological theories, Dougherty (1992) showed that participants from different functional backgrounds often reside in different thought worlds. due to differences in their training, socialization and expertise.

2.3 Mutual Respect:

Mutual respect increases the likelihood that participants will be receptive to communication from their colleagues irrespective of their relative status, thus increasing the opportunity for shared knowledge and problem solving. Finally, effective coordination depends upon participants having respect for other participants in the same work process. Disrespect is one of the potential sources of division among those who play different roles in a given work process. Occupational identity serves as a source of pride, as well as a source of invidious comparison. Members of distinct occupational communities often have different status and may bolster their own status by actively cultivating disrespect for the work performed by others (Van Maanen and Barley, 1984). When members of these distinct occupational communities are engaged in a common work process, the potential for these divisive relationships to undermine coordination is apparent. By contrast, respect for the competence of others creates a powerful bond, and is integral to the effective coordination of highly interdependent work (Eisenberg, 1990; Rubenstein, Barth and Douds, 1971). The Fig5 shown below portrays the mutual reinforcement that is expected to occur between the communication and relationship dimensions of relational coordination, illustrating how this mutual reinforcement can occur in either a positive or negative direction. (JOURNAL OF APPLIED BEHAVIOUR SCIENCE-2008) and fig 5 shows the meditation model of RC

Figure 4: Mutual Reinforcement among Relational Coordination Dimensions

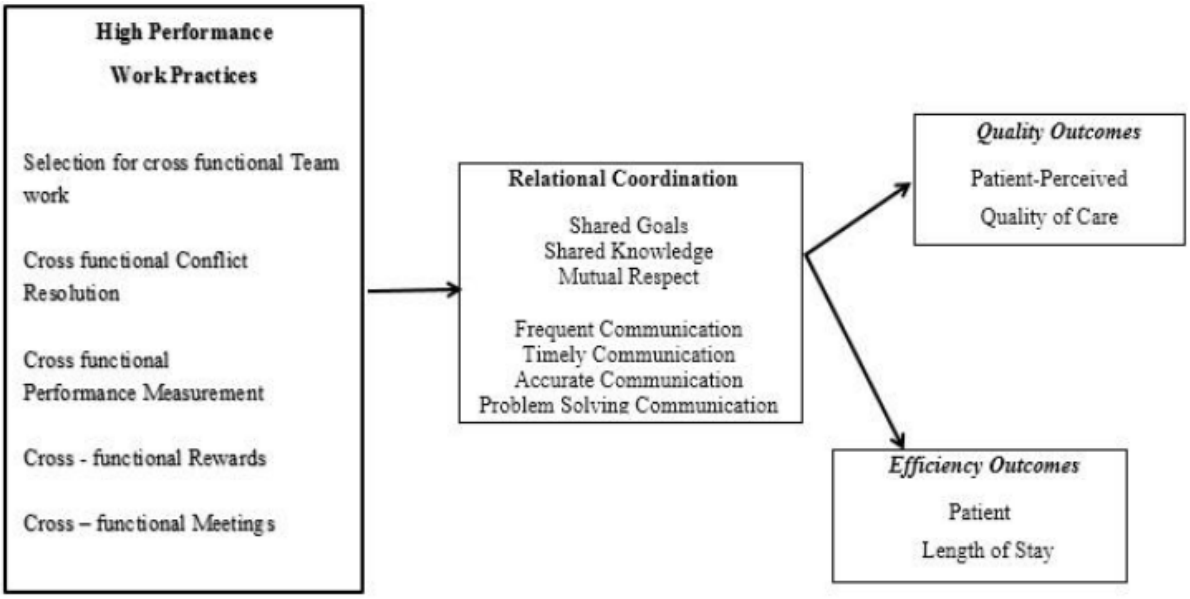

Figure 5: A Modified Model of How High-Performance Work Systems Work

3. MEASURING RC

Relational coordination is measured by surveying participants in a particular work process about their communication and relationships with other participants in that work process. The first step in measuring relational coordination is to identify a work process that serves a client population of interest the focal work process then to identify the roles or functional groups that are involved in carrying out that focal work process. It is helpful to conduct informational interviews to identify all functional groups that are expected to impact the quality and efficiency outcomes of that focal work process. The set of

functional groups involved in a patient care process, for example, may include Doctors, Nurses, Technicians, Pharmacy and Receptionist. These functional groups are listed in the relational coordination survey instrument below each of the seven relational coordination questions enabling the survey respondent to answer each of the questions about their coordination with members of each of these functional groups.

3.1 Survey Items

The relational coordination measures are aggregated from seven survey questions including four questions about communication (frequency, timeliness, accuracy, problem-solving) and three questions about relationships (shared goals, shared knowledge, mutual respect). Respondents from each of the functions believed to be most central to the focal work process are asked to answer each of the following questions with respect to each of the other functions, with responses recorded on a seven-point Likert-type scale.

To lessen the problem of socially desirable responses to survey questions, the relational coordination survey asks respondents to report the behaviors of others as opposed to being asked to report their own behaviors. For example we ask : "Do people in this group communicate with you in timely way about work process?" Due to social desirability bias, respondents are likely to overestimate the extent to which they communicate in a timely way with other employees, for example, but less likely to overestimate extent to which other employees communicate with them in a timely way. In addition, relational coordination questions are asked to elicit respondents, perceptions of typical patterns rather than specific incidents. Finally, in order to reduce the problem of retrospective response error, the questions do not ask for past reports, rather they ask respondents to describe current working conditions

The unit of observation for relational coordination is the individual participant in the work process, represented by the individual survey respondent. These individual respondents are then aggregated into a larger unit of analysis in order to construct a measure of relational coordination. That unit of analysis will depend on the hypothesis you are exploring. If you are studying an intervention that is expected to improve relational coordination of a particular work process, and the performance of that work process, your unit of analysis will be different periods of time, i.e. before and after the intervention has been implemented. If you are doing a cross sectional study in which multiple sites that independently carry out the same work process are expected to have different levels of relational coordination, which are expected to result in different level of performance, your unit of analysis will be the site.

4. ANALYZING RC

4.1 Validating a questionnaire

First, we should test the validity of aggregating the seven dimensions of RC into a single index. Using individual survey responses as your unit of observation, test Cronbach alpha among the seven dimensions of RC to see if they constitute a valid index. For index validity, Cronbach alpha should be greater than 0.70 for an exploratory study, and greater than 0.80 for a non-exploratory study (Nunnally, 1978).

Cronbach's Alpha

If we want to evaluate a survey, it would be nice to know that the instrument we are using will always elicit consistent and reliable response even if questions were replaced with other similar questions.

When we have a variable generated from such a set of questions that return a stable response, Then the questions designed for questionnaire are said to be reliable. Cronbach's alpha is an index of reliability associated with the variation accounted for by the true score of the "underlying construct." Alpha coefficient ranges in value from 0to1 and may be used to describe the reliability of factors extracted from dichotomous (that is, questions with two possible answers) and/or multi-point formatted questionnaires or scales (i.e., rating scale:1=poor, 7=excellent). The higher the score, the more reliable the generated scale is. Nunnally (1978) has indicated 0.7 to be an acceptable reliability coefficient but lower thresholds are sometimes used in the literature. Cronbach (alpha)is a coefficient of internal consistency. It is commonly used as an estimate of the reliability of a psychometric test for a sample of examinees. It was first named alpha by Lee Cronbach in 1951, as he had intended to continue with further coefficients. The measure can be viewed as an extension of the Kuder–Richardson Formula.

It is given by,

$$\alpha = \frac{K}{K-1}\left(1 - \sum_{i=1}^{K} \frac{\sigma^2 Y}{\sigma^2 X}\right)$$

Fig.6 Formula for calculating Cronbach alpha

where,
K-No of questions

σ_X^2 is the variance of the observed total test scores,

$\sigma_{Y_i}^2$ the variance of component i for the current sample of persons

Cronbach alpha is the most common measure of internal consistency ("reliability"). It is most commonly used when we have multiple Likert questions in a survey/questionnaire that form a scale. The acceptable alpha values are shown in the table

ALPHA VALUE	INTERNAL CONSISTENCY
$\alpha \geq 0.9$	EXCELLENT
$0.7 \leq \alpha < 0.9$	GOOD
$0.6 \leq \alpha < 0.7$	ACCEPTABLE
$0.5 \leq \alpha < 0.6$	POOR

Table4.1 Acceptable alpha values for checking internal consistency

5.DESIGN OF QUESTIONNAIRE

The questionnaire design process starts with the formulation of survey objectives and information requirements and continues with the following steps:

- Knowledge of the respondents
- Reviewing previous questionnaires
- Draft the questionnaire
- Validate the questionnaire
- Review & revise questionnaire
- Finalize questionnaire.

Questionnaire is framed based on the meditation model of RC.32 questions have been framed under High Performance Work Practices and 14 questions from Relational Coordination.

CONCLUSION:

The theory of RC is a promising framework for improving communication and relationships among health care providers. To summarize, the theory of relational coordination states that the coordination of work is most effectively carried out through frequent, high quality communication and through high quality relationships among participants. Furthermore, the theory of relational coordination argues that relationships of shared goals, shared knowledge and mutual respect support frequent, high quality communication and vice versa and that these dimensions work together to enable participants to effectively coordinate their work. Shared goals motivate participants to move beyond sub-goal optimization and to act with regard for the overall work process. Shared knowledge informs participants of how their own tasks and the tasks of others contribute to the overall work process, enabling them to act with regard for the overall work process. Respect for the work of others encourages participants to value the contributions of others and to consider the impact of their actions on others, further reinforcing the inclination to act with regard for the overall work process. This web of relationships reinforces, and is reinforced by, the frequency, timeliness, accuracy and problem-solving nature of communication, enabling participants to effectively coordinate the work processes in which they are engaged. I propose the above work practices: selection and training for cross-functional teamwork, the use of conflict resolution to build relationships between workers, feedback and rewards that are oriented toward contributions to shared goals, and information sharing or coordinating mechanisms like team meetings and boundary spanners. I call these work practices relational work practices and suggest that when they are adopted together, they form a relational work system. Thus, the Relational Coordination parameters were selected for measuring the coordination among workers in health care sectors. Based on the relational work practices, questionnaire will be prepared for surveying purpose. The data obtained from the survey will be used for measuring the quality of health care.

REFERENCES:

[1] KOT S.: Supply Chain Management in Global Scale. Enterprises, Entrepreneurship in the Face of Globalization and European Integration - Economic, Legislative and Social Challenges. Ed.by Ryszard Borowiecki, Ofic. Wyd. ABRYS, Kraków2006

[2] RUTKOWSKI K. (ed.): Zintegrow any łańcuchdostaw. Doświadczeniaglobalne I Polskie, red. AGH, Warszawa 2002

[3] COYLE J.J, BARDI E.J., LANGLEY C.J: Zarządzanielogistyczne, PWE Warszawa 2003

[4] Gittell: Relational Coordination: Guidelines for Theory, Measurement and Analysis, February 7, 2012

[5] Gittell and Anthony L. Suchman: An overview of Relational Coordination,2011

[6] Gittell, JH, Fairfield, K, Bierbaum, B, Jackson, R, Kelly, M, Laskin, R, Lipson, S, Siliski, J, Thornhill, T, Zuckerman, J. Impact of relational coordination on quality of care, post-operative pain and functioning, and length of stay: a nine-hospital study of surgical patients. Medical Care. 2000; 8(8): 807-819.

[7] Bae, SH, Mark, B, Fried, B. Impact of nursing unit turnover on patient outcomes in hospitals. Journal of Nursing Scholarship 2010; 42(1): 40-49.

[8] Gittell JH. Relationships between service providers and their impact on customers. Journal of Serv Res. 2002; 4(4): 299-311.

[9] Gittell, JH, Weinberg, DB, Pfefferle, S, Bishop, C. Impact of relational coordination on job satisfaction and quality outcomes: A study of nursing homes. Human Resource Management Journal2008;18(2):154-170.

[10] Weinberg, DB, Lusenhop, W, Gittell, JH, Kautz, C. Coordination between formal providers and in for all care givers. Health Care Management Review. 2007; 32(2): 140-150.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

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

Manuscript Guidelines

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

Title

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

Letterhead Title

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

Author's Name

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

Contact Details

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

Type of Articles

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

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

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

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

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

Keywords

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

Acknowledgements

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

Tables and Illustrations

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

Citation in the Text

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

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Address of the Editorial Office:

Enriched Publications Pvt. Ltd.
S-9,IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-45525005