INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 225-231

ADAPTIVE SOFT COMPUTING STRATEGIES FOR DYNAMIC MULTI MODAL ENVIRONMENTS

Pagidi Premchand[1], Yerravarapu V V Durga Prasad[2]

[1]Asst. Professor, Department of CSE-CS, Chalapathi Institute of Technology, Guntur, A.P, India.

[2]Asst. Professor, Department of CSE-AIML&DS, Aditya College of Engineering and Technology, Surampalem, A.P, India.

ABSTRACT

In this paper we will describe an intelligent multi-modal interface for a large workforce management system called the smart work manager. The main characteristics of the smart work manager are that it can process speech, text, face images, gaze information and simulated gestures using the mouse as input modalities, and its output is in the form of speech, text or graphics. The main components of the system are a reasoner, a speech system, a vision system, an integration platform and an application interface. The overall architecture of the system will be described together with the integration platform and the components of the system which include a non-intrusive neural network based gaze tracking system. Fuzzy and probabilistic techniques have been used in the reasoner to establish temporal relationships and learn interaction sequences.

Keywords: Soft Computing, Multi-modal system, Smart work manger.

1. INTRODUCTION

Soft computing techniques are beginning to penetrate into new application areas such as intelligent interfaces, information retrieval and intelligent assistants. The common characteristic of all these applications is that they are human-centred. Soft computing techniques are a natural way of handling the inherent flexibility with which humans communicate, request information, describe events or perform actions. A multi-modal system is one that uses a variety of modes of communication between human and computer either in combination or isolation. Typically research has concentrated on enhancing standard devices such as keyboard and mouse, with non-standard ones such as speech and vision An Intelligent multi-modal system is defined as a system that combines, reasons with and learns from, information originated from different modes of communication between human and the computer.

The main reason for using multi-modality in a system is to provide a richer set of channels through which the user and computer can communicate.

The necessary technologies for such systems are

- AI and Soft Computing for representation and reasoning
- User interfaces for effective communication channels between the user and the system

In this paper we will describe the development of an intelligent multi-modal system referred to as the smart work manger (SWM) for a large-scale work force scheduling application.

2. RELATED WORK

Research in human-computer interactions has mainly focused on natural language, text, speech and vision primarily in isolation. Recently there have been a number of research projects that have concentrated on the integration of such modalities using intelligent reasoners. The rationale is that many inherent ambiguities in single modes of communication can be resolved if extra information is available. Among the projects reviewed in the references are CUBRICON from Calspan-UB Research Centre, XTRA from German Research Centre for AI and the SRI system from SRI International.

The main characteristics of the SWM are that it can process speech, text, face images, gaze information and simulated gestures using the mouse as input modalities, and its output is in the form of speech , text or graphics. The main components of the system are the reasoner, a speech system, a vision system, an integration platform and the application interface.

3. ENGINEERING ISSUES

Intelligent multi-modal systems use a number of input or output modalities to communicate with the user, exhibiting some form of intelligent behavior in a particular domain. The functional requirements of such systems include the ability to receive and process user input in various forms such as:

- Typed text from keyboard,
- Mouse movement or clicking,

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 225-231

- Speech from a microphone,
- Focus of attention of human eye captured by a camera,

The system must be also able to generate output for the user using speech, graphics, and text.

A system, which exhibits the above features, is called a multi-modal system. For a multi-modal system to be also called intelligent, it should be capable of reasoning in a particular domain automating human tasks, facilitating humans to perform tasks more complex than before or exhibiting a behavior which can be characterized as intelligent by the users of the system.

Given these requirements for intelligent multi-modal systems, it becomes obvious that such systems, in general, are difficult to develop. A modular approach is therefore necessary for breaking down the required functionality into a number of sub-systems which are easier to develop or for which software solutions already exist. Other requirements for such systems are concurrency, a communication mechanism and distribution of processes across a network.

4. RESEARCH METHODOLOGY

The overall architecture of SWM with the various modules and the communications between them is given in Fig. 1.

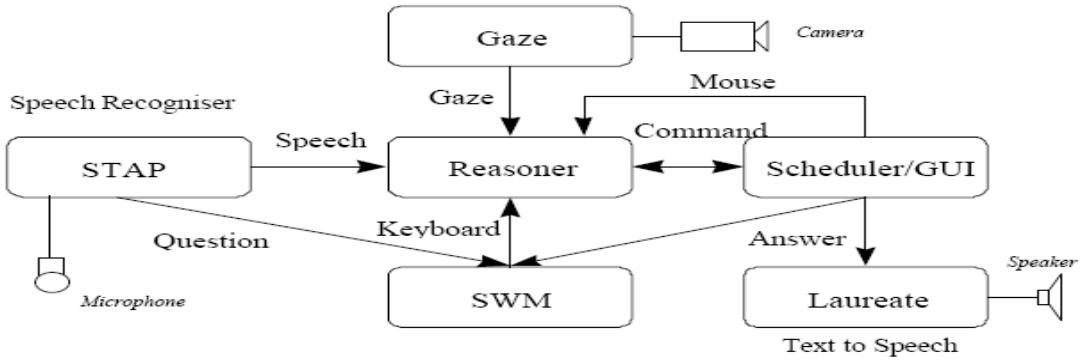

Figure 1 - An overview of SWM Architecture

The Reasoner:

The main functions of the reasoner are two folds. First it must be able to handle ambiguities such as give me this of that.Second it must have the capabilities to deal with often-conflicting information arriving from various modalities. The capabilities of the reasoner are to a large extent dependant upon the capabilities provided by the platform on which the reasoner is implemented. The platform used for the reasoner is CLIPS, which is a well known expert system shell developed by NASA with object oriented, declarative and procedural programming capabilities and the fuzzy CLIPS extension. The reasoner handles ambiguities by using a knowledge base that is being continually updated by the information arriving from various modalities. The structure of the reasoner is shown in Fig. 2. There are five modules in the reasoner: fuzzy temporal reasoning, query pre-processing, constraint checking, resolving ambiguities (WIZARD) and post-processing.

The fuzzy temporal reasoning module receives time-stamped events from various modalities and determines the fuzzy temporal relationship between them. It determines to what degree two events have a temporal relationship, such as before, during or overlapping. Using the certainty factors (CF) of *fuzzy* CLIPS the temporal reasoner can answer questions such as

What is the CF that event 1 took place before event 2

What is the CF that event 1 took place just before event 2

What is the CF that event 1 took place during event 2

What is the CF that event 1 is overlapping with event 2

The relationship with the highest CF will be chosen as the most likely relationship between the two events. This relationship can be used later by the reasoner to resolve conflicts between, and checking dependency of, the modalities.

In the query pre-processing module a sentence in natural language form is converted to a query which conforms to the system's pre-defined grammar.Redundant words are removed, keywords are placed in the right order and multiple word attributes are converted into single strings.

The constraint checking module examines the content of the queries. If individual parts of the query do not satisfy pre-defined constraints then they are replaced by ambiguous terms (*this, that*) to be resolved later, otherwise the query is passed on to the next module.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 225-231

The WIZARD is at the heart of the reasoner and is the module that resolves ambiguities. The ambiguities in this application take the form of reserved words such as *this* or *that*,and they refer to objects that the user is or has been talking about, pointing at or looking at. The ambiguities are resolved in a hierarchical manner as shown in Fig3.

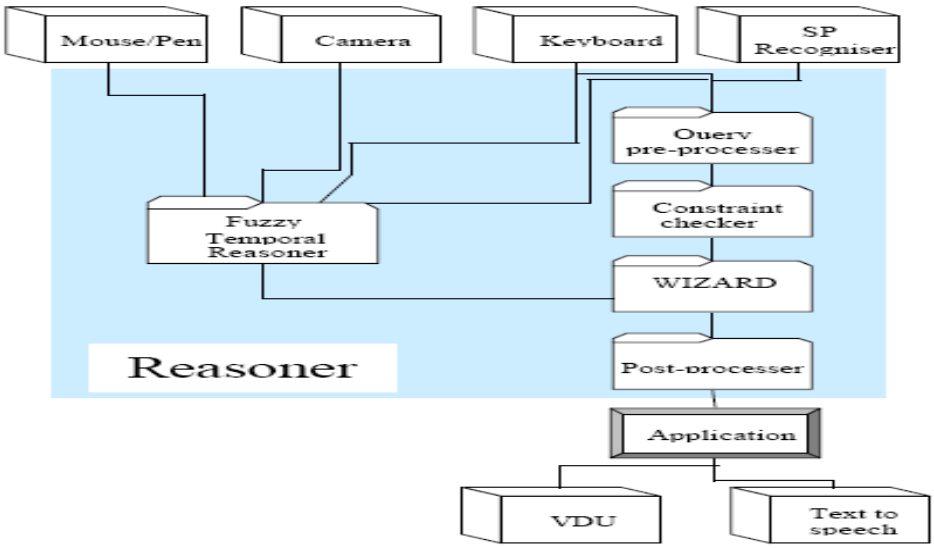

Figure 2 – The structure of the Reasoner

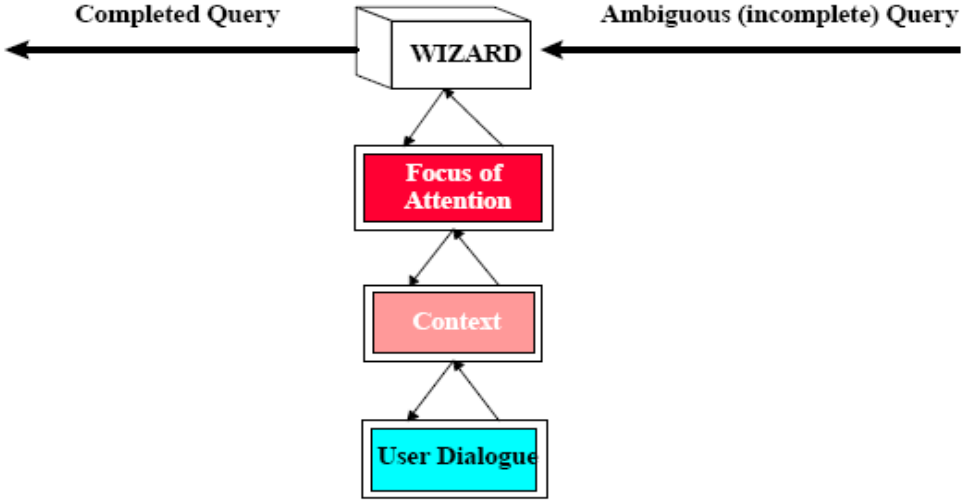

Figure 3 - Resolving ambiguities in the Reasoner

The context of the interactions between the user and the system, if it exists, is maintained by the reasoner in the knowledge base. When a new query is initiated by the user, it is checked against the context. If there are ambiguities in the query and the context contains relevant information then the context will be used to create a complete query that will be sent to the application interface for processing.Another mechanism used to resolve ambiguities is the focus of attention, which is obtained from the user when pointing with the mouse or gazing at an object on the screen. At the same time there could be dialogue with the user through the speech recognizer or the keyboard. CLIPS is mainly used in a declarative mode in this system and therefore all the modules work in parallel. This can cause conflict between information arriving from various modalities. The conflict resolution strategy used in the reasoner is hierarchical as shown in Fig. 3, with the focus of attention having the highest priority and the dialogue system the lowest. This means that the dialogue system will act as a safety net for the other modalities if all fails, or if inconsistent information is received from the modalities. In cases where text input is required however, the dialogue system is the only modality that will be called upon. In all other cases the dialogue system will be redundant unless all others fail in which case a simple dialogue in the form of direct questions or answers will be initiated by the system. The WIZARD sends the completed queries to the post-processing module.

The post-processing module simply converts the completed queries in a form suitable for the application. This involves simple operations such as formatting the query or extracting key words from it.

Table 4 contains some examples of interactions with the reasoner and how it works.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 01, January 2024, pp : 225-231

Query	User Action	Reasoning Process
show me technician ab123 job locations on the map	none	Complete query, process command and sent to application
tell me the duration of this job	mouse is clicked or eyes are focused on a job	this job is ambiguous it is resolved using focus context is updated
show me this of that technician	no focus context is technician *ab123 end of day*	two ambiguities context is used to solve ambiguities
read me this	no focus no context	**everything** is ambiguous the user is asked to repeat the missing parts the context is updated

Table 1 -Examples of Interactions with the Reasoner

Gaze tracking: Technical Issues:

The initial objective for the gaze system is the capability of tracking the eye movements within slightly restricted environments. More specifically, the scenario is a user working in front of a computer screen viewing objects on display while a camera is pointed at the user's face. The objective is to find out where (which object) on the screen the user is looking at, or to what context he is paying attention. This information in combination with other inputs provided by speech and other modalities would be most useful to resolve some real application tasks.

The general difficulties we have to face to build a gaze tracking system are as follows;

- Imprecise data, or the head may pan and tilt resulting in many eye images (relative to the viewing camera) corresponding to the same co-ordinates on the screen.
- Noisy images, mainly due to change of lighting. This is typical in an uncontrolled open plan office environment.
- Possibly infinitely large image set, in order to learn the variations of the images and make the system generalize better.
- Accuracy and speed compromise, for a real-time running system, more complicated computation intensive algorithms have to give way to simple algorithms.

Gaze Tracking: System description

Neural networks have been chosen as the core technique to implement the gaze tracking system

- A three-layer feed-forward neural network is used. The net has 600 input units, one divided hidden layer of 8 hyperbolic units each and corresponding divided output layer with 40 and 30 units, respectively, to indicate the positions along x- and y- direction of the screen grid.
- Grey-scale images of the right eye are automatically segmented from the head images inside a search window, the images of size 40 □□15 are then normalized, and a value between -1 and 1 is obtained for each pixel. Each normalized image comprises the input of a training pattern.
- The pre-planned co-ordinates of this image, which is the desired output of the training pattern, is used to stimulate two related output units along the x and y output layer respectively.
- The training data are automatically grabbed and segmented when one tracks with the eyes a cursor movement following a pre-designed zigzag path across or up and down the screen. A total of 2000 images is needed to train the system to achieve better performance.

The neural network described has a total of 11,430 connection weights. The off-line training takes about half an hour on the Ultra-1 workstation. Once trained, the system works in real-time.

5. CONCLUSION

In this paper I have shown how such flexibility can be exploited within the context of an intelligent multi-modal interface. Soft computing techniques have been used at the interface for temporal reasoning, approximate query matching, learning action sequences and gaze tracking.

It is important to note that soft computing has been used in conjunction with other AI- based systems performing dynamic scheduling, logic programming, speech recognition, and natural language understanding. I believe that soft computing in combination with other AI techniques can make a significant contribution to human-centred computing in terms of development time, robustness, cost, and reliability.

We plan to investigate the following improvements to the reasoner in the near future:

- The context can be extended to have a tree like structure such that the user is able to make reference to previously used contexts.
- The temporal reasoner can be used more extensively in conflict resolution.
- The grammar can be extended to include multiple format grammars.
- The dialogue system can be improved to become more user friendly.
- Different approaches can be used for resolving ambiguities such as competition between modalities or bidding.

6. REFERENCES

[1] Kommineni, K. K. ., & Prasad, A. . (2023). A Review on Privacy and Security Improvement Mechanisms in MANETs. International Journal of Intelligent Systems and Applications in Engineering, 12(2), 90–99. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/4224

[2] Vellela, S.S., Balamanigandan, R. Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023- 15926-5

[3] Vellela, S. S., Reddy, B. V., Chaitanya, K. K., & Rao, M. V. (2023, January). An Integrated Approach to Improve E-Healthcare System using Dynamic Cloud Computing Platform. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 776-782). IEEE.

[4] K. N. Rao, B. R. Gandhi, M. V. Rao, S. Javvadi, S. S. Vellela and S. Khader Basha, "Prediction and Classification of Alzheimer's Disease using Machine Learning Techniques in 3D MR Images," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 85-90, doi: 10.1109/ICSCSS57650.2023.10169550.

[5] VenkateswaraRao, M., Vellela, S., Reddy, V., Vullam, N., Sk, K. B., & Roja, D. (2023, March). Credit Investigation and Comprehensive Risk Management System based Big Data Analytics in Commercial Banking. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 2387-2391). IEEE [6]

[6] S Phani Praveen, RajeswariNakka, AnuradhaChokka, VenkataNagarajuThatha, SaiSrinivasVellela and UddagiriSirisha, "A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques" International Journal of Advanced Computer Science and Applications(IJACSA), 14(6), 2023. http://dx.doi.org/10.14569/IJACSA.2023.01406128

[7] Vellela, S. S., & Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.

[8] Vullam, N., Vellela, S. S., Reddy, V., Rao, M. V., SK, K. B., & Roja, D. (2023, May). Multi-Agent Personalized Recommendation System in E-Commerce based on User. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1194-1199). IEEE.

[9] Vellela, S. S., Balamanigandan, R., & Praveen, S. P. (2022). Strategic Survey on Security and Privacy Methods of Cloud Computing Environment. Journal of Next Generation Technology (ISSN: 2583-021X), 2(1).

[10] Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07), 2020.

[11] Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2022). A New Multi-Level Semi-Supervised Learning Approach for Network Intrusion Detection System Based on the 'GOA'. Journal of Interconnection Networks, 2143047.

[12] Madhuri, A., Praveen, S. P., Kumar, D. L. S., Sindhura, S., &Vellela, S. S. (2021). Challenges and issues of data analytics in emerging scenarios for big data, cloud and image mining. Annals of the Romanian Society for Cell Biology, 412-423.

[13] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., &Swapna, D. (2022, November). An Adaptive Load Balancing Technique for Multi SDN Controllers.In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1403-1409).IEEE.

[14] Vellela, S. S., Basha Sk, K., & Yakubreddy, K. (2023). Cloud-hosted concept-hierarchy flex-based infringement checking system. International Advanced Research Journal in Science, Engineering and Technology, 10(3).

[15] Rao, M. V., Vellela, S. S., Sk, K. B., Venkateswara, R. B., & Roja, D. (2023). SYSTEMATIC REVIEW ON SOFTWARE APPLICATION UNDERDISTRIBUTED DENIAL OF SERVICE ATTACKS FOR GROUP WEBSITES. Dogo Rangsang Research Journal UGC Care Group I Journal, 13(3), 2347-7180.

[16] Venkateswara Reddy, B., Vellela, S. S., Sk, K. B., Roja, D., Yakubreddy, K., & Rao, M. V. Conceptual Hierarchies for Efficient Query Results Navigation. International Journal of All Research Education and Scientific Methods (IJARESM), ISSN, 2455-6211.

[17] Sk, K. B., Roja, D., Priya, S. S., Dalavi, L., Vellela, S. S., & Reddy, V. (2023, March). Coronary Heart Disease Prediction and Classification using Hybrid Machine Learning Algorithms. In 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA) (pp. 1-7). IEEE.

[18] Sk, K. B., & Vellela, S. S. (2019). Diamond Search by Using Block Matching Algorithm. DIAMOND SEARCH BY USING BLOCK MATCHING ALGORITHM", International Journal of Emerging Technologies and Innovative Research (www. jetir. org), ISSN, 2349-5162.

[19] Yakubreddy, K., Vellela, S. S., Sk, K. B., Reddy, V., & Roja, D. (2023). Grape CS-ML Database-Informed Methods for Contemporary Vineyard Management. International Research Journal of Modernization in Engineering Technology and Science, 5(03).

[20] Vellela, Sai Srinivas and Chaganti, Aswini and Gadde, Srimadhuri and Bachina, Padmapriya and Karre, Rohiwalter, A Novel Approach for Detecting Automated Spammers in Twitter (June 24, 2023). Mukt Shabd Journal Volume XI, Issue VI, JUNE/2022 ISSN NO : 2347-3150, pp. 49-53 , Available at SSRN: https://ssrn.com/abstract=4490635

[21] Vellela, Sai Srinivas and Pushpalatha, D and Sarathkumar, G and Kavitha, C.H. and Harshithkumar, D, ADVANCED INTELLIGENCE HEALTH INSURANCE COST PREDICTION USING RANDOM FOREST (March 1, 2023). ZKG International, Volume VIII Issue I MARCH 2023, Available at SSRN: https://ssrn.com/abstract=4473700

[22] D, Roja and Dalavai, Lavanya and Javvadi, Sravanthi and Sk, Khader Basha and Vellela, Sai Srinivas and B, Venkateswara Reddy and Vullam, Nagagopiraju, Computerised Image Processing and Pattern Recognition by Using Machine Algorithms (April 10, 2023). TIJER International Research Journal, Volume 10 Issue 4, April 2023, Available at SSRN: https://ssrn.com/abstract=4428667

[23] Vellela, Sai Srinivas and Basha Sk, Khader and B, Venkateswara Reddy and D, Roja and Javvadi, Sravanthi, MOBILE RFID APPLICATIONS IN LOCATION BASED SERVICES ZONE (June 14, 2023). International Journal of Emerging Technologies and Innovative Research, Vol.10, Issue 6, page no. ppd851-d859, June2023, http://www.jetir.org/papers/JETIR2306410.pdf

[24] Vellela, Sai Srinivas and Sk, Khader Basha and B, Venkateswara Reddy, Cryonics on the Way to Raising the Dead Using Nanotechnology (June 18, 2023). INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS), Vol. 03, Issue 06, June 2023, pp : 253-257,

[25] Vellela, Sai Srinivas and D, Roja and B, Venkateswara Reddy and Sk, Khader Basha and Rao, Dr M Venkateswara, A New Computer-Based Brain Fingerprinting Technology (June 18, 2023). International Journal Of Progressive Research In Engineering Management And Science, Vol. 03, Issue 06, June 2023, pp : 247-252 e-ISSN : 2583-1062.,

[26] Gajjala, Buchibabu and Mutyala, Venubabu and Vellela, Sai Srinivas and Pratap, V. Krishna, Efficient Key Generation for Multicast Groups Based on Secret Sharing (June 22, 2011). International Journal of Engineering Research and Applications,Vol. 1, Issue 4, pp.1702-1707, ISSN: 2248-9622

[27] Kiran Kumar Kommineni, Ratna Babu Pilli, K. Tejaswi, P. Venkata Siva,Attention-based Bayesian inferential imagery captioning maker,Materials Today: Proceedings,2023,ISSN 2214-7853,https://doi.org/10.1016/j.matpr.2023.05.231.

[28] Venkateswara Reddy, B., &KhaderBashaSk, R. D. Qos-Aware Video Streaming Based Admission Control And Scheduling For Video Transcoding In Cloud Computing. In International Conference on Automation, Computing and Renewable Systems (ICACRS 2022).

[29] Reddy, N.V.R.S., Chitteti, C., Yesupadam, S., Desanamukula, V.S., Vellela, S.S., Bommagani, N.J. (2023). Enhanced speckle noise reduction in breast cancer ultrasound imagery using a hybrid deep learning model. Ingénierie de Systèmesd'Information, Vol. 28, No. 4, pp. 1063-1071. https://doi.org/10.18280/isi.280426

[30] Vellela, S.S., Balamanigandan, R. An intelligent sleep-awake energy management system for wireless sensor network. Peer-to-Peer Netw. Appl. (2023). https://doi.org/10.1007/s12083-023-01558-x

[31] Rao, D. M. V., Vellela, S. S., Sk, K. B., &Dalavai, L. (2023). Stematic Review on Software Application Under-distributed Denial of Service Attacks for Group Website. DogoRangsang Research Journal, UGC Care Group I Journal, 13.

[32] S. S. Priya, S. SrinivasVellela, V. R. B, S. Javvadi, K. B. Sk and R. D, "Design And Implementation of An Integrated IOT Blockchain Framework for Drone Communication," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205659.

[33] N. Vullam, K. Yakubreddy, S. S. Vellela, K. BashaSk, V. R. B and S. SanthiPriya, "Prediction And Analysis Using A Hybrid Model For Stock Market," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205638.

[34] K. K. Kumar, S. G. B. Kumar, S. G. R. Rao and S. S. J. Sydulu, "Safe and high secured ranked keyword searchover an outsourced cloud data," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 20-25, doi: 10.1109/ICICI.2017.8365348.

[35] Sk, K. B., Vellela, S. S., Yakubreddy, K., &Rao, M. V. (2023). Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation. KhaderBashaSk, Venkateswara Reddy B, SaiSrinivasVellela, KancharakuntYakub Reddy, M VenkateswaraRao, Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation, 10(3).

[36] Vellela, S. S., Sk, K. B., Dalavai, L., Javvadi, S., &Rao, D. M. V. (2023). Introducing the Nano Cars Into the Robotics for the Realistic Movements. International Journal of Progressive Research in Engineering Management and Science (IJPREMS) Vol, 3, 235-240.

[37] Kumar, K. & Babu, B. & Rekha, Y.. (2015). Leverage your data efficiently: Following new trends of information and data security. International Journal of Applied Engineering Research. 10. 33415-33418.

[38] S. S. Vellela, V. L. Reddy, R. D, G. R. Rao, K. B. Sk and K. K. Kumar, "A Cloud-Based Smart IoT Platform for Personalized Healthcare Data Gathering and Monitoring System," 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIANCON58793.2023.10270407.