Asian Journal of Research in Computer Science

10(1): 34-46, 2021; Article no.AJRCOS.70315
ISSN: 2581-8260

Paralinguistic Speech Processing: An Overview

Rozin Majeed Abdullah[1*], Siddeeq Y. Ameen[1], Dindar Mikaeel Ahmed[1], Shakir Fattah Kak[1], Hajar Maseeh Yasin[1], Ibrahim Mahmood Ibrahim[1], Awder Mohammed Ahmed[2], Zryan Najat Rashid[2], Naaman Omar[1] and Azar Abid Salih[1]

[1]Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq.
[2]Sulaimani Polytechnic University, Sulaimani, Kurdistan Region, Iraq.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJRCOS/2021/v10i130233
Editor(s):
(1) Prof. M.A.Jayaram, Siddaganga institute of Technology, India.
Reviewers:
(1) S.Kannadhasan, Cheran College of Engineering, India.
(2) Antonio Luque De La Rosa, Universidad De Almería, Spain.
Complete Peer review History: http://www.sdiarticle4.com/review-history/70315

Review Article

Received 12 April 2021
Accepted 21 June 2021
Published 22 June 2021

ABSTRACT

Clients can adequately control PCs and create reports by speaking with the guide of innovation, discourse acknowledgment empowers records to be delivered all the more effectively in light of the fact that the program normally produces words as fast as they expressed, which is typically a lot faster than a human can compose. Discourse acknowledgment is an innovation that consequently finds the words and expressions that best match the contribution of human discourse. The most normal use of discourse acknowledgment is correspondence, where discourse acknowledgment can be utilized to create letters/messages and different reports. Point of discourse handling: - to comprehend discourse as a mechanism of correspondence, to reflect discourse for transmission and propagation; - to inspect discourse for robotized data discovery and extraction-to find some physiological highlights of the speaker. In discourse combination, there are two significant capacities. The first is the interpretation of voice to message. The second is for the content to be converted into human voice.

Keywords: Speech processing; speech recognition; speech coding; machine learning; application.

**Corresponding author: E-mail: rozin.abdullah@dpu.edu.krd;*

1. INTRODUCTION

The information on discourse handling is beforehand bit of our standard lives, yet is as yet restricted to generally straightforward orders until further notice, As the innovation propels, specialists will have the option to make more keen frameworks that comprehend conversational discourse, you ought to address your machine the manner in which you converse with any person then it will be skilled to communicate contemplated reactions back to you [1-4].Technologies for signal processing will make all of this possible. In this field, the number of specialists required is increasing, furthermore, numerous organizations are searching for skilled individuals who might want to be important for it. Preparing, deciphering, and comprehension of a discourse signal is the way to some amazing [5-6].

- No visual contact possible.
- No requirement for special equipment.
- It can be done when doing other things.
- Phones, AT&T.
- Mobile Phones (1G ,2G,3G).

Types of Speech processing: 1-Speech Coding. 2- Speech Synthesis. 3- Speech Recognition. 4-Low Bit Rate Speech Coding. 5-Speech Signal. A control subsystem (CS) and a recognition subsystem (RS) interconnected by a CS/RS interface provide a real-time speech processing production framework [7-10]. A control processor, an administrator interface, a UI, and a control programming module are utilized in the control subsystem to stack any of a large number of control programs that utilization discourse acknowledgment measures [11-13] .The acknowledgment framework RS incorporates an expert processor, discourse signal processor, and layout coordinating processors all interconnected on a typical transport which speaks with the control subsystem through the intervention of the CS/RS interface [1,14] . The two-section design permits the control subsystem to be gotten to by the administrator for non-continuous framework capacities, and the acknowledgment subsystem to be gotten to by the client for constant discourse handling capacities. An exemplification of A speaker validation technique requires the enrollment of models, layout preparing, acknowledgment by format link and time arrangement, quietness and filler layout age, and speaker observing modes [15].

To find patterns in vast volumes of data, machine learning algorithms use statistics [16]. And here, data contains a lot of things, figures, phrases, photos, clicks, everything you have. Machine learning is the mechanism that drives all of the utilities we use today: recommendation platforms such as those on Netflix, YouTube, and Spotify; search engines such as Google and Baidu; social media feeds such as Facebook and Twitter; voice assistants such as Siri and Alexa. If it can be processed digitally, it can be fed into a machine learning algorithm. The list keeps running [17].

The test of wily hand-designed acoustic highlights (include designing) Truly, it has been utilized as a particular issue from the undertaking of planning successful AI (ML) models to settle on expectation and grouping choices [18]. This procedure has two key downsides: first, manual element designing is wasteful and includes human information; and the second fabricated highlights probably won't be best for the current objective [19]. This has started the groundwork of another improvement in the discourse world to utilize portrayal learning techniques that can naturally gain proficiency with a moderate portrayal of the information signal [20].

For as long as twenty years, robotized discourse acknowledgment (ASR) has been one of the essential developments in discourse union, is still habitually utilized in applications, for example, voice control and discourse to-message [21,22]. Be that as it may, the programmed acknowledgment of discourse debased by acoustic clamor and resonation is as yet viewed as an extremely testing issue [23,24]. By creating pre-preparing calculations like dereverberation and furnishing the ASR motor with perceptible data on the reverberant climate, we plan to upgrade ASR vigor in loud and reverberant conditions [25,26].

The processing of speech as an explicit event sequence is ubiquitous in automated recognition of speech (linguistic events) but, despite its potential for wide acoustic event sequences, has gotten comparatively little study in paralinguistic speech categorization [27,28]. This study provides a framework for analysis and use of speech to the identification of depression as a sequence of acoustic events [29,30]. Acoustic space regions are in this context referred to as 'words' which reflect speech occurrences at fixed or irregular intervals [31,32]. This tokenization enables the use of acoustic word characteristics using known processing methods in natural language [33,34]. A major feature of this framework is its capacity to accept diverse forms of event: here, we mix auditory words with language landmarks linked to articulation [35,36]. The option of confusing such diverse events to several levels, including the degree of integration, is another advantage [37,38].

Due to the limited availability of speech and linguistic pathologists, subjective screening of children with speech problems is costly and time consuming and not practical (SLPs) [39,40]. Therefore, automatic speech analyzes of children with speech problems are becoming increasingly important as they might give a realistic alternative to human evaluation [41,42]. Paralinguistic characteristics are a set of low level descriptors often employed to recognize speech emotions [43,44]. However, children's speech sound problems such as apraxia, phonological and articulation problems have not yet been investigated [45,46]. In this article we have examined the efficacy of paralinguistic characteristics in distinguishing between children usually developing and children with various types of sounds of speech [47,48]. The Geneva Minimalist Acoustic Parameter Set (GeMAPS) and its extensive variant, (eGeMAPS), have been examined in two types of conventional paralinguistic characteristics [49,50].

In paralinguistic, emotional speech played a very important part in the virtual assistants' intelligent discourse on human connection [51,52]. This paper presents a thorough picture of the study of emotional speech and paralinguistic noises carried out to far [53,54].

Translation from language to language (S 2ST) is used to generate a spoken word in a language other than a spoken language [55,56]. The typical method to S2ST has solely been focused on linguistic information by directly translating the spoken word from the source language into the target language without taking paralyzing and non language information such as the emotional conditions in the source language into consideration [56,57].

There is almost little understanding on the impact of speech coding and recognition on the transmission of narrow-band speech signals within specific frequency areas, particularly with regard to the recognition of paralinguistic indications in speech [58,59]. We thus studied the influence on the mechanical categorization of affecting vocalizations and clinical voice recordings of the narrow-band conventional speech-coders [60,61]. Further we investigated the effect of low-pass speech filtering with a set of varied cutting frequencies, chosen as static values of 0.5 to 5 kHz, or dynamically determined by the various upper limits of the first five speakers (F1-F5) [62,63]. Speech coding and recognition were evaluated primarily by utilizing the Geneva multimodal emotional portrayals according to short-term speaker moods [64,65]. Secondly, with regard to the long-term characteristics of speech speakers, the Child Pathological Speech Database has evaluated the vocal recording of clinical populations with speech impairment [66,67].

The vocal tract is the world's most dexterous and skilful device for manufacturing speech to communicate rich linguistic and spoken information [68,69]. The knowledge of how people differ in their language articulation because of variations in their physical vocal instrument's form and size, and its acoustic implications are not well understood [70,71]. Knowing how people differ in their speech output may assist build enhanced technology to recognize speakers and guide technology design for solid language-based access to persons and information [72,73]. The lecture focuses on steps for improving science in how the morphology and articulation of the vocal tract interact, and explains the various and invariant features of speech signals across speakers [74,75]. The nature of articulation techniques chosen by persons to attain phonetic equivalence in the midst of structural disparities among them is of special scholarly interest [76,77]. The elements of vocal distinctions morphologically reflecting and how they are reflected in the acoustic speech signal are equally interesting and if these differences from speech acoustics can be calculated [78,79]. A major element in this objective is to develop forward and reverse computational models that link voice tract features, speech acoustics and the construction of robust speaker acknowledgement systems to shed light on the variations between speakers [80,81]. The major focus of speech research is on surface-speech acoustic characteristics; concerns remain on how language features vary across speech, linguistic and paralinguistic contexts. However, the fundamental features of the acoustic signal alone can only be uncovered [82,83].

A methodology of extraction of computer paralinguistically effective recently presented is the Bag-of-Audio-Words (BoAW), where each spoken utterance is based on its frames cluster and is composed of a frame level training vector [84,85]. Several enhancements to the original BoAW technique have been presented in recent years but none of them has studied the effect of the stochastic character of the clustering process [86,87]. We show experimentally in this work that the random component in the BoAW clustering phase really spreads to the following classification stage, eventually leading to poor classification [88,89].

In recent years, interest in the categorization of paralinguistic information has increased significantly [90]. However, there are almost no standardized companies and test circumstances to evaluate performance in the exact same settings compared with similar speech processing tasks such as automatic speech and speaker reconnaissance [91]. For assessing the performance of statistical classifiers the consecutive challenges presented at the largest automated speech processing conference, namely the INTERSPEECH conferences, are crucial [92].

In Computational Paralinguistics data sparsity is one of the main bottlenecks. Partially monitored techniques to learning can assist to address this problem without costly human-labeling efforts. We are thus investigating the possibility of co-workout for example tasks of paralinguistic language processing over the time-continuum: from short-term emotions through sleep in the middle term to long-term sex [93]. The semi-completed method to co-learning chooses instances with high confidence values in each view by partitioning the acoustic feature with two as independent and adequate views as feasible and agglomerates them with their forecasts into initial training sets via iteration [94].

Speech analysis might give an indication of Alzheimer's disease and aid in the development of clinical instruments to identify and monitor the progression of the illness automatically [61]. While prior research used auditory characteristics (discussion) to characterize Alzheimer's dementia, these research concentrated on certain common prosodic characteristics, frequently in conjunction with transcript lexical and syntaxic characteristics [95]. We offer a comprehensive investigation of the predictive usefulness of solely auditory elements, automatically derived from the computational paralinguistics perspective from the spontaneous speech for Alzheimer's diagnosis of dementia [96]. On a balanced sample of spontaneous speech data from DementiaBank's Pitt, patients with matches of sex and age were evaluated for the efficacy of numerous state-of-the-art paralinguistic features for Alzheimer [97]. The evaluated feature sets included an expanded minimum acoustic parameter set for Geneva (eGeMAPS), emobase, ComParE 2013 and a new Multi-Resolution cochleagram [98].

2. SPEECH PROCESSING

There are two major tasks in the ASR system: phoneme detection and whole-word decoding. The relationship between the voice signal and telephones is formed in two distinct phases in ASR [99]. On the basis of prior knowledge, the first step, useful features are derived from the speech signal. This stage is known as the process of collection of data or reduction of dimensionality [100,101] . In this, by choosing the information based on task-specific awareness, the dimensionality of the speech signal is minimized. In conventional ASR systems, highly advanced features like MFCC are the chosen option. In the second level, discriminative models estimate the likelihood of each phoneme.

2.1 Speech Recognition

Discourse acknowledgment is the change cycle. into a machine fathomable organization of expressed words, which essentially implies into a paired language to make an interpretation of Discourse to on-screen text or request for an assistance, a framework needs to experience some convoluted advances [102]. He makes vibrations noticeable all around, which fundamentally are comparable, when talking humanly. This simple wave is changed over by a simple to-advanced converter (ADC) into computerized information that the PC can comprehend. It tests or digitizes the sound to do this by taking right wave estimations at customary spans [103]. The machine channels the digitized sound to wipe out pointless clamor to handle discourse, and furthermore to isolate it into different groups of pitch. It likewise standardizes the tone or sets it to a recurrence range that is steady. It will likewise must be adjusted transiently [104].

2.2 Speech Coding

Discourse coding is the information encoding technique for discourse containing advanced sound signs [105]. Utilizing sound sign preparing strategies to display the discourse signal, discourse coding utilizes discourse explicit boundary assessment, along with related information encoding calculations to speak to the subsequent demonstrated boundaries in a compacted bit-stream [106]. Discourse coding is principally used to fortify in telecom organizations, signal transmission and gathering and to improve the sign to-commotion level.

2.3 Low Bit Rate Speech Coding

Speech coders were initially utilized for voice signal encryption and they are as yet utilized for ensured voice correspondences today. In any case, their most outstanding utilization is bit rate saving in a systems administration medium, for example, a cell phone cell or a parcel network association with permit more clients. Then again, to guarantee more excellent playback, a high-goal coder or a more detailed coding framework may be required [107]. Truth be told, the accessibility of ever more extensive band associations and bigger limit media has made some view discourse encryption as unnecessary, yet the developing populace of transmitters and the undeniably more extravagant substance have taken up the "data transfer capacity" made available by the approach of broadband administrations. An adequately fine computerized portrayal might be called 'straightforward' or practically identical to the first sign, since human hearing has a limited affectability [108,109]. A piece pace of 706 kbit/s per line, conservative plate (Album) quality, is usually called straightforward on account of general sound sign, while 64 kbit/s for phone discourse is taken as cost quality (Table 1). Despite the fact that it is somewhat slippery to im-represent a reach for low piece rate discourse coding as it is a mov-in objective, it appears to be that these days it is best limited by4 kbit/s from above, given the enduring exertion to agree to a cost quality discourse coder at that rate at the ITU-T (1), (2), And it is limited by roughly 1 kbit/s from underneath, principally considering the anticipated range of driving cod-in procedures in the lower low-rate locale and the upper area. extremely low-rate area (3). The low rate between 2.4 kbit/s and 8 kbit/s was a solid and point by point manual for discourse coding (4) just a few years prior.

2.4 Speech Synthesis

PCs play out their positions in three distinct stages called input (where you feed information in, some of the time with a console or mouse) [110]. Creation (when the framework alludes to your criticism, for example, embeddings any numbers you composed in or upgrading the tones on a checked photograph), and yield where you can perceive how your info has been taken care of by the program, ordinarily on a screen or worked out on paper [111]. Discourse amalgamation is basically a sort of execution where, with a characteristic or counterfeit voice playing over an amplifier, a PC or other gadget recites words to you for all to hear; the product Text-to-discourse is otherwise called (TTS) [112].

2.5 Speech Encoding

Its recurrence reaction is characterized as the capacity of a sound handling gadget to imitate frequencies, and its dynamic reach is known as its capacity to accomplish legitimate commotion and non-abrasiveness [113]. Together, these words are here and there alluded to as the reliability of a sound framework. In its least complex structure, encoding is a method of recreating sound utilizing these two key standards, just as having the option to handily store and transport that information. Sound comprises of waveforms made out of the intervention of different frequencies and amplitudes of waves [114]. To describe these waveforms inside computerized designs the waveforms should be tested at frequencies that can in any event speak to hints of the most elevated recurrence that you need to repeat, and they should likewise store enough piece

Table 1. Bit rates of typical acoustic signals [109]

	Bandwidth	Sampling frequency	Bits per sample	Bit rate
Narrowband speech	300 Hz-34kHz	8.0 kHz	8	64 kbit/s
Wideband speech	50 Hz-7.0 kHz	16.0 kHz	14	224 kbit/s
Wideband audio (DAT format)	10 Hz- 20.0 kHz	48.0 kHz	16	768 kbit/s
Wideband audio (CD format)	10 Hz- 20.0 kHz	44.1 kHz	16	706 kbit/s

profundity to speak to the legitimate sufficiency tumult and delicate quality of the waveforms inside the sound example. A sound arrangement doesn't compare to a sound encoding design for instance, a typical record configuration, for example, .WAV indicates the header organization of a sound document, yet it isn't simply the sound encoding. Frequently, however not generally, .WAV sound records utilize direct PCM encoding; don't assume that a WAV document has any unique encoding except if you examine its header. FLAC is similarly a document encoding which now and again prompts certain mistaken assumptions [115]. The lone encoding inside the Discourse to-Text Programming interface that requires sound information to have a header is FLAC; for any remaining sound encodings, header less sound information is determined. We regularly apply to the codec inside the Discourse to-Text Programming interface once we notice to FLAC. We can utilize the configuration 'a .FLAC document' as we add it to the FLAC record design.

3. APPLICATION OF SPEECH PROCESSING

The best program for speech recognition offers you the capacity to streamline your workflow [116]. This is a big factor in our constantly busy world why it is gaining attention. Well-designed applications for voice recognition will help you improve efficiency significantly at work as well as at home [99]. A text can be written at about three times the speed at which it is typed. And you can do so with much more accuracy with the proper program [117]. Customized voice commands make hands-free dictation possible. You can perform tasks like asking your machine to open and edit a certain file, depending on the software, so you don't have to dig through files to find it. You can unlock your voice with a simple voice order.

3.1 Application of Low Bit Rate Speech Coding

Speech coders were first utilized for scrambling the discourse signal as they actually are today for secure voice correspondences [39]. In any case, their most significant use is bit rate saving to oblige more clients in an interchanges station, for example, a cell phone cell or a bundle network connect. Then again, a high-goal coder or a more detailed coding technique might be needed to accommodate a higher loyalty playback [118,119]. Basically, the accessibility of ever more extensive band associations and bigger limit media has provoked others to view discourse coding as unnecessary, however the "data transmission" made conceivable by the approach of broadband frameworks has been absorbed by the number of inhabitants in transmitters and the consistently more extravagant substance.

3.2 Application of Speech Coding

Speech coding is the strategy for accomplishing a convenient portrayal of voice signals as well as capacity for effective conveyance over wired and remote band-restricted organizations [120]. Discourse coders have been significant segments today in media communications and in sight and sound organizations [121]. Cell organizing, voice over web convention VOIP, video conferencing, electronic toys, chronicling, and computerized concurrent voice and information DSVD, just as different PC-based games and mixed media applications, are business advancements that depend on precise discourse encryption.

3.3 Application of Speech Signal

Numerous examinations in discourse insight necessitate that acoustic discourse boosts be introduced to subjects in an accurately controlled design [122]. For this reason, some of the time everything necessary is to record badge of common discourse from human speakers, however ordinarily, before these can be utilized as improvements, they must be altered and prepared differently, If specific boundaries of the acoustic discourse signal should be controlled for the boosts may must be made by methods for a discourse synthesizer [123]. A discourse synthesizer takes into account the conscious control and orderly variety of, for example major recurrence and the frequencies and data transfer capacities of the formants of the discourse signal. Be that as it may, the synthesizer boundaries should be controlled by estimations of their qualities in examples of normal discourse [124]. In this way, the arrangement and running of examinations using discourse improvements requires programming that incorporates methods for recording, proliferation, investigation, and preparing of discourse signals.

3.4 Application of speech Synthesis

In different implementations, robotic speech can be used. From poor quality speaking calculators to advanced 3D applications, such as speaking heads, contact aids have evolved [125]. The implementation approach mostly relies on the program used. Unrestricted language is not needed in some circumstances, such as announcing or alert systems, and the best outcome is typically obtained with some basic messaging method. With appropriate application, such funds can also be saved [126]. On the other hand, such systems, such as blind or electronic-mail reader reading devices, need limitless vocabulary and a TTS device is required. While the synthetic speech application area is expanding rapidly.

4. MOBILE SPEECH CODING

Speech coding is at the core of advanced remote communication. It comprises of diminishing the quantity of pieces expected to speak to the discourse signal while keeping up adequate quality. Computerized cell communication started in the last part of the 1980s when discourse coding had sufficiently developed to make it conceivable. Discourse coding has made computerized communication an alluring suggestion by compacting the discourse signal, in this way permitting a limit increment over simple frameworks [127]. Discourse coding principles are important to permit gear from various producers to effectively interoperate, subsequently giving a brought together arrangement of remote administrations to however many clients as could reasonably be expected. Guidelines bodies indicate all parts of the whole correspondence framework, including the air interface, balance strategies, correspondence conventions, numerous entrance innovations, flagging, and discourse coding and related channel blunder control systems [128]. Regardless of the target of accomplishing boundless interoperability, political and financial real factors just as mechanical components have prompted the development of a few provincial guidelines bodies around the world. Thus, we have seen the expansion of various contrary principles, here and there even in a similar geographic territory.

5. CONCLUSION

A speech processing system's key objective and benefit is the amount of protection it offers. Although speech recognition is largely stable, there are still bugs in it. This biometric framework can be paired with more conventional security features to offer an extra layer of security to support its adoption. The use of other biometrics or protection mechanisms, such as RSA, PINS or a combination of many different mechanisms, can be used. With more growth, voice recognition may be one of the most effective and largest biometrics technologies in the future.

DISCLAIMER

The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Rabiner L. "Fundamentals of speech recognition," Fundamentals of speech recognition; 1993.
2. Juang BH, Rabiner LR. "Hidden Markov models for speech recognition," Technometrics. 1991; 33:251-272.
3. Bristow G. Electronic speech recognition: techniques, technology, and applications: McGraw-Hill, Inc; 1986.
4. Roe DB, Wilpon JG. Whither speech recognition: the next 25 years," IEEE communications magazine. 1993;31:54-62.
5. Rudnicky AI, Hauptmann AG, Lee KF. "Survey of current speech technology," Communications of the ACM. 1994;37:52-57.
6. Baker J. "Milestones in speech technology-past and future," Speech Technology Magazine. 2005;10.
7. Dudley H, Tarnoczy TH. "The speaking machine of Wolfgang von Kempelen," The Journal of the Acoustical Society of America. 1950;22:151-166.
8. Fant G. Acoustic theory of speech production: Walter de Gruyter; 1970.
9. Dudley H. "The carrier nature of speech," Bell System Technical Journal. 1940;19:495-515.
10. Rabiner LR, Schafer RW. Digital processing of speech signals: Prentice-hall; 1978.

11. Furui S, Sondhi M. "Advances in speech signal processing Marcel Dekker," Inc.(USA); 1992.
12. Elias P. "Predictive coding--I," IRE Transactions on Information Theory. 1955;1:16-24.
13. Jelinek F. Statistical methods for speech recognition: MIT press; 1997.
14. Atal BS. The history of linear prediction," IEEE Signal Processing Magazine. 2006;23:154-161.
15. Rabiner LR. "A tutorial on hidden Markov models and selected applications in speech recognition," Proceedings of the IEEE. 1989;77:257-286.
16. Hazan H, Hilu D, Manevitz L, Ramig LO, Sapir S. "Early diagnosis of Parkinson's disease via machine learning on speech data," in 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel. 2012;1-4.
17. Van der Maaten L, Hinton G. "Visualizing data using t-SNE," Journal of Machine Learning Research. 2008;9.
18. Nair V, Hinton GE. "Rectified linear units improve restricted boltzmann machines," in Icml; 2010.
19. Bengio Y. Learning deep architectures for AI: Now Publishers Inc; 2009.
20. Bengio Y, Courville A, Vincent P. "Representation learning: A review and new perspectives," IEEE transactions on pattern analysis and machine intelligence. 2013;l35:1798-1828.
21. Chiang ML, Cheng HS, Liu HY, Chiang CY. "SDN-based server clusters with dynamic load balancing and performance improvement," Cluster Computing. 2021;24:537-558.
22. Othman A, Ameen SY, Al-Rizzo H. "A new channel quality indicator mapping scheme for high mobility applications in LTE systems," Journal of Modeling and Simulation of Antennas and Propagation. 2015;1:38-43.
23. Khalid LF, Ameen SY. "Secure Iot integration in daily lives: A review," Journal of Information Technology and Informatics. 2021;1:6-12.
24. Othman A, Othman SY, Al-Omary A, Al-Rizzo H. "Comparative Performanceof Subcarrier Schedulers in Uplink LTE-A under High Users' Mobility," International Journal of Computing and Digital Systems. 2015;4.
25. Hasan BMS, Ameen SY, Hasan OMS. "Image Authentication Based on Watermarking Approach," Asian Journal of Research in Computer Science. 2021;34-51.
26. Othman A, Ameen SY, Al-Rizzo H. "An Energy-Efficient MIMO-Based 4G LTE-A Adaptive Modulation and Coding Scheme for High Mobility Scenarios," International Journal of Computing and Network Technology. 2015;3.
27. Sharif KH, Ameen SY. "A Review of Security Awareness Approaches With Special Emphasis on Gamification," in 2020 International Conference on Advanced Science and Engineering (ICOASE). 2020;151-156.
28. Ameen SY, Nourildean SW. "Firewall and VPN investigation on cloud computing performance," International Journal of Computer Science and Engineering Survey. 2014;5:15.
29. Yahia HS, Zeebaree SR, Sadeeq MA, Salim NO, Kak SF, Adel AZ, et al. "Comprehensive survey for cloud computing based nature-inspired algorithms optimization scheduling," Asian Journal of Research in Computer Science. 2021;1-16.
30. Al-Khayat ON, Ameen SY, Abdallah MN. "WSNs Power Consumption Reduction using Clustering and Multiple Access Techniques," International Journal of Computer Applications. 2014;87.
31. Fawzi LM, Alqarawi SM, Ameen SY, Dawood SA. "Two Levels Alert Verification Technique for Smart Oil Pipeline Surveillance System (SOPSS)," International Journal of Computing and Digital Systems. 2019;8:115-124.
32. Ameen SY, Yousif MK. "Decode and forward cooperative protocol enhancement using interference cancellation," Int. J. Elect., Comput., Electron. Commun. Eng. 2014;8:273-277.
33. Ageed ZS, Zeebaree SR, Sadeeq MM, Kak SF, Rashid ZN, Salih AA, et al. "A survey of data mining implementation in smart city applications," Qubahan Academic Journal. 2021;1:91-99. 2
34. Elzanati WM, Ameen SY. "Cost effective air-conditioning for bahrain domestic applications," in 2013 7th IEEE GCC Conference and Exhibition (GCC). 2013;535-540.
35. Al-Sultan MR, Ameen SY, Abduallah WM. "Real Time Implementation of Stegofirewall System," International Journal of

Computing and Digital Systems. 2019;8:498-504.
36. Ameen SY, Al-Badrany MR. "Optimal image steganography content destruction techniques," in International Conference on Systems, Control, Signal Processing and Informatics, 2013;453-457.
37. Haji SH, Ameen SY. "Attack and Anomaly Detection in IoT Networks using Machine Learning Techniques: A Review," Asian Journal of Research in Computer Science. 2021;30-46.
38. Ameen SY, Ahmed IM. "Design and implementation of e-laboratory for information security training," in Fourth International Conference on e-Learning" Best Practices in Management, Design and Development of e-Courses: Standards of Excellence and Creativity". 2013;310-317.
39. Ageed ZS, Zeebaree SR, Sadeeq MA, Abdulrazzaq MB, Salim BW, Salih AA, et al. "A state of art survey for intelligent energy monitoring systems," Asian Journal of Research in Computer Science. 2021;46-61.
40. Ameen SY, Al-Jammas MH, Alenezi AS. "FPGA implementation of modified architecture for adaptive Viterbi decoder," in Saudi International Electronics, Communications and Photonics Conference (SIECPC), 2011;1-9.
41. Al Janaby AO, Al-Omary A, Ameen SY, Al-Rizzo HM. "Tracking High-Speed Users Using SNR-CQI Mapping in LTE-A Networks," in 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). 2018;1-7.
42. Ameen SY, Mahdi AH. "AES cryptosystem development using neural networks," International Journal of Computer and Electrical Engineering. 2011;3:309.
43. Abdulqadir HR, Zeebaree SR, Shukur HM, Sadeeq MM, Salim BW, Salih AA, et al. "A study of moving from cloud computing to fog computing," Qubahan Academic Journal. 2021;1:60-70.
44. Ameen SY, Ibrahimi IA. "MANET routing protocols performance evaluation with TCP Taho, Reno and new-Reno," International Journal of u-and e-Service, Science and Technology. 2011; 4:37-49.
45. Salih AA, Zeebaree SR, Abdulraheem AS, Zebari RR, Sadeeq MA, Ahmed OM. "Evolution of Mobile Wireless Communication to 5G Revolution," Technology Reports of Kansai University. 2020;62:2139-2151.
46. [46] S. Y. Ameen and L. J. Saud, "COMPUTING NODES AND LINKS APPEARANCES ON GEODESICS IN NETWORK TOPOLOGIES USING GRAPH THEORY," in Full Papers Conference Proceeding, 2011, p. 235.
47. Abdulla AI, Abdulraheem AS, Salih AA, Sadeeq MA, Ahmed AJ, Ferzor BM, et al. "Internet of Things and Smart Home Security," Technol. Rep. Kansai Univ. 2020;62:2465-2476.
48. Taher KI, Saeed RH, Ibrahim RK, Rashid ZN, Haji LM, Omar N, et al. "Efficiency of Semantic Web Implementation on Cloud Computing: A Review," Qubahan Academic Journal. 2021;1:1-9.
49. Othman A, Ameen SY, Al-Rizzo H. "Dynamic Switching of Scheduling Algorithm for," International Journal of Computing and Network Technology. 2018;6.
50. Sadeeq MM, Abdulkareem NM, Zeebaree SR, Ahmed DM, Sami AS, Zebari RR. "IoT and Cloud computing issues, challenges and opportunities: A review," Qubahan Academic Journal. 2021;1:1-7.
51. Abdulraheem AS, Salih AA, Abdulla AI, Sadeeq MA, Salim NO, Abdullah H, et al. Home automation system based on IoT; 2020.
52. Kareem FQ, Zeebaree SR, Dino HI, Sadeeq MA, Rashid ZN, Hasan DA, et al. "A survey of optical fiber communications: challenges and processing time influences," Asian Journal of Research in Computer Science. 2021;48-58.
53. Izadeen GY, Ameen SY. "Smart Android Graphical Password Strategy: A Review," Asian Journal of Research in Computer Science. 2021;59-69.
54. Omer MA, Zeebaree SR, Sadeeq MA, Salim BW, Mohsin SX, Rashid ZN, et al. "Efficiency of malware detection in android system: A survey," Asian Journal of Research in Computer Science. 2021;59-69.
55. Ameen SY, ALSH Ali. A Comparative Study for New Aspects to Quantum Key Distribution," Journal of Engineering and Sustainable Development. 2018;11:45-57.
56. Rashid ZN, Zeebaree SR, Sengur A. "Novel Remote Parallel Processing Code-Breaker System via Cloud Computing; 2018.

57. Haji SH, Zeebaree SR, Saeed RH, Ameen SY, HM. Shukur, N. Omar, et al., Comparison of Software Defined Networking with Traditional Networking," Asian Journal of Research in Computer Science, pp. 1-18, 2021.
58. Ageed ZS, Zeebaree SR, Sadeeq MM, Kak SF, Yahia HS, Mahmood MR, et al. Comprehensive survey of big data mining approaches in cloud systems," Qubahan Academic Journal. 2021;1:29-38.
59. Rashid ZN, Zeebaree SR, Shengul A. "Design and analysis of proposed remote controlling distributed parallel computing system over the cloud," in 2019 International Conference on Advanced Science and Engineering (ICOASE). 2019;118-123.
60. Abdulrahman LM, Zeebaree SR, Kak SF, Sadeeq MA, Adel AZ, Salim BW, et al., "A state of art for smart gateways issues and modification," Asian Journal of Research in Computer Science. 2021;1-13.
61. Rashid ZN, Zebari SR, Sharif KH, Jacksi K. "Distributed cloud computing and distributed parallel computing: A review," in International Conference on Advanced Science and Engineering (ICOASE). 2018;167-172.
62. Yazdeen AA, Zeebaree SR, Sadeeq MM, Kak SF, Ahmed OM, Zebari RR. "FPGA implementations for data encryption and decryption via concurrent and parallel computation: A review," Qubahan Academic Journal. 2021;1:8-16.
63. Rashid ZN, Sharif KH,. Zeebaree S "Client/Servers Clustering Effects on CPU Execution-Time, CPU Usage and CPU Idle Depending on Activities of Parallel-Processing-Technique Operations," Int. J. Sci. Technol. Res. 2018;7:106-111.
64. Fawzi LM, Ameen SY, Alqaraawi SM, Dawwd SA. "Embedded real-time video surveillance system based on multi-sensor and visual tracking," Appl. Math. Infor. Sci. 2018;12:345-359.
65. Jijo BT, Zeebaree SR, Zebari RR, Sadeeq MA, Sallow AB, Mohsin S, et al. "A comprehensive survey of 5G mm-wave technology design challenges," Asian Journal of Research in Computer Science. 2021;1-20.
66. Hassan RJ, Zeebaree SR, Ameen SY, Kak SF, Sadeeq MA, Ageed ZS, et al., "State of art survey for iot effects on smart city technology: challenges, opportunities, and solutions," Asian Journal of Research in Computer Science. 2021;32-48.
67. Sadeeq MA, Zeebaree S. "Energy management for internet of things via distributed systems," Journal of Applied Science and Technology Trends. 2021;2:59-71.
68. Malallah H, Zeebaree SR, Zebari RR, Sadeeq MA, Ageed ZS, Ibrahim IM, et al. "A Comprehensive Study of Kernel (Issues and Concepts) in Different Operating Systems," Asian Journal of Research in Computer Science. 2021;16-31.
69. Maulud DH, Zeebaree SR, Jacksi K, Sadeeq MAM, Sharif KH. "State of art for semantic analysis of natural language processing," Qubahan Academic Journal. 2021;1:21-28.
70. Ibrahim IM. "Task scheduling algorithms in cloud computing: A review," Turkish Journal of Computer and Mathematics Education (TURCOMAT). 2021;12:1041-1053.
71. Shukur H, Zeebaree SR, Ahmed AJ, Zebari RR, Ahmed O, Tahir BSA, et al. "A State of Art Survey for Concurrent Computation and Clustering of Parallel Computing for Distributed Systems," Journal of Applied Science and Technology Trends. 2020;1:148-154.
72. Zebari IM, Zeebaree SR, Yasin HM. "Real Time Video Streaming From Multi-Source Using Client-Server for Video Distribution," in 4th Scientific International Conference Najaf (SICN). 2019;109-114.
73. Jacksi K, Ibrahim RK, Zeebaree SR, Zebari RR, Sadeeq MA. "Clustering Documents based on Semantic Similarity using HAC and K-Mean Algorithms," in International Conference on Advanced Science and Engineering (ICOASE), 2020;205-210.
74. Yasin HM, Zeebaree SR, Zebari IM. "Arduino Based Automatic Irrigation System: Monitoring and SMS Controlling," in 2019 4th Scientific International Conference Najaf (SICN), 2019;109-114.
75. Sadeeq MA, Abdulazeez AM. "Neural Networks Architectures Design, and Applications: A Review," in International Conference on Advanced Science and Engineering (ICOASE). 2020;199-204.
76. Zeebaree S, Yasin HM. "Arduino based remote controlling for home: power saving, security and protection," International Journal of Scientific & Engineering Research. 2014;5:266-272.

77. Ageed ZS, Ibrahim RK, Sadeeq MA. "Unified Ontology Implementation of Cloud Computing for Distributed Systems," Current Journal of Applied Science and Technology. 2020;82-97.
78. Zeebaree S, Zebari I. "Multilevel Client/Server Peer-to-Peer Video Broadcasting System," International Journal of Scientific & Engineering Research. 2014;5.
79. Sulaiman MA, Sadeeq M, Abdulraheem AS, Abdulla AI. "Analyzation Study for Gamification Examination Fields," Technol. Rep. Kansai Univ. 2020;62:2319-2328.
80. Zebari SR, Yaseen NO. "Effects of Parallel Processing Implementation on Balanced Load-Division Depending on Distributed Memory Systems," J. Univ. Anbar Pure Sci. 2011;5:50-56.
81. Sadeeq M, Abdulla AI, Abdulraheem AS, Ageed ZS. "Impact of Electronic Commerce on Enterprise Business," Technol. Rep. Kansai Univ. 2020;62:2365-2378.
82. Yasin HM, Zeebaree SR, Sadeeq MA, Ameen SY, Ibrahim IM, Zebari RR, et al. "IoT and ICT based Smart Water Management, Monitoring and Controlling System: A Review," Asian Journal of Research in Computer Science. 2021;42-56.
83. Alzakholi O, Shukur H, Zebari R, Abas S, Sadeeq M. "Comparison among cloud technologies and cloud performance," Journal of Applied Science and Technology Trends. 2020;1:40-47.
84. Ageed Z, Mahmood MR, Sadeeq M, Abdulrazzaq MB, Dino H. "Cloud computing resources impacts on heavy-load parallel processing approaches," IOSR Journal of Computer Engineering (IOSR-JCE). 2020;22:30-41.
85. Sallow A, Zeebaree S, Zebari R, Mahmood M, Abdulrazzaq M, Sadeeq M. "Vaccine tracker," SMS reminder system: Design and implementation; 2020.
86. Ali ZA, Ameen SY. "Detection and Prevention Cyber-Attacks for Smart Buildings via Private Cloud Environment," International Journal of Computing and Network Technology. 2018;6:27-33.
87. Sadeeq MA, Zeebaree SR, Qashi R, Ahmed SH, Jacksi K. "Internet of Things security: a survey," in International Conference on Advanced Science and Engineering (ICOASE), 2018;162-166.
88. Abdullah SMSA, Ameen SYA, Sadeeq MA, Zeebaree S. "Multimodal emotion recognition using deep learning," Journal of Applied Science and Technology Trends. 2021;2:52-58.
89. Abdulazeez AM, Zeebaree SR, Sadeeq MA. "Design and Implementation of Electronic Student Affairs System," Academic Journal of Nawroz University. 2018;7:66-73.
90. Sallow AB, Sadeeq M, Zebari RR, Abdulrazzaq MB, Mahmood MR, Shukur HM, et al. "An Investigation for Mobile Malware Behavioral and Detection Techniques Based on Android Platform," IOSR Journal of Computer Engineering (IOSR-JCE). 2020;22:14-20.
91. Zeebaree S, Ameen S, Sadeeq M. "Social media networks security threats, risks and recommendation: A case study in the kurdistan region," International Journal of Innovation, Creativity and Change. 2020;13:349-365.
92. Aziz ZAA, Ameen SYA. "Air pollution monitoring using wireless sensor Networks," Journal of Information Technology and Informatics. 2021;1:20-25.
93. Ameen SY, Hadi W. "Investigation of using turbo code to improve the performance of DWT-OFDM system over selective fading channel," in 2nd International Conference on Information & Communication Technologies. 2006;2365-2369.
94. Amanuel SVA, Ameen SYA. "Device-to-device communication for 5G security: A Review," Journal of Information Technology and Informatics. 2021;1:26-31.
95. Ameen SY, Abdala MA, Ali SH. "Modeling and Simulation of Schematic Quantum Cryptography System Based Entangled Photons," Iraqi Journal of Computers, Communication, Control & Systems Engineering. 2006;6.
96. Fawzi LM, Ameen SY, Dawwd SA, Alqaraawi SM. "Comparative Study of Ad-hoc Routing Protocol for Oil and Gas Pipelines Surveillance Systems," International Journal of Computing and Network Technology. 2016;4.
97. Farhan FY, Ameen SY. "Improved hybrid variable and fixed step size least mean square adaptive filter algorithm with application to time varying system identification," in 2015 10th System of Systems Engineering Conference (SoSE). 2015;94-98.

98. Abdullah DM, Ameen SY. "Enhanced Mobile Broadband (EMBB): A review," Journal of Information Technology and Informatics. 2021;1:13-19.
99. Herff C, Schultz T. "Automatic speech recognition from neural signals: a focused review," Frontiers in neuroscience. 2016;10:429.
100. Purwins H, Li B, Virtanen T, Schlüter J, Chang SY, Sainath T. "Deep learning for audio signal processing," IEEE Journal of Selected Topics in Signal Processing. 2019;13:206-219.
101. Jaitly N, Hinton G. "Learning a better representation of speech soundwaves using restricted boltzmann machines," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2011;5884-5887.
102. Gales M, Young S. "The application of hidden Markov models in speech recognition; 2008.
103. Hori T, Cho J, Watanabe S. "End-to-end speech recognition with word-based RNN language models," in 2018 IEEE Spoken Language Technology Workshop (SLT). 2018;389-396.
104. Tran DT. "Fuzzy approaches to speech and speaker recognition," university of Canberra; 2000.
105. Kleijn WB, Krasinski DJ, Ketchum RH. "Fast methods for the CELP speech coding algorithm," IEEE transactions on acoustics, speech, and signal processing. 1990;38:1330-1342.
106. Chen JH, Cox RV, Lin YC, Jayant N, MJ. Melchner, "A low-delay CELP coder for the CCITT 16 kb/s speech coding standard," IEEE Journal on Selected Areas in Communications. 1992;10: 830-849.
107. Gottesman O, Gersho A. "Enhanced waveform interpolative coding at low bit-rate," IEEE transactions on speech and audio processing. 2001;9:786-798.
108. Xydeas CS, Chapman TM. "Segmental prototype interpolation coding," in IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258). 1999;2311-2314.
109. Dimolitsas S, Ravishankar C, Schroder G. "Current objectives in 4-kb/s wireline-quality speech coding standardization," IEEE Signal Processing Letters. 1994;1:157-159.
110. Flanagan JL. "Speech synthesis," in Speech Analysis Synthesis and Perception, ed: Springer; 1972;204-276.
111. Burkhardt F. "Simulation emotionaler sprechweise mit sprachsyntheseverfahren," Reihe Berichte aus der Kommunikationstechnik; 2001.
112. Cox R. "Speech coding standards," Speech coding and Synthesis. 1995;49-78.
113. Kondoz AM. Digital speech: coding for low bit rate communication systems: John Wiley & Sons; 2005.
114. Hanzo L, Somerville F, Woodard J. Voice compression and communications: Principles and applications for fixes and wireless channels; 2001.
115. Quatieri TF. Discrete-time speech signal processing: principles and practice: Pearson Education India; 2006.
116. ITU-T RG. "191,“Software tools for speech and audio coding standardization,” Int. Telecom. Union, Geneva, Switzerland, ed; 2005.
117. Kinnunen T, Li H. "An overview of text-independent speaker recognition: From features to supervectors," Speech communication. 2010;52:12-40.
118. Huang CW, Narayanan S. "Characterizing types of convolution in deep convolutional recurrent neural networks for robust speech emotion recognition," arXiv preprint ar. 2017;Xiv:1706.02901.
119. Thyssen J, Gao Y, Benyassine A, Shlomot E, Murgia C, Su HY, et al., "A candidate for the ITU-T 4 kbit/s speech coding standard," in IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221). 2001;681-684.
120. Lee KS, Cox RV. "A very low bit rate speech coder based on a recognition/synthesis paradigm," IEEE Transactions on Speech and Audio Processing. 2001;9:482-491.
121. Spanias AS. "Speech coding: A tutorial review," Proceedings of the IEEE. 1994;82:1541-1582.
122. Diesch E. "SpeechLab: PC software for digital speech signal processing," ed: Springer; 1997.
123. Agbinya JI. "Discrete wavelet transform techniques in speech processing," in Proceedings of Digital Processing Applications (TENCON'96). 1996; 514-519.
124. Gold B, Morgan N, Martin JCD. "Books on Telecommunications: Reviews and New

Releases: Speech and Audio Signal Processing John Wiley & Sons, Inc. 1999," European Transactions on Telecommunications. 2000;11:237-237.
125. Wu Z, Watts O, King S. "Merlin: An Open Source Neural Network Speech Synthesis System," in *SSW*; 2016.
126. Tokuday K, Zen H. "Directly modeling speech waveforms by neural networks for statistical parametric speech synthesis," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015;4215-4219.
127. Abro FI, Rauf F, Chowdhry B, Rajarajan M. "Towards security of GSM voice communication," Wireless Personal Communications. 2019;108:1933-1955.
128. Sun X. "Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio," in 2002 IEEE international conference on acoustics, speech, and signal processing. 2002;I-333-I-336.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/70315