Evolutionary Intelligence (2022) 15:1025–1036
https://doi.org/10.1007/s12065-019-00274-x

SPECIAL ISSUE

Research outlook and state-of-the-art methods in context awareness data modeling and retrieval

S. G. Gollagi[1] · M. M. Math[2] · U. P. Kulkarni[3]

Received: 24 November 2018 / Revised: 27 May 2019 / Accepted: 15 July 2019 / Published online: 16 August 2019

Abstract
As the data or information gets increased in various applications, it is very much essential to make the retrieval and modeling easier and simple. Number of modeling aspects already exists for this crisis. Yet, context awareness modeling plays a significant role in this. However, there requires some advancement in modeling system with the incorporation of advanced technologies. Hence, this survey intends to formulate a review on the context-aware modeling in two aspects: context data retrieval and context data modeling. Here, the literature analyses on diverse techniques associated with context awareness modeling. It reviews 60 research papers and states the significant analysis. Initially, the analysis depicts various applications that are contributed in different papers. Subsequently, the analysis also focuses on various features such as web application, time series model, intelligence models and performance measures. Moreover, this survey gives the detailed study regarding the chronological review and performance achievements in each contribution. Finally, it extends the various research issues, mainly the adoption of Evolutionary algorithms, which can be useful for the researchers to accomplish further research on context-aware system.

Keywords Big data · Context awareness · Data modelling · Data retrieval, applications · Performance measures

1 Introduction

One of the current analyses has shown the determinant future of context-aware applications and that has revealed the 6 important areas. Five of those components have been focussed on context-aware retrieval (CAR) [68–70] of data or information. It is believed that the significant role of CAR applications plays with two major CAR nature. They are the renowned fields like information retrieval (IR) and information filtering (IF). In general, the common CAR applications include the usage of mobile. In the sense, the user's context [61–63] is changing. Further, the important aim or objective of these two factors is to be effective and efficient. In this, the efficiency means that the delivery of data or information rapidly, which should be happened without the need of any resources even if there has huge count of information to be retrieved. In fact, the efficiency is prominently the related denying information that is often the useless data.

The applications of CAR [71–73] are often helpful in the environs like resource-hungry environment, in which the less storage usage as well as another resource are very critical. Moreover, the effectiveness is considered with the retrieval of related documents. This means, the document that is termed to be more useful by the user. The important count of IR/IF research work is considered with the derivation of weighing approaches for accurate ranking of the documents as per the apparent related requirement of user. Still, it moves over the CAR [15, 64, 74, 75, 81–83] as it is the needed factor on information delivery. Further, the delivery accuracy is also important and most likely in the CAR system, in which the display space is limited and that is peripheral to the present task of user. Practically, the empirical review reveals that if the CAR user receives the doubtful relevance information, then the system gets abandon. Overall suggestion is that the IR/IF techniques are closer to the CAR application since there has great potential benefits and insights in terms of relating them.

✉ S. G. Gollagi
sggollagi@gmail.com

[1] Hirasugar Institute of Technology, Nidasoshi, Belagavi, Karnataka, India

[2] Gogte Institute of Technology, Belagavi, Karnataka, India

[3] SDM College of Engineering and Technology, Dharawad, Karnataka, India

This paper contributes a review related to Context aware [64–67] data retrieval and modeling. The survey focuses on different retrieval and modelling models exploited in diverse applications that can offer better improvements. In this, review stands on various parameters like web application, time series model, performance measures. The rest of the paper is organized as follows: Section II describes the related works done under this topic. Section III demonstrates the study on context data modelling and retrieval and section IV depicts the models and performance measures enabled with this topic. Accordingly, section V portrays the research gaps and challenges, and section VI concludes the paper.

2 Literature review

2.1 Related works based on context data retrieval

In 2016, Foschini et al. [1] have presented a new context data distribution infrastructure for CRHWNs that was for attaining data distribution more effectively and reliably. They have also shown how the proposed solution has increased the scalability of retrieval and the context data distribution as well. This was mainly by the self-adapting data distribution paths. Finally, the investigational outcomes have validated the assumptions and have also demonstrated how the proposed solutions limit the overhead of runtime overhea... In 2001, Brown and Jones [2] have considered the close relation of distinct, information retrieval and the ...ring concept to the context-aware based retrieval. This paper has aimed to make a new initialization towards the concept. In 2007, Ioanna et al. [3] have elaborated a novel context management model, which has been ...delled for coping with both cases location based and free-text based context retrieval and also for effective control in context consistency. Moreover, the developed model has been included in the platform of multi-function pervasive services, whereas the model has been empirically assessed. In 2011, Stefano and Luca [4] have presented a common purpose solution for the purpose of web content in terms of mobile devices, named Social Context-Aware Browser, which was a new model to access the information on the basis of context of user context. This has exploited both social as well as collaborative methods for overtaking the restrictions of conventional solutions. Finally, the performance was analysed by the benchmark based modality.

In 2018, Singh and Kumar [5] have presented a new model and technique to retrieve the multimedia content of user through the context of the media/image file. They have applied the Logical Itemset mining that includes textual data, which are related to the images. Further, the model has addressed paraphrase issue to describe the media file content. Finally, the performance of the proposed model was compared over other methods, and has proven the superiority of the work. In 2014, Ricardo [6] have described the functionality of a profile of the protocol namely Search/Retrieve URL protocol. The implementation of SRU was a major component in the distributed environ, in which the archivists could place the records of authority and could reuse those record under access points. In 2006, John and Verma [7] have presented a new context-based model for the retrieval of ad-hoc of web documents. Further, the objective of this research work was the generation of document measure term at the time of retrieval, which could be utilized as the co-contributor or substitute. They have shown the expansion terms of generation through the typical query expansion model. Finally, the proposed model has shown the high retrieval accuracy. In 2011, Golitsyna and Maksimov [8] have presented a generalized iterative query-response model, which was the effective retrieval process. They have also stated the systematic characterization of retrieval approaches, which was aimed in the retrieval task support under different types. They have also developed the multidimensional quantitative analysis that was on the basis of coordinate indexing, which was for performing external feedback.

In 2015, Harvey et al. [9] have taken a fresh look at the issues in information retrieval (IR) in the extraction of documents of recall balancing along precision balancing. They have examined the construction of IR in terms of context, uncertainty and relevance by developing a novel process approach. Finally, they have conducted a sequence of investigation for evaluating the artifact and model, and have shown the performance improvement. In 2014, Rahat [10] have aimed to evaluate the collaborative computing models and services. Based upon these problems, they have made a great objective for balancing the innovation of technology with required consideration at International Workshop on Ubiquitous and Collaborative Computing (iUBICOM) forum. Evaluations were made with the proposed model with respect to transparency, speed, efficiency and quality of service. In 2014, Yu et al. [11] have introduced a new pairwise 3-D shape context to match the partial object and its retrieval. This was for extracting the 3-D light poles from mobile laser scanning (MLS) point clouds. This algorithm could concurrently model the global and local geometric structures. Finally, the algorithm has been analysed over other methods in terms of accuracy.

In 2016, Li et al. [12] have investigated the issue of prediction of tag importance, in which the aim was the automatic prediction of tag importance and its utility in image retrieval. For attaining this, they have initially developed a model for measuring the relative significance of object and the corresponding scene tags to the description of image sentence. They have made the investigation with 3 real time datasets, and has proven the performance of proposed work. In 2017, Amara et al. [13] have proposed an approach of

image annotation, which have incorporated the contextual cues that were gathered from both extrinsic and intrinsic for bridging the semantic gap. Further, the major objective of this research work was the collection of huge real time dataset of news images. Finally, they have evaluated the performance of proposed model by comparing other conventional methods. In 2017, Yu et al. [48] have developed a new algorithm along better complexity of runtime, and has then extended to a MapReduce style. This was for taking the benefits of popular platform of distributed computing. The experimental results have shown the betterments of proposed model over other conventional methods. In 2018, palazzo et al. [49] has presented an integration approach that has related the context as well as the domain information for permitting the relational information to be retrieved in context. The results have shown the significant reduction of tuples.

2.2 Related works based on context data modelling

In 2016, Sheikh et al. [14] have proposed a Neural Bag-of-Weighted Words (NBOW2) approach that has learned for assigning higher weights for the words that were significant for OOV PN retrieval. From the investigation, the authors have proven that the NBOW and NBOW2 models outperformed the conventional methods that were on the basis of raw embeddings from Skip-gram and LDA approaches.

In 2016, Lei et al. [15] have presented a context-associative model for recognizing the human-object interaction. Here, the developed model could attain the recognition of incoming visual content, which was on the basis of preceding activities. They have parsed the high level activity into successive sub-activities, and have constructed the context cluster for modelling the temporal relations. Finally, they have evaluated the performance of proposed model under 3 data sets and have proven the efficiency of proposed model over other conventional models. In 2006, Tim Erickson [16] has focused on certain details. How the students have modelled the data, the way the students complete the challenging task, the reason behind this usage and so on. These were collected from the students of San Francisco, California area. The authors have documented them more thoroughly. Finally, they have discussed some of the issues that could be focused in this concept.

In 2019, Zhou et al. [17] have proposed a Multi-Context Trajectory Embedding approach, namely MC-TEM for exploring contexts in the systematic manner. They have developed the MC-TEM in the framework of distributed representation learning, which was more flexible in characterizing different context kinds under varied applications. Further, it was found that this was the first time used in the application of trajectory data. Exclusive investigations were conducted and have proven the performance of proposed model over other conventional methods. In 2015, Tri Huynh et al. [18] have proposed a learning-based method for tackling the issues in this field. Initially, they have partitioned a MR image into some patch sets. For each patch, they have used the structured random forest, which was for directing the computed tomography prediction. Further, a novel ensemble approach was used for assuring the prediction robustness. At last, the authors have proven the betterments of proposed model other conventional methods.

In 2015, Wenrui Dai et al. [19] have developed the generalized context modelling (GCM) to compress the heterogeneous data. The GCM have constructed the model graph along combinatorial structuring. Moreover, the model has extended the suffix element of predicted series in context modelling (classic). Finally, the investigational outcome has proven that the performance of proposed model has outperformed other conventional modelling models. In 2011, Damon et al. [20] have introduced 4 bandit tasks that were of two-armed, which was for investigating the information influence related to group members. They have also employed soft-max choice approach, which have emerged from the process of drift-diffusion that was generally utilized on modelling the perceptual decision making. It was found that the proposed model was fit for every task and has made a comparison over the allocation of choice across subjects.

In 2003, En-hui et al. [21] have developed the data compression models (lossless), hierarchical context-dependent model along with the enhanced sequential contest dependent model. All these were on the basis of developed greedy CDG transform. It was observed that the developed context-dependent models could attain best expected redundancies over other conventional algorithms. In 2000, Yang and John [22] have presented a greedy grammar transform and the respective construction of grammar transform has performed sequentially, from where the original data sequence could highly recover. On the basis of this, the authors have developed hierarchical algorithm. Further, it was proven that the redundancy (worst case) between each distinctive series were upper-bounded by log. The results have proven the superiority of proposed model. In 2017, Prabha et al. [23] have developed a wavelet based sampling model to choose the least sampling rate to assure the reliability of data. The model results along the domain knowledge were utilized for developing the data collection models (context aware), which could improve the system's lifetime. The model has also developed the mathematical model for Context Aware Data Management (CAD), and has evaluated under real-world databases. In 2017, Milne and David [24] have tried to attain 3 major objectives (i) For identifying the 'special' about big data (ii) for provoking debate about the future association among big data and transport planning and (iii) identifying corresponding themes in this application. It was identified that the some of the challenges were needed to be rectified on the basis of reducing change from status.

In 2016, Olegas et al. [25] have developed 6 needs of dynamic business process (DBP) and also developed a model to model the rule- and context-based DBP. The respective model was on the basis of business rule changing along with the series at the process runtime. The prototype simulation tool was also developed based on these rules. Finally, the outcomes have revealed the need of dynamic changes. In 2016, Marko et al. [26] have presented a context approach, which has included 13 contextual factors that has captured different situations where the developers have direct interaction to IDE. Further, the proposed context model could be utilized on supporting and improving interaction of user with an integrated development Environment (IDE), which would enhance the accuracy rate as well. Finally, they have discovered the results on this interaction. In 2017, Chen and Xi [27] have proposed certain semi parametric estimators that were without the improvement of monotonicity restriction. On these regularity constraints, the proposed were founded to be asymptotically normal and consistent. Moreover, the investigational results have shown that the proposed estimators could work well in finite samples.

In 2013, José et al. [28] have developed a textual Domain-Specific Language (DSL) that was particularly tailored to model the context data. This has been implemented through the applications of Model-Driven Development (MDD) models for generating software artifacts from context approaches. They have applied the proposed model in 2 varied middleware platforms. In 2016, Erfani et al. [29] have introduced a context-aware meta-model that was on the basis of ontology, which has taken the benefit of Semantic Web models for capturing and formalizing the context data. They have given various case studies that were associated to the domain of software evolution for illustrating the advantages of sharing as well as reusing context. In 2017, Neven et al. [30] have developed a context-aware design model for positioned analytics namely Extended Model View Controller. The developed approach was designed the event-driven that has permitted the seamless transmission among information and physical space while usage. They have demonstrated the utilization of design pattern with the interactions like menus, pinch zoom as well as details-on-demand.

In 2000, Coakes and Coakes [31] have shown how the social and organizational context could take into account while the development process of information system. For producing the utilization of information system, there needs the approaches of socio-technical models, the authors have recommended the modelling method. This has resulted the computerised information model. In 2010, Agelos et al. [32] have used different chemometric models for analysing and modelling the potable water quality information or data. They have measured 20 water quality parameters at various sites and have applied the Factor analysis (FA) for standardizing the data sets (log-transformed). The efficiency was determined by confirming the FA analysis. In 2014, Sylvain et al. [33] have introduced the approach of context-aware application CAA, where, they have specified the adaptation of context explicitly as the model transformations. They have exploited the graph transformation techniques, and the evaluation of the proposed model has proven its superiority.

In 2017, Oscar et al. [34] have presented three-level context ontology (3LConOnt), which could be reused more easily; it could be extended as well as adapted for distinct generic purposes. Further, the developed method could consolidate the context knowledge that was already exist. Finally, the usability was illustrated as well. In 2010, Guy et al. [35] have considered the utility of varied Markov chains orders for modelling the dependence at sequence of ancestral root. They have found that the strong support was Bayes factors, which was for utilizing a 2nd order Markov chain. It was also identified that the Hwang and Green's continuous time approximation have greatest support with respect to Bayes factors.

In 2001, Daniela et al. [36] have introduced the architecture that utilized in 2 context-aware museum guides. They have discussed the designed model for triggering the system action. It was identified that the rule-based system has offered its flexible characteristics in rapid prototyping models. In 2012, Guotao et al. [37] have empirically investigated the possibility and authority of non-EUT models in the context, revealed preference (RP). The findings have pointed the significance of exploring the non-EUT approaches in exposed performance context for modelling the difficult choice in real world. In 2006, Marian et al. [38] have utilised a modelling scheme for combining data on the efficiency and cost of 4 separate approaches for changing the behaviours of GP for estimating their relevant cost efficiency. Finally, it was evident that the computerised reminder system was more efficient in attaining minimized standard drink counts.

In 2000, Keith et al. [39] have analysed the expected model error context of the evaluation of wind energy. Initially, the study has used the statics of single point for showing the development of mean value along time. Then, they have used the discrete Fourier transforms for showing the important spectral energy amount. Finally, it was reviewed that the inaccuracy is because of lacking of scale separation. In 2014, Manuela et al. [40] have made the evaluation of groundwater flow approach to analyse the aquifer response under different pumping models. Firstly, the simulation of groundwater flow was standardized adequately. Finally, the simulations have granted the maximized piezometric level. In 2018, Malicki et al. [41] have developed a model that has linked the results of 5 regional climate approaches. The approach was practiced in a representative village. Further, the respective integrated approach has reviewed the

development of 3 socio-economic groups. Finally, it was observed that the vulnerability model could apply in any of the situation.

In 2016, João et al. [42] have presented the framework of multi-scale spatial model for including strong modelling architecture (Multimodel Inference and Ensemble Modelling), which was for rectifying certain limits. This model was viewed by Hakea sericea Schrad with the concern of 2 studies. Finally, it was identified that the climate and landscape were the important determinants. In 2004, TOM and WOLFGANG [43] have developed the concept of progressing the awareness data in terms of awareness contexts. Aiding this concept, the authors have addressed the issues of contextualising event notifications. They have also reported the investigational study and made the conclusion in context-awareness modelling. In 2009, Brian and Seamus [44] have constructed the approach of multimedia systems development MSD, and it was on the basis of broader concept formulation of context difficulties. The authors have also explained the strategy and designing of the respective complexities. It was concluded that the interventions must be directed for the complexity situations.

In 2010, Daniel and Maria [45] have developed the Markov random field along the parameters of Markov, which was for modelling the relationships of spatial and topological among scene objects. Finally, the evaluation was done in terms of different 100 hand segmented images. In 20[illegible] Moroni et al. [46] have committed for the management of model for the explanation of anatomical structures, which was on the basis of geometrical and topological features. They have argued that the 3D model might do the enhancement in other non-geometrical pieces. Finally, the investigations have made with different medical applications. In 2007, Syteev et al. [47] has describes the Geographic Information System (GIS) oriented models for creating the atlas's 1st volume that has looked at the spatial allocation of five hazards (flood, landslide, wind speed, heat and seismic hazard). They have provided a novel tool set for refining the analytical abilities during the examination of natural hazards.

In 2017, Emma et al. [50] have presented the models and results of combining population projections over multi-spatial scales along the model of urban growth. The investigational results have discussed by experimenting different urban conditions.

In 2015, Hasti et al. [51] has extended the work of Semantic and Time-dependent Expertise Profiling (STEP) model to capture the expertisation from micro-contributions. Moreover, the authors have extended the experimentation for structured micro-contributions, which has emerged from the environment of ontology engineering. Finally, they have investigated the baseline profile coverage. In 2009, John et al. [52] have described the present knowledge regarding the potential of thermo–hydro–mechanical– chemical modelling. Initially, they have described the disturbances related to excavation, and then the results were discussed. Finally, they have verified the performance of proposed strategy. In 2017, Sunoj and Vipin [53] have studied the properties of provisional partial moments exist in weighted models context. They have attained the ratio of income gap for the respective characterizations. At last, they have introduced the non parametric estimators, which were subsequently evaluated by simulation works using real-time data set.

In 2014, Vivek and thakur [54] have discussed the design of pattern warehouse and have discussed their quality issues. The authors have focussed on both logical and conceptual design through the introduction of context. The snowflake system was also extended in this research work. In 2001, Wong and lam [55] have defined the dimension operators that were on the basis of multidimensional logic. This could be aided in many of the applications that identified in data warehouses. At last, they have prototyped the multidimensional XML database system and has also given the clear description. In 2016, Hang and Uta [56] have used the systematic model for identifying the extension of data sharing and the related factors that were influencing the willingness of distinct individuals. However, the results have shown that there is the lack of national regulations in Vietnam context regarding the data sharing.

In 2018, Hu et al. [57] have developed a novel context dependent random walk family along the tree pattern graph matching kernel. Experimental evaluation has proven the superiority of proposed model over other conventional methods. In 2016, Giorgio and Edoardo [58] have developed a new variation model, which could model the impacts of perceptual context-driven. This might have extended to the judgement of non-physical characteristics. Finally, the proposed model has proven its superiority by means of flexibility. In 2018, Jerzy et al. [59] have presented the attempt of industrial, real, engineering modelling that were used in companies. They have focussed on the team knowledge concepts. They have proven that the issues with the high degree difficulties might be solved by the proposed approach. In 2017, Joerg and Heiner [60] have investigated the queries that were gathered from different public SPARQL. The respective analysis has reviewed various quires in practically and has shown its safe section. Moreover, it was applicable in terms of scalability for huge data sets.

2.3 Chronological review

This section reviews the chronological review and its percentage of contributions for the respective years, which is shown in Fig. 1. This survey analyses the papers published in various years. At first, 13.33% of papers are taken from the year, 2018. Then, 20% of papers are taken from the year, 2017. 15% of papers are gathered from the year 2016. Only

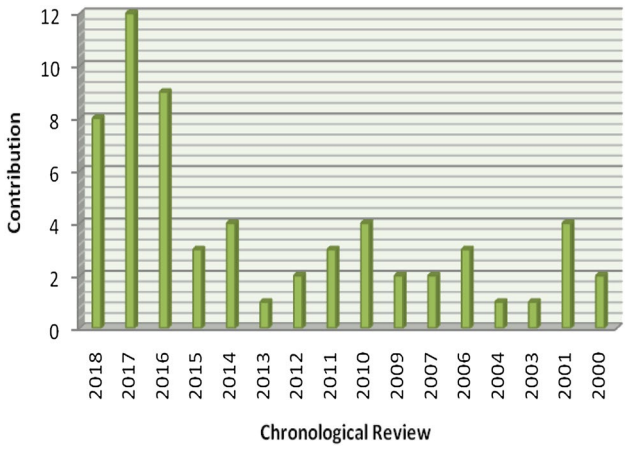

Fig. 1 Bar chart representing chronological review

5% of papers are taken from the year 2015, 2006 and 2011, respectively. Similarly, 6.66% of papers are taken from the years 2014, 2010, and 2001, respectively. 1.66% of papers are from the years 2013, 2004 and 2003, respectively. 3.33% of contribution under the context data retrieval and data modelling is from the years 2012, 2009, 2007 and 2000, respectively.

3 Study on context data modellling and retreival

3.1 Applications

This section explains the applications that used in the reviewed papers. As said before, the analysis is made in two concepts: Context data retrieval and Context data modelling, in which 15 papers are related to retrieval concept (Fig. 2) and the remaining 45 papers are focussing on data modelling. The applications used by the retrieval systems are as follows: the author in [1] have commenced with the Natural disaster calamities. [2] have worked on Sensors in network. [3, 4], and [5] have focussed on Pervasive computing, Web applications, and Multimedia application applications as well. The civil concepts were focused in [6]. Subsequently, the web applications are focused in [7]. Clustering, medical applications are focused in [8, 9]. Health care application, Mobile application, multimedia application, image processing, mobile traffic and ontology applications are focused in [10, 48, 49].

Then the applications focused in the context data modelling are as follows (Fig. 3): In this, the speech recognition application is focussed in [14]. The human activity is context modelled in the papers like [2, 15, 36, 57]. Education application is concentrated in [16]. The mobile application is the main focus of papers like [17, 33]. The papers like [18, 44–46] have concentrated on imaging techniques. Data compression is the main focus of [19, 21, 22]. The paper [23] has

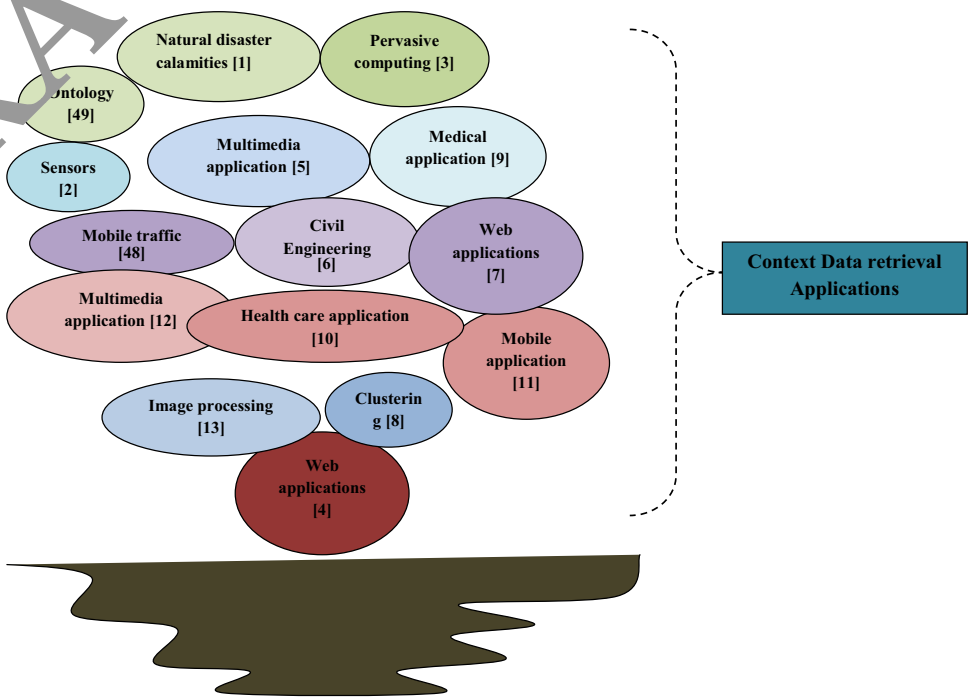

Fig. 2 Applications focused in reviewed papers that is of context data retrieval

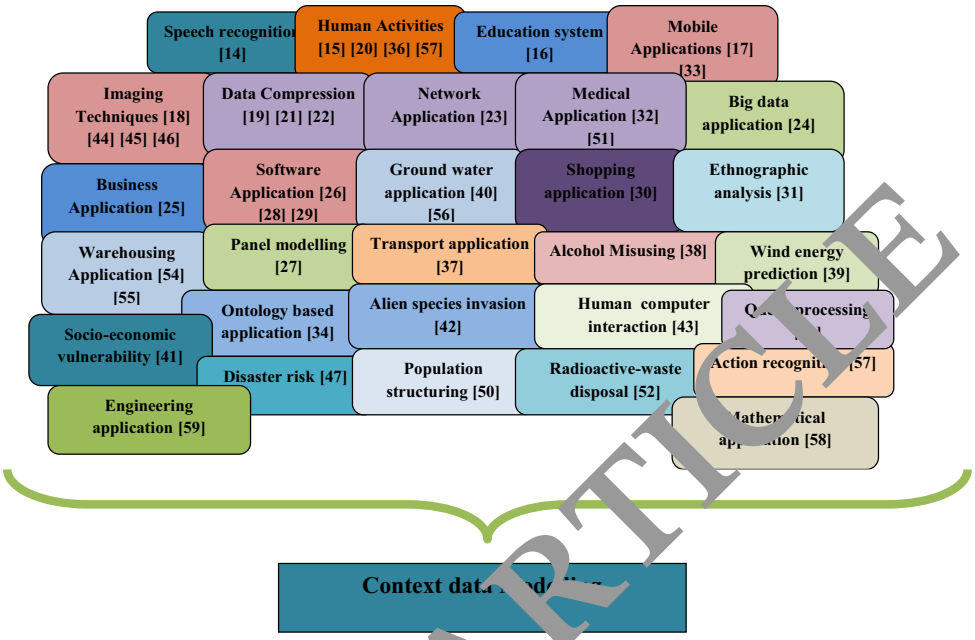

Fig. 3 Applications focused in reviewed papers that is of context data modelling

concentrated the network application. The big data application is focused in [24]. The business application is focused in [25]. The papers like [26, 28, 29] have concentrated on software application. The ground water or fresh water based context modelling is preceded in [40, 56]. Papers, [2 51] have worked under medical applications. The data wa housing is the application that concentrated i 54, 55] Panel modelling is concentrated in [27]. Th paper [ ] has worked under shopping application. Ethnographic analysis is made in [31]. Ontology based applic tion is focused in [34]. The applications like human applic n, transport application and Alcohol Misusing s g in [36–38]. Wind energy prediction is the application focussed in [39]. Socio-economic vulnerability Ali n spe ies invasion, human computer i teraction, Disaster risk, Population structuring, dioactive-waste disposal, mathematical application, Engi ering application and query processing are focused [4 –43, 47, 50, 52, 58–60], respectively.

3.2 Web application

This section explains the papers contribution under web application for both cases: Data Retrieval and Data Modeling. In this, Fig. 4a shows the graphical representation of Web application process in data retrieval. Totally, 15 papers are reviewed in data retrieval, and as per the graph, 66.66% of total papers (data retrieval) have worked or process under the web application. 33.33% of papers have not concentrated

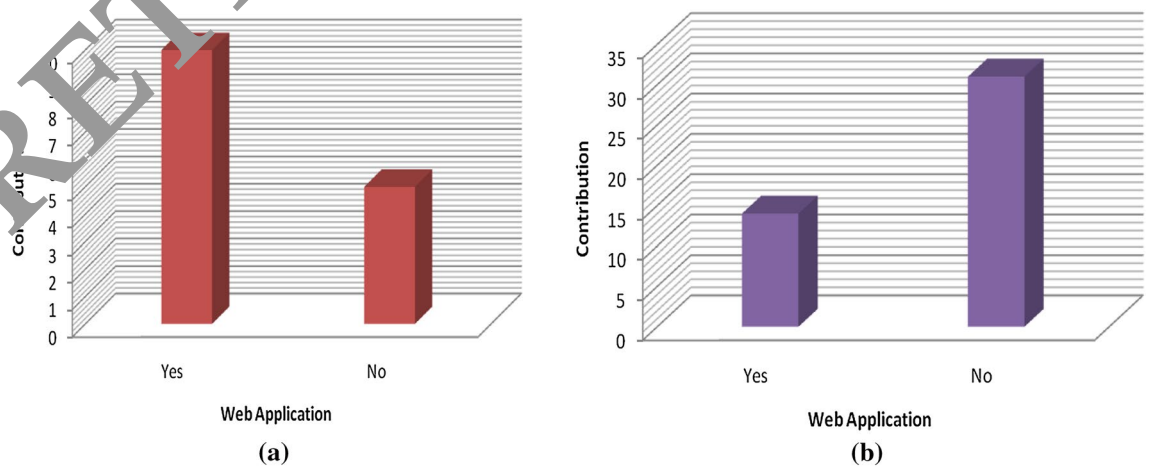

Fig. 4 Analysis on Web application **a** context data retrieval, **b** context data modelling

on the web application, and they were processed with some other applications.

Similarly, the context data modelling papers were also analysed to identify the contributions of web application (Fig. 4b. Here, 45 papers are reviewed under context data modelling. In this, it is identified that the 31.11% of contributions have worked under the web applications and 68.88% of total contributions have not processed with web, but worked under other applications.

3.3 Time series model

This section explains the contribution of time series model from all the reviewed papers, which is shown in Table 1. Totally, 60 papers have been reviewed in retrieval and modeling cases. In this, 10% of total contributions have used the time series model and 90% of total contributions have not processed under time series model.

4 Review on models and performance measures

4.1 Intelligence or other models

This section reviews the intelligence and other models that are used in the reviewed papers (Table 2). From the total reviewed papers only minimal percentage of

Table 1 Research status on contribution related to time series model

S. no	Author [citation]	Time series model	S. no	Author [citation]	Time series model
1	Foschini et al. [1]	No	34	Oscar et al. [34]	No
2	Brown and Jones [2]	No	35	Guy et al. [35]	No
3	Ioanna et al. [3]	No	36	Daniela et al. [36]	No
4	Stefano and Luca [4]	No	37	Guotao et al. [37]	No
5	Singh and Kumar [5]	No	38	Marian et al. [38]	No
6	Ricardo [6]	No	39	Keith et al. [39]	Yes
7	John and Verma [7]	No	40	Manuela et al. [40]	No
8	Golitsyn and Maksimov [8]	Yes	41	Malicki et al. [41]	No
9	Harvey et al. [9]	No	42	João et al. [42]	No
10	Pahat [10]	No	43	Wolfgang [43]	No
11	Yu et al. [11]	No	44	Brian and Seamus [44]	No
12	Li et al. [12]	No	45	Daniel and Maria [45]	No
13	Amara et al. [13]	No	46	Moroni et al. [46]	No
14	Sheikh et al. [14]	No	47	Syteev et al. [47]	No
15	Lei et al. [15]	No	48	Yu et al. [48]	No
16	Tim Erickson [16]	No	49	palazzo et al. [49]	No
17	Zhou et al. [17]	No	50	Emma et al. [50]	No
18	Tri Huynh et al. [18]	Yes	51	hasti et al. [51]	No
19	Wenrui Dai et al. [19]	Yes	52	John et al. [52]	No
20	Damon et al. [20]	No	53	Sunoj and Vipin [53]	No
21	En-hui et al. [21]	No	54	Vivek and thakur [54]	No
22	Yang and John [22]	No	55	Wong and lam [55]	Yes
23	Prabha et al. [23]	Yes	56	Hang and Uta [56]	No
24	Milne and David [24]	No	57	Hu et al. [57]	No
25	Olegas et al. [25]	No	58	Giorgio and Edoardo [58]	No
26	Marko et al. [26]	No	59	Jerzy et al. [59]	No
27	Chen and Xi [27]	No	60	Joerg and Heiner [60]	No
28	José et al. [28]	No			
29	Erfani et al. [29]	No			
30	Neven et al. [30]	No			
31	J.M.Coakes and Coakes [31]	No			
32	Agelos et al. [32]	No			
33	Sylvain et al. [33]	No			

Table 2 Research status on contribution on time series model

Models or system	Contribution
Optimization	[1–3, 13, 23, 59]
Ranking algorithm	[5]
Content Matching technique	[7]
Pairwise 3D shape context	[11]
Support vector machine	[12]
Mapreduce	[48]
Neural Bag-of-WeightedWords (NBOW2) model	[14]
Multi-context trajectory embedding model	[17]
Auto-context model	[18]
Semiparametric estimators	[27]
Semantic web technology	[29]
Mapping technique	[33]
Bayes network	[35]
Modelling approach	[38]
Multimedia system development model	[44]
Markov model	[45]
Geographic information system	[47]
Semantic modelling	[51]
Multidimensional logic	[55]
Kernel learning algorithm	[57]

contributions has mentioned the respective intelli ce. In this, the intelligence model like optimizatio concept is used by the papers like [1–3, 13, 23, 59], w ch is the 10% of total contribution. 1.66% of tota contri tions have used the Support Vector Machi e. Papers like [5, 7, 11] have used the Ranking algorith Conte t Matching technique and Pairwise 3D shape c xt models. Mapreduce, Neural Bag-of-Wei hted Words (NBOW2) model and Multi-Context Traject ry Embedding model are the used models in [48, 4, 17]. Then, the remaining papers have used so e o er models, which are clearly given in Table II.

4.2 Performance measures

The attained performance measure is analysed from all the reviewed papers and the result is shown here. In this, the context data retrieval papers have analysed under the retrieval accuracy rate and that is shown in Fig. 5a, whereas the context data modelling papers have reviewe under the error rate, which is shown in Fig. 5b. Obvious , most of the reviewed haven't mentioned the value of perfo ance measures, in both cases: retrieval and m delling. First in case of accuracy rate, among 15 retrieval pa rs 13.33% of papers have attained below 10% of accuracy ate (retrieval accuracy) and 13.33% of tot con ibuti ns have attained the accuracy of above 10%. Co ing to the error rate analysis, 13.33% of total con ibution (papers), have attained the error rate that is the r ge of 0.1–0.9. 6.66% of total contributions ha e ained th error rate that is in the range of 1.0–10.0. Si ila 2.22% of total contributions have attained the error a above 10.

5 Effect of evolutionary algorithms

In va ous fields like production systems, economics, engi- ering designs, etc., optimization problems have been acutely noticed, and hence, there is a prerequisite to implement an efficient optimization algorithm with less computaional complexity. For solving the problems in context-aware data modeling and retrieval, optimization algorithm has deployed in [1–3, 13, 23, 59]. Moreover, Evolutionary computation is suitable in solving the complex real world problems in datamining, medical, industry, etc. Among different categories of optimization algorithms, Evolution-based algorithms have been implemented, which can be applied to problems where heuristic solutions are not available or generally lead to unsatisfactory results. As a result, evolutionary algorithms have recently received an increased interest, particularly with regard to the manner in which they may be applied for practical problem-solving. As the Evolutionary algorithms hold the beneficial characteristics like

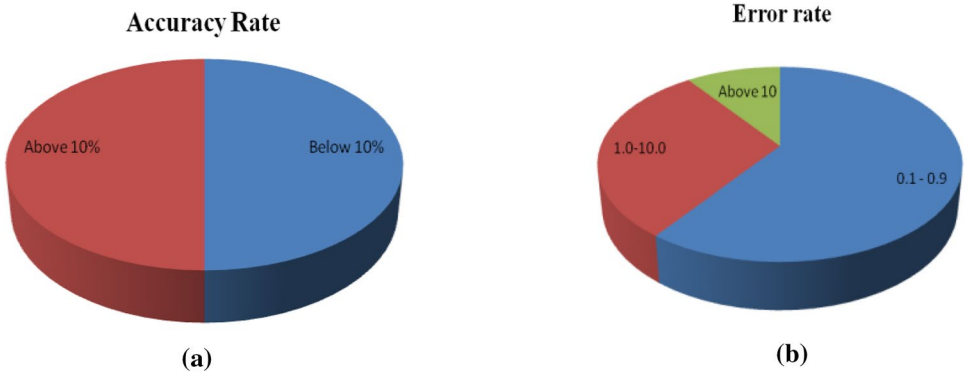

Fig. 5 a ary nder performance m sures **a** retrieval accuracy b error rate

simplicity, robustness, and flexibility, it is urged to merge the concept of Evolutionary algorithms on context-aware data modeling and retrieval. A lot of Evolutionary Algorithms are available like Genetic Algorithm (GA) [76], Differential Evolution (DE) [77], Evolutionary Programming (EP) [78], Evolution Strategy (ES) [79], and Biogeography-Based Optimizer (BBO) [80]. Since there is no other contribution with these algorithms for effectually improving the performance of the current research wok, different approaches of Evolutionary algorithms can be combined, which will generate various refinement to solve the problems related with context-aware data modeling and retrieval. This suitable execution, in turn, achieves remarkable attention from diverse researchers for extending the future works.

6 Research gaps and challenges

As the usage of context aware fact is very crucial in different factors, so many aspects are needed to be fulfilled that are often fails in practical. Even though, the concept is more helpful, some improvement and advancement is needed in case of concerning certain characteristics that are relevant to the context modelling, handling and adaptation. These characteristics are as follows: Context aggregation, consistency, acquisition, discovery, context query, adaptation, reasoning, integration, quality indicators and so on. In this, the query mechanism poses design issues in terms of triggering messages, query optimizations, as well as definitions of constraints, which is to be addressed in future. Moreover, the complexity is also a important issue with this concept, and that should be transparent to end users. The retrieval and modelling concepts must need the care in context reasoning regarding consistency, deduction of high level context.

Some other angle of limitations reviewed in computing context aware is as follows: the system is highly required to rectify the drawbacks and to leverage the greatest potential of system with high performance rate. In order to increase the performance, the system must have to concentrate on different modules. (i) Context Definition: this is the vast criteria that emphasis all the feasible parameters. Moreover, the application should define and find the related parameters of the current situation. (ii) Sensing of Context information: This must realise the improvement and modelling of personalized context sensors. (iii) Representation and storage of context data: The representation approach must be tackled with the sharing and interpretation process that to in standardize structure. Further, the most critical issue is the adaptation of system behaviour and the interpretation process is based on adaptation. This must be concerned in the future works. The privacy and security control issues must be solved so that the context awareness system becomes stronger. This would be the main factor since the information must be more secured from the entry of malicious entities and so on.

Some of the major future work is as follows (i) Modelling of context information with numerous requirements like distributed composition, partial validation, quality of information and so on. (ii) Real time application is also a major fact since every context modelling approaches is not realizable. The modelling must be implemented and realized within the existing infrastructure, which will be one of the vital concerns in the future work. Going further, integration is needed in CBIR framework for adding more meaningful context descriptions (Metadata) to the images. Also, the intelligent algorithms like Evolutionary Algorithms are needed to cluster the images based on time and location and find the context based on the metadata of the clustered images, which are to be noticed in future work.

7 Conclusion

This paper has presented a detailed review on context data retrieval and modeling systems that were enumerated in the above sections. In this, various models along with their better achievements were analyzed and described. From the review, it was known that several context data modeling and retrieval systems were really in the enabling view of better results of information retrieval and modeling. In conclusion,

- This paper reviewed about 60 research papers and declared the significant analysis.
- Initially, the analysis also focused on various context data retrieval that were reviewed in this paper.
- Subsequently, the analysis also focused on different context data modelling systems that were revealed in this paper.
- The analysis also reviewed the different applications used the concept of context aware data modelling and retrieval.
- The analysis also reviewed various intelligence methods that followed in the retrieval and modelling process.
- The accuracy rate of data retrieval and the error rate of retrieval was also analysed from all the reviewed papers.
- At last, this paper has presented various research issues that can be helpful for the researchers to accomplish further research on the features of context based systems.

References

1. Foschini L, Montanari R, Boukerche A, Corradi A (2016) Scalable and mobile context data retrieval and distribution for

community response heterogeneous wireless networks. IEEE Commun Mag 54(4):101–107
2. Brown PJ, Jones GJF (2001) Context-aware retrieval: exploring a new environment for information retrieval and information filtering. Pers Ubiquit Comput 5(4):253–263
3. Roussaki I, Strimpakou M, Pils C (2007) Distributed context retrieval and consistency control in pervasive computing. J Netw Syst Manag 15(1):57–74
4. Mizzaro S, Vassena L (2011) A social approach to context-aware retrieval. World Wide Web 14(4):377–405
5. Singh S, Kumar P (2018) User specific context construction for personalized multimedia retrieval. Multimed Tools Appl 77(11):13459–13486
6. Eito-Brun R (2016) Remote access to EAC-CPF context and authority records for metadata indexing: a solution based on open information retrieval standards. Arch Sci 16(2):149–165
7. Zakos J, Verma B (2006) A novel context-based technique for web information retrieval. World Wide Web 9(4):485–503
8. Golitsyna OL, Maksimov NV (2011) Information retrieval models in the context of retrieval tasks. Autom Doc Math Linguist 45(1):20–32
9. Hyman Sincich T, Will R, Agrawal M, Padmanabhan B, FridyIII W (2015) A process model for information retrieval context learning and knowledge discovery. Artif Intell Law 23(2):103–132
10. Iqbal R (2014) Information retrieval, decision making process and user needs in the context of ubiquitous and collaborative computing. J Ambient Intell Humaniz Comput 5(1):91–92
11. Yu Y, Li J, Yu J, Guan H, Wang C (2014) Pairwise three-dimensional shape context for partial object matching and retrieval on mobile laser scanning data. IEEE Geosci Remote Sens Lett 11(5):1019–1023
12. Li S, Purushotham S, Chen C, Ren Y, Kuo CCJ (2017) Measuring and predicting tag importance for image retrieval. IEEE Trans Pattern Anal Mach Intell 39(12):2423–2436
13. Tariq A, Foroosh H (2017) a context-driven extractive framework for generating realistic image descriptions. IEEE Trans Image Process 26(2):619–632
14. Sheikh I, Fohr D, Illina I, Linarès G (2017) Modelling semantic context of OOV words in large vocabulary continuous speech recognition. IEEE/ACM Trans Audio Speech Lang Process 25(3):598–610
15. Wang L, Zhao X, Si Y, Cao L, Liu Y (2017) Context-associative hierarchical memory model for human activity recognition and prediction. IEEE Trans Multimed 19(3):646–659
16. Erickson Tim (2010) Stealing from physics: modeling with mathematical functions in data-rich contexts. Teach Math Appl Int J IMA 25(1):25–32
17. Zhou N, Zhao WX, Zhang X, Wen JR, Wang S (2016) A general multi-context embedding model for mining human trajectory data. IEEE Trans Knowl Data Eng 28(8):1945–1958
18. Huynh T, Gao Y, Kang J, Wang L, Zhang P, Lian J, Shen D (2016) Estimating CT Image from MRI data using structured random forest and auto-context model. IEEE Trans Med Imaging 35(1):174–183
19. Dai W, Xiong H, Wang J, Cheng S, Zheng YF (2015) Generalized context modeling with multi-directional structuring and mdl-based model selection for heterogeneous data compression. IEEE Trans Signal Process 63(21):5650–5664
20. Nedic A, Tomlin D, Holmes P, Prentice DA, Cohen JD (2012) A decision task in a social context: human experiments, models, and analyses of behavioral data. Proc IEEE 100(3):713–733
21. Yang E-H, He D-K (2003) Efficient universal lossless data compression algorithms based on a greedy sequential grammar transform. 2. With context models. IEEE Trans Inf Theory 49(11):2874–2894
22. Yang E-H, Kieffer JC (2000) Efficient universal lossless data compression algorithms based on a greedy sequential grammar transform. I. Without context models. IEEE Trans Inf Theory 46(3):755–777
23. Prabha R, Ramesh MV, Rangan VP, Ushakumari PV, Hemalatha T (2017) Energy efficient data acquisition techniques using context aware sensing for landslide monitoring systems. IEEE Sens J 17(18):6006–6018
24. Milne D, Watling D (2018) Big data and understanding change in the context of planning transport systems. Elsevier, Amsterdam
25. Vasilecas O, Kalibatiene D, Lavbic D (2016) Rule- and context-based dynamic business process modelling and simulation. J Syst Softw 22:1–15
26. Gasparic M, Murphy C, Ricci F (2017) A context model for IDE-based recommendation systems. J Syst Softw 128:200–219
27. Chen S, Wang X (2018) Semiparametric estimation of panel data models without monotonicity or separability. J Econom 206:515–530
28. Hoyos JR, García-Molina J, Botía JA (2013) A domain-specific language for context modeling in context-aware systems. J Syst Softw 86(11):2890–2905
29. Erfani M, Zamani M, Rilling J, Keivanloo I (2016) Context-awareness in the software domain—a semantic web enabled modeling approach. J Syst Softw 121:345–357
30. ElSayed N, Smith RT, Marriott K, Thomas BH (2018) Context-aware design pattern for situated analytics: blended model view controller. J Vis Lang Comput 44:1–12
31. Coakes JM, Coakes EW (2000) Specifications in context: stakeholders, systems and modelling of conflict. Requir Eng 5(2):103–113
32. Papaioannou A, Dovriki E, Rigas N, Plageras P, Rigas I, Kokkora M, Papastergiou P (2010) Assessment and modelling of groundwater quality data by environmetric methods in the context of public health. Water Resour Manag 24(12):3257–3278
33. Degrandsart S, Demeyer S, Van den Bergh J, Mens T (2014) A transformation-based approach to context-aware modelling. Softw Syst Model 13(1):191–208
34. Cabrera O, Franch X, Marco J (2017) 3LConOnt: a three-level ontology for context modelling in context-aware computing. Softw Syst Model 18:1–34
35. Baele G, Van de Peer Y, Vansteelandt S (2010) Modelling the ancestral sequence distribution and model frequencies in context-dependent models for primate non-coding sequences. BMC Evol Biol 10:244
36. Petrelli D, Not E, Zancanaro M, Strapparava C, Stock O (2001) Modelling and adapting to context. Pers Ubiquit Comput 5(1):20–24
37. Guotao H, Sivakumar A, Polak JW (2012) Modelling travellers' risky choice in a revealed preference context: a comparison of EUT and non-EUT approaches. Transportation 39(4):825–841
38. Shanahan M, Shakeshaft A, Mattick RP (2006) Modelling the costs and outcomes of changing rates of screening for alcohol misuse by gps in the australian context. Appl Health Econ Health Policy 5(3):155–166
39. Ayotte Keith W, Davy Robert J, Coppin Peter A (2001) A simple temporal and spatial analysis of flow in complex terrain in the context of wind energy modelling. Bound-Layer Meteorol 98(2):275–295
40. Lasagna M, Caviglia C, De Luca DA (2014) Simulation modelling for groundwater safety in an overexploitation situation: the Maggiore Valley context (Piedmont, Italy). Bull Eng Geol Environ 73(2):341–355
41. Zorom M, Barbier B, Gouba E, Somé B (2018) Mathematical modelling of the dynamics of the socio-economic vulnerability of rural Sahelian households in a context of climatic variability. Model Earth Syst Environ 4:1–11

42. Martins J, Richardson DM, Henriques R, Marchante E, Marchante H, Alves P, Gaertner M, Honrado JP, Vicente JR (2016) A multi-scale modelling framework to guide management of plant invasions in a transboundary context. For Ecosyst 3:17
43. Gross T, Prinz W (2004) Modelling shared contexts in cooperative environments: concept, implementation, and evaluation. Comput Support Coop Work (CSCW) 13(3–4):283–303
44. Webb B, Gallagher S (2009) Action in context and context in action: modelling complexity in multimedia systems development. J Inf Technol 24(1):126–138
45. Heesch D, Petrou M (2010) Markov random fields with asymmetric interactions for modelling spatial context in structured scene labelling. J Signal Process Syst 61(1):95–103
46. Moroni D, Salvetti M, Salvetti O (2010) Shape analysis, semantic annotation and context modelling for the retrieval of 3D anatomical structures. Pattern Recognit Image Anal 20(1):86–93
47. El Morjani ZEA, Ebener S, Boos J, Abdel Ghaffar EA, Musani A (2007) Modelling the spatial distribution of five natural hazards in the context of the WHO/EMRO atlas of disaster risk as a step towards the reduction of the health impact related to disasters. Int J Health Geogr 6:8
48. Yu Z, Wong RK, Chi CH (2017) Efficient role mining for context-aware service recommendation using a high-performance cluster. IEEE Trans Serv Comput 10(6):914–926
49. Maran V, Machado A, Machadoa GM, Augustin I, de Oliveira JP (2018) Domain content querying using ontology-based context-awareness in information systems. Data Knowl Eng 115:152–173
50. Terama E, Clarke E, Rounsevell MDA, Fronzek S, Carter TR (2017) Modelling population structure in the context of urban land use change in Europe. Reg Environ Change 19:1–11
51. Ziaimatin H, Groza T, Tudorache T, Hunter J (2016) Modelling expertise at different levels of granularity using semantic similarity measures in the context of collaborative knowledge-curation platforms. J Intell Inf Syst 47(3):469–490
52. Hudson John A, BäckströmJ A, Rutqvist L, Jing T, HackersM, Chijimatsu R, Christiansson X-T, Feng A, Kobayashi T, Koyama H-S, Lee I, Neretnieks P-Z, Pan M Rinne, Shen B-T (2009) Characterising and modelling the excavation damaged zone in crystalline rock in the context of radioactive waste disposal. Environ Geol 57(6):1275–1297
53. Sunoj SM, Vipin N (2017) Some properties of conditional partial moments in the context of stochastic modelling. Stat Pap. https://doi.org/10.1007/s00362-017-0904-x
54. Tiwari V, Thakur RS (2015) Contextual snowflake modelling for pattern warehouse logical design. Sadhana 40(1):15–33
55. Wong RK, Lam F, Fergun MA (2001) Modelling and manipulating multidimensional data in semistructured databases. World Wide Web 4(1–2):73–99
56. Thu HN, Wehn U (2016) Data sharing in international transboundary contexts: the Vietnamese perspective on data sharing in the Lower Mekong Basin. J Hydrol 536:351–364
57. Hu W, Wu B, Wang P, Yuan C, Li Y, Maybank S (2018) Context-aware random walk graph kernels and tree pattern graph matching kernels with applications to action recognition. IEEE Trans Image Process 27(10):5060–5075
58. Gronchi G, Provenzi E (2017) A variational model for context-driven effects in perception and cognition. J Math Psychol 77:124–141
59. Pokojski J, Oleksinski K, Pruszynski J (2018) Knowledge based processes in the context of conceptual design. J Ind Inf Integr. https://doi.org/10.1016/j.jii.2018.07.002
60. Schoenfisch J, Stuckenschmidt H (2017) Analyzing real-world SPARQL queries and ontology-based data access in the context of probabilistic data. Int J Approximate Reasoning 90:374–388
61. Choi S, Seo J, Kim M, Kang S, Han S (2017) Chrological big data curation: a study on the enhanced information retrieval system. IEEE Access 5:11269–11277
62. Chen Z, Zhong F, Min G, Leng Y, Ying Y (2018) Supervised intra- and inter-modality similarity preserving hashing for cross-modal retrieval. IEEE Access 6:27796–27808
63. Yang J, Jiang B, Li B, Tian K, Lv Z (2017) A fast image retrieval method designed for network big data. IEEE Trans Ind Inf 13(5):2350–2359
64. Chi M, Plaza A, Benediktsson JA, Sun Z, Shen J, Zhu Y (2016) Big data for remote sensing: challenges and opportunities. Proc IEEE 104(11):2207–2219
65. Tang J, Wang K, Shao L (2016) Supervised matrix factorization hashing for cross-modal retrieval. IEEE Trans Image Process 25(7):3157–3166
66. Chen Z et al (2015) A survey of bitmap index compression algorithms for Big Data. Tsinghua Sci Technol 20(1):100–115
67. Ayala-Romero JA, Alcaraz JJ, Vales-Alonso J (2018) Data-driven configuration of interference coordination parameters in HetNets. IEEE Trans Veh Technol 67(6):4974–5
68. Shi L, Wu Y, Liu L, Sun X, Jiang (2018) Event detection and identification of influential spreaders in social media data streams. Big Data Min Anal 1(1):34–
69. Li Y, Zhang Y, Huang X, Zhu, Ma J (2018) Large-scale remote sensing image retrieval by deep hashing neural networks. IEEE Trans Geosci Remote Sens 56(2):950–965
70. Müller H, Unay D (2017) Retrieval from and understanding of large-scale multi-modal medical datasets: a review. IEEE Trans Multimed 19(9):2093–21
71. Shao Z, Cai J, Wang Z (2018) Smart monitoring cameras driven intelligent processing to big surveillance video data. IEEE Trans Big Data 4(1):105–116
72. Zhao G, Jia Q, Guo L, Li M, Li P (2017) Privacy-preserving verifiable set operation in big data for cloud-assisted mobile crowdsourcing. IEEE Internet Things J 4(2):572–582
73. Wu Y et al (2016) CAMP: a new bitmap index for data retrieval in traffic archival. IEEE Commun Lett 20(6):1128–1131
74. Jiang S, Qian X, Mei T, Fu Y (2016) Personalized travel sequence recommendation on multi-source big social media. IEEE Trans Big Data 2(1):43–56
75. Zhu Q, Shyu ML (2015) Sparse linear integration of content and context modalities for semantic concept retrieval. IEEE Trans Emerg Topics Comput 3(2):152–160
76. Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge
77. Storn Rainer, Price Kenneth (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359
78. Yao X, Liu Y, Lin Guangming (1999) Evolutionary programming made faster. IEEE Trans Evol Comput 3(2):82–102
79. Rechenberg I (1973) Evolution strategy. Springer, Berlin, pp 83–114
80. Simon D (2008) Biogeography-based optimization. IEEE Trans Evol Comput 12(6):702–713
81. Müller H, Guadagni S (2008) Regional chemotherapy for carcinoma of the lung. Surg Oncol Clin N Am 17(4):895–917. https://doi.org/10.1016/j.soc.2008.04.012
82. Fiorentini G, Rossi S, Bernardeschi P, Cantore M, Guadagni S (2015) Is there a new drug beyond floxuridine for intra-arterial hepatic chemotherapy in liver metastases from colorectal cancer. J Clin Oncol 23(9):2105
83. Guadagni S, Fiorentini G, Clementi M, Palumbo G, Masedu F, Deraco M, De Manzoni G, Chiominto A, Valenti M, Pellegrini C (2017) MGMT methylation correlates with melphalan pelvic perfusion survival in stage III melanoma patients. A pilot study. Melanoma Res 27:439–447