SEARCH FOR SOCIAL SMART OBJECTS CONSTITUTING SENSOR ONTOLOGY, SOCIAL IoT AND SOCIAL NETWORK INTERACTION

Vaibhava lakshmi R
Department of Computer Technology
Madras Insitute of Technology
Chennai, India
vaibhavi18092002@gmail.com

Gerard Deepak
[2]Manipal Institute of Technology
Bengaluru, Manipal Academy of Higher Education,
Manipal, India
gerard.deepak.christuni@gmail.com

Santhanavijayan A
Department of Computer Science and Engineering,
National Institute of Technology,
Tiruchirappalli, India
vijayana@nitt.edu

Radha S
School of Advanced Sciences,
Vellore Institute of Technology
Chennai, India
radha.s@vit.ac.in

***Abstract*— An emerging constituent of Internet of Things is the Social IoT, which aids creation of Social relationships amongst interacting objects. SIoT attempts to moderate the shortcomings of IoT in the areas of trust, resource discovery and scalability by taking a cue from social computing. In this paper, we have proposed the OntoSSSO framework for recommending Socially Similar Smart objects to users, which is knowledge-centric, ontology-driven and dataset-driven. It incorporates Semantic Intelligence. The proffered model is compared for performance along with the baseline models using sundry performance metrics. Our model outperforms the other models, yielding a precision of 95.83 %.**

***Keywords*— Social Computing, Social Internet of Things, Semantic Intelligence**

I. INTRODUCTION

Social web is the social aspect of the World Wide Web or the Web 3.0 which is a community by itself. The social web constitutes web services and interfaces that aid human interactions. People have social media accounts and they use it for a plenitude of reasons like sharing pictures, connecting with people, messaging etc. Hence, it is fostering better communication between people. In any network, entities connect and form groups and collaborate in attaining certain goals. The users and entities by themselves have interactions hence forming social relationships

Internet of Things refers to physical objects with sensors, software and other technologies that link to other devices and send/receive data over the Internet. Digitization is conversion of physical information into digital data. All companies can transform into digital companies by using an IoT platform that can digitize their products. IoT based digitization is comprehensible as well as intriguing enough to make employees of a company work on it.

Social Internet of Things is a constituent of IoT which helps humans to create social relationships with other objects. SIoT aims to overcome the demerits of IoT in the fields of resource discovery, scalability and trust. It aids humans to machine communication in comparison to IoT which just includes machine to machine communication. The use of SIoT can be a great boon as it helps to establish a level of trustworthiness in order to leverage the interactions betwixt humans and objects. Also, incorporation of SIoT can help models developed to study social networks be reused to look into IoT related concerns.

A. Motivation

Objects and entities are present along with the collaborating users in an IoT infrastructure. Discovery of smart objects is essential. Since, social web is tending to become highly cohesive and data-dense, semantic labeling of entities must be done. Owing to the cohesiveness, it is very difficult to search and annotate entities. Hence, semantic social web compliant models are required. Hence, the proposed OntoSSSO framework can help to overcome the aforementioned drawbacks owing to the knowledge-centric and semantically inclined strategy.

B. Contribution

In this paper, we propose the knowledge-centric and dataset-driven paradigm for recommending socially similar smart objects to the users. It includes the following steps: 1) Preprocessing of the dataset followed by Latent Dirichlet Allocation for enriching the dataset. 2) Aggregation of SSN, SOSA and SemSNI ontologies to form a hybrid, for Ontology alignment. 3) Classification of metadata using the Vanilla LSTM classifier. 4) Aggregation of the top 50 % of the classified instances with upper ontology. 5) Merging the formalized ontology with 100 % of classified instances to obtain the detailed ontology. 6) Aligning the detailed ontology with the initially aggregated hybrid ontology. 7) Ranking and recommending based on the increasing order of Jaccard similarity value.

C. Organization

This paper has the following sections: Section II is on related works and the extant methodologies. In section III, we have elucidated the architecture of the proffered framework. In section IV, the results for the proposed framework are discussed and compared with the other baseline models. Finally, section V contains the conclusion and future works to be done.

II. RELATED WORKS

Roopa M.S. et. al. [1] have put forth a new methodology for searching objects using propinquity of physical location and social milieu of users in social groups, so as to ameliorate the performance of searching over the SIoT. In terms of the mean path length, this technique outperforms extant methods. Kumaran P. et. al. [2] have reviewed a multitude of techniques, real life applications, interpretations and difficulties related to Social IoT. Sahraoui Dhelim et. al. [3] have discussed the significance of IoT in management of social relationships and the issue of social relationships surge in IoT. They have used Artificial Social Intelligence to assess the proposed methodologies including the Deep Learning and Social-oriented Machine learning techniques.

Mozhgan Malekshahi Rad et. al. [4] have reviewed the articles on SIoT published between 2011 and 2019 in order to study the characteristics, parameters, advantages and disadvantages associated with the components of SIoT. Along with this, the merits and demerits of the articles were examined and the predominantly used simulation tools are mentioned. Pratibha Mahajan et. al. [5] have proffered a new smart object recommendation framework called SORec for SBSN. Transition probabilities betwixt two nodes in SBSN are determined that help to efficiently represent objects and users. Zhang et. al. [6] have put forth a smart object recommendation method that draws out and determines topics and features of text pertinent to service, connected to smart objects. Also, it introduces the “thing-thing” relationship information in the IoT to ameliorate the efficacy of recommendation.

Jike Ge et. al. [7] have proffered an ontology-based technique for personalized recommendation of knowledge. It provides users an automatic tool that cuts down redundant retrieved information. Li Ling et. al. [8] have proposed an ontology-based service recommendation system for social networks. They have ameliorated the Hownet-based semantic similarity algorithm by considering the density of sememe tree. Also, for more precise retrieval of user interests, they have improvised the TF-IDF algorithm in accordance to the features of micro logs by combining with text rank algorithm. Sajjad Ali et. al. [9] have put forward a framework providing a basis for development of lightweight microservices based on web objects that are socially connected. A semantic ontology model is being constructed to realize the operations of interoperability service.

Yuanyi Chen et. al. [10] have proffered a time-aware smart object recommendation model by taking into account, the user’s choice over time and the social similarity between smart objects. Subash Rajendran et. al. [11] have proffered the Object Recommendation based Friendship Selection (ORFS) model for navigability of networks and handling of social relationships for smart objects. A User Object Selection based on Grey Wolf algorithm for recommending smart objects is incorporated. The maximum rank neighbourhood technique is adopted for Object Friendship Selection through navigability of networks. Daniel Defiebre et. al. [12] have put forth the DANOS framework for SIoT that includes three focal factors: dynamicity, decentralization and anthropomorphism. In DANOS, based on the job, smart

objects themselves adjust their social neighbourhood. Based on the idiosyncrasy of users, smart objects adopt anthropomorphic behaviour.

Amar Khelloufi et. al. [13] have proffered a service recommendation system based on the social relationships that owners of devices share. The connection between the service requester and service provider determines the kind of recommendation. Ashish Pandharipande [14] has reviewed diverse IoT applications facilitated by social data. Sundry primary social sensing methodologies and concomitant challenges are also elucidated. A case study of connected landmark driving based on social sensing is presented to show how social sensing methods can be an application of great significance.

D. Sivaganesan [15] has put forth an interest-based algorithm for coextensive social action that escalates the influence in social sites. An efficient influential nodes set is derived by implementing the machines with CPUs with parallelism via community framework. This helps in having reduced time of execution and overcoming real-world difficulties pertaining to social networks. Jennifer S Raj [16] has proposed a secure proxy decryption model with improved publishing and subscribing facility in the mobile social networks. The credentials of the users and confidentiality of the data are secured by ingress control methodologies that aid privacy protection in a self-sufficient way. Public-key encryption based on keyword searching with encryption based on ciphertext policy attribute is used in this model. At the end users, ciphertext decryption is performed to bring down the energy expenditure by the protected proxy decryption framework.

In [17-22] sundry literatures pertinent to the proffered framework have been depicted.

III. Proposed Architecture

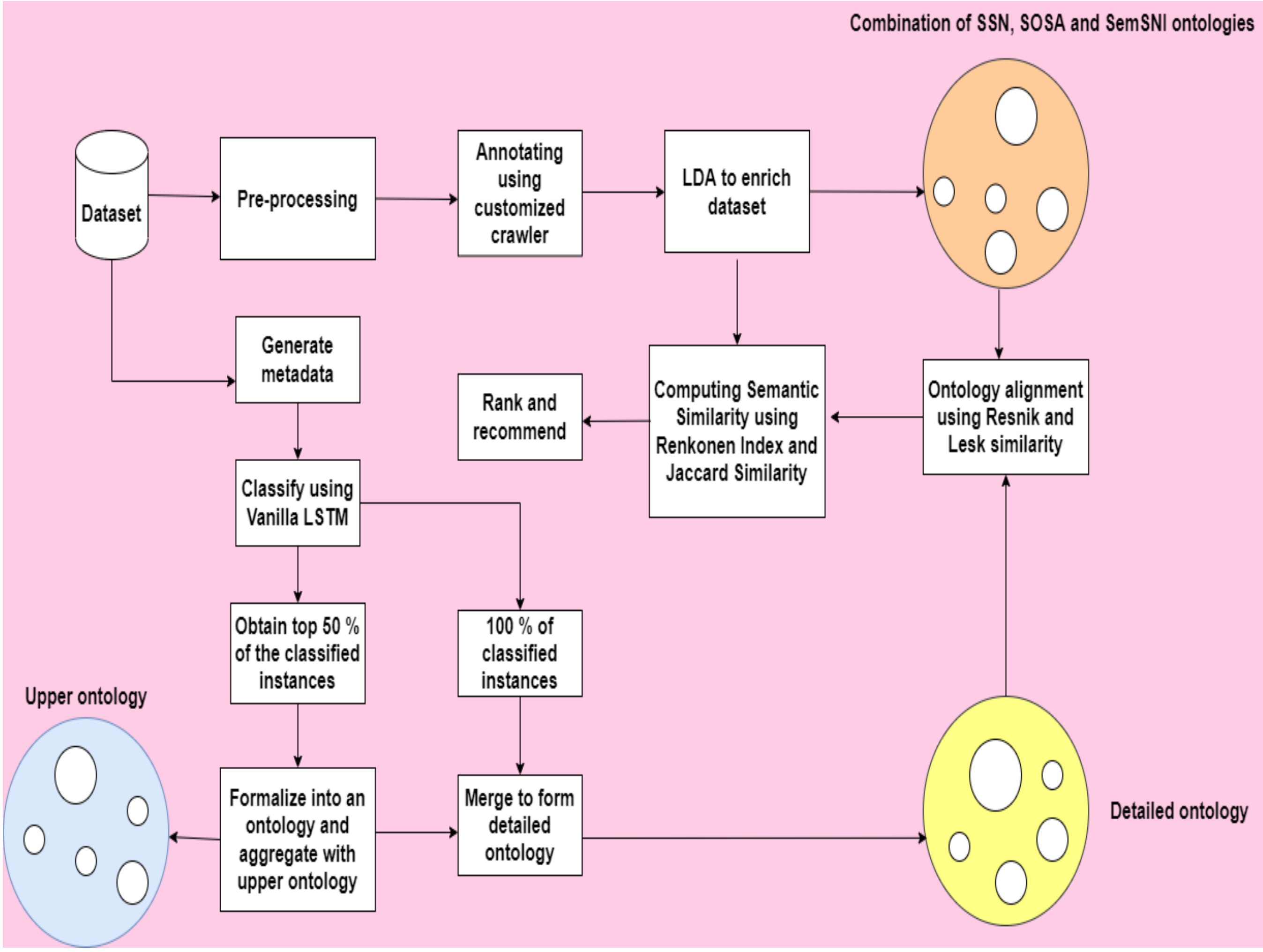

Fig. 1. Architecture of the OntoSSSO framework

A. Dataset

We have used the Brightkite dataset obtained from the Stanford Large Network Dataset Collection. The Stanford Large Network Dataset Collection is a set of 50 large network datasets. It comprises of graphs that depict social networks, web graphs, citation networks, online reviews, online communities and more. The Brightkite was earlier a service provider of social networking based on locations, where locations were shared by users by registering. The friendship network includes 58228 nodes with 214078 edges. A total of 4,491,143 user entries were recorded from April,2008 to October,2010. Actually, the network was directed, but then, a network with undirected edges was created denoting friendship from both sides. It has an average clustering coefficient value of 0.1723. The clustering coefficient of a node denotes how absolute the neighbourhood of the node is. When a whole network of nodes is considered, the average clustering coefficient is computed.

B. Implementation

Fig. 1. depicts a dataset-driven, knowledge-centric paradigm for recommendation of socially connected SIoT objects by the means of ontological, knowledge-centric and semantically inclined strategy. Initially, the dataset is subjected to pre-processing so as to eliminate inconsistencies in the form of punctuations, stop words and typos. Pre-processing involves tokenization, lemmatization, stop word removal and named entity recognition for the text in the dataset. For pre-processing, Python's NLTK (Natural Language Tool Kit) was used and for lemmatization, the WordNet 3.0 Lemmatizer was used. A white space special character and period punctuated tokenizer was used. For stop word removal, RegEx based stop word matching algorithm was incorporated. Thesaurus based NER (Named Entity Recognition) was achieved. The dataset is annotated. Annotating or labeling is mandatory because of the limited textual content in the dataset. Annotations are obtained using automatic and customized crawlers. After that, the Latent Dirichlet Allocation is done in order to furnish the LDAenriched dataset. LDA is a topic modeling approach used for identifying and recommending pertinent topics which are hidden or uncovered. It uses matrix factorization method. In LDA, documents are composed of collections of topics and topics constitute words. The LDA makes sure that there is lot of variational heterogeneity and gives more depth to the categories in the dataset, hence making it enriched. The

enriched dataset is further used for Ontology Alignment using three distinct Ontologies namely, the SSN, SOSA and SemSNI Ontology. The SSN (Semantic Sensor Network) Ontology is a standard ontology for describing about sensors, the incorporated procedures, observed properties and the actuators. The SemSNI is an ontology of Social Network interactions. The SOSA (Sensor, Observation, Sample and Actuator) ontology provides a diaphanous and general-purpose specification for depicting the interaction between the objects involved in actuation, sampling and observation.These three ontologies are aggregated together. The aggregation is done by computing the node similarity using the Shannon's entropy by using step deviation of 0.45. The Shannon's entropy is a measure of extent of aberrancy amongst the entities of a dataset. It is given as:

$$H(x) = -\sum_{i=1}^{n} p(x_i)\, log\,_b p(x_i) \quad (1)$$

where $p(x_i)$ is the probability of occurrence for the term/node x_i and the $\sum_{i=1}^{n} operator$ sums the probabilities for 'i' nodes ranging from 1 to n in a network. The Shannon's entropy measure has sundry units. It is based on the base of the logarithm 'b'. We have used base 2 for the calculation. Hence, the node similarity between two nodes x_i and x_j are computed as:

$$S_{i,j} = |H(x_i)| - |H(x_j)| \quad (2)$$

The reason for setting the step deviation value for similarity as 0.45 is to establish a single link between the terminal nodes of the three ontologies. Node similarity is computed randomly across the nodes in order to arbitrarily align the nodes between the SemSNI, SSN and SOSA ontology to construct the hybrid of large scale connected sensor network ontology and the social ontology which is perfect for Social IoT. The dataset categories are used for generating the metadata using the MetaTag Harvester and the RDF Distiller tool. The RDF subjects and objects are considered. If the RDF predicate is a link, it is expelled. If it is not a URL and is textual, then it is aggregated into the metadata. Since this metadata is exponentially large, it is subjected to classification using the Vanilla LSTM classifier. It is a deep learning classifier which supports automatic feature selection and classification, hence parameters are not explicitly mentioned. It has a single hidden layer of LSTM units. The Vanilla LSTM uses gating procedure that influences memorization. It consists of 3 gates: input, forget and output gates. The memory is stored by these gates in analog format. The forget gate determines what information to be kept and which one to be ignored. The input gate regulates the amount of information to be written onto the internal cell state. The output gate decides the value of the succeeding hidden state. This state contains details on antecedent inputs. We considered a stack of 7 Vanilla LSTMs, each with one hidden layer having 18 nodes each. A dropout value of 0.5, decay rate of 0.97 and a momentum value of 0.7 was used. Weights were initialized between 0 and 0.25. Sigmoid activation function was made use of. Batch size was set to 64 for a total of 52 epochs.

The top 50 % of the classified instances are yielded. Subsequently, an upper ontology is formulated using the terms of the dataset and also by discovering similar terms relevant to the dataset by using customized crawlers. The deviations are kept minimal in order to support the upper ontology. Upper ontology is generated using OntoCollab tool. The top 50 % of the classified instances from the Vanilla LSTM are again formalized into an ontology and aggregated with the upper ontology. This upper ontology is aggregated with the final rank terms for enriching the expanded click of the rank terms. This formalized ontology is merged with the 100 % classified instances obtained from the Vanilla LSTM to yield the detailed ontology. This detailed ontology is aligned with that of the initially built ontologies namely, the SSN, SOSA and SemSNI. This alignment is done using the Resnik's similarity and the Lesk similarity with a threshold of 0.5 and 0.75 respectively. The nodes which are common are collated into the final ontology. This final ontology is computed for semantic similarity with the LDA enriched dataset by using Renkonen Index and the Jaccard similarity.

Jaccard similarity is a vicinage measure for computing the correspondence amongst entities. The mathematical representation is given as:

$$J(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (3)$$

where A and B are two sets respectively. |A|, |B| and |A∩B| represent cardinalities of the sets A, B and A∩B respectively.

The Renkonen Similarity Index is used to measure the dissimilarity between two communities based on commensurate abundances of individuals of compound entities. The mathematical equation is as below:

$$p_i = \frac{n_i}{\sum_i n_i} \quad (4)$$

where n_i is the abundance of a species 'i' and $\sum_i n_i$ is the sum of abundances of all species.

The Renkonen Index is set with a step deviation of 0.5 and the Jaccard similarity is set to a threshold of 0.80. The reason for lapsing the step deviations of Renkonen Index is owing to the strong feature of Renkonen Index and to allow initial aggregation of ontological instances, nodes and individuals with liberation. However, this is made strong by enhancing the threshold of Jaccard Similarity to 0.80. The instances are ranked and recommended in the increasing order of the Jaccard similarity and is yielded to the user. If the user is satisfied, the search haults, otherwise, there will be new clicks.

IV. PERFORMANCE EVALUATION AND RESULTS

The proffered OntoSSSO framework, is a social smart objects search scheme which constitutes social ontology, social IoT and social networks interaction. It is evaluated using Precision, Recall, F-measure percentages and False Discovery Rate (FDR) as potential metrics. Precision, Recall and F-measure percentages indicate the pertinence of the results whereas the FDR quantifies the number of False Positives rendered by the model. In order to compare the performance of the proposed OntoSSSO framework, it is baselined with ORFS (Object Recommendation-based Friendship Selection), HDAR (Human-centered Decentralized Architecture and Recommendation) and SRSRS (Social-Relationships-based Service Recommendation System) models put forth in [11], [12] and [13] respectively. The baseline models are evaluated in the same environment as the proposed OntoSSSO model and the results are tabulated.

TABLE I. COMPARISON OF PERFORMANCE OF THE PROPOSED ONTOSSS FRAMEWORK WITH OTHER APPROACHES TABLE TYPE STYLES

Model	Average Precision %	**Average Recall %**	**F-Measure %**	**FDR (False Discovery Rate)**
ORFS (Object Recommendation-based Friendship Selection)	90.44	92.47	91.44	0.10
HDAR (Human-centered Decentralized Architecture and Recommendation)	91.63	93.18	92.40	0.09
SRSRS (Social-Relationships-based Service Recommendation System)	91.97	93.29	92.62	0.09
Proposed OntoSSSO	**95.83**	**97.47**	**96.64**	**0.05**

From Table. 1., it can be observed that the proposed model yields the highest Precision of 95.83 %, Recall of 97.47 %, F-measure of 96.64 % and the least FDR of 0.05, whereas, the ORFS model yields the least Precision of 90.44 % , Recall of 92.47 % and F-measure of 91.44 % respectively. The ORFS model produces the highest FDR value of 0.10.

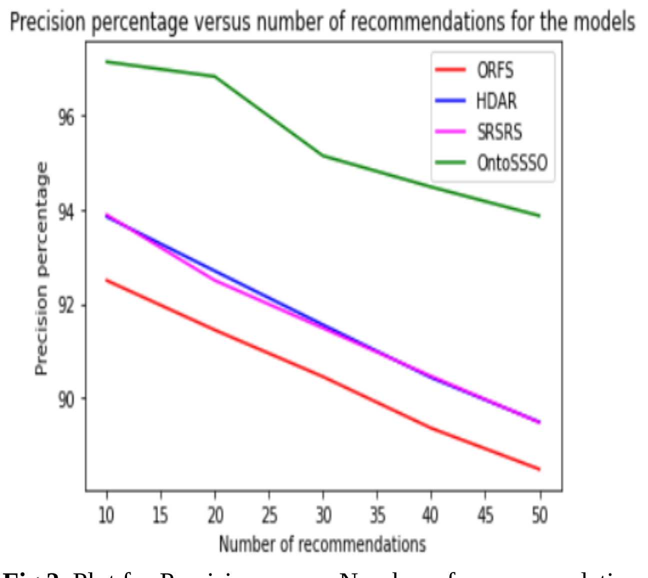

Fig 2. Plot for Precision versus Number of recommendations for all the models.

The reason why the proffered OntoSSSO model outperforms the baseline models is that the OntoSSSO architecture incorporates a semantically inclined, knowledge-driven and dataset-driven strategy. Also, it is enriched by means of annotations, Latent Dirichlet Allocation, the integration of three distinct Ontologies and the formalizing of Upper ontologies. Apart from this, large-scale metadata is generated. Moreover, the Vanilla LSTM, which is a Deep Learning classification algorithm was used for classification of the metadata. Also, Ontology alignment was achieved by using the Resnik and Lesk Similarities with a defined threshold. The Resnik Similarity is computed using Renkonen Index and Jaccard Similarity with differential thresholds and step deviation measures owing to the differential threshold scheme and heterogeneity in the relevance computation mechanisms. Therefore, integration of various knowledge generation mechanisms like metadata generation, Ontology alignment and Ontology generation ensure large degree of auxiliary knowledge to the framework and makes sure that the proposed model transcends the baseline models.

Fig.2. depicts the performance versus the number of recommendations distribution curve. In order to obtain this curve, the top 10, 20, 30, 40 and 50 recommendations were considered for every query in the implementation. For all the queries, ground-truth was collected and for all the ground-truths, the deviants were computed between the ground truths and the top 10, 20, 30, 40 and 50 recommendations respectively. However, the ground-truth was not prioritized. It was rather considered as a whole irrespective of the order of order of occurrence, whether it was top 10 or bottom 10 queries. The number of queries were taken to be equal to number of recommendations. For instance, for 20 queries, top 20 recommendations were considered. The proposed OntoSSSO framework occupies the uppermost position despite the number of recommendations. It is immediately followed by the HDAR and SRSRS models respectively. The lowest in the hierarchy is the ORFS model regardless of the number of recommendations.

V. CONCLUSION

In this paper, we have discussed about the OntoSSSO framework which amalgamates various knowledge generation mechanisms like metadata generation, ontology alignment and ontology generation to ensure large degree of auxiliary knowledge. Also, our model is enriched by means of annotations, Latent Dirichlet Allocation, the integration of three distinct ontologies and the formalizing of upper ontologies. Hence, our model outperforms the extant techniques. The OntoSSSO model yielded a Precision of 95.83 %, transcending the baseline models. As a part of future work, we plan to incorporate further hybridizations and a more efficient scheme for aggregation of ontologies with fewer extraneous sources.

REFERENCES

[1] Roopa, M. S., Valla, D., Buyya, R., Venugopal, K. R., Iyengar, S. S., & Patnaik, L. M. (2018, December). SSSSS: search for social similar smart objects in SIoT. In 2018 Fourteenth International Conference on Information Processing (ICINPRO) (pp. 1-6). IEEE.

[2] Kumaran, P., & Sridhar, R. (2020, June). Social Internet of Things (SIoT): techniques, applications and challenges. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 445-450). IEEE.

[3] Dhelim, S., Ning, H., Farha, F., Chen, L., Atzori, L., & Daneshmand, M. (2021). IoT-enabled social relationships meet artificial social intelligence. *IEEE Internet of Things Journal*, *8*(24), 17817-17828.

[4] Rad, M. M., Rahmani, A. M., Sahafi, A., & Qader, N. N. (2020). Social Internet of Things: vision, challenges, and trends. *Human-centric Computing and Information Sciences*, *10*(1), 1-40.

[5] Mahajan, P., & Kaur, P. D. (2021). Smart object recommendation (SORec) architecture using representation learning in Smart objects-Based Social Network (SBSN). The Journal of Supercomputing, 77(12), 14180-14206.

[6] Zhang, H., Zhu, L., Dai, T., Zhang, L., Feng, X., Zhang, L., & Zhang, K. (2022). Smart object recommendation based on topic learning and joint features in the social internet of things. Digital Communications and Networks.

[7] Ge, J., Chen, Z., Peng, J., & Li, T. (2012, August). An ontology-based method for personalized recommendation. In 2012 IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing (pp. 522-526). IEEE.

[8] Ling, L., He, C., & Yingwei, S. (2015, September). Ontology based service recommendation system for social network. In 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC) (pp. 1640-1644). IEEE.

[9] Ali, S., Kibria, M. G., Jarwar, M. A., Lee, H. K., & Chong, I. (2018). A model of socially connected web objects for IoT applications. Wireless Communications and Mobile Computing, 2018.

[10] Chen, Y., Zhou, M., Zheng, Z., & Chen, D. (2019). Time-aware smart object recommendation in social internet of things. *IEEE Internet of Things Journal*, *7*(3), 2014-2027.

[11] Rajendran, S., & Jebakumar, R. (2021). Object Recommendation based Friendship Selection (ORFS) for navigating smarter social objects in SIoT. Microprocessors and Microsystems, 80, 103358.

[12] Defiebre, D., Sacharidis, D., & Germanakos, P. (2022). A human-centered decentralized architecture and recommendation engine in SIoT. User Modeling and User-Adapted Interaction, 1-57.

[13] Khelloufi, A., Ning, H., Dhelim, S., Qiu, T., Ma, J., Huang, R., & Atzori, L. (2020). A social relationships based service recommendation system for SIoT devices. IEEE Internet Things J 8: 1859–1870.

[14] Pandharipande, A. (2021). Social sensing in IoT applications: A review. IEEE Sensors Journal, 21(11), 12523-12530.

[15] Sivaganesan, D. "Novel Influence Maximization Algorithm for Social Network Behavior Management." Journal of ISMAC 3, no. 01 (2021): 60-68.

[16] Raj, Jennifer S. "Secure Data Sharing Platform for Portable Social Networks with Power Saving Operation." Journal of IoT in Social, Mobile, Analytics, and Cloud 3, no. 3 (2021): 250-262.

[17] Cho, E., Myers, S. A., & Leskovec, J. (2011, August). Friendship and mobility: user movement in location-based social networks. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1082-1090).

[18] Shreyas, K., Deepak, G., & Santhanavijayan, A. (2020). GenMOnto: a strategic domain ontology modelling approach for conceptualisation and evaluation of collective knowledge for mapping genomes. Journal of Statistics and Management Systems, 23(2), 445-452.

[19] Deepak, G., & Priyadarshini, J. S. (2018). Personalized and enhanced hybridized semantic algorithm for web image retrieval incorporating ontology classification, strategic query expansion, and content-based analysis. Computers & Electrical Engineering, 72, 14-25.

[20] Deepak, G., & Santhanavijayan, A. (2020). OntoBestFit: a best-fit occurrence estimation strategy for RDF driven faceted semantic search. Computer Communications, 160, 284-298.

[21] Deepak, G., Teja, V., & Santhanavijayan, A. (2020). A novel firefly driven scheme for resume parsing and matching based on entity linking paradigm. Journal of Discrete Mathematical Sciences and Cryptography, 23(1), 157-165.

[22] Pushpa, C. N., Gerard Deepak, J. Thriveni, and K. R. Venugopal. "Onto Collab: Strategic review oriented collaborative knowledge modeling using ontologies." In 2015 Seventh International Conference on Advanced Computing (ICoAC), pp. 1-7. IEEE, 2015.