Wireless Communications and Mobile Computing

Algorithm Optimization for Wireless Mobile Applications of Smart Cities

Lead Guest Editor: Donghyun Kim
Guest Editors: Michele Nogueira and Ravanasamudram N. Uma

Algorithm Optimization for Wireless Mobile Applications of Smart Cities

Algorithm Optimization for Wireless Mobile Applications of Smart Cities

Lead Guest Editor: Donghyun Kim
Guest Editors: Michele Nogueira and Ravanasamudram N. Uma

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 1069824, 9 pages
https://doi.org/10.1155/2020/1069824

Research Article

Optimizing Resource Discovery Technique in the P2P Grid Systems

Wang Tun,[1] J. Pourqasem,[2] and S. A. Edalatpanah[2]

[1]*Department of Computer, Changzhi University, Changzhi, Shanxi 046011, China*
[2]*Department of Research Center, Ayandegan Institute of Higher Education, Tonekabon, Iran*

Correspondence should be addressed to S. A. Edalatpanah; saedalatpanah@gmail.com

Received 6 April 2019; Revised 22 August 2019; Accepted 14 November 2019; Published 4 January 2020

Guest Editor: Michele Nogueira

Distributed discovery service is a main concept in the scalable and dynamic grid environments. In this paper, based on the super-peer technique, we propose a new topology for the grid discovery service. The model is designed in such a way that each super-peer within the cluster has the routing indices (RIs) based on cobweb and uses the hop-count routing index (HRI) to select the best neighbor. Besides, each super-peer includes a cache table, which stores the query and the query results. Furthermore, from the point of view of the response time and the number of submitted messages, we compare the new model with an existing method. An illustrative simulation is also presented to show the efficiency and validation of the new technique.

1. Introduction

Grid systems generally solve the science and engineering problems in the large-scale environment and integrate the high volume of computing and storage resources, data, services, and applications that are distributed geographically [1]. Discovery service in grid is directly affected by the diversity, dynamic, heterogeneity, and distribution features, which can notably increase the grid efficiency [2]. The classical approaches that are utilized for grid discovery are based on the centralized [3–7] or hierarchical [8, 9] architectures. These methods collect the grid resource information and use the central index server to sustain them. The grid environments suffer from two main drawbacks: bottleneck and single point of failure [10–12]. To avoid the mentioned problems, the P2P model has been adopted [3, 13–15]. The P2P topology in grid environments consists of the structured and unstructured models [16–23]. The super-peer mechanism is a category of P2P technology, which has been proposed in [24], for more details see [25–28].

The modified breadth-first-search (BFS) and intelligent search (IS) methods have been introduced in [29]. Breadth-first-search [30] technique is an extended protocol of Gnuttela that is implemented locally and is based on keyword searching, which selects the neighbor peers randomly. In the IS method, first for each peer, a profile is built and then this profile is used for the sending query. Ghorbani et al. [31] presented a self-adaptive resource discovery in the unstructured P2P environment based on the feedback of the network nodes. They also used the learning automata (LA) algorithm [32] to educate peers in the discovery process for finding the best neighbor peer. Löser et al. [33], using the semantic clustering method [34, 35] in super-peer network, presented the semantic overlay clustering (SOC) to link the information provider peer to each cluster super-peer semantically.

Some approaches [2, 36–39] also index and summarize the resource features to reduce the amount of information that is kept and transferred in the discovery service.

Routing indices (RIs) as a P2P technique organize the indexed and summarized resource information of nodes [40]. The benefit of this approach is that queries are disseminated and forwarded only among the locations of the network where resources existed, thus avoiding to flood query requests to the nodes which are not useful. The main drawback of this technique is that this indexing system comes from the presence of cycles in the network graph. As an extension of the RIs, hop-count routing index (HRI) is produced that keeps the resource information in a table structure at different hops [41–43].

Marzolla et al. [44] proposed a discovery technique in which the resource information of each domain and the summarized information of the neighbor domains are maintained by the brokers. This method uses the k-bit vector to represent the indices and summarize the information of resources. Puppin et al. [41] implemented a grid information service (GIS) that utilized the HRI. This method has two significant entities: (1) the agent that is in control of a super-peer builds the node's resource information and (2) the aggregator that receives the resource information from the mentioned agent and indexes it.

Caminero et al. [42] proposed the use of HRI as a means to route jobs within a P2P system. The main drawback of this model is that it only considers numeric parameters (such as effective bandwidth, the number of total machines in the domain, and the domain workload) to perform the resource discovery. Caminero et al. [43] presented a model that uses the HRI to construct the summary information based on cobweb tree [37]. This approach calculates the *goodness function* [40] for routing and forwarding the query to the neighbor peers, which likely have more probability for that query. Furthermore, to adapt the summaries with RIs, they proposed the technique so-called n-level summaries to filter attributes that have the same values, and their probability is less than a threshold.

In this paper, we present a new discovery method based on super-peer network in which each super-peer within the cluster has the routing indices (RIs) based on cobweb and uses the hop-count routing index (HRI) to select the best neighbor. Furthermore, each super-peer utilizes a cache table to store the query and the query results.

Remaining of this paper has been structured as follows: Section 2 expresses the configuration of the grid based on the super-peer network and Section 3 introduces the proposed resource discovery method in detail. Section 4 evaluates the simulation results based on the GridSim [45] that shows the effective performance of our approach. The final section concludes the paper and presents the guideline for the later works.

2. Configuration of Grid

Grid systems can use the super-peer topology for implementation of their infrastructure. As shown in Figure 1, the super-peers (SP) in each virtual organization (VO) or cluster operate as a server. Super peers receive the query from the client peer (CP) or the neighbor super-peers and answer them. The SPs communicate with each other and make an overlay network [46]. Generally, an SP has high capability among other CP in each VO/cluster that processes the query. When a CP connects to a VO/cluster, the local SP indexes the information of its resources. At leaving, the indexed information related to that CP is removed from the VO/cluster. In addition, if the resource information of a CP is changed, the updated information is sent to the mentioned SP. When a CP needs resources, it creates a query and submits it to the SP. This SP seeks among indexed information to find appropriate resources. If the requested resources are found, the IP address of a CP, that is, the owner of the result, is sent to the requested peer; otherwise, the SP creates a copy of the request and forwards it to the overlay network. The SP of each cluster searches its domain to find the requested resources.

3. New Discovery Method

We present a new discovery technique based on the super-peer network. First, we create some clusters and organize the grid nodes. In each cluster, a super-peer maintains the summary of resource information that is owned by client peers. We use cobweb [37–39] to cluster and summarize the resource information. Each super-peer uses the RIs [40] to structure and keep the summary information of peers within the cluster. Table 1 is a sample of RIs based on Figure 2 that shows attributes and their probability values provided by the cobweb in each cluster. When a peer joins or leaves the cluster, the RIs entry related to it is created or deleted. Each cluster super-peer sends the maximum probability of each attribute to the neighbor super-peers. When the resource information of a peer changes, that peer sends the updates to its super-peer. Then the super-peer creates the summary again and updates the probabilities of RIs which has higher than the threshold. Next, the maximum probability of attributes is resubmitted to the neighbor super-peers. The super-peers use HRI [41, 43] to organize the summary information of neighbor super-peers in the grid environment. Table 2 shows the aggregated probabilities of the neighbor super-peers at different hops. It is noteworthy that the neighbor super-peers are selected based on the HRI information and not by random or flooding methods. This selection is based on *goodness function* [40] at the predefined hop-count. Consequently, the queries are forwarded to super-peers that are likely to match the query requirements and have nodes in their cluster that own the requested resources.

The goodness of neighbor super-peer (i) with respect to the query requirement is calculated as follows:

$$goodness\ (neighbor_i, \mathrm{R}eq) = \sum_{j=0\cdots h} goodness\,(N_i\,[j], \mathrm{R}eq), \quad (1)$$

where $neighbor_i$ is the i-th neighbor super-peer; h is the horizon of HRI (number of hops); $N_i\,[j]$ refers to the HRI entry for $neighbor_i$ at j-th hop; Req is the requested requirement; and *goodness* points to the probability of a neighbor super-peer in the (i, j) position of the HRI table with respect to Req. The cluster super-peer selects the best neighbor super-peer based on the maximum result of the goodness function and forwards the query to it.

Moreover, in our approach, each super-peer caches the submitted queries and their results in a cache table. The cache table can improve the grid performance by decreasing the delay, bandwidth usage, traffic, and the number of messages, which are sent in the discovery process. Table 3 is

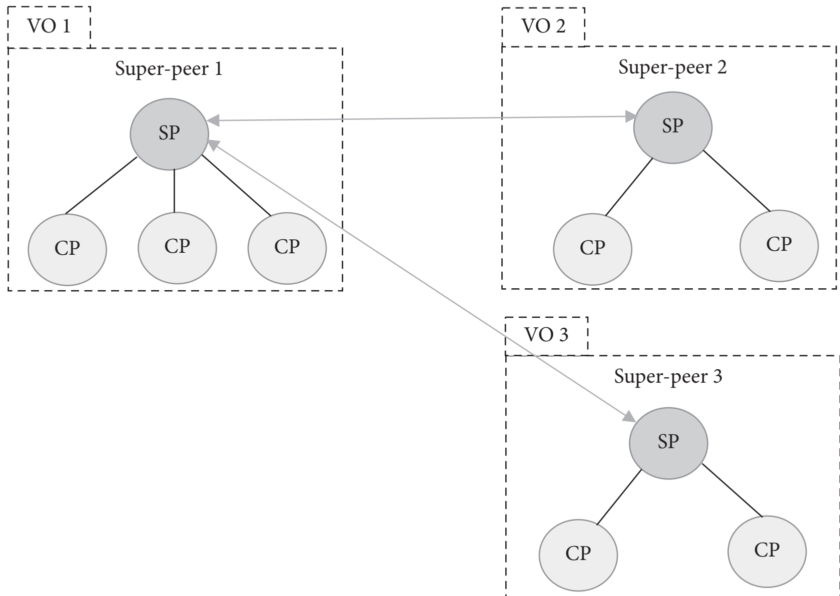


Figure 1: The configuration of a grid.

Table 1: A sample of RIs for resource information.

Attribute / Value	OS name / Windows	Available memory / 2-3 GB	Hard disk / >500 GB	Architecture / x64
P1	—	25%	60%	40%
P2	30%	50%	45%	74%
P3	63%	80%	—	85%
P4	52%	25%	25%	64%

the cache table of a super-peer that contains the requirements of three queries in the form of attributes and the address of nodes that are the owner of requirements (result nodes). Each entry in this table is made in the local domain or is sent by the neighbor super-peers in other clusters.

3.1. Discovery Algorithm. In Figure 2, consider the cluster 1; let P1 need some computing resources to execute its project. The cluster super-peer receives a query and searches the local domain. If the responding peer finds, the super-peer sends the address of the result node back to the requested node and inserts an entry for that query or updates its cache table. Furthermore, it sends the update to the neighbor super-peers. If no results are found, the super-peer looks up the cache table before forwarding the query. If there is a related entry for that request, the super-peer answers the query locally and returns the query result to the requesting node. If not, the super-peer of the cluster 1 calculates the goodness function of its neighbor super-peers (cluster 2 and cluster 3) for that query. We assume that the goodness function of cluster 2 is higher than cluster 3. Therefore, the super-peer of cluster 2 first seeks locally to find the result and then looks up the cache table. If the query result is not in the cache table of cluster 2, the first best neighbor (say cluster 4) is selected and the query is forwarded to it. If the finding process is not successful, the query is bounced back to the parent super-peer (cluster 2). Then, cluster 2 sends the query to the second best neighbor (cluster 5). If the search process is unsuccessful and as a result of cluster 2 has no other neighbor domain except cluster 1, the query is bounced back and cluster 1 forwards query to cluster 3 (the second best neighbor). When the algorithm reaches a response, the cache table is updated and forwarded to the neighbor clusters. Algorithm 1 shows our discovery approach.

3.2. Stability of New Approach. In this section, we compare our model with the existing model in [43]. To this end, consider Figure 3. This is a sample of P2P topology based on the model in [43], in which each peer is responsible for the discovery process. We can see in the mentioned model in [43] that there is a tree structure with 3-hop-count deep and 15 connected peers. For using our model, first, we eliminate all connections and reconfigure this topology based on super-peer as shown in Figure 4. This figure shows that there are three clusters in topology with 1-hop-count deep. The new configuration increases the scalability of discovery as shown in Figure 3.

Consider Figure 3; let a user in the domain of P0 send a query, and the node P0 searches locally to response it. If there is no required resource, P0 must send the query to the best neighbor peer. We consider two scenarios (optimistic and pessimistic statuses).

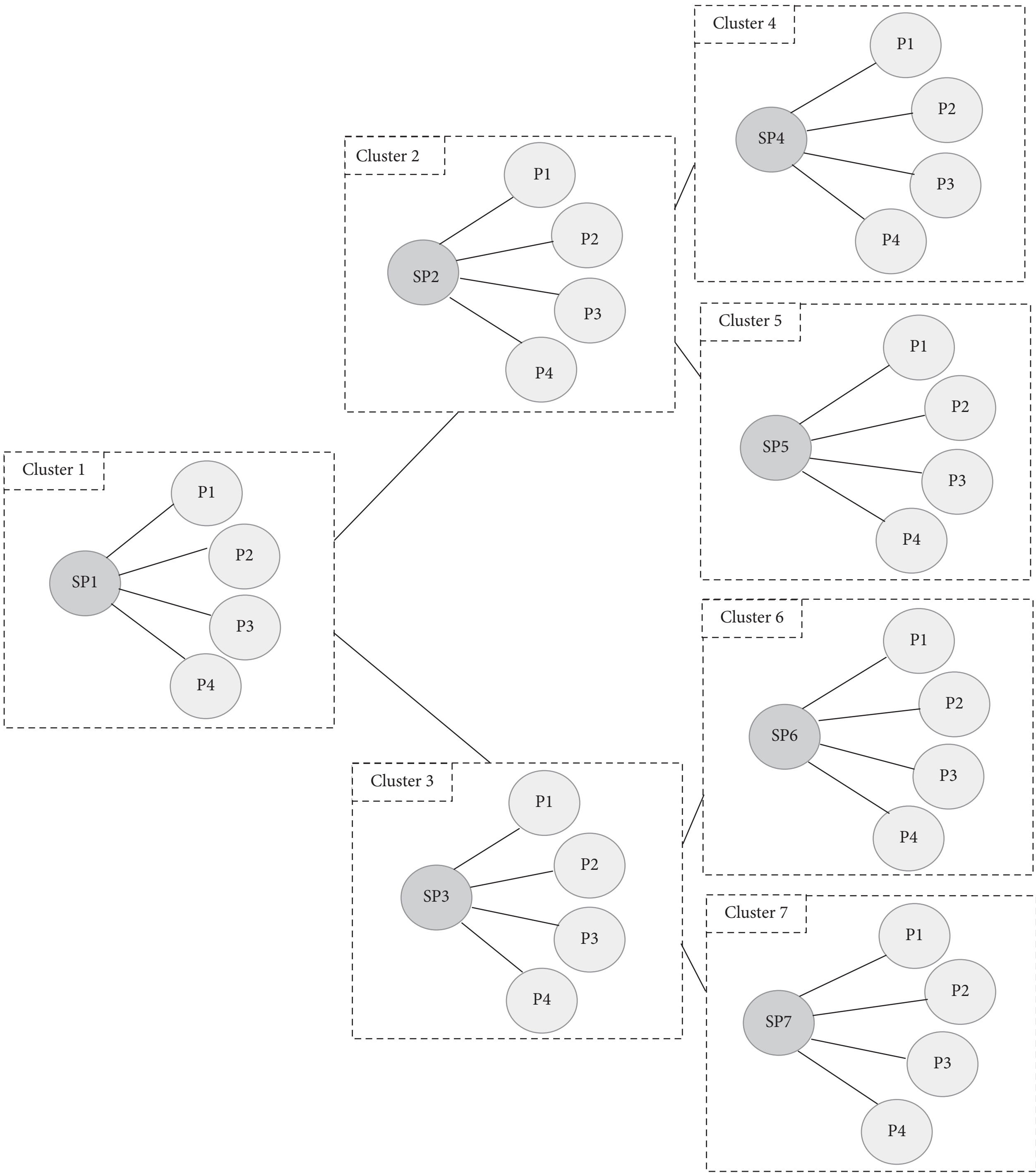


Figure 2: The P2P overlay connection between clusters.

Table 2: HRI for cluster 1 (SP1).

	1hop				2hop			
Attribute								
	OS name	Available memory	Hard disk	Architecture	OS name	Available memory	Hard disk	Architecture
Value								
	Windows	2-3 GB	>500 GB	x64	Windows	2-3 GB	>500 GB	x64
SP2	—	25%	45%	52%	39%	85%	50%	—
SP3	65%	26%	28%	32%	36%	54%	39%	47%

Table 3: A sample of the cache table.

Attribute	OS name	Available memory	Hard disk	Architecture	Query result
Query1	Windows	2-3 GB	—	—	URL
Query2	—	4–6 GB	>500 GB	X64	URL
Query3	Windows	—	—	X86	URL

```
q: new incoming query
ClusterResource: a resource in cluster
Cache: a query result in cache table
BestNeighbor: a neighbor super-peer selected by goodness function
Neighbor: next neighbor super-peer
for incoming q do
ClusterResource = MatchQueryClusterResource (q)
if (ClusterResource = = null) then
    Cache = MatchQueryCache (q)
        if (Cache = = null) then
        BestNeighbor = HRI (q, Neighbor)
            if (BestNeighbor = = null) then
            Receiver = Sender (q)
            else
            Receiver = BestNeighbor
            end if
        ForwardQueryToReceiver (q, Receiver)
        else
        SendResponseToRequester (q)
        end if
else
SendResponseToRequester (q)
Store/UpdateResultToCache (q)
Send update to Neighbor
end if
end for
```

Algorithm 1: Pseudocode of the new discovery algorithm.

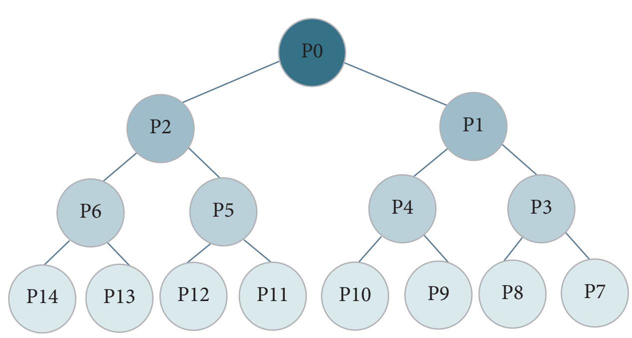


Figure 3: Tree topology of P2P network.

In the optimistic status, the response node is at 1-hop-count level (P1 or P2). We assume that P1 is the best neighbor for P0, then P0 sends the query to P1 and the discovery process successfully finishes. In the pessimistic status, P1 receives the query from P0 and searches locally. If P1 cannot find the requested resources, the query is sent to the best neighbor (say P3). We assume that node P3 is not the owner of the needed resources, and then it sends the query to the best neighbor peer (say P7). If P7 cannot respond to the request, the query is bounced backed. Then P3 sends the query to the second best neighbor (P8). We assume that P8 is not the owner of the needed resource, and then the query is bounced back to P1. P1 sends the query to the second best neighbor (P4), and P4 sends it first to P9, and if there is no response, then it submits it to the P10. Let the requested resources are not found in P10, so the query is bounced back to P0. Similarly, for the left side of P0, in the pessimistic status, we through the following routes

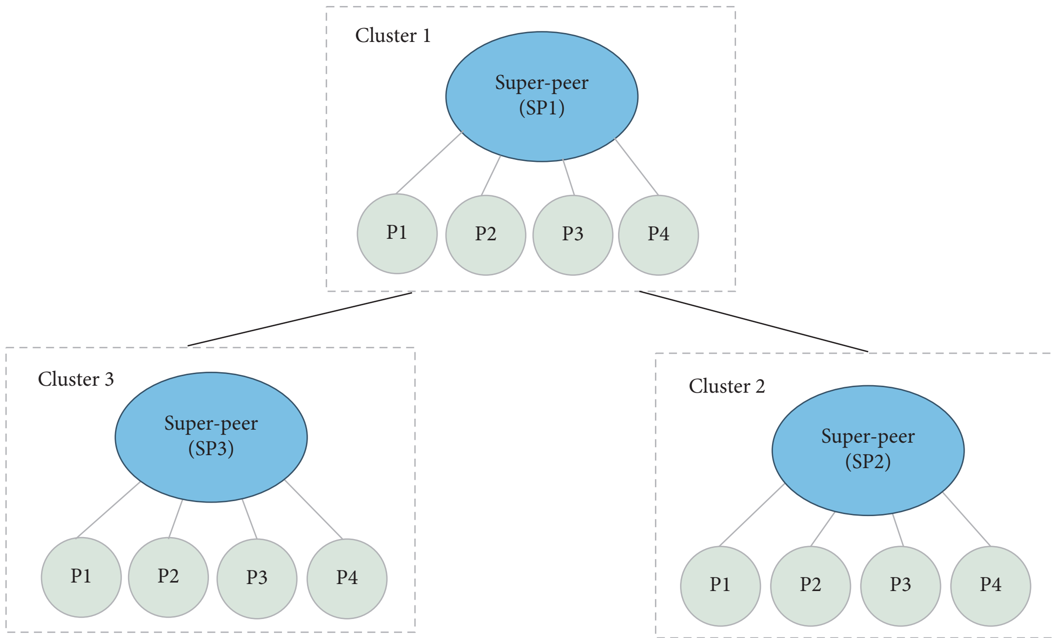


Figure 4: A sample of the clustered diagram of Figure 3.

P2 ⟶ P5 ⟶ P11 ⟶ P12 ⟶ P6 ⟶ P13 ⟶ P14. Therefore, considering P14 has the requested resources, the discovery process of the method [43] can be finished.

Next, consider Figure 4; let us assume that a user in cluster 1 sends a query to the cluster super-peer (SP1). The query process is begun, and the SP1 searches its domain locally. If it finds the requested resource, obviously the process is finished. Except for this situation, we consider two scenarios (optimistic and pessimistic statuses). Let SP1 forwards the query to the best neighbor (say cluster 2); for the optimistic status, the discovery process in cluster 2 is successful and the response node is found. In the pessimistic status, the search in cluster 2 is not successful, so the query is bounced back and forwarded to cluster 3. If this cluster finds the requested recourse, the discovery process is finished successfully. It is well-known that the resource discovery scalability is related to the cost of the discovery process in terms of time and the query messages that are sent to reach the response. Therefore, based on the above demonstration, it is easy to see that the new approach is superior to the existing model in [43] from the scalability point of view. In the next section, we will simulate and compare the above two topologies (Figures 3 and 4) based on the scalability feature.

4. Simulation Results

In this section, we simulate the discovery process in the grid system. First, we simulate the discovery process based on the new method (Algorithm 1) and evaluate the query response time and the submitted messages in comparison with the discovery approach presented by Caminero et al. [43]. Second, we compare the scalability of discovery related to the two topologies presented in Section 3.2. Furthermore, the grid environment has been simulated by using the GridSim toolkit [45].

Suppose that 100 users corporate in our simulation environment; the connection bandwidth is 100 Mbs; the propagation delay is 10 second, and the packet are transmitted by 1500 packet/sec maximally. The resources contain one machine that each of them has four processing elements and "Intel" architecture, and their operating systems are "Linux." The network routers use the RIP protocol, and their scheduling method is FIFO. The first part of our simulation consists of three scenarios that the responding cluster selected randomly: (1) the result node is in the neighbor cluster at 1-hop count, (2) the result node is in the neighbor cluster at 2-hop count, and (3) the result node is in the neighbor cluster at 3-hop count. The second part of the simulation consists of two scenarios: (1) the result node is found in the optimistic status and (2) the result node is found in the pessimistic status.

Figures 5 and 6 compare the query response time and the number of forwarded discover messages in two grid environments: (1) the grid environment simulated based on the new model in which all super-peers use the cache table to response the queries and (2) the grid environment simulated with the model in [43].

The simulation results show that the using of the cache table improves the query response time and decreases the discovery delay. Furthermore, the results illustrate that the number of submitted query messages also decreases. We can see that when the number of hop count increases, the result of the new model is much better than the existing model in [43].

In the second part, the discovery processes are simulated in the optimistic status and the pessimistic status with two mentioned topologies. The simulated results show that using of new topology compared with the existing topology, in both statuses (optimistic status and pessimistic status), decreases the discovery time and improves the scalability of grid discovery (Figure 7).

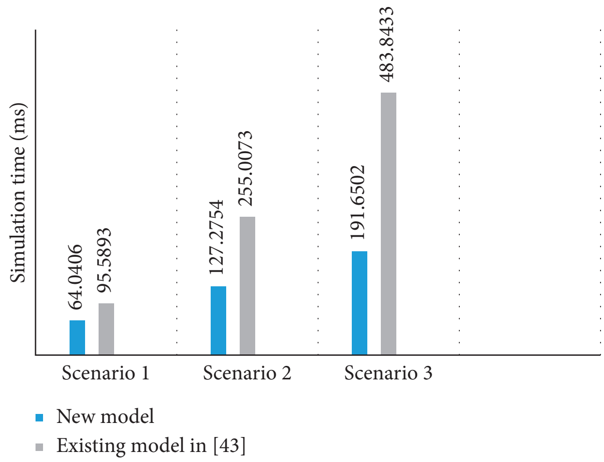


Figure 5: The comparison of the response time for the discovery process.

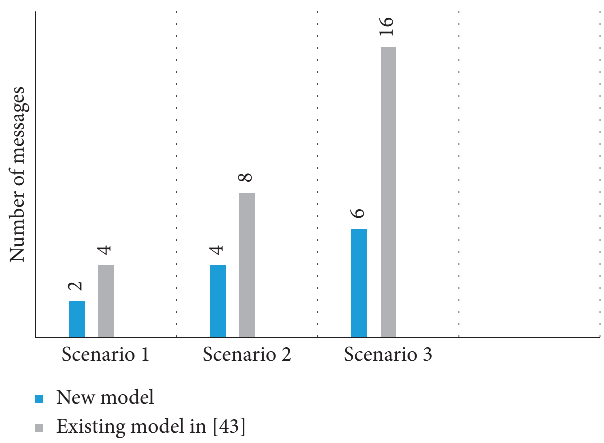


Figure 6: The comparison of the submitted messages in the discovery process.

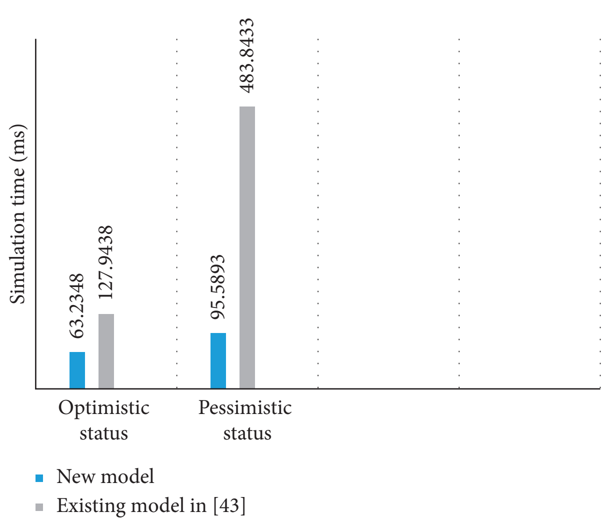


Figure 7: The comparison of scalability of the discovery process under optimistic status and the pessimistic status.

5. Conclusions and Future Works

This paper introduced a new model of grid discovery using P2P technology. For improvement of the scalability of the discovery process, the new model was configured based on the super-peer approach. Furthermore, to reduce the discovery delay, a cache table has been used in the model. The simulated results show that the new model compared with the existing method optimized the query response time and decreased the submitted query messages. For the future, we intend to use some intelligent clustering approaches in this topology and choose the super-peers based on the network attributes such as traffic and bandwidth of the network.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] V. Muppavarapu and S. M. Chung, "Semantic-based access control for grid data resources in open grid services architecture-data access and integration (OGSA-DAI)," in *Proceedings of the 20th IEEE International Conference on Tools with Artificial Intelligence*, pp. 315–322, IEEE, Dayton, OH, USA, November 2008.

[2] L. Mohamad Khanli and S. Kargar, "FRDT: footprint resource discovery tree for grids," *Future Generation Computer Systems*, vol. 27, no. 1, pp. 148–156, 2011.

[3] V. Nagarajan and M. Mohamed, "A decentralized two phase resource discovery model for peer-to-peer grid environments," *International Journal of Advanced Engineering Technology*, vol. 1092, p. 1095, 2016.

[4] J. Zarrin, R. L. Aguiar, and J. P. Barraca, "Resource discovery for distributed computing systems: a comprehensive survey," *Journal of Parallel and Distributed Computing*, vol. 113, pp. 127–166, 2018.

[5] P. Trunfio, D. Talia, H. Papadakis et al., "Peer-to-peer resource discovery in grids: models and systems," *Future Generation Computer Systems*, vol. 23, no. 7, pp. 864–878, 2007.

[6] R. Guharoy, S. Sur, S. Rakshit et al., "A theoretical and detail approach on grid computing a review on grid computing applications," in *Proceedings of the 8th Annual Conference Industrial Automation and Electromechanical Engineering (IEMECON)*, pp. 142–146, Bangkok, Thailand, August 2017.

[7] P. Bellavista, M. Cinque, A. Corradi, L. Foschini, F. Frattini, and J. Povedano-Molina, "GAMESH: a grid architecture for scalable monitoring and enhanced dependable job scheduling," *Future Generation Computer Systems*, vol. 71, pp. 192–201, 2017.

[8] M. Mollamotalebi, G. Sadeghi, and M. Mehdi, "Weighted-HR: an improved hierarchical grid resource discovery," *Journal of Computer & Robotics*, vol. 11, no. 2, pp. 7–16, 2018.

[9] E. Huedo, R. S. Montero, and I. M. Llorente, "A recursive architecture for hierarchical grid resource management,"

Future Generation Computer Systems, vol. 25, no. 4, pp. 401–405, 2009.

[10] Y. Deng, F. Wang, and A. Ciura, "Ant colony optimization inspired resource discovery in P2P grid systems," *The Journal of Supercomputing*, vol. 49, no. 1, pp. 4–21, 2009.

[11] S. Asghari and N. J. Navimipour, "Resource discovery in the peer to peer networks using an inverted ant colony optimization algorithm," *Peer-to-Peer Networking and Applications*, vol. 12, no. 1, pp. 129–142, 2019.

[12] M. Hasanzadeh and M. R. Meybodi, "Grid resource discovery based on distributed learning automata," *Computing*, vol. 96, no. 9, pp. 909–922, 2014.

[13] L. Wang, W. Jie, and J. Chen, *Grid Computing: Infrastructure, Service, and Applications*, CRC Press, Boca Raton, FL, USA, 2018.

[14] B. Gueye, O. Flauzac, C. Rabat, and I. Niang, "A self-adaptive structuring for large-scale P2P Grid environment: design and simulation analysis," *International Journal of Grid and Utility Computing*, vol. 8, no. 3, pp. 254–267, 2017.

[15] C. Mastroianni, D. Talia, and O. Verta, "Designing an information system for grids: comparing hierarchical, decentralized P2P and super-peer models," *Parallel Computing*, vol. 34, no. 10, pp. 593–611, 2018.

[16] N. J. Navimipour and F. S. Milani, "A comprehensive study of the resource discovery techniques in peer-to-peer networks," *Peer-to-Peer Networking and Applications*, vol. 8, no. 3, pp. 474–492, 2015.

[17] A. Ahmadian, M. Zavvar, A. Saeedi, and F. Ramezani, "Resource discovery in non-structured peer to peer grid systems using the shuffled frog leaping algorithm," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 10, no. 4, pp. 9–14, 2018.

[18] N. J. Navimipour, A. M. Rahmani, A. H. Navin, and M. Hosseinzadeh, "Resource discovery mechanisms in grid systems: a survey," *Journal of Network and Computer Applications*, vol. 41, pp. 389–410, 2014.

[19] P. Felber, P. Kropf, E. Schiller, and S. Serbu, "Survey on load balancing in peer-to-peer distributed hash tables," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 473–492, 2014.

[20] Y.-H. Chen, E. J.-L. Lu, Y.-T. Chang, and S.-Y. Huang, "RDFchord: a hybrid PDMS for P2P systems," *Computer Standards & Interfaces*, vol. 43, pp. 53–67, 2016.

[21] J. Furness, M. Kolberg, and M. Fayed, "An evaluation of Chord and Pastry models in OverSim," in *Proceedings of the Modelling Symposium (EMS)*, pp. 509–513, Manchester, UK, November 2013.

[22] B. Yin, X. Lu, J. Huang, and Y. Kang, "Analysis of topology dynamics for unstructured P2P networks," *Computer Communications*, vol. 80, pp. 72–81, 2016.

[23] A. Hameurlain, D. Cokuslu, and K. Erciyes, "Resource discovery in grid systems: a survey," *International Journal of Metadata, Semantics and Ontologies*, vol. 5, no. 3, pp. 251–263, 2010.

[24] B. B. Yang and H. Garcia-Molina, "Designing a super-peer network," in *Proceedings of the 19th International Conference on Data Engineering (Cat. No. 03CH37405)*, pp. 49–60, IEEE, Bangalore, India, March 2003.

[25] A. M. Saghiri and M. R. Meybodi, "An adaptive super-peer selection algorithm considering peers capacity utilizing asynchronous dynamic cellular learning automata," *Applied Intelligence*, vol. 48, no. 2, pp. 271–299, 2018.

[26] S. Hashemi Najafi, J. Pourqasem, and S. A. Edalatpanah, "The use of super node to process query in peer-to-peer networks," *Science and Education*, vol. 1, no. 1, pp. 28–32, 2015.

[27] J. Pourqasem and S. A. Edalatpanah, "Verification of super-peer model for query processing in peer-to-peer networks," in *Innovative Research and Applications in Next-Generation High Performance Computing*, pp. 306–332, IGI Global, PL, USA, 2016.

[28] Y.-H. Tan, k. Lü, and Y.-P. Lin, "Organisation and management of shared documents in super-peer networks based semantic hierarchical cluster trees," *Peer-to-peer Networking and Applications*, vol. 5, no. 3, pp. 292–308, 2012.

[29] V. Kalogeraki, D. Gunopulos, and D. Zeinalipour-Yazti, "A local search mechanism for peer-to-peer networks," in *Proceedings of the Eleventh International Conference on Information and Knowledge Management*, pp. 300–307, ACM, McLean, VI, USA, November 2002.

[30] A. A. Jamal and W. J. Teahan, "Alpha multipliers breadth-first search technique for resource discovery in unstructured peer-to-peer networks," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 7, no. 4, 2017.

[31] M. Ghorbani, M. R. Meybodi, and A. M. Saghiri, "A novel self-adaptive search algorithm for unstructured peer-to-peer networks utilizing learning automata," in *Proceedings of the 3rd Joint Conference of AI & Robotics and 5th RoboCup Iran Open International Symposium*, pp. 1–6, IEEE, Tehran, Iran, April 2013.

[32] M. Ghorbani, M. R. Meybodi, and A. M. Saghiri, "An adaptive topology management algorithm in P2P networks based on learning automata," in *Proceedings of the 7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, pp. 1–4, IEEE, Tehran, Iran, January 2019.

[33] A. Löser, F. Naumann, W. Siberski et al., "Semantic overlay clusters within super-peer networks," in *Proceedings of the International Workshop on Databases, Information Systems, and Peer-t-Peer Computing*, pp. 33–47, Springer, Berlin, Heidelberg, September 2003.

[34] A. Kutuzov and E. Kuzmenko, "Neural embedding language models in semantic clustering of web search results," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pp. 3044–3048, Portorož, Slovenia, May 2016.

[35] Y. He, C. S. Yang, L. C. Yu et al., "Sentiment classification of short texts based on semantic clustering," in *Proceedings of the 2015 International Conference on Orange Technologies (ICOT)*, pp. 54–57, IEEE, Hong Kong, China, December 2015.

[36] R.-S. Chang and M.-S. Hu, "A resource discovery tree using bitmap for grids," *Future Generation Computer Systems*, vol. 26, no. 1, pp. 29–37, 2010.

[37] D. H. Fisher, "Knowledge acquisition via incremental conceptual clustering," *Machine Learning*, vol. 2, no. 2, pp. 139–172, 1987.

[38] A. C. Caminero, E. Huedo, O. Rana et al., "Summary creation for information discovery in distributed systems," in *Proceedings of the 19th International Euromicro Conference on Parallel, Distributed and Network-Based Processing (PDP)*, pp. 167–171, Ayia Napa, Cyprus, February 2011.

[39] R. Brunner, A. C. Caminero, O. F. Rana, F. Freitag, and L. Navarro, "Network-aware summarisation for resource discovery in P2P-content networks," *Future Generation Computer Systems*, vol. 28, no. 3, pp. 563–572, 2012.

[40] A. Crespo and H. Garcia-Molina, "Routing indices for peer-to-peer systems," in *Proceedings of the 22nd International*

Conference on Distributed Computing Systems, pp. 23–32, Vienna, Austria, July 2002.

[41] D. Puppin, S. Moncelli, R. Baraglia et al., "A grid information service based on peer-to-peer," in *Proceedings of the 11th International European Conference on Parallel Processing*, pp. 454–464, Springer, Lisbon, Portugal, August 2005.

[42] A. Caminero, O. Rana, B. Caminero, and C. Carrión, "Network-aware heuristics for inter-domain meta-scheduling in grids," *Journal of Computer and System Sciences*, vol. 77, no. 2, pp. 262–281, 2011.

[43] A. C. Caminero, A. Robles-Gómez, S. Ros, R. Hernández, and L. Tobarra, "P2P-based resource discovery in dynamic grids allowing multi-attribute and range queries," *Parallel Computing*, vol. 39, no. 10, pp. 615–637, 2013.

[44] M. Marzolla, M. Mordacchini, and S. Orlando, "Resource discovery in a dynamic grid environment," in *Proceedings of the Sixteen International Workshop on Database and Expert Systems Applications*, pp. 356–360, Copenhagen, Denmark, July 2005.

[45] A. Sulistio, U. Cibej, S. Venugopal, B. Robic, and R. Buyya, "A toolkit for modelling and simulating data grids: an extension to GridSim," *Concurrency and Computation: Practice and Experience*, vol. 20, no. 13, pp. 1591–1609, 2008.

[46] C. Mastroianni, D. Talia, and O. Verta, "A super-peer model for building resource discovery services in grids: design and simulation analysis," in *Proceedings of the European Grid Conference on Advances in Grid Computing*, pp. 132–143, Springer, Amsterdam, Netherlands, February 2005.

Hindawi
Wireless Communications and Mobile Computing
Volume 2019, Article ID 2430656, 12 pages
https://doi.org/10.1155/2019/2430656

Review Article

A Survey on Radio Resource Allocation for V2X Communication

Ahlem Masmoudi, Kais Mnif, and Faouzi Zarai

NTS'COM Research Unit, National School of Electronics and Telecommunications, University of Sfax, Sfax, Tunisia

Correspondence should be addressed to Ahlem Masmoudi; ahlem.massmoudi@gmail.com

Received 1 May 2019; Revised 22 September 2019; Accepted 27 September 2019; Published 24 October 2019

Guest Editor: Donghyun Kim

Thanks to the deployment of new techniques to support high data rate, high reliability, and QoS provision, Long-Term Evolution (LTE) can be applied for diverse applications. Vehicle-to-everything (V2X) is one of the evolving applications for LTE technology to improve traffic safety, to minimize congestion, and to ensure comfortable driving which requires stringent reliability and latency requirements. As mentioned in the 3rd Generation Partnership Project (3GPP), LTE-based Device-to-Device (D2D) communication is an enabler for V2X services to meet these requirements. Therefore, radio resource management (RRM) is important to efficiently allocate resources to V2X communications. In this paper, we present the V2X communications, their requirements and services, the V2X-based LTE-D2D communication modes, and the existing resource allocation algorithms for V2X communications. Moreover, we classify the existing resource allocation algorithms proposed in the literature and we compare them according to selected criteria.

1. Introduction

Internet of Things (IoT) include billions of intelligent objects which are regarded as a part of the future Internet [1, 2]. IoT is a growing number of things (i.e., physical objects) that are wirelessly connected to the Internet via smart sensors. IoT enables these things to interact and exchange data in an efficient way without human intervention. The 3rd Generation Partnership Project (3GPP) Long-Term Evolution (LTE) is one of the main strengths of the IoT which seeks to cover all the IoT applications. Vehicle-to-Everything (V2X) communication is an important main area of IoT. Recently, 3GPP has made Release 14 for LTE-based V2X communication (cellular V2X) in order to improve the safety, the efficiency, and the comfort of intelligent transportation systems (ITS). The term V2X refers to the vehicle to vehicle (V2V), vehicle to infrastructure (V2I), vehicle to pedestrian (V2P), and vehicle to Network (V2N) communications [3] which require a high reliability and low latency.

On the other hand, in Release 16, 3GPP is working near the growth of New Radio (NR) V2X, constructing over 5G NR that was standardized in Release 15 of 3GPP. NR V2X should support advanced V2X services that necessitate much more stringent QoS assurances compared to services that may be supported by C-V2X [4, 5]. For both evolutionary, the 5G NR and C-V2X are designed to progress lower end-to-end latency and reliability services and to support services that require high throughput. But, their design methodologies meaningfully vary. 3GPP does not carry out a similar constraint on NR V2X. V-UEs equipped with NR V2X can interconnect with C-V2X devices. Nevertheless, this will be reached through a dual-radio access system—one radio for NR V2X and another for C-V2X [6].

The proximity service (ProSe) known as a Device-to-Device (D2D) communication, defined in Release 12, refers to a direct communication between two or more devices in proximity to each other rather than travelling through the eNodeB [7]. Therefore, several kinds of advantages can be offered due to the proximity, reuse, and hop gain. D2D communications were initially proposed to improve network performance (i.e., enhancing spectrum utilization, improving UEs throughput, increasing cellular capacity, and extending UEs battery lifetime) in cellular technologies [8].

Since V2X technologies have stringent reliability and latency requirements [9], it was declared in Release 14 that the D2D communication can be applied in vehicular technologies to support V2V communications. As V2X is based on the D2D communication, resources are allocated in the V-UEs either in the overlay mode or in the underlay mode. For that reason, radio resource management (RRM) plays an important role in V2X system performances.

In this paper, we provide a wide review of available literature for resource allocation algorithms in V2X services based on D2D communication. Moreover, we provide a comprehensive comparison and classification of V2X resource allocation algorithms in terms of numerous aspects. The paper is structured as follows: The V2X services are described in Section 2. The V2X communication modes and the LTE-V-based D2D communication are introduced, respectively, in Section 3 and Section 4. A review and classification of the existing resource allocation algorithms for V2X communication are addressed in Section 5. We discuss and conclude this paper in Section 6 and Section 7.

2. V2X Requirements and Services

3GPP defines 27 use cases for V2X services in Release 14 that can be divided into safety and nonsafety V2X services [10]. Therefore, the transmission of V2X messages should be classified according to the message type (e.g., safety vs. nonsafety). The safety services-related use cases aim to avoid automobile accidents and to protect the property and life, which require high reliability and short delay (e.g., vehicle platooning and automated driving). The nonsafety use cases aim to improve the driving experience to be more comfortable and efficient (e.g., mobile high data rate entertainment and traffic management). Therefore, the non-safety and the safety V2X services have different latency and packet size requirements [11].

Five categories of service requirements of this use cases are defined as follows [9, 10]:

(i) Speed: the maximum absolute velocity 160 km/h and the maximum relative velocity 280 km/h shall be supported. The maximum relative speed of 500 km/h is supported for the possible scenario without speed limitation.

(ii) Communication range: the effective distance is larger than the distance calculated as the ample response time (e.g., 4 seconds) for the drivers to avoid collision according maximum relative speed.

(iii) Latency/reliability: the maximum end-to-end latency between two UEs supporting V2V/V2P/V2I applications shall be 100 ms. The minimum radio layer reliability should be supported without retransmissions of application-layer messages in the effective distance and the restricted latency.

(iv) Message size: the variable length of a periodically message between two UEs supporting V2X applications is about 50–300 bytes. The message size of event-triggered messages can be up to 1200 bytes.

(v) Message generation period: the minimum message generation period can be 100 ms.

2.1. Safety-Critical V2X Services. The safety-critical V2X services aim to reduce the potential accident and the risk to pedestrians and to deteriorate the possibility of life dead for V-UEs. The information shared between pedestrians, Road Side Unit (RSU), and vehicles (i.e., vehicle, speed, position, and distance) is usually short broadcast messages which impose stringent requirements on latency, packet loss, and reliability. Two types of safety V2X transmissions messages have been standardized, the periodic transmission and event-triggered transmission. The periodic transmissions among V2V-UEs (e.g., the cooperative awareness messages (CAMs) are short broadcast messages periodically transmitting aware information of position, presence, speed, and directions among vehicles and their neighbors. The event-triggered transmission among V2X-UEs (i.e., decentralized environmental notification messages (DENMs)) is short broadcast messages to alert road users about road status. Both periodic transmission and event-triggered transmission can be operated over the LTE-Uu and the PC5 interfaces [12].

2.2. Nonsafety V2X Services. The nonsafety V2X services essentially focus on management application, control congestion, traffic efficiency, and entertainment; these services enable more comfortable and efficient driving experience. Nonsafety V2X applications aim to improve traffic coordination, assistance, and vehicle traffic flow. In addition, they also provide entertainment video, maps, and updated local information. Therefore, a large amount of V-UEs sensed data sent to neighbors or to infrastructure. These services have no strict requirements on reliability and latency.

2.3. V2X Requirements

2.3.1. Reliability Requirement. The reliability requirement defined in different works is generally interpreted from either the outage probability (OP) or the packet reception ratio (PRR).

The PRR is the ratio of successful reception among the total number of the vehicle transmitter neighbors, which are calculated as follows:

$$\mathrm{PRR}_u(t) = \frac{1}{N_u(t)} \sum_{j=1}^{N_u(t)} 1\left\{\gamma_j^{(u)}(t) \geq \gamma_{\mathrm{th}}\right\},$$
$$\gamma_j^{(u)}(t) = \sum_{k=1}^{N_{RB}} \frac{s_k^u(t) P_u(t) \left|H_{uj}^k(t)\right|^2}{\sigma^2 + \sum_{m=1}^{N_L} s_m^u(t)\, P_j(t) \left|H_m^k(t)\right|^2}, \tag{1}$$

where $N_u(t)$ denotes the number of the vehicles neighbors of the u^{th} vehicle at time t, $\gamma_j^{(u)}(t)$ is the receiving signal-to-interference-plus-noise ratio (SINR) of the j^{th} vehicle among $N_u(t)$, γ_{th} is the SINR threshold, $P_u(t)$ is the transmit power of u^{th} vehicle, σ^2 is the power of additive white Gaussian noise, N_L is the number of interfering vehicles, $s_k^u(t)$ is the RBs allocation at time t, and $s_k^u(t) = \begin{cases} 1, & \text{if } k^{\text{th}}\ RB \text{ is allocated to user } u \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases}$

The OP is the most utilized which is the probability that N_u error-free bits cannot be sent by any coding pattern which is computed as follows:

$$p_u^{\text{out}} \triangleq \Pr\left\{\sum_{i=1}^{E_u^{\text{all}}} \rho \log\left(1+\gamma_i\right) < N_u\right\}, \tag{2}$$

where E_u^{all} is the RBs number that are allocated to the u^{th} V-UE, ρ is the number of complex symbols in the RB, and γ_i is the SINR of the i^{th} RB.

2.3.2. Latency Requirement. The latency requirement in some studies is interpreted by the latency constraint that allows V-UEs to allocate their required RBs before the maximum tolerable latency according to scheduling time unit number. Most of the studies designate that C-V2X can support safety services reliably that required an end-to-end latency of nearby 100 milliseconds (msec).

3. V2X Communication Modes

To support both safety/nonsafety V2X applications, the 3GPP standard provides two radio interfaces for LTE-Vehicle (LTE-V) [13, 14]: the Uu cellular interface that supports V2I communications (e.g., enhanced Multimedia Broadcast Multicast Service (eMBMS) and the PC5 interface (direct LTE sidelink) that supports V2V communications as shown in Figure 1.

3.1. LTE-Cellular Communication Mode Using Uu Interface. LTE-Cellular communication refers to the communication mode between V-UEs and the eNB, which is suitable for nondelay tolerant use cases (e.g., mobility services and situational awareness) using the LTE technologies. The eNB receives unicast messages from V-UEs and rebroadcasts them for all V-UEs receivers in the pertinent area using eMBMS. Based on coordination between multiple cells, MBMS aims to reduce the latency and to improve the performance of the cell edge.

3.2. LTE-D2D Communication Mode Using PC5 Interface. LTE-D2D communication refers to the communication mode between two V-UEs in proximity to each other, bypassing the eNB. This mode is appropriate for V2V safety services that require low latency delay (e.g., advanced driver-assistance systems (ADAS). This communication mode is based on Release 12 proximity services (ProSe) which exploit direct communication between neighboring devices. Two transmission modes are introduced in Release 12 (mode 1 and mode 2) for LTE sidelink (or D2D communication) public safety. These modes are designed in order to prolong the devices' battery lifetime at the cost of increasing the latency. Therefore, mode 1 and mode 2 are not convenient for V2X applications since connected vehicles require low

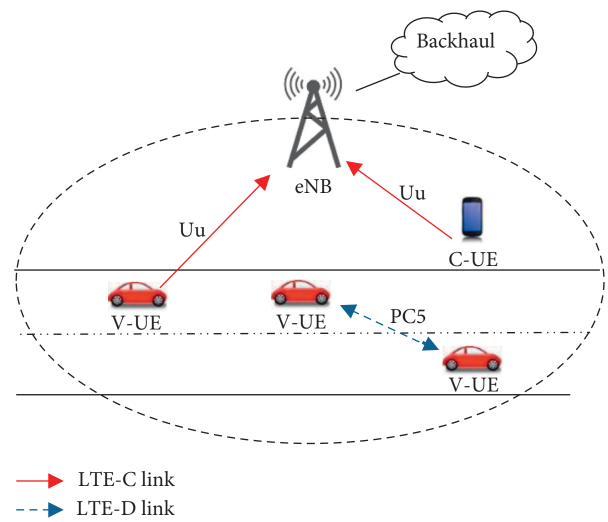


Figure 1: Radio interfaces for V2X communications.

latent and highly reliable V2X communications. Recently, two new sidelink communication modes are introduced in Release 14, which are typically intended for V2V communications, the mode 3 and the mode 4. In mode 3, the radio resources used by the direct V2V communications and the interference management are managed and assisted by the cellular infrastructure (e.g., eNB). However, vehicles autonomously select and assign the radio resources in mode 4 for their direct V2V communications without infrastructure assistance (i.e., this mode can be operated out of cellular coverage).

In this paper, we investigate the radio resource allocation in mode 3 for the V2X services based on D2D communication in LTE cellular network.

4. LTE-V-Based D2D Communication

Since there is a similarity between the localized nature of V2V and D2D communications, it is mentioned in Release 14 that the D2D communication can be useful in vehicular networks. The V2V/D2D communication can allocate the cellular resources in orthogonal or nonorthogonal way.

4.1. V2V-Based D2D Communication Modes. There are two V2V communication modes: the overlay mode (which is the orthogonal resource allocation) and the underlay mode (which is the nonorthogonal one) as shown in Figure 2.

In the overlay mode, specific radio resources from cellular resources are dedicated to the D2D/V2V communication. Thus, the C-UEs cannot achieve the full capacity of the eNB; consequently, this mode decreases the spectrum utilization. The advantage of the overlay mode is that the interference between C-UEs and V2V-UEs does not need to be managed.

In the underlay mode, the eNB allows V2V-UEs and C-UEs to share the same radio resources, which can achieve a greatest spectrum efficiency. However, the eNB needs to manage the strong interference among V2V-UEs communications and the C-UEs communications.

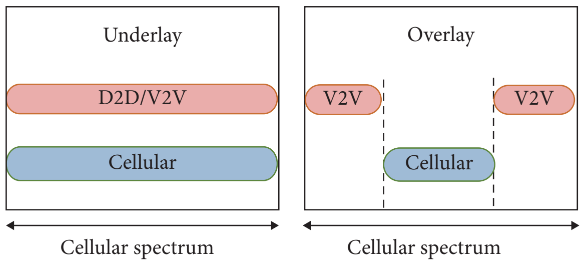


Figure 2: Underlay vs overlay modes.

Both of the overlay and the underlay modes can utilize either the Downlink (DL) or the Uplink (UL) subframe. The set of resources that will be allocated to V2V communication are chosen from the UL subframe owing to their minor peak-to-average power ratio (PAPR) and because it is less utilized than the DL subframe.

4.2. Nonorthogonality among V-UEs. Orthogonal multiple access (OMA) techniques are established by all of the current cellular mobile networks (i.e., LTE, LTE-A) such as orthogonal frequency division multiple access (OFDMA) and time-division multiple access (TDMA) where a single user can be served in each orthogonal resource block (RB) in OFDMA subcarriers. Nevertheless, these techniques cannot meet the great demands of the future radio access networks [15]. Therefore, the nonorthogonal multiple access (NOMA) is suggested as a candidate for 5G cellular mobile systems which becomes a significant principle for the 5G radio access techniques [16]. NOMA can be deployed in the future and the existing mobile systems due to its compatibility with other technologies. Unlike OMA and without needing any modifications to the LTE RBs, NOMA can serve multiple users on the same RB (in the same frequency and time domain) as shown in Figure 3. Therefore, the NOMA system throughput can be expressively greater than that of the OMA. The NOMA technique is also applied to cellular vehicular network in order to achieve low latency and high reliability and to reduce resource collision [17].

5. RRM for V2X Services over LTE

RRM in LTE includes a diversity of techniques and procedures, such as radio resource allocation and packet scheduling. Nowadays, resource allocation in LTE-V is one of the most discussed topics. The V-UEs can communicate using either direct link or cellular link. They can also operate either in the underlay or in the overlay modes. The majority of existing related works suggest using the cellular spectrum for both V-UEs and C-UEs. Since in the overlay mode, a set of RBs are dedicated to V-UEs using the direct link, the CUEs cannot achieve the full spectrum capacity. Therefore, studies based on the overlay mode aim to avoid the waste of resources. On the other hand, C-UEs and V-UEs share the same RBs in the underlay mode, which reach a best spectrum efficiency but the huge interference could occur among C-UEs and V-UEs. Several V2X resource allocation algorithms have been investigated in the literature. The majority

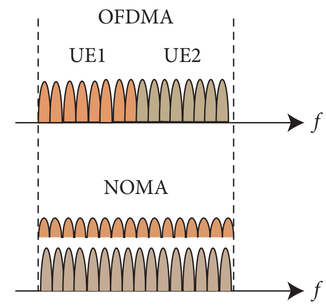


Figure 3: Difference between OMA and NOMA.

of these algorithms underlaying C-UEs need to deal interferences among C-UEs and V-UEs. In this section, we demonstrate an overview classification on V2X resource allocation algorithm existing in the literature as shown in Figure 4.

Then, we provide a comprehensive classification and comparison of the V2X resource allocation algorithms existing in the literature in terms of numerous aspects (e.g., communication mode and allocation process). Moreover, we study the existing effort in LTE-V2X resource allocation algorithms in-band communication over the underlay mode and the overlay mode where the allocation of RBs is done by the eNB (mode 3).

5.1. Underlaying Resource Allocation in V2X Services. Early-proposed V2X radio resource allocation algorithms in LTE-V environments suggest reusing cellular spectrum for V2X communications. Nowadays, the majority of existing resource allocation work is dedicated to the underlaying V2X communications in cellular networks. These works typically study the interference problems between V-UEs and C-UEs due to RBs shared between them. The existing resources allocation algorithms designed in the underlay mode can be categorized according to several criteria. We classify these algorithms according to RBs sharing process, the RBs sharing based on user pairing, user clustering/grouping, and user geographic location.

5.1.1. RBs Sharing-Based User Pairing. Most of the underlaying resource allocation algorithms allow UEs to share the same RBs based on user pairing where the eNB finds the optimal combination between at least two UEs to share the same RB. The main objective is to allow at least one V-UE using direct link to share the same RB with only one C-UE or with only one V-UE using cellular link. The comparison of these algorithms is summarized in Table 1.

In [18], Zhang et al. proposed a novel resource sharing of the underlaying communication mode for vehicular networks which aims to increase throughput of the vehicular network through efficient interference management protocols. Different V2I and V2V communication links are allowed to access the same resources for their data transmission. The resource-sharing problem was formulated as a resource allocation optimization problem,

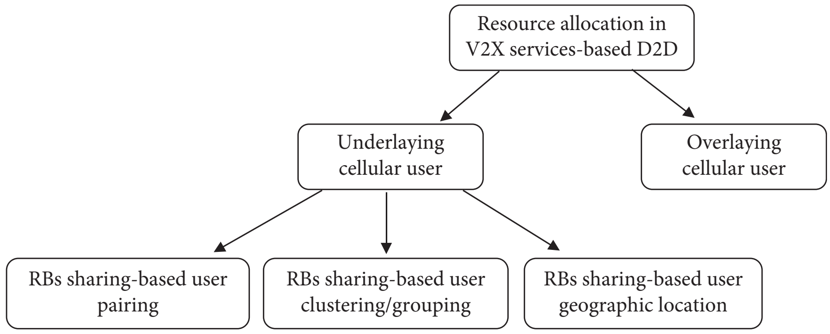


Figure 4: Resources allocation in V2X services classification.

Table 1: Comparison of the existing underlaying RBs sharing-based user pairing V2X services.

Ref.	Unicast/ broadcast	Objectives	Scenarios	User types	Allocation constraints	Power control/ allocation	Allocation process	Methods/theory	RB sharing
[18]	Broadcast	Increase throughput	Road	V2V-UEs V2I-UEs	PF	—	Orthogonal RBs are allocated to each user type	Graph theory	1 V2V-UE 1 V2I-UE
[19]	Unicast	Maximize sum rate fairness; minimize latency reliability	Urban	C-UEs V-UEs	OP SINR Latency	PC	Orthogonal RBs are allocated to each user type	An interior point method Hungarian algorithm	1 C-UE 1 V-UE
[20]	Unicast	Maximize sum rate fairness; minimize latency reliability	Urban	C-UEs V-UEs	OP SINR Latency	PA	Orthogonal RBs are allocated to each user type	Karush–Kuhn–Tucker theory Dual decomposition method Hungarian algorithm	1 C-UE 1 V-UE
[21]	Unicast	Maximize sum rate fairness; minimize latency reliability	Urban	C-UEs V-UEs	OP SINR Latency	PA	Orthogonal RBs are allocated to C-UEs; nonorthogonal RBs are allocated to V-UEs	Perron–Frobenius theory Interior point method	1 C-UE N V-UEs
[22]	Unicast	Maximize sum rate fairness; minimize latency reliability	Urban	C-UEs V-UEs	OP SINR Latency	PA	Orthogonal RBs are allocated to C-UEs; nonorthogonal RBs are allocated to V-UEs	Matching theory Interior point method	1 C-UE N V-UEs
[23]	Broadcast	Maximize the number of concurrent V2V transmissions	Urban freeway	V-UEs	SINR	—	Nonorthogonal RBs are allocated to V-UEs	Perron–Frobenius theory	N V-UEs
[24]	Unicast	Maximize throughput; minimize latency	Urban	C-UEs Safety and nonsafety V-UEs	SINR Latency	—	Orthogonal RBs are allocated to each user type	Hypergraph matching theory	1 C-UE 1 safety V-UE 1 nonsafety V-UE
[25]	Unicast	Maximize throughput reliability	Multilane freeway	V2V-UEs V2I-UEs	OP	PA	Orthogonal RBs are allocated to each user type	Hungarian method	1 V2V-UE 1 V2I-UE

Table 1: Continued.

Ref.	Unicast/ broadcast	Objectives	Scenarios	User types	Allocation constraints	Power control/ allocation	Allocation process	Methods/theory	RB sharing
[26]	Unicast	Reliability; maximize the ergodic capacity	Multi-lane freeway	V2V-UEs V2I-UEs	OP	PA	Orthogonal RBs are allocated to each user type	Hungarian method	1 V2V-UE 1 V2I-UE
[27]	Broadcast	Maximize the C-UEs information rate; guarantee reliability/ latency requirement of V-UEs	Urban	C-UEs V-UEs	Latency OP SINR	PA	Orthogonal RBs are allocated to each user type	Lagrange dual decomposition method Binary search method Subgradient iteration method	1 C-UE 1 V-UE
[28]	Unicast	Maximize throughput (C-UEs/nonsafety V-UEs); guarantee QoS demand (W-UEs/safety V-UEs)	Freeway	W-UEs C-UEs Safety and nonsafety V-UEs	SINR	—	Orthogonal RBs are allocated to each user type	Kuhn–Munkres algorithm Gale–Shapley algorithm	1 C-UE 1 nonsafety V-UE
[29]	Unicast	Maximize the total throughput for C-UEs and V-UEs	Freeway	W-UEs C-UEs Safety and nonsafety V-UEs	SINR	—	Orthogonal RBs are allocated to each user type	An interior point method Kuhn–Munkres method	1 C-UE 1 nonsafety V-UE
[30]	Unicast	Maximize V2I-UEs sum rate; guarantee V2V-UEs reliability requirement	Freeway	V2I-UEs V2V-UEs	SINR Buffer size Packet delay	—	Orthogonal RBs are allocated to each user type	—	1 V2I-UE 1 V2V-UE
[31]	Unicast	Maximize V2I-UEs sum rate; guarantee V2V-UEs reliability requirement	Freeway	V2I-UEs V2V-UEs	SINR Buffer size Packet delay	PC	Orthogonal RBs are allocated to each user type	—	1 V2I-UE 1 V2V-UE
[32]	Unicast	Maximize C-UEs sum rate; guarantee V-UEs reliability requirement	Freeway	GBR C-UEs NGBR C-UEs V-UEs	PDR SINR Buffer size Packet delay	—	Orthogonal RBs are allocated to each user type	—	1 C-UE 1 V-UE
[33]	Unicast	Maximize sum rate and respect constraint delay of C-UEs; guarantee V-UEs reliability and latency	Freeway	C-UEs Safety V-UE Nonsafety V-UE	PDOR SINR Delay	—	Orthogonal RBs are allocated to each user type	—	1 C-UE 1 safety V-UE 1 nonsafety V-UE

taking into account the interference between diverse communication links. The resource-sharing problem has been solved using the graph theory that develops two interference graph-based resource-sharing schemes: the interference-aware and the interference-classified graph-based resource-sharing scheme.

The authors in [19–22] proposed a radio resources and power allocation algorithm for safety-critical vehicular communications. This algorithm aims to maximize the C-UE sum rate with proportional bandwidth fairness under the constraint of satisfying the V-UEs' requirements on latency and reliability. First, they mathematically formulate the requirement of V2X communication into optimization constraints that compute with only slowly varying CSI. Then, they propose a RRM to decide which users can share the same RB.

In [19], a two-step resources allocation algorithm was investigated. Firstly, resources are allocated to both C-UEs and V-UEs in an optimal way by allowing equal power allocation. To communicate with the eNB and among V-UEs, orthogonal RBs are used by C-UEs and V-UEs, respectively. So, both V-UE and C-UE can use the same RB that will produce intracell interference among each other. By transforming the problem of RB allocation into a maximum weight matching (MWM) problem for bipartite graphs, the interference will be resolved. Secondly, the transmit power is optimally adjusted for each C-UE and V-UE. In [20], the C-UEs sum rate maximized as much as possible and the V-UEs transmit power is minimized.

In [21], a heuristic radio resources allocation scheme was designed considering the fast fading effects in order to support a significantly higher number of V-UEs by allowing nonorthogonality among V-UEs. Resource sharing can take place not only between vehicles and cellular users but also among different vehicles on condition that the SINR constraints of V-UEs and the best rate of C-UEs are satisfied. To do this, they will first introduce the underlaying mode using the Perron–Frobenius theory to design an RB sharing metric. Second, they propose a heuristic RB sharing scheme that associates in a sequential fashion each V-UE with a C-UE. Finally, the power is allocated based on the RB allocation. Then, they extend in [22] to improve more strict latency and reliability requirement designed based on matching theory in order to obtain higher performance especially for high load scenario.

In [23], a radio resource allocation algorithm based on RBs sharing was proposed in order to maximize the concurrent V2V transmissions number unlike other authors who maximize the sum rate, where one RB can be shared by multiple V-UEs by allowing nonorthogonal access among them. Firstly, the reliability requirement is transformed into constraint of spectral radios matrix to limit the interference among V-UEs. Secondly, they mathematically formulate the RB sharing problem that maximizing the number of V-UEs, which is equivalent to minimize the occupied RBs number. To better improve the spectrum efficiency, they use the spectral radius estimation theory.

In [24], Wei et al. proposed a 3D-matching-based radio resources allocation algorithm for V2X communication. The objective is to maximize the total throughput of nonsafety V-UEs on condition of satisfying the SINR requirements on C-UEs and on safety V-UEs. They proposed three stages to allow these UEs to share the same RB on condition that one RB cannot be shared by more than one UE in the same type. In the first stage, they obtain the data rate for each nonsafety V-UE in all possible combinations for each RB. In the second stage, they construct the hypergraph model that represents all the combinations of these users to find the set of the largest sum weight of hyperedge obtained from the first stage. Then, in the third stage, the resources allocation matrix was founded based on k-claw.

In [25], the authors proposed a robust radio resource and power allocation scheme for V2X communication in order to maximize the sum throughput of all V2I links while guaranteeing the reliability of each V2V link. A low-complexity algorithm was designed to find the optimal strategy of spectrum sharing among V2V and V2I links while adjusting their transmit powers. Firstly, they mathematically formulate the optimized problem to meet the V2V and V2I requirements where one V2I-UE shares spectrum with one V2V-UE. Secondly, a theorem is described in order to obtain the optimal power allocation that maximizes the capacity for V2I-UEs when it shares RB with V2V-UEs, while guaranteeing the minimum capacity requirement of V2I. Then, the Hungarian method is used to find the optimal resources reuses.

In [26], Liang et al. proposed a spectrum sharing resources and power allocation algorithm based only on slowly varying large-scale fading information of wireless channels over the fast fading. The algorithm objective is to maximize the V2I-UE ergodic capacity while guaranteeing the reliability requirement for each V2V-UEs. This algorithm was proposed to support both types of vehicular connections, i.e., V2I and V2V links, where the resource sharing happens between V2V-UEs and V2I-UEs. First, each V2V-UE is paired with each corresponding V2I-UE that satisfies the minimum capacity requirement. Then, the optimal spectrum sharing is found between the V2V-UEs and V2I-UEs sets by constructing a bipartite graph using the Hungarian method.

In [27], Mei et al. investigated a power and resource allocation scheme to jointly optimize the V2V communications. This scheme aims at maximizing the C-UEs information rate and at guaranteeing the reliability and the latency requirements of V-UEs. The latency packets are considered as the most important requirement rather than the data rate. Firstly, they mathematically formulate the end-to-end latency packets and transform it into data rate constraint. To guarantee this requirement, a minimum amount of RBs must be assigned to each V-UE. Then, the Lagrange dual decomposition method is applied to find the optimal solution of RBs sharing where at most one V-UE can share the RB that is assigned to C-UE.

In [28], Wei et al. proposed a joint resource sharing, power control and resource allocation for V2X communication over unlicensed spectrum aiming at guaranteeing fair coexistence among C-UEs, WiFi-UEs, and V-UEs. The latter is classified into nonsafety and safety V-UEs. The safety V-UEs require high reliable services in terms of latency and rate, so they are allowed to use the licensed spectrum. In case of licensed spectrum shortage, they can use the unlicensed spectrum, whereas the nonsafety V-UEs can select to use unlicensed spectrum over Content-Free Period (CFP) LTE-U or CP-based WiFi modes. Therefore, a low complexity scheme is designed in order to maximize the total throughput for nonsafety V-UEs and C-UEs.

In [29], the authors proposed a joint resource sharing and power allocation scheme for heterogeneous vehicular environment over LTE-U. This scheme aims at maximizing the total throughput for C-UEs and V-UEs, where V-UEs are categorized into nonsafety and safety V-UEs. The safety V-UEs are allowed to allocate orthogonal RB from the licensed spectrum without sharing with C-UEs in order to guarantee the safety V-UEs reliability. However, the nonsafety V-UEs can allocate RB into two modes. With slow vehicle speed, the nonsafety V-UEs compete RB with W-UEs from unlicensed spectrum during the CP interval. With high vehicle speed, the nonsafety V-UEs use the reserved Content-Free Period (CFP) based on LTE-U.

In our previous work [30], radio resource management was investigated for V-UEs where both V2V and V2I communication exist. A resource allocation algorithm is proposed aiming at promising the V2V-UEs reliability requirement and at maximizing the V2I-UEs sum rate. Firstly, V-UEs are separated into two user types; the V2V-UEs and the V2I-UEs, which are sorted in the TD scheduler according to its corresponding metric. Then, in the FD scheduler, RBs are allocated to V2I-UEs by maximizing their sum rate and to V2V-UEs by ensuring their SINR constraint in condition that at most, one V-UEs from each category can share the same resources. In [31], we add a power control mechanism to minimize the interference caused by the V2V-UEs when sharing RBs with V2I-UEs. Therefore, not only the SINR constraint is considered but also the V2V-UEs power is controlled.

In [32], we designed an efficient scheduling and resources allocation scheme for both V-UEs and C-UEs communications in order to improve C-UEs sum rate and respecting V-UEs latency constraint and C-UEs Packet Drop Rate (PDR) constraint. We classify users into three classes: the V-UEs class, the GBR C-UEs class, and the NGBR C-UEs class. Firstly, all classes' packets are prioritized according to their QoS requirement. Secondly, based on the PDR ratio, resources are dynamically adjusted for GBR and NGBR C-UEs. Then, resources that already allocated to C-UEs are reused by the V-UEs.

In [33], we proposed a mixed resource allocation and traffic sharing among C-UEs, nonsafety V-UEs, and safety V-UEs in order to promise the reliability and latency requirements. The main goal of our proposed is to regard the safety V-UEs and the C-UEs delay constraint and the SINR of all users while maximizing the C-UEs sum rate. After allocating RBs to C-UEs, resources are reused by the nonsafety and safety V-UEs where at most one RB can be shared by three users from different classes taking into consideration each class requirements.

5.1.2. RBs Sharing-Based User Clustering/Grouping. In this section, we investigate the underlaying resource allocation allowing V-UEs to share the same RBs based on user grouping/clustering. The comparison of these algorithms is summarized in Table 2.

In [34], a novel proximity-aware resource allocation-based QoS was designed for V2V-UEs in order to reduce the total power transmission considering the reliability and the queuing latency requirements. The authors exploited the spatial-temporal aspects of V-UEs based on their physical proximity and traffic demands. First of all, clustering mechanism is demonstrated to gather V-UEs into several zones based on their physical proximity. Therefore, dedicated RBs are allocated to each zone based on their traffic demands and QoS requirements. Secondly, within each zone, a power minimization solution based on the leveraging Lyapunov optimization techniques is proposed for each V2V-UE pair.

In [35], the authors proposed a graph-based resource allocation algorithm for broadcast V2V communications aiming at maximizing the sum-rate capacity of the system. Many broadcast communication clusters are gathered. Within any cluster, V-UEs should transmit in orthogonal RBs to avoid conflicts but they cannot receive and transmit simultaneously. To prevent conflicts, resources should be allocated in different subframes. However, V-UEs in different clusters can share the same RBs. Therefore, they introduce a bipartite graph-based solution aiming to assign every V-UE with a RB that attains the maximum sum rate.

In [36], the authors proposed a hybrid-scheduling algorithm for geographical zone aiming at maximizing the sum rate of the transmitting V-UEs while taking into account the reliability for all receiving V-UEs. They mathematically model their objective function problem to meet each V-UEs service requirements and to resolve it by adopting the greedy algorithm. Firstly, a hybrid scheduling is applied to allocate resource for V-UEs. The V-UEs in each geographical zone are gathered into groups based on their geographical locations. Secondly, the reuse patterns are defined where resources are reused in each reuse pattern. Then, specific RBs are allocated to V-UEs in each zone, and they will be reused in each reuse pattern.

In [37], Liang et al. studied and presented a graph-based resource allocation for V2V-UEs-based direct link and V2I-UEs-based cellular link to improve the sum V2I-UEs and to guarantee theV2V-UEs reliability. Firstly, they mathematically model the resource and the power allocation problem that meets the QoS requirements for V2I-UEs and V2V-UEs. Secondly, a graph partitioning algorithm is exploited to address this problem which can be converted into a weighted 3-dimensional matching problem. The main objective is to gather V2V-UEs into many clusters based on their mutual interferences. Then, all V2V-UEs belonging to the same cluster are allowed to share the same RBs with the corresponding V2I-UEs, whereas V2V-UEs in different clusters are not allowed to share the same RBs.

5.1.3. RBs Sharing-Based User Geographic Location. In this section, we investigate the underlaying resource allocation algorithms that allow V-UEs to share the same RBs based on user geographic location. The comparison of these algorithms is summarized in Table 3.

In [38], Yang et al. proposed a two-stage resource allocation in a dense urban in order to satisfy the data rate and the reliability requirements for both nondelay sensitive services and delay-sensitive services while minimizing the delay-sensitive services latency. The delay-sensitive V-UEs utilize LTE-based D2D link, whereas the nondelay-sensitive V-UEs can utilize LTE-C link. The

Table 2: Comparison of the existing underlaying RBs sharing based-user clustering/grouping in V2X services.

Ref.	Unicast/ broadcast	Objectives	Scenarios	User types	Allocation constraints	Power control/ allocation	Allocation process	Methods/theory	RB sharing
[34]	Unicast	Minimize the total power transmission and latency reliability	Urban	V-UEs	Queue length	PA	Nonorthogonal RBs are allocated for V2V-UEs in different clusters	Karush–Kuhn–Tucker theory	N V2V-UEs
[35]	Broadcast	Maximizing the sum rate capacity of the system	—	V-UEs	Capacity	—	Nonorthogonal RBs are allocated to user in different clusters	Kuhn–Munkres method	N V-UEs
[36]	Broadcast	Maximize the sum rate of the tr. V-UEs reliability for rx. V-UEs	Highway	V2V-UEs	OP	—	Orthogonal RBs are allocated in each zone Nonorthogonal RBs are allocated in different zones	Greedy algorithm	N V2V-UEs
[37]	Unicast	Maximize the sum V2I reliability of V2V	Multilane freeway	V2V-UEs V2I-UEs	OP	PA	Nonorthogonal RBs are allocated to V2V-UEs in the same cluster Orthogonal RB is allocated to V2I-UEs	Graph theory	1 V2I-UEs N V2V-UEs

Table 3: Comparison of the existing underlaying resource allocation in V2X services.

Ref.	Unicast/ broadcast	Objectives	Scenarios	User types	Allocation constraints	Power control/ allocation	Allocation process	Methods/theory	RB sharing
[38]	Broadcast	Minimize the delay and reliabilit Maximize data rate	Dense urban Intersection	V-UEs (nondelay- and delay-sensitive)	Avg. Que. Leng. PRR	—	Orthogonal RBs are allocated to each user type	Traffic flow theory	1 nondelay-sensitive 1 delay-sensitive
[39]	Unicast	Minimize network cost	Freewayurban	V-UEs	SINR	—	Nonorthogonal RBs are allocated to V-UEs	Hare–Niemeyer method Matching theory	N V-UEs
[40]	Unicast	Reduce the signaling overhead and interference	Urban	C-UEs V-UEs	SINR	—	Orthogonal RBs are allocated to each user type	—	1 C-UE 1 V-UE
[41]	Unicast	Reduce the signaling overhead and interference	Urban	C-UEs V-UEs	SINR	PC	Orthogonal RBs are allocated to each user type	Graph theory	1 C-UE 1 V-UE

intersection is split into four subregions in order to reduce the complexity. In the first stage, for each subregion, resources are allocated based on the Traffic Density Information (TDI), where orthogonal resources are allocated in different subregions. In the second stage, based on the Channel/Queue State Information (CSI/QSI), reusable resources are used among subregions.

In [39], a dynamic resource allocation in V2V communications is proposed with proximity awareness. This proposed algorithm aims to minimize the total network cost and to maintain successful transmissions while satisfying the QoS requirements for V-UEs. The total network cost is calculated according to traffic load and successful transmission. First, they proposed a dynamic clustering scheme to group the V-UE pairs with similar characteristics into sets of zones based on traffic load and mutual interference using the Hare–Niemeyer method to calculate the set of V-UEs in each zone, where V-UEs can reuse resources in each zone while simultaneously satisfying their QoS. Each zone has a dynamic size and changes over time according to the traffic load and proximity information. Then, a novel intrazone coordination mechanism is described based on a matching game in order to allocate resources among V-UEs in each zone.

Table 4: Comparison of the existing overlaying resource allocation in V2X services.

Ref.	Unicast/ broadcast	Objectives	Scenarios	User types	Allocation constraints	Allocation process	Methods/theory	RB sharing
[42]	Broadcast	Improve resource utilization efficiency	Highway	V-UEs	Reuse distance	Orthogonal resource is allocated to users whose relative distance is less than resource reuse distance	—	N V-UEs
[43]	—	—	Urban/ freeway	Periodic traffic of V-UEs	—	Orthogonal RB are allocated for each user	—	1 V-UEs
[44]	Unicast	Minimize the total latency	Multilane highway	V2I-UEs	SINR	Orthogonal RBs are allocated to each V-UEs using D2D communication over C-V2X or 802.11p	Greedy algorithm	1 V2I-UE
[45]	Unicast	Improve network sum rate	Freeway	C-UEs V-UEs	Fairness index OP	Orthogonal RBs are allocated to C-UEs and V-UEs	Ant colony optimization mechanism	1 V-UE or 1 C-UE

In [40], Botsov et al. designed a centralized scheduling mechanism based on the position of the vehicles within a single cell. They aim to guarantee continuous transmission for V2V link while reducing the interference and the signaling overhead within the primary network and to satisfy the V2V safety services requirements. The principle is to divide the cell coverage into several zones, and a set of RBs were assigned to each zone and is dedicated for D2D communication. These sets of RBs are reused by C-UEs while maintaining a minimum distance zone and are chosen according to RB availability and data rate requirements of the V2V service. Then, they extended to cover multicellular deployments in [41]. The zone layout and RB sets are then fixed and do not change over time.

5.2. Overlaying Resource Allocation in V2X Services. Different from the underlaying resource allocation algorithms studied in the previous section, the authors in [42–45] proposed to dedicate specific resources for V2V communications. These proposed resources remove the concerns of interference among V-UEs and C-UEs communications and on the other hand decrease the total achievable resources for cellular communications and accordingly lead in resources starvation. The comparison of the overlaying V2X resource allocation algorithms is summarized in Table 4.

In [42], Zhang et al. proposed two resource allocation schemes based on V-UEs locations for V2V broadcast services in order to improve RBs utilization and transmission precision in an efficient way and to minimize the time delay. The first scheme is the centralized scheduler, where RBs are allocated to V-UEs in orthogonal way to avoid the cofrequency interference. These RBs are reused in condition that the V-UEs distance is less than resource reuse distance. The distributed scheduler is the second one, where the highway is divided into many areas and RBs are gathered into groups in order to allow V-UEs to select RBs from a specific group in each area.

In [43], Kim et al. proposed a resource allocation scheme based on vehicle direction, position, speed, and density for V2V communication. This scheme includes two resource allocation strategies according to vehicle location, the freeway case, and the urban case. A specific resources pool is assigned for each geometric area. For the urban case, high vehicle density occurs in the intersection region, so a special resource was allocated in this region based on traffic density. For the freeway case, resources are allocated based on vehicle direction and position. Each zone of the freeway has a specific resources pool, and when a vehicle enters a zone, it must allocate resources of this zone.

In [44], 802.11p technology and C-V2X communication were investigated in a hybrid resource allocation aiming to improve the reliability of V-UEs and to minimize the total latency. Either using 802.11p or C-V2X interface, a V-UE can transmit packets that will improve the reliability. If the eNB requests a D2D link, V-UEs can transmit using C-V2X interface, if not they will use the 802.11p interface. In order to improve the latency performance, a set of RBs, periodically, are assigned for V-UEs based D2D link.

In our previous work [45], a swarm intelligence resource allocation algorithm is proposed in order to improve network sum rate while satisfying the QoS requirements for both V-UEs and C-UEs. Firstly, we mathematically express the outage probability as the requirement of the V-UEs and the user fairness index as the requirement of the C-UEs. Secondly, we adaptively (each TTI) assign a set of RBs to the V-UEs and C-UEs where orthogonal RBs are allocated among them. Then, an ant colony optimization (ACO) mechanism is adopted to the resources allocation algorithm to reduce the complexity and to getting a satisfactory performance.

6. Discussion

It seems that the V2X resource allocation based-D2D communication algorithms underlaying C-UEs is more popular than the overlaying one. Most of the existing works were proposed in the underlay mode, which are typically designed to avoid the interference problems. However, allocating dedicated RBs to V-UEs is not efficient in terms of spectral efficiency compared to the RBs sharing (underlay mode). The latter increases the spectral efficiency whereas RBs might be wasted in the overlay mode. In terms of spectrum sharing, RBs are shared between one C-UE (or V2I-UEs) and at least one V-UE or between more than two V-UEs. Therefore, in the existing resources allocation for V2X based-D2D, nonorthogonal resources are allocated to V-UEs using direct link (V2V-UEs), whereas orthogonal resources are allocated to C-UEs (or V-UEs using direct link (V2I-UEs)).

The focus of the proposed V2X resource allocation algorithms aims to maximize the number of V2V-UEs that share the same RB with only one C-UE or with only one V2I-UE. Then, allocating nonorthogonal resources to C-UEs (or V2I-UEs) can increase the spectral efficiency. Therefore, allocating nonorthogonal access to C-UEs (or V2I-UEs) and nonorthogonal access to V2V-UEs to share the same RB will be more challenging.

7. Conclusion

LTE-D2D is a promising candidate for V2X services which can meet the V2X requirements in terms of latency and reliability. Under LTE-D2D in-band communication, V-UEs can reuse the cellular resources in the underlay mode or in the overlay mode. Hence, it is essential to design a resource allocation algorithm in a way that V-UEs do not affect the C-UEs. V-UEs can communicate using the direct link and share resources with the C-UEs (or V2I-UEs-based cellular link) which may affect the cellular network performance. Consequently, the interference should be managed by resource allocation and power control/allocation algorithms. In this paper, we provide a classification and a comparison of the recent existing resource allocation algorithms for V2X communications.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Cao, M. Ma, H. Li, Y. Zhang, and Z. Luo, "A survey on security aspects for LTE and LTE-A networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 283–302, 2014.

[2] K. J. Ahmed and M. J. Lee, "Secure LTE-based V2X service," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3724–3732, 2018.

[3] 3GPP TR 22.885, v14.0.0, *Study on LTE Support for Vehicle to Everything (V2X) Services (Release 14)*, 3GPP, Valbonne, France, 2015.

[4] 3GPP TR 38.801, V14.0.0, *Study on New Radio Access Technology: Radio Access Architecture and Interfaces*, 3GPP, Valbonne, France, 2017.

[5] 3GPP TR 38.802, *Study on New Radio Access Technology Physical Layer Aspects*, 3GPP, Valbonne, France, 2017.

[6] G. Naik, B. Choudhury, and J.-M. Park, "IEEE 802.11 Bd & 5G NR V2X: evolution of radio access technologies for V2X communications," *IEEE Access*, vol. 7, pp. 70169–70184, 2019.

[7] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.

[8] P. Mach, Z. Becvar, and T. Vanek, "In-band device-to-device communication in OFDMA cellular networks: a survey and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1885–1922, 2015.

[9] S. Chen, J. Hu, Y. Shi et al., "Vehicle-to-everything (v2x) services supported by LTE-based systems and 5G," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 70–76, 2017.

[10] 3GPP TS 22.185, v14.3.0, *LTE; Service Requirements for V2X Services*, 3GPP, Valbonne, France, 2017.

[11] X. Wang, S. Mao, and M. X. Gong, "An overview of 3GPP cellular vehicle-to-everything standards," *GetMobile: Mobile Computing and Communications*, vol. 21, no. 3, pp. 19–25, 2017.

[12] 3GPP TS 33.185 V15.0.0, *Security Aspect for LTE Support of Vehicle-To-Everything (V2X) Services (Release 15)*, 3GPP, Valbonne, France, 2018.

[13] R. Molina-Masegosa and J. Gozalvez, "LTE-V for sidelink 5G V2X vehicular communications: a new 5G technology for short-range vehicle-to-everything communications," *IEEE Vehicular Technology Magazine*, vol. 12, no. 4, pp. 30–39, 2017.

[14] M. Gonzalez-Martín, M. Sepulcre, R. Molina-Masegosa, and J. Gozalvez, "Analytical models of the performance of C-V2x mode 4 vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1155–1166, 2019.

[15] B. Di, L. Song, Y. Li, and G. Y. Li, "Non-orthogonal multiple access for high-reliable and low-latency V2X communications in 5G systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2383–2397, 2017.

[16] 3GPP R1-163111, *Initial Views and Evaluation Results on Non-orthogonal Multiple Access for NR Uplink*, 3GPP, Valbonne, France, 2016.

[17] B. Di, L. Song, Y. Li, and Z. Han, "V2X meets NOMA: non-orthogonal multiple access for 5G-enabled vehicular networks," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 14–21, 2017.

[18] R. Zhang, X. Cheng, Q. Yao, C.-X. Wang, Y. Yang, and B. Jiao, "Interference graph-based resource-sharing schemes for vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 8, pp. 4028–4039, 2013.

[19] W. Sun, E. G. Strom, F. Brannstrom, Y. Sui, and K. C. Sou, "D2D-based V2V communications with latency and reliability constraints," in *Proceedings of the 2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 1414–1419, Austin, TX, USA, December 2014.

[20] W. Sun, E. G. Strom, F. Brannstrom, K. C. Sou, and Y. Sui, "Radio resource management for D2D-based V2V communication," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6636–6650, 2016.

[21] W. Sun, D. Yuan, E. G. Strom, and F. Brannstrom, "Resource sharing and power allocation for D2D-based safety-critical V2X communications," in *Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 2399–2405, London, UK, June 2015.

[22] W. Sun, D. Yuan, E. G. Strom, and F. Brannstrom, "Cluster-based radio resource management for D2D-supported safety-critical V2X communications," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2756–2769, 2016.

[23] S. Zhang, Y. Hou, X. Xu, and X. Tao, "Resource allocation in D2D-based V2V communication for maximizing the number

of concurrent transmissions," in *Proceedings of the 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, Valencia, Spain, September 2016.

[24] Q. Wei, W. Sun, B. Bai, L. Wang, E. G. Strom, and M. Song, "Resource allocation for V2X communications: a local search based 3D matching approach," in *Proceedings of the 2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, Paris, France, May 2017.

[25] L. Liang, J. Kim, S. C. Jha, K. Sivanesan, and G. Y. Li, "Spectrum and power allocation for vehicular communications with delayed CSI feedback," *IEEE Wireless Communications Letters*, vol. 6, no. 4, pp. 458–461, 2017.

[26] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for D2D-enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186–3197, 2017.

[27] J. Mei, K. Zheng, L. Zhao, Y. Teng, and X. Wang, "A latency and reliability guaranteed resource allocation scheme for LTE V2V communication systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3850–3860, 2018.

[28] Q. Wei, L. Wang, Z. Feng, and Z. Ding, "Wireless resource management in LTE-U driven heterogeneous V2X communication networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7508–7522, 2018.

[29] Q. Wei, L. Wang, Z. Feng, and Z. Ding, "Cooperative co-existence and resource allocation for V2X communications in LTE-unlicensed," in *Proceedings of the 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, Las Vegas, NV, USA, January 2018.

[30] A. Masmoudi, S. Feki, K. Mnif, and F. Zarai, "Radio resource allocation algorithm for device to device based on LTE-V2X communications," in *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications*, pp. 265–271, Porto, Portugal, 2018.

[31] A. Masmoudi, S. Feki, K. Mnif, and F. Zarai, "Efficient radio resource management for D2D- based LTE-V2X communications," in *Proceedings of the IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, Aqaba, Jordan, November 2018.

[32] A. Masmoudi, S. Feki, K. Mnif, and F. Zarai, "Efficient scheduling and resource allocation for D2D-based LTE-V2X communications," in *Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 496–501, Tangier, Morocco, June 2019.

[33] A. Masmoudi, S. Feki, K. Mnif, and F. Zarai, *A Mixed Traffic Sharing and Resource Allocation for V2X Communication*, CCIS 1118, Chapter 11, 2019.

[34] M. I. Ashraf, C.-F. Liu, M. Bennis, and W. Saad, "Towards low-latency and ultra-reliable vehicle-to-vehicle communication," in *Proceedings of the 2017 European Conference on Networks and Communications (EuCNC)*, pp. 1–5, Oulu, Finland, June 2017.

[35] L. F. Abanto-Leon, A. Koppelaar, and S. H. de Groot, "Graph-based resource allocation with conflict avoidance for V2V broadcast communications," in *Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–7, Montreal, Canada, October 2017.

[36] C.-Y. Wei, A. C.-S. Huang, C.-Y. Chen, and J.-Y. Chen, "QoS-aware hybrid scheduling for geographical zone-based resource allocation in cellular vehicle-to-vehicle communications," *IEEE Communications Letters*, vol. 22, no. 3, pp. 610–613, 2018.

[37] L. Liang, S. Xie, G. Y. Li, Z. Ding, and X. Yu, "Graph-based radio resource management for vehicular networks," 2018, http://arxiv.org/abs/1801.02679.

[38] H. Yang, L. Zhao, L. Lei, and K. Zheng, "A two-stage allocation scheme for delay-sensitive services in dense vehicular networks," in *Proceedings of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1358–1363, Paris, France, May 2017.

[39] M. I. Ashraf, M. Bennis, C. Perfecto, and W. Saad, "Dynamic proximity-aware resource allocation in vehicle-to-vehicle (V2V) communications," in *Proceedings of the 2016 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Washington, DC, USA, December 2016.

[40] M. Botsov, M. Klugel, W. Kellerer, and P. Fertl, "Location dependent resource allocation for mobile device-to-device communications," in *Proceedings of the 2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1679–1684, Istanbul, Turkey, April 2014.

[41] M. Botsov, M. Klugel, W. Kellerer, and P. Fertl, "Location-based resource allocation for mobile D2D communications in multicell deployments," in *Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 2444–2450, London, UK, June 2015.

[42] X. Zhang, Y. Shang, X. Li, and J. Fang, "Research on overlay D2D resource scheduling algorithms for V2V broadcast service," in *Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, Montreal, Canada, September 2016.

[43] J. Kim, J. Lee, S. Moon, and I. Hwang, "A position-based resource allocation scheme for V2V communication," *Wireless Personal Communications*, vol. 98, no. 1, pp. 1569–1586, 2018.

[44] F. Abbas and P. Fan, "A hybrid low-latency D2D resource allocation scheme based on cellular V2X networks," in *Proceedings of the 2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Kansas City, MO, USA, May 2018.

[45] S. Feki, A. Masmoudi, A. Belghith, F. Zarai, and M. S. Obaidat, "Swarm intelligence-based radio resource management for D2D-based V2V communication," *International Journal of Communication Systems*, p. e3817, 2018.

Hindawi
Wireless Communications and Mobile Computing
Volume 2019, Article ID 2121850, 10 pages
https://doi.org/10.1155/2019/2121850

Research Article

Recommendation of Crowdsourcing Tasks Based on Word2vec Semantic Tags

Qingxian Pan ,[1,2] Hongbin Dong ,[1] Yingjie Wang,[2] Zhipeng Cai ,[3] and Lizong Zhang[4]

[1] *College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China*
[2] *School of Computer and Control Engineering, Yantai University, Yantai 264005, China*
[3] *Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA*
[4] *School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*

Correspondence should be addressed to Hongbin Dong; donghongbin@hrbeu.edu.cn

Received 1 November 2018; Revised 18 February 2019; Accepted 3 March 2019; Published 24 March 2019

Guest Editor: Michele Nogueira

Crowdsourcing is the perfect show of collective intelligence, and the key of finishing perfectly the crowdsourcing task is to allocate the appropriate task to the appropriate worker. Now the most of crowdsourcing platforms select tasks through tasks search, but it is short of individual recommendation of tasks. Tag-semantic task recommendation model based on deep learning is proposed in the paper. In this paper, the similarity of word vectors is computed, and the semantic tags similar matrix database is established based on the Word2vec deep learning. The task recommending model is established based on semantic tags to achieve the individual recommendation of crowdsourcing tasks. Through computing the similarity of tags, the relevance between task and worker is obtained, which improves the robustness of task recommendation. Through conducting comparison experiments on Tianpeng web dataset, the effectiveness and applicability of the proposed model are verified.

1. Introduction

Deep learning was proposed by Geoffrey Hinton et al. in 2006. This method simulates human brain neural network to model and realize multiple level abstraction [1, 2]. In 2006, Jeff Howe of American Wired magazine reporter proposed crowdsourcing concept [3]. As a new kind of business model, crowdsourcing has been widespread concern in various fields and becomes the new hot point of computer research fields. Task requester, crowdsourcing platform, and worker make up crowdsourcing system [4]. The process of crowdsourcing includes designing task, publishing task, selecting task, sensing task, submitting solution, and integrating solution. Among them, task selection is the key phase in the process of crowdsourcing. This is the key to complete crowdsourcing task that the appropriate worker selects appropriate task in appropriate time [5].

The popular crowdsourcing platforms use task searching to get the favourite task by keyword searching [6]. However, with the rapid development of crowdsourcing, the problem of information overload is more and more serious. In addition, it is more and more difficult to get the favourite crowdsourcing task for worker. Recommender system is an effective medium to solve the problem, which is used on many E-Commerce Platforms, such as Alibaba, Amazon, and Netflix [7]. But there are many problems which are not solved in recommender systems, such as similarity calculation, the lower recommended accuracy, data sparseness, and cold boot. In brief, improving the accuracy and reliability of recommender systems has been paid more attention by scholars.

However, individual recommendation research of the task is lesser in crowdsourcing, and task selection is relied on hobbies and expertise. Few crowdsourcing platforms can actively recommend task. This paper researches the crowdsourcing tasks recommendation model based on Word2vec semantic tags in order to achieve individual recommendation of crowdsourcing tasks [8].

The main contributions of this paper include following three contents:

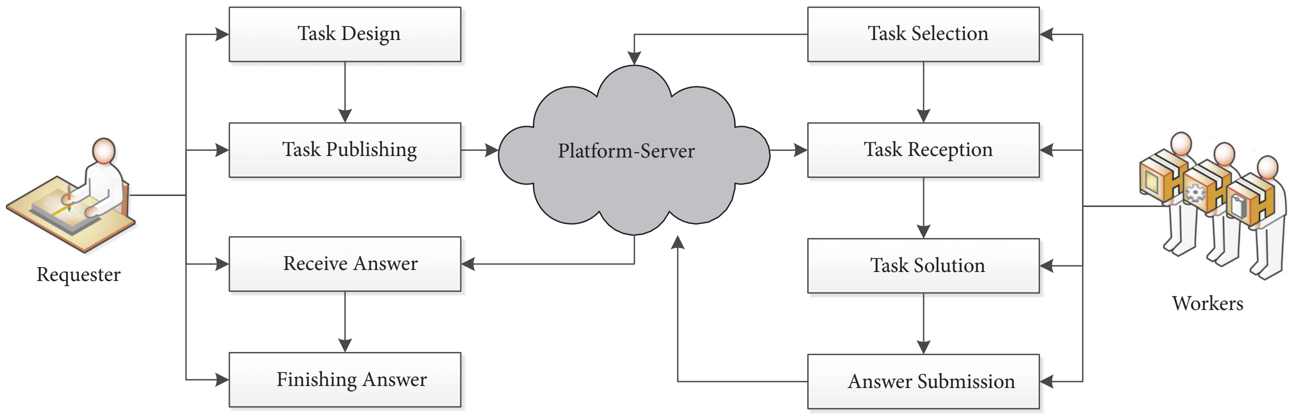


Figure 1: The workflow of crowdsourcing.

(1) Compute the similarity of word vectors and build the semantic tags similar matrix database based on the Word2vec deep learning.

(2) Research the task recommending model based on semantic tags to achieve the individual recommendation of crowdsourcing tasks. This paper computes similarity of tasks and workers based on the semantic tag similar matrix.

(3) Utilizing the Tianpeng Web dataset, the experiments are conducted. The experimental results show that the model is feasible and effective. The model can be used in other fields according to the different semantic databases.

This paper is organized as follows. Section 2 reviews the related works. The Work2vec is discussed in Section 3. In addition, the tasks recommendation model and realization method based on semantic tags are researched in Section 4. The comparison experiments, as well as the analysis for the experimental results, are introduced in Section 5. The conclusion is presented in Section 6.

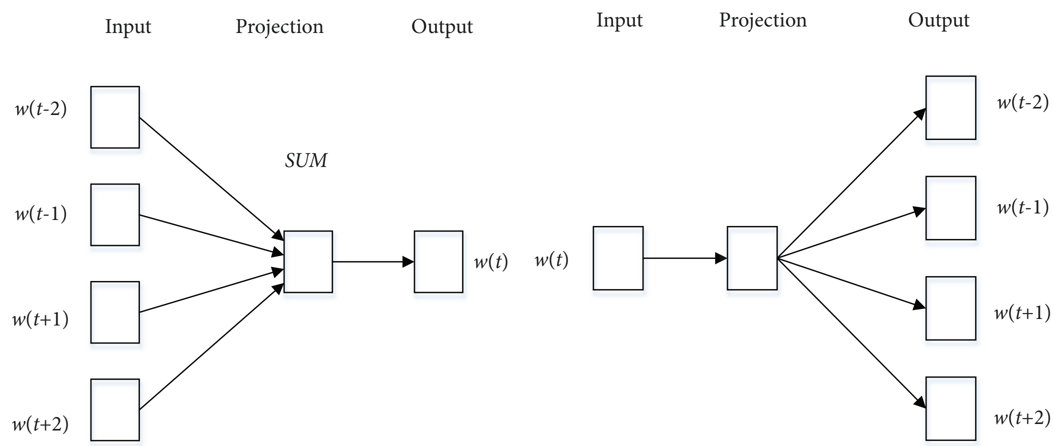
2. Related Works

In order to discuss the related works for recommendation of crowdsourcing, we, respectively, introduce the related works of crowdsourcing and recommendations.

2.1. Crowdsourcing. In 2006, Jeff Howe proposed crowdsourcing concept firstly [3]: a company or an institution outsources the tasks performed by an employee in the past to an unspecific public network in a free and voluntary manner. With the development of crowdsourcing technology, the different crowdsourcing concepts appeared. Chen et al. [9] summarized 40 different crowdsourcing definitions. Feng et al. [10] gave the definition of crowdsourcing according to the basic features of crowdsourcing. According to the definition, crowdsourcing is a distributed problem-solving mechanism opening to the Internet public, and it completes the tasks that are difficult to complete by a computer through integrating computers and the unknown public on the Internet [11].

Crowdsourcing is successfully applied in language translation, image recognition, intelligent transportation, software development, entry interpretation, tourism photography, and other fields, which has become the perfect embodiment of group wisdom [12, 13]. Crowdsourcing is made up of the task requester, crowdsourcing platform, and workers. The crowdsourcing workflow includes designing tasks by task requester, publishing tasks, selecting tasks by workers, solving tasks, submitting answer, and arranging answer. The workflow of crowdsourcing is shown by Figure 1. The public participation is the basis of crowdsourcing. And the key to high-quality complete crowdsourcing tasks is to recommend appropriate tasks to appropriate worker in appropriate time [14].

2.2. Recommender Systems. With the arrival of big data era, the problem of information overload is more and more serious and that finding the useful and best information is more and more difficult. Recommender Systems is an effective medium to solve the above problems [15]. However, there are some inherent defects in recommendation systems, such as low accuracy, data sparseness, cold boot, the defects of the centralized system, similarity calculation, and being easy to be attacked. In addition, many recommender systems applied to business systems, whose purpose is to sell more goods and seek the maximum benefits, rather than to recommend the best commodities to users. In brief, the credibility and accuracy of recommendation systems need to be improved, which has attracted the attention of scholars. Yang et al. [16] proposed a recommender system based on transfer learning. Chen et al. [17] proposed a recommender system based on bind context. Tang et al. [18] researched recommender system based on crossing knowledge. Liu [19] and Zhou et al. [20] researched recommender systems for social recommendation. Combining Markov and social attributes of users, Wang et al. [21] proposed a probability-based recommendation model to recommend items for users.



(a) CBOW model

(b) Skip-gram model

Figure 2: CBOW model and Skip-gram model.

Crowdsourcing task recommendation is mainly from the perspective of crowdsourcing platform. Based on the task discovery model, crowdsourcing platform recommends related tasks according to the preferences of workers [5]. The main crowdsourcing platforms basically adopt the way of task search and rarely adopt the method of task recommendation [22]. Some task recommendation methods were researched based on traditional recommendation methods, including content-based recommendation, collaborative filtering, and mixed recommendation algorithms. Ambati et al. [23] proposed the use of task and workers' historical information for task recommendation. Yuen et al. [24] proposed a worker-task recommendation model through combining the historical information of workers and browsing history. Deng et al. [25] researched the problem of maximizing task selection for spatiotemporal tasks.

3. Word2vec

In 2003, Bengio et al. [26] proposed Neural Network Language Model-NNLM based on 3 levels. NNLM is used to compute the probability $p(w_t = i \mid context)$ of the next word w_t of a context, and word vector is the byproduct during training. Word2vec is a tool based on deep learning to compute the similarity of word vector which was proposed by Google company in 2013 [27]. It converts the word into word vector and computes similarity according to the cosine between word vectors. When using the tool, the texts after segmentation are input, and the output-word vector can be used to do a lot of Natural Language Processing (NLP) related work, such as clustering, looking for synonyms, and part of speech analysis.

Word2vec uses word vector presentation mode based on Distributed representation. Distributed representation is proposed by Hinton in 1986 [28]. Its basic thought is to map each word into a k-dimension real vector by training (k is a hyperparameter in the model) and to judge the semantic similarity between them according to the distance between words (such as cosine similarity, Euclidean distance). It uses a '3 layers neural network', input layer-hidden layer-output layer. Its core technology is to use Huffman code according to word frequency, which makes the activated content basically consistent of all word frequency similar words in hidden layer. The higher the frequency of the word, the less the number of hidden layers they activate, which effectively reduces the computational complexity.

Compared with Latent Semantic Index-LSI and Latent Dirichlet Allocation-LDA, Word2vec uses the context of words and makes the semantic information richer. There are two kinds of training model-CBOW (Continuous Bag-of-Words) and Skip-gram in Word2vec, which are shown by Figure 2. Two models both include input layer, projection layer, and output layer. CBOW model predicts the current words according to the known context, and Skip-gram model predicts context according to the current words.

In this paper, the objective optimization function of CBOW is expressed by

$$p(w \mid Context(w)) = \prod_{j=2}^{l^w} p\left(d_j^w \mid x_w, \theta_{j-1}^w\right) \quad (1)$$

where x_w means the word vector of the root node in the Hoffman tree, $Context(w)$ represents the context of word w, that is, the collection of w peripheral words, l^w represents the nodes number of the path p^w, and $d_j^w \in \{0,1\}$ represents Huffman code of the word w; $\theta_1^w, \theta_2^w, \ldots, \theta_{j-1}^w \in R^m$ represents the vectors corresponding to nonleaf nodes of the path p^w. Therefore, the logistic regression probability $p(d_j^w \mid x_w, \theta_{j-1}^w)$ that w passes a node j in the Hoffman tree is shown by (2). The corresponding parameter $\sigma(x_w^T \theta_{j-1}^w)$ is shown by (3).

$$p\left(d_j^w \mid x_w, \theta_{j-1}^w\right) = \begin{cases} \sigma\left(x_w^T \theta_{j-1}^w\right), & d_j^w = 0; \\ 1 - \sigma\left(x_w^T \theta_{j-1}^w\right), & d_j^w = 1. \end{cases} \quad (2)$$

$$\sigma\left(x_w^T \theta_{j-1}^w\right) = \frac{1}{1 + e^{-x_w^T \theta_{j-1}^w}} \quad (3)$$

In order to clearly represent the meaning of logistic regression probability $p(d_j^w|x_w, \theta_{j-1}^w)$, we combine (2) and (3) to obtain the value of $p(d_j^w|x_w, \theta_{j-1}^w)$, which is shown by

$$p\left(d_j^w \mid x_w, \theta_{j-1}^w\right) = \left[\sigma\left(x_w^T\theta_{j-1}^w\right)\right]^{1-d_j^w} \cdot \left[1-\sigma\left(x_w^T\theta_{j-1}^w\right)\right]^{d_j^w} \quad (4)$$

For avoiding the value of $p(w \mid Context(w))$ too small, logarithm Likelihood function is used to represent the objective function; thus, (1) can be converted into

$$L = \sum_{w\in C} \log p\left(w \mid Context\left(w\right)\right) \quad (5)$$

Through combining (4) and (5), the objective function L is shown by

$$\begin{aligned} L &= \sum_{w\in C} \log \prod_{j=2}^{l^w} \left\{\left[\sigma\left(x_w^T\theta_{j-1}^w\right)\right]^{1-d_j^w} \cdot \left[1 - \sigma\left(x_w^T\theta_{j-1}^w\right)\right]^{d_j^w}\right\} \\ &= \sum_{w\in C}\sum_{j=2}^{l^w} \left\{\left(1-d_j^w \cdot \log\left[\sigma\left(x_w^T\theta_{j-1}^w\right)\right] + d_j^w \cdot \log\left[1-\sigma\left(x_w^T\theta_{j-1}^w\right)\right]\right)\right\} \end{aligned} \quad (6)$$

Therefore, (6) is the object function of CBOW in this paper. Word2vec uses random gradient ascent method to optimize the object function of CBOW.

4. The Tasks Recommendation Model and Realization Method Based on Semantic Tags

4.1. Basic Model Frame and Mathematical Computation Model. The results and discussion may be presented separately, or in one combined section, and may optionally be divided into headed subsections.

The core of the model is the research of tag similar matrix. The model uses tag similar matrix to compute the similarity of workers and tasks, produces worker-tag similar matrix, and realizes tasks recommendation or workers recommendation. In model, tag similar matrix is obtained by Word2vec computing. Worker-tag matrix is got according to history work information of the worker, registration information, etc. And task-tag matrix is got according to task description, task classification, etc.

Define tag similar matrix $L \in R^{m\times m}$, $\begin{bmatrix} l_{11} & \cdots & l_{1m} \\ \vdots & \ddots & \vdots \\ l_{m1} & \cdots & l_{mm} \end{bmatrix}$, L is a symmetric matrix, that is, $l_{ij} = l_{ji}$, l_{ij} represents the similarity of tag i and tag j, $l_{ij} \in [0,1]$, and its value is got through using Word2vec tool to compute. Define worker-tag matrix $W \in R^{n\times m}$, $\begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nm} \end{bmatrix}$, and, among them, $w_{ij} = \{1,$ worker i has tag j; $0,$ worker i has not tag $j\}$.

We define the task-tag matrix $T \in R^{p\times m}$, $\begin{bmatrix} t_{11} & \cdots & t_{1m} \\ \vdots & \ddots & \vdots \\ t_{p1} & \cdots & t_{pm} \end{bmatrix}$, and, among them, $t_{ij} = \{1,$ task i has tag j; $0,$ task i has not tag $j\}$.

Therefore, the worker-task similar matrix W_T is obtained by (7), where W is the worker-tag matrix, L is the tag similar matrix, and T^T means the task-tag transposed matrix. Through (7), the relationship between workers and tasks can be obtained.

$$\begin{aligned} W_T &= W \times L \times T^T \\ &= \begin{bmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nm} \end{bmatrix} \times \begin{bmatrix} l_{11} & \cdots & l_{1m} \\ \vdots & \ddots & \vdots \\ l_{m1} & \cdots & l_{mm} \end{bmatrix} \times \begin{bmatrix} t_{11} & \cdots & t_{1m} \\ \vdots & \ddots & \vdots \\ t_{p1} & \cdots & t_{pm} \end{bmatrix}^T \end{aligned} \quad (7)$$

4.2. Basic Flow. The main steps of the process of the proposed recommendation model are shown as follows: (1) compute the word vectors based on Word2vec; (2) computing the similarity of word vectors; (3) generating the tag similar matrix; (4) obtaining the worker-tag matrix and task-tag matrix; (5) computing the worker-task similarity matrix; (6) L^2 standardization and normalization; (7) tasks and workers recommendation. Tag similar matrix generation uses Word2vec tool. Worker-task similarity computation uses mathematical methods introduced in the previous section. The section mainly introduces standardization and normalization method.

L^2 standardization method: the L^2 norm definition of vector $x(x_1, x_2, \ldots, x_n)$ is shown as follows: $norm(x) = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

In order to make x normalized to the unit L^2 norm, the mapping between x and x' is established, so that the L^2 norm of x' is 1, and the proof is shown as follows:

$$\begin{aligned} 1 = norm\left(x'\right) &= \frac{\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}}{norm\left(x\right)} \\ &= \sqrt{x_1'^2 + x_2'^2 + \cdots + x_n'^2} \\ &= \sqrt{\left(\frac{x_1}{norm\left(x\right)}\right)^2 + \left(\frac{x_2}{norm\left(x\right)}\right)^2 + \cdots + \left(\frac{x_n}{norm\left(x\right)}\right)^2} \end{aligned} \quad (8)$$

where the value of x_i' is shown by

$$x_i' = \frac{x_i}{norm\left(x\right)} \quad (9)$$

In order to get the standardization and generality of data, the standardization data of L^2 is normalized, so that the data fall in the interval $[0,1]$, the conversion formula is shown by (10), where $\min(X)$ means the minimum in X, and $\max(X)$ is the maximum in X.

$$x_i' = \frac{x_i - \min\left(X\right)}{\max\left(X\right) - \min\left(X\right)} \quad (10)$$

Table 1: Word2vec parameter setting.

Parameter	Value	Parameter	Value
window	8	hs	1
size	100	cbow	yes
threads	20	alpha	0.001
binary	0	negative	25

Table 2: Tag similar matrix L of simulation dataset.

	L1	L2	L3	L4	L5	L6	L7	...
L1	1.000	0.407	0.124	0.119	0.126	0.434	0.075	...
L2	0.407	1.000	0.766	0.917	0.993	0.642	0.546	...
L3	0.124	0.766	1.000	0.930	0.477	0.526	0.744	...
L4	0.119	0.917	0.930	1.000	0.909	0.531	0.394	...
L5	0.126	0.993	0.477	0.909	1.000	0.636	0.860	...
L6	0.434	0.642	0.526	0.531	0.636	1.000	0.166	...
L7	0.075	0.546	0.744	0.394	0.860	0.166	1.000	...
...	...	...	...	...	...	...	...	...

Table 3: Worker-tag matrix W.

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	...
W1	0	0	0	1	0	0	0	0	0	0	0	...
W2	0	0	0	0	0	0	0	1	0	0	0	...
W3	0	1	0	0	1	0	1	0	0	1	0	...
W4	1	0	0	0	0	0	0	0	1	0	0	...
W5	0	0	0	0	0	1	0	0	0	0	0	...
...	...	...	...	...	...	...	...	...	...	...	...	...

Table 4: Task-tag matrix T.

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	...
T1	0	0	0	0	0	0	0	1	0	0	0	...
T2	0	0	1	0	0	0	1	0	1	0	0	...
T3	0	0	0	0	0	0	0	0	0	0	0	...
T4	0	0	1	0	1	0	0	0	0	0	0	...
T5	0	0	0	0	0	0	1	0	0	0	0	...
...	...	...	...	...	...	...	...	...	...	...	...	...

Table 5: Worker-task similar matrix.

	T1	T2	T3	T4	T5	T6	T7	...
U1	0.754	0.410	0.369	0.438	0.365	0.420	0.396	...
U2	0.680	0.694	0.712	0.706	0.682	0.747	0.720	...
U3	0.387	0.378	0.407	0.403	0.385	0.678	0.351	...
U4	0.405	0.304	0.681	0.731	0.733	0.835	0.704	...
U5	0.279	0.278	0.696	0.284	0.294	0.265	0.324	...
...	...	...	...	...	...	...	...	...

5. Experiment and Simulation

In this section, we conduct the comparison experiments on the simulation dataset and real dataset, respectively. The real dataset is the dataset crawled from Tianpeng web site.

In the experiment, text8 is corpora training set, and experimental environment is Intel Core (TM) i5-337U CPU @1.8GHz dual-core, and 8GB memory.

5.1. The Experiments Conducted on Simulation Dataset. In this group of comparison experiments, the training parameters are shown in Table 1.

In addition, the tag similar matrix after training is shown in Table 2. In the matrix, the elements indicate the similarities between tags.

In this group of experiments, there are 100 workers, 50 tasks, 2000 tags in the experiment. The worker-tag matrix is generated randomly, which is shown in Table 3. The elements in Table 3 represent the similarities between workers and tags. The task-tag matrix is shown in Table 4. The elements in Table 4 indicate the similarities between tasks and tags. After computing the worker-task matrix, the standardization and normalization of worker-task matrix are shown in Table 5. The elements in Table 5 mean the similarities between workers and tasks.

Recall, precision, and F-measure are commonly used evaluation indexes [29]. The computing methods for the three evaluation indexes are shown by (11), (12), and (13). According to (11), (12), and (13), it can be seen that F-measure index is the comprehensive measure index through considering both recall and precision.

$$\text{Recall} = \frac{\text{the quantity of related information retrieved}}{\text{the quantity of related information in system}} \tag{11}$$

$$\text{Precision} = \frac{\text{the quantity of related information retrieved}}{\text{the quantity of all information retrieved}} \tag{12}$$

$$\text{F-measure} = \frac{2 \times \text{Precision} \times \text{Pecall}}{\text{Precision} + \text{Recall}} \tag{13}$$

The threshold values are 0.55, 0.6, and 0.65, respectively, and the recall, precision, and F-measure of the 50 tasks are obtained. The comparison experimental results on recall, precision, and F-measure indexes are shown by Figures 3, 4, and 5, respectively. In these experiments, x-coordinate indicates the Task-tag matrix T, and y-coordinates are recall rate, precision rate, and F-measure rate, respectively. From the experimental results, it can be seen that *threshold*=0.6 has better performance than other two thresholds comprehensively.

In addition, we compare the proposed method with the method of tasks research. The experimental result is shown in Figure 6, where x-coordinate indicates the Task-tag matrix T and y-coordinate means the number of workers. The method used in this paper is better than the method used in tasks research, which proves the effectiveness of the method of this paper. In addition, the potential workers can be found

Table 6: Tag similar matrix L of Tianpeng dataset.

	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	...
L1	1.000	0.004	-0.041	0.100	0.018	0.048	0.020	-0.040	-0.049	-0.029	-0.009	0.038	-0.026	...
L2	0.004	1.000	0.803	0.261	0.882	0.225	0.493	0.610	0.391	0.315	0.817	0.666	0.601	...
L3	-0.041	0.803	1.000	0.231	0.761	0.166	0.393	0.533	0.351	0.259	0.722	0.603	0.609	...
L4	0.100	0.261	0.231	1.000	0.268	0.134	0.234	0.176	0.191	0.229	0.248	0.251	0.173	...
L5	0.018	0.882	0.761	0.268	1.000	0.218	0.475	0.571	0.352	0.228	0.753	0.659	0.583	...
L6	0.048	0.225	0.166	0.134	0.218	1.000	0.133	0.135	0.101	0.095	0.198	0.222	0.181	...
L7	0.020	0.493	0.393	0.234	0.475	0.133	1.000	0.334	0.190	0.192	0.504	0.459	0.296	...
L8	-0.040	0.610	0.533	0.176	0.571	0.135	0.334	1.000	0.295	0.258	0.556	0.480	0.515	...
L9	-0.049	0.391	0.351	0.191	0.352	0.101	0.190	0.295	1.000	0.248	0.386	0.239	0.277	...
L10	-0.029	0.315	0.259	0.229	0.228	0.095	0.192	0.258	0.248	1.000	0.238	0.288	0.236	...
L11	-0.009	0.817	0.722	0.248	0.753	0.198	0.504	0.556	0.386	0.238	1.000	0.616	0.535	...
L12	0.038	0.666	0.603	0.251	0.659	0.222	0.459	0.480	0.239	0.288	0.616	1.000	0.455	...
L13	-0.026	0.601	0.609	0.173	0.583	0.181	0.296	0.515	0.277	0.236	0.535	0.455	1.000	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

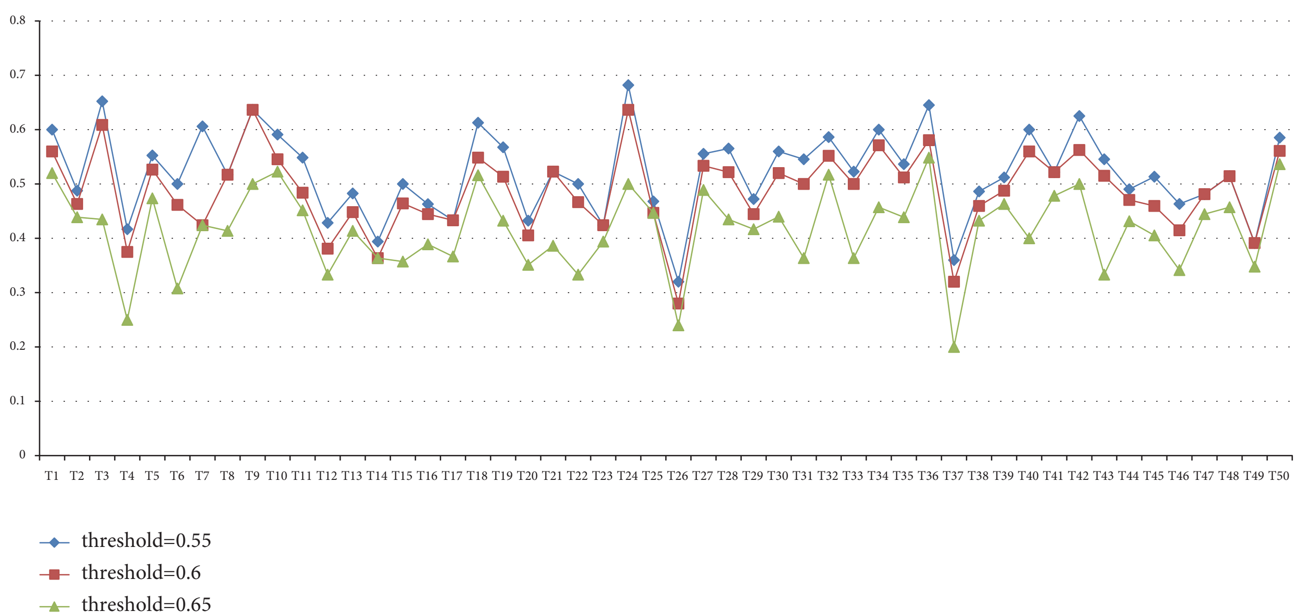


Figure 3: Recall of different thresholds.

by lowering the threshold, which can be used to analyze the potential users.

5.2. The Experiments Conducted on Tianpeng Dataset. The data collected from the Tianpeng web site were collected to form a corpus for training, and the tag similarity matrix was obtained as shown in the Table 6.

We select 510 workers and 371 tasks from Tianpeng dataset as experimental objects. Utilizing the dataset, we conduct the comparison experiments to verify the effectiveness of the proposed model. In the comparison experiments, 0.6 is taken as the threshold, and 20 tasks are randomly selected as recommended objects. The experimental results were compared with binary map matching and greedy algorithm in terms of recall rate, accuracy rate, and F-value measure indexes.

According to the recall measure index, the comparison experimental result is shown by Figure 7. The x-coordinate indicates the Task-tag matrix T, and y-coordinate presents the recall rate. From the experimental result, it can be seen that the proposed recommendation model has the best performance on recall rate through compared with greedy algorithm and bipartite graph matching. In addition, the proposed recommendation model has better stability with the changing of T.

Figure 8 shows the experimental result on precision rate. Similarly, the x-coordinate indicates the Task-tag matrix T, and y-coordinate means the precision rate. In experimental

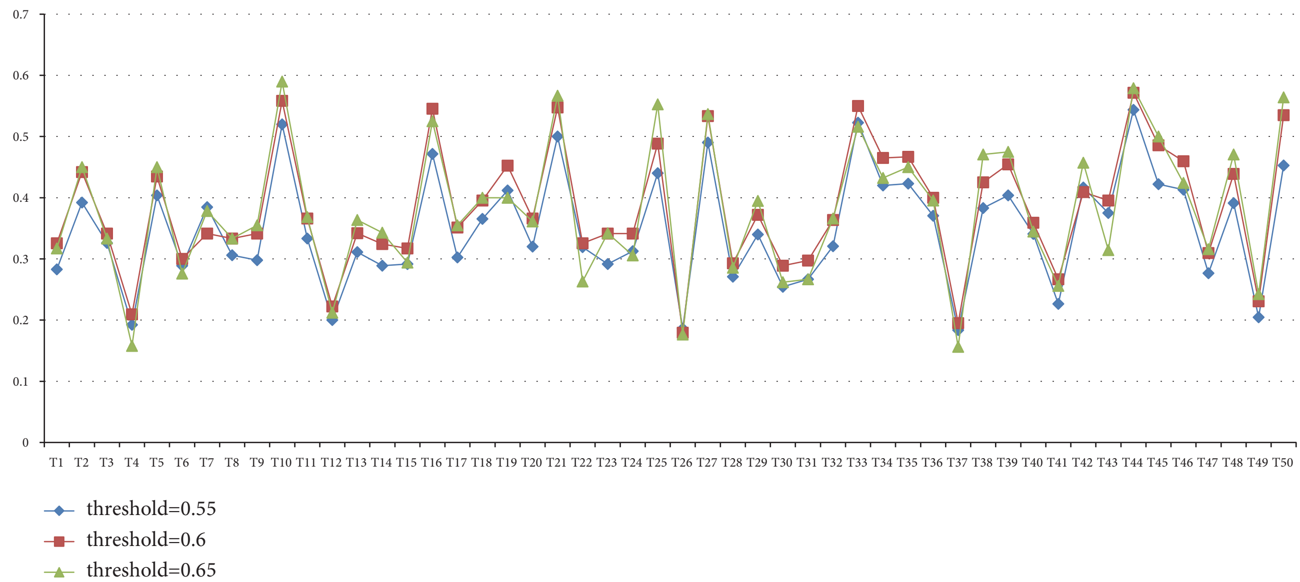


Figure 4: Precision of different thresholds.

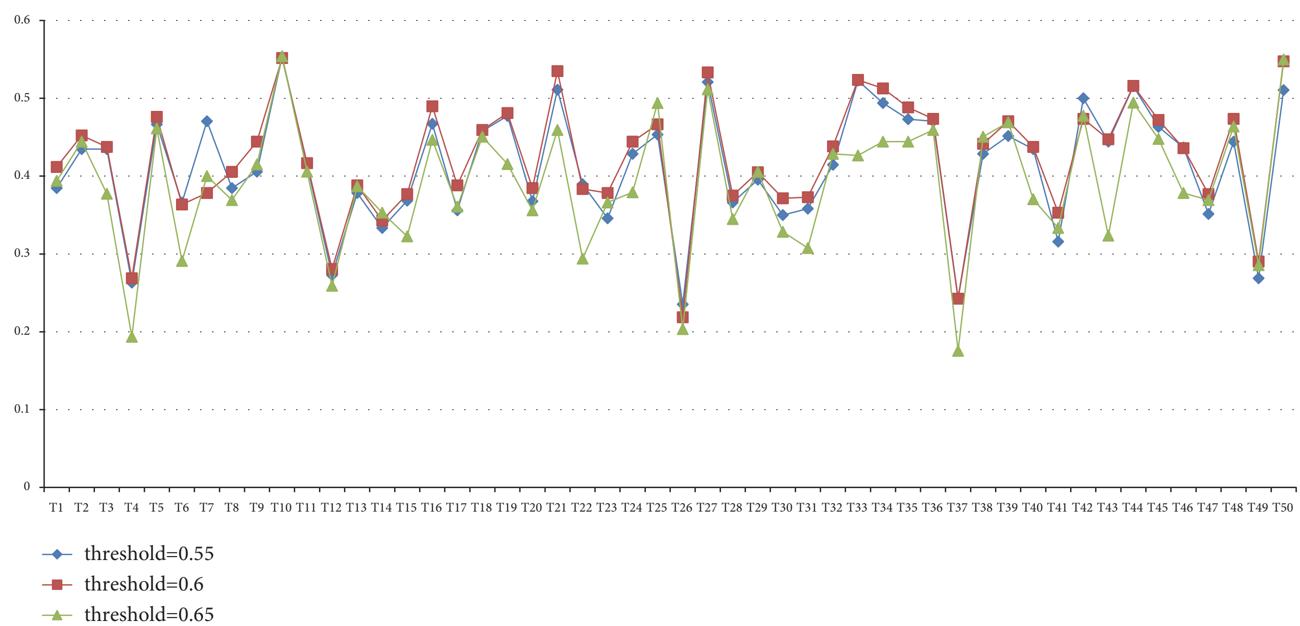


Figure 5: F-measure of different thresholds.

result, the average precision rate of the proposed recommendation is better than other two algorithms. From Figure 7, it can be seen that the proposed recommendation has the best performance on precision rate through compared with greedy algorithm and bipartite graph matching.

According to the experimental result on F-measure shown by Figure 9, we can see that the proposed recommendation also has the best performance on F-measure. In addition, F-measure index is the comprehensive measure index through considering both recall and precision. Therefore, we can infer that the proposed recommendation has the best performance through compared with greedy algorithm and bipartite graph matching algorithm.

Through the comparison shows that the proposed methods than the binary map matching method, greedy algorithm in the recall, F-measure index significantly, in terms of accuracy with high and low, because to make the task would be able to complete the task of recommended for workers as much as possible, including the potential of workers, so the accuracy index can be put lower in the recommended requirements. It can be seen that the method proposed in this paper has higher practical significance and application value.

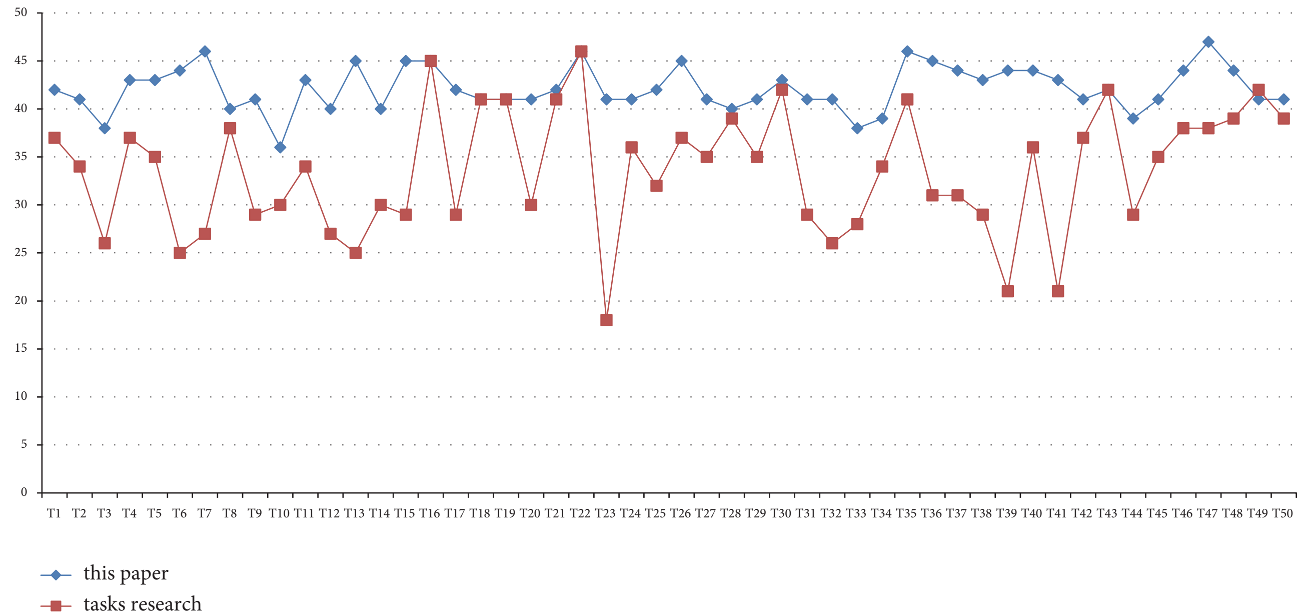


Figure 6: Comparison of experimental results.

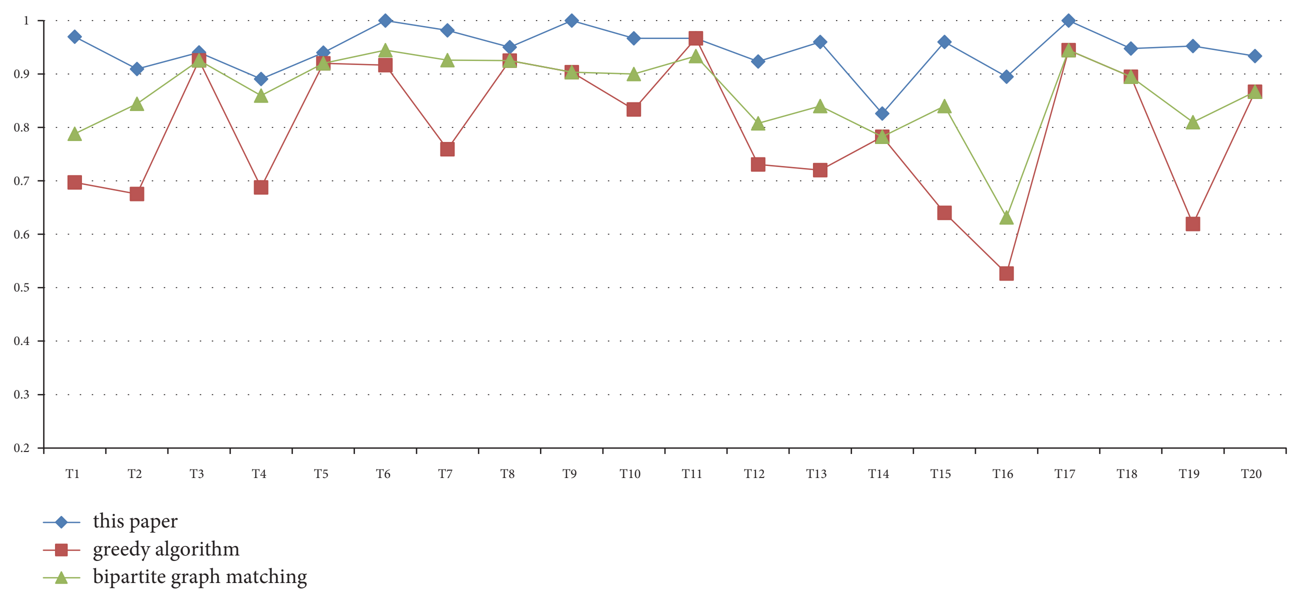


Figure 7: Recall of different methods.

6. Conclusion

Crowdsourcing is the prefect shown of group wisdom. It was applied in many fields as a new business model. In recent years, it has become the new hot research in computer science. The success key of crowdsourcing is to recommend task to appropriate worker. The recommendation method based on tag similar matrix is proposed in this paper. The method uses Word2vec technology to generate tag similar matrix and then computes the similarity of worker and task. According to the comparison experiments, it proves that the method is effective and feasible. The recommendation method can be extended to other fields with the different corpora.

Because the success key of crowdsourcing is the participate rate of workers, it has become a hot topic in crowdsourcing research, such as reputation mechanism, preference evolution, and privacy protection of workers. It will be the focus of future research to improve the accuracy of recommender systems by combining recommender systems with reputation, preference evolution and historical information.

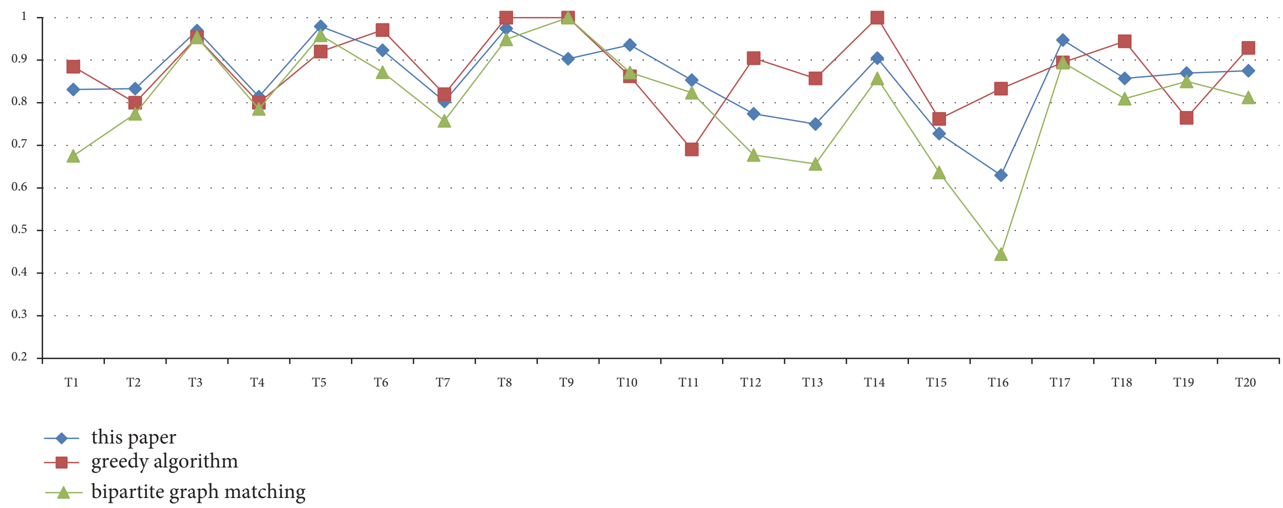


Figure 8: Precision of different methods.

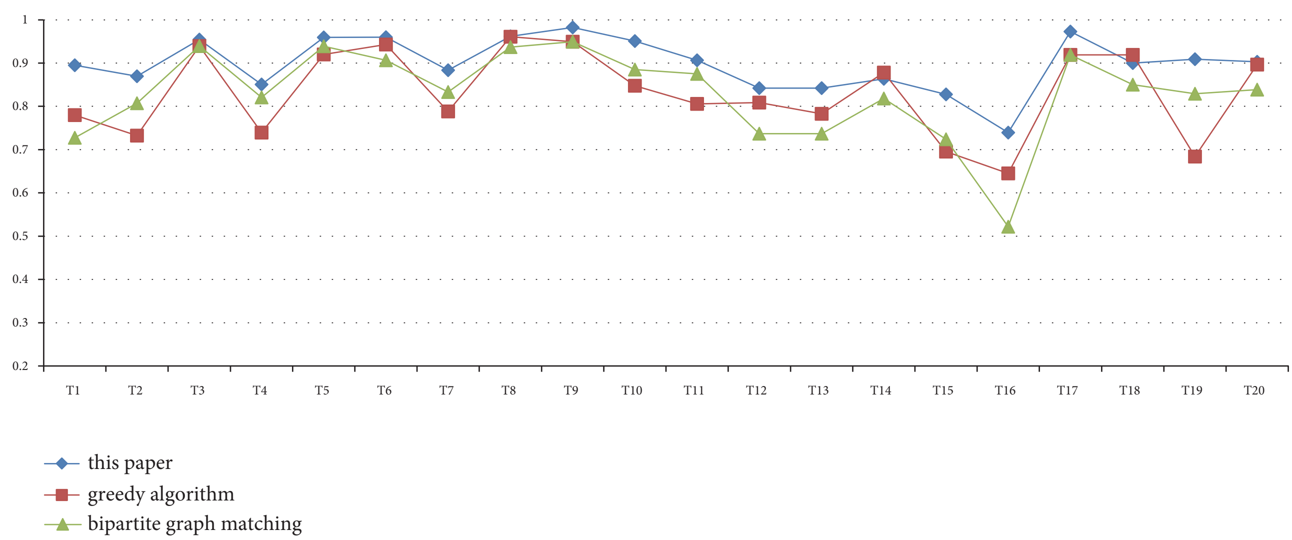


Figure 9: F-measure of different methods.

Data Availability

The [Tianpeng] dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants No. 61472095, No. 61502410, and No. 61572418, the China Postdoctoral Science Foundation under Grant No. 2017M622691, the National Science Foundation (NSF) under Grants No. 1704287, No. 1252292, and No. 1741277, and the Natural Science Foundation of Sichuan Province under Grant No. 2018HH0075.

References

[1] Y. Cun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and G. Wu, "An incentive mechanism with privacy protection in mobile crowdsourcing systems," *Computer Networks*, vol. 102, pp. 157–171, 2016.

[3] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, pp. 1–4, 2006.

[4] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, p. 1, 2018.

[5] Y. Hu, Y. Wang, Y. Li, and X. Tong, "An incentive mechanism based on multi-attribute reverse auction in mobile crowdsourcing," *Sensors*, vol. 18, no. 10, p. 3453, 2018.

[6] J. Li, Z. Cai, J. Wang, M. Han, and Y. Li, "Truthful incentive mechanisms for geographical position conflicting mobile crowdsensing systems," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 324–334, 2018.

[7] R. Katarya and O. P. Verma, "Recent developments in affective recommender systems," *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 182–190, 2016.

[8] K. W. Church, "Emerging trends: Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.

[9] X. Chen, P. N. Bennett, K. Collins-Thompson, and E. Horvitz, "Pairwise ranking aggregation in a crowdsourced setting," in *Proceedings of the Sixth ACM International Conference*, pp. 193–202, Rome, Italy, Feburary 2013.

[10] J. Feng, G. Li, and J. Feng, "A survey on crowdsourcing," *Chinese Journal of Computers*, vol. 38, pp. 1713–1726, 2015.

[11] Z. Duan, W. Li, and Z. Cai, "Distributed auctions for task assignment and scheduling in mobile crowdsensing systems," in *Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pp. 635–644, GA, USA, June 2017.

[12] Y. Wang, Z. Cai, X. Tong, Y. Gao, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems," *Computer Networks*, vol. 135, pp. 32–43, 2018.

[13] Y. Wang, Y. Li, Z. Chi, and X. Tong, "The truthful evolution and incentive for large-scale mobile crowd sensing networks," *IEEE Access*, vol. 6, pp. 51187–51199, 2018.

[14] J. L. Cai, M. Yan, and Y. Li, "Using crowdsourced data in location-based social networks to explore influence maximization," in *Proceedings of the 35th Annual IEEE International Conference on Computer Communications*, 2016.

[15] P. Resnick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[16] W. Pan and Q. Yang, "Transfer learning in heterogeneous collaborative filtering domains," *Artificial Intelligence*, vol. 197, pp. 39–55, 2013.

[17] G. Chen and L. Chen, "Recommendation based on contextual opinions," *UMAP 2014, LNCS 8538*, pp. 61–73, 2014.

[18] L. Liu, J. Tang, J. Han, and S. Yang, "Learning influence from heterogeneous social networks," *Data Mining and Knowledge Discovery*, vol. 25, no. 3, pp. 511–544, 2012.

[19] J. Tang, X. Hu, and H. Liu, "Social recommendation: a review," *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1113–1133, 2013.

[20] L. Lü, M. Medo, C. H. Yeung, Y. Zhang, Z. Zhang, and T. Zhou, "Recommender systems," *Physics Reports*, vol. 519, no. 1, pp. 1–49, 2012.

[21] Y. Wang, G. Yin, Z. Cai, Y. Dong, and H. Dong, "A trust-based probabilistic recommendation model for social networks," *Journal of Network and Computer Applications*, vol. 55, pp. 59–67, 2015.

[22] L. Zhang, Z. Cai, and X. Wang, "FakeMask: a novel privacy preserving approach for smartphones," *IEEE Transactions on Network and Service Management*, vol. 13, no. 2, pp. 335–348, 2016.

[23] V. Ambati, S. Vogel, and J. Carbonell, "Towards task recommendation in micro-task markets," in *Proceedings of the 25th AAAI Workshop in Human Computation*, pp. 80–83, CA, USA, 2011.

[24] M. C. Yuen, I. King, and K. S. Leung, "Probabilistic matrix factorization in task recommendation in crowdsourcing systems," in *Proceedings of the 19th International Conference on Neural Information Processing*, pp. 516–525, Springer, Doha, Qatar, 2012.

[25] D. Deng, C. Shahabi, and U. Demiryurek, "Maximizing the number of worker's self-selected tasks in spatial crowdsourcing," in *Proceedings of the 21st ACM SIGSPATIAL International Conference*, pp. 1–10, FL, USA, November 2013.

[26] J. Turian, L. Ratinov, and Y. Bengio, "Word representations: a simple and general method for semi-supervised learning," in *Proceedings of the 8th Annual Meeting of the Association for Computational Linguistics*, pp. 384–394, Uppsala, Sweden, July 2010.

[27] Y. Yao, X. Li, X. Liu et al., "Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model," *International Journal of Geographical Information Science*, vol. 31, no. 4, pp. 825–848, 2017.

[28] R. Wang, H. Zhao, B.-L. Lu, M. Utiyama, and E. Sumita, "Bilingual continuous-space language model growing for statistical machine translation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 23, no. 7, pp. 1209–1220, 2015.

[29] L. Li, G. Liu, and Q. Liu, "Advancing iterative quantization hashing using isotropic prior," in *Proceedings of the International Conference on Multimedia Modelling*, pp. 174–184, Springer International Publishing, 2016.