

RESEARCH

Open Access



A crowdsourcing method for online social networks security assessment based on human-centric computing

Zhiyong Zhang^{1*} , Junchang Jing¹, Xiaoxue Wang¹, Kim-Kwang Raymond Choo² and Brij B. Gupta³

*Correspondence:
xidianzzy@126.com

¹ Information Engineering
College, Henan University
of Science and Technology
and Henan International Joint
Laboratory of Cyberspace
Security Applications,
Luoyang 471023, People's
Republic of China
Full list of author information
is available at the end of the
article

Abstract

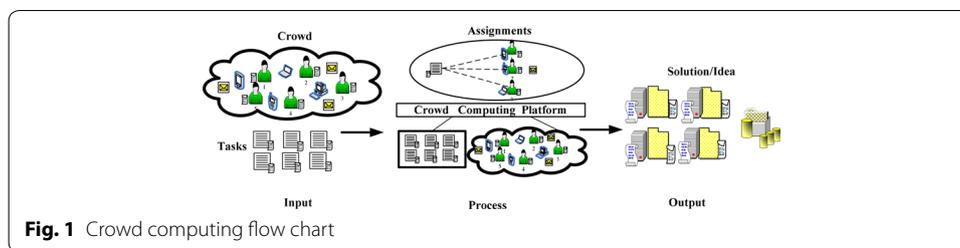
Crowdsourcing and crowd computing are a trend that is likely to be increasingly popular, and there remain a number of research and operational challenges that need to be addressed. The human-centric computational abstraction called situation may be used to cope with these difficulties. In this paper, we focus on one such challenge, which is how to assign crowd assessment tasks about security and privacy in online social networks to the most appropriate users efficiently, effectively and accurately. Specifically, here we propose a novel task assignment method to facilitate crowd assessment, which improves the security and trustworthiness of social networking platforms, as well as a task assignment algorithm based on SocialSitu, which is a social-domain-focused situational analytics. Findings from our crowd assessment experiments on a real world social network Shareteches show that the precision and recall of the proposed method and algorithm are 0.491 and 0.538 higher than those of a random algorithm's, as well as 0.336 and 0.366 higher than users' theme-aware algorithm's, respectively. Moreover, these results further suggest that our experimental evaluation enhance the security and privacy of online social networks.

Keywords: Online social networks, Security and privacy, Crowd computing, Human-centric computing, *SocialSitu*, Task assignment, Artificial intelligence

Introduction

Mobile devices are often used by online social network (OSN) users to access, share, and exchange information, whenever and wherever possible [1–3]. One relatively recent trend is crowdsourcing for information, share resources, and/or to get certain tasks done, as evidenced by the increasing number and variety of crowdsourcing applications, such as Amazon Mechanical Turks, chats (e.g., Firechat), knowledge sharing (e.g., Fig. 1), ridesharing (e.g., Uber) and accommodation sharing (e.g., Airbnb).

Crowdsourcing [4] often refers to a solution pattern which outsources the tasks that were previously performed by full-time employees to non-specific solution providers via public online platforms. Such crowdsourcing can be voluntary (unpaid) or paid. Crowd computing [5] is a computing model, which aims to integrate numerous users who may know each other (i.e., the crowd) and computing resources (i.e.,



the computer cluster) on the Internet, to handle complex tasks that are difficult to accomplish with existing computing technologies. In addition, the human-centric computational abstraction called situation can be used into Human-Embedded Computing or End-User Embedded Computing [6].

To maximize human individual or collective intelligence in crowd computing, researchers have studied task assignment that allows one to assign real-time crowd tasks to appropriate user crowd for execution [7–9]. The majority of existing task assignment approaches in the crowd computing literature consider user behaviors as homogenous [10]; thereby, completely ignoring the influence of such behaviors on the recycling results. Existing approaches also often assume static scenes [11], although tasks occur dynamically in practical applications and results should be returned immediately within a specific time. Therefore, such approaches are generally not practicable in real-world applications.

In this paper, we focus on the task itself. This is because in crowd computing, user crowds are often passive, and user multi-dimensional information, functional experience, and integrity are minimally considered. The *SocialSitu* theory [12], for example, analyzes the *SocialSitu(t)* sequence of users' access behaviors and draws out the patterns of such behaviors under different intentions (i.e., frequent functional experience). Hence, in this paper we propose a task assignment algorithm to determine user task suitability, based on *SocialSitu*. After finding user crowds who frequently use a certain function, the task of calculating task suitability and assigning the related crowd tasks based on user crowds' multi-dimensional information will become considerably accurate.

The contributions of this paper are twofold. (1) We design a task assignment method that can be used to perform crowd assessment for the security and privacy of online social networks. (2) Moreover, we propose a task assignment algorithm to determine user task suitability based on *SocialSitu*, in order to achieve efficient and accurate assignment of crowd tasks about security and privacy.

The remainder of this paper is structured as follows. “[Related work](#)” section briefly introduces the extant literature. “[Task assign method of crowd assessment for online social networks](#)” and “[Task assignment algorithm of user task suitability based on SocialSitu](#)” sections present our proposed task assignment method and algorithm, respectively. In “[Experimental setup and analysis](#)” section, we describe the crowd assessment experiments and discuss the findings which show that the proposed approach facilitates efficient task assignment and improves the task precision and recall. The last section concludes this paper.

Related work

Context-Aware (CA) [13–15] was first proposed by Schilit et al. (1994), in which a scene is defined as a location, a collection of people and objects in the vicinity, and changes in these objects. Chang et al. [16, 17] described situation analysis theory and the significance and influence of Situ architecture in software engineering. They provided a detailed description of the Situ architecture, which updates services in real time by identifying users' new intentions in software engineering; thereby, providing users with truly personalized services. To discover users' intentions in social media in a timely manner and provide additional personalized services to users, Zhang et al. [12] developed the *SocialSitu* theory based on the Situ theory of Chang et al. [16, 17], and designed a discovery method for user behavior patterns in multimedia social networks. This method analyzes users' *SocialSitu(t)* sequences and draws their behavior patterns under different intentions. Consequently, users' current intentions can be predicted by comparing their current behavior sequences with those in the database.

In crowd computing, existing task assignment structures remain relatively simple and a few cases involve tasks that are randomly assigned (i.e., randomly assigns tasks to user crowds). Therefore, most of the tasks cannot be handled by suitably qualified individuals and the advantages of crowd computing human-computer collaboration cannot be completely realized. An et al. [18] proposed an algorithm for the discovery and selection of service nodes based on the analytic hierarchy process. This takes into consideration users' mobility, complexity, real-time performance and other characteristics. In [19], Zhang proposed a simple and effective framework, MacroWiz, to manage the wisdom of crowds on mobile social networks. MacroWiz encourages online users to contribute their own knowledge or opinions through an incentive mechanism. This framework also assists task requester in collecting answers, choosing the reliable answers, and making final decisions. In the respect of spatial crowdsourcing, in order to cope with some problems about the multi-skill aware task assignment, Song et al. [20] presented Online-Exact algorithm and Online-Greedy algorithm. For the uncertain mobile crowdsourcing scenario, Guo et al. [21] found out that the results of mobile crowdsourcing largely depend on the quality of location-related users. Subsequently, Sun et al. [22] formulated an optimization problem of mobile crowdsourcing task allocation by using the trustworthiness of workers and movement distance costs. Then, they present a Markov decision process according to mobile crowdsourcing model to solve the problem of dynamic trust-aware task allocation. In [23], Mao et al. proposed an optimal user crowd algorithm. This algorithm assigns tasks to the least number of users based on users' historical task completion status; thereby, minimizing the total cost and ensuring the completion of a task. Zhang et al. [24] designed a task assignment algorithm with user theme awareness. The algorithm is designed to solve the blind randomness of random task assign algorithm and obtain users' themes (i.e., their fields of specialization) by analyzing their historical task information. In this manner, the completeness of the overall tasks is substantially improved. However, users' themes are constantly changing. Hence, obtaining these themes solely by historical data is no longer possible. Evidently, such themes are inaccurate and not real-time. In [25], Kim et al. presented a multi-layered information analysis approach based on crowdsourcing theory and effectively uses topic analysis to track scientific issues. In social networks privacy preservation and trust ways,

Yuan et al. [26] investigated privacy protection in spatial crowdsourcing and presented a privacy-preserving framework. Specifically, they also proposed a grid-based location protection method to protect the locations of workers and tasks. Moreover, Li et al. [27] exploited a secure grid-based index method to solve the problem of privacy-aware spatial crowdsourcing. This method can not only protect workers location privacy but also improve the spatial task processing time. In [28], Sharma et al. presented a novel trust relaying and privacy preservation architecture which included a distributed query system for social Internet of Things by using edge-crowdsourcing techniques. Ma et al. [29] utilized some advanced blockchain technologies to study security, privacy, and trust in the field of crowdsourcing services. Wang et al. [30] combined incentive mechanism and the techniques of location privacy-preserving to enhance the validity of mobile crowdsourcing systems. In [31], Chi et al. developed an effective location privacy protection method and solved the problem between location privacy protection and service quality, to better assess the reliability of workers for task allocation, Jiang et al. [32] developed a context-aware reliable and efficient crowdsourcing technique in simple social networks and multiple social networks. In [33], Huang et al. put forward an efficient reputation evaluation technique for crowdsourcing participant, which is based on some important machine learning methods and multidimensional evaluation index mechanism.

In summary, most current assignment methods are unable to find suitable crowds for certain tasks; consequently, leading to low completeness, precision, and recall. To solve these problems, this study defines the decision factors that affect task assignment and suitability of user tasks in online social networks. The current research completely considers users' functional experience and historical information, which are used as bases to design a task assign method of crowd assessment for the security of online social networks. This study also proposes a task assign algorithm based on *SocialSitu* user task suitability, which can assign crowd tasks efficiently and accurately in real time.

Task assign method of crowd assessment for online social networks

Crowd assessment system architecture

Crowd computing has five basic elements, namely: user crowds C , man-machine interaction M , purposeful crowd activity A , tasks T , and collective intelligence I . We use these characteristics as bases to propose a crowd computing application scene (i.e., crowd assessment) and discuss the function and architecture of different roles within scenes (see Fig. 1). Accordingly, crowd assessment has three steps, and each step is based on a crowd computing system. First, we discuss the input. Many users with certain purposes gather to form a user crowd. Interactive devices serve as media for user crowds and the system. Then, task publishers have many uncompleted tasks that should be solved by users with different backgrounds. This way, the relationship is established among the publication- \rightarrow allocation- \rightarrow execution. Second, we focus on processing. Crowd computing system analyzes the information of task collection T and user collection U and employs the task assign algorithm to select the suitable user crowd u_i for task t_j . Then, the output is discussed. After the processing of crowd computing system, the corresponding results or strategies will be obtained and data will be stored.

Three roles are mainly involved in crowd assessment, which are user collection $U=(id, o, \chi, \delta, \lambda)$, crowd computing system platform P , and task collection $T=(id, \omega,$

ϕ, θ, γ). Task publishers publish task t_1 ($t_1 \in T$) to system P. Users log into the system to query and receive task t_1 . The system finds user crowd u_1 ($u_1 \in U$) suitable for t_1 through the analysis of tasks and users. Moreover, id is the unique identifier of the two and ω is the description of t_1 , which is the general basic information. ϕ refers to the category of t_1 and θ and γ are the number of users required for t_1 and deadlines, respectively. o is the devices for user crowd u_1 , including mobile devices m and fixed devices w . χ is the field of specialization of users, δ is the historical information, such as degree of completion and degree of correlation, and λ is the users' situation information (i.e., *SocialSitu(t)* sequence).

The purpose of a crowd assessment system is to assign appropriate assessment tasks to user crowd. Accordingly, the system architecture must meet the following three basic requirements: (1) user crowd who is best suitable for some tasks that can be found; (2) the server can independently design, publish, analyze, and process the result data; and (3) interface for users to perform tasks can be provided. Therefore, a hierarchical architecture is designed to meet these requirements. Figure 2 shows the crowd assessment system architecture. The specific layered architecture is designed as follows

Social platform facilities layer

The first layer is the social platform facilities layer, where social media platform server designs and publishes crowd assessment tasks. The main function of this layer is to process data returned to it from users who have completed tasks. In addition, the user crowds and task information are simultaneously updated.

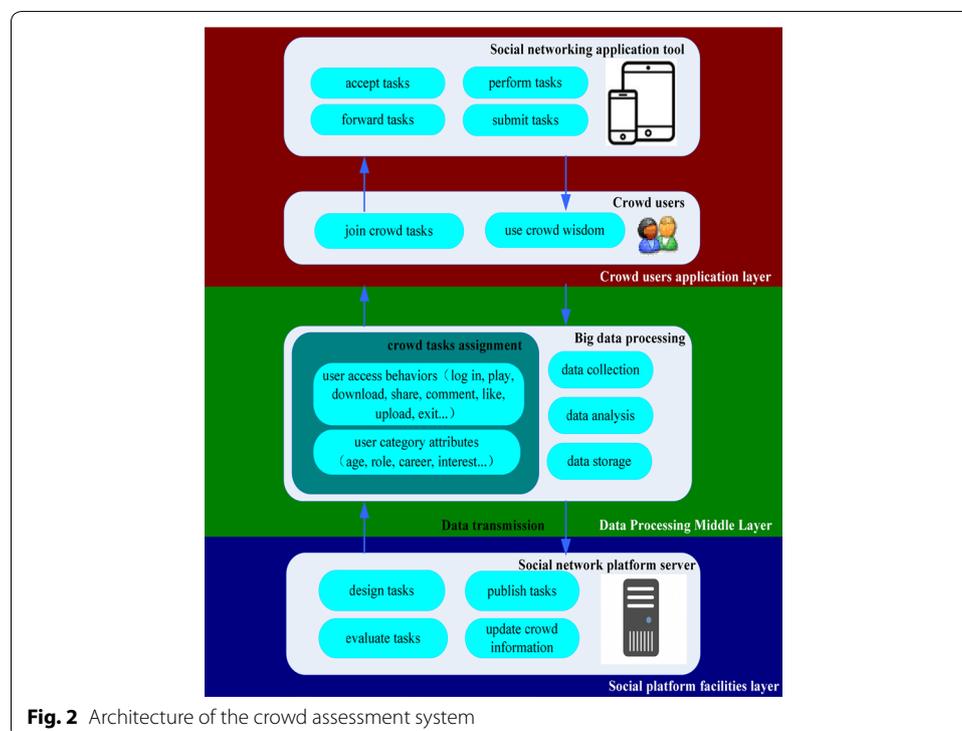


Fig. 2 Architecture of the crowd assessment system

Data processing middle layer

The second layer is the data processing middle layer. Data transmission passes the data of crowd users of social platform facilities layer to the data processing middle layer and assigns the appropriate tasks to crowd users based on the task assign method of crowd assessment proposed in this study. Moreover, this layer organizes, analyzes, and stores the result data from the crowd user application layer.

Crowd user application layer

The third layer is the crowd user application layer, which provides the corresponding user interfaces. After users accept the crowd assessment tasks, they can forward these tasks to friends who have common interests. After the tasks are executed, the result data will be returned to the data processing middle layer for processing.

Hierarchical assign structure

Due to online social network task assignment crowd selection existing many optional schemes, among the degree of completion, service response time, the degree of correlation, operation behaviors, and coincidence for each schemes have strong fuzziness. Hence, it is difficult to determine the importance of these attributes that affect decision making. Aim at the complexity and fuzziness of task assignment crowd selection, a fuzzy analytic hierarchy process [34] (FAHP) is adopted for decision-making. The FAHP breaks down elements that are constantly related to decision-making into target, criteria, attributes, and other layers. FAHP performs qualitative and quantitative analysis based on the comprehensive consideration of various factors that affect things. Accordingly, this process is a scientific assessment and decision-making method. This study defines task assignment factors and their attributes by analyzing that affects task assignment and builds a hierarchical assign structure based on this analysis.

Task assign factor

In OSNs, users' intentions may change constantly, which are ultimately reflected in changes in users' behaviors. Moreover, changes in behaviors directly reflect users' experience of certain functions. Therefore, task assignment patterns in online social networks are changed in response to such changes of users. Task assignment should consider users' service capability and focus on their functional experience. This study analyzes the operation behaviors and historical information of online social users and defines three task assign factors (TAF) that affect task assignment. This study also defines and calculates user-task-suitability (UTS) through the following three factors: service feedback quality (SFQ), operation behavior (OB), and coincidence degree (CD).

Service feedback quality

Definition 1 Service feedback quality is an objective assessment of the degree of completion of social user tasks by the platform by assigning them to completion. SFQ of user U_i is related to historical information δ of the user. Moreover, the historical service parameters of the user include completeness S_{succ} , average delay S_{del} , and degree of correlation S_{rel} .

Definition 2 Degree of completion S_{succ} is the proportion of successful completion of historical tasks by user U_i , such as Eq. (1), where N_{succ} is the number of successful historical services and N_{total} is the total number of historical services.

$$S_{succ}(U_i) = \frac{N_{succ}}{N_{total}} \tag{1}$$

Definition 3 Average delay S_{del} is the time interval between a system’s task assignment and task accepted by users, such as Eq. (2), where $d(h_j)$ is the duration historical task h_j ($h_j \in H$), $S_{succ}(U_i)$ is a task’s assigning time, and $e(U_i)$ is the time when user U_i accepted a task.

$$S_{del}(U_i) = \frac{\sum_{h_j \in H} d(h_j)}{|S(U_i) - e(U_i)|} \tag{2}$$

Definition 4 Degree of correlation S_{rel} is the degree of correlation for users’ skilled fields, such as Eq. (3), where $s(h_j)$ is the users’ degree of correlation ($h_j \in H$) of historical tasks h_j and ρ (Eq. (4)) is the attenuation factor. It is an amount that dynamically attenuates over time. That is, the longer the interval, the lesser the influence of users’ degree of correlation on service decisions.

$$S_{rel}(U_i) = \sum_{h_j \in H} \rho_{h_j} s(h_j) \tag{3}$$

$$\rho(h_j) = \begin{cases} 1, & h_j = H \\ \rho(h_j - 1) = \rho(h_j) - 1/H, & 1 \leq h_j \leq H \end{cases} \tag{4}$$

Operation behavior

Definition 5 Operation behaviors are users’ operation process in social networks. This study uses the user-intentional serialization algorithm based on situational analysis proposed in [12] to analyze the *SocialSitu(t)* sequence of users’ operation behaviors which draws out their behavior patterns under different intentions (i.e., frequent functional experience).

SocialSitu(t): Indicates the situational information at time t . *SocialSitu(t)* is a tetrad, $SocialSitu(t) = \{ID, d, A, E\}$, where d represents users’ intentions at time t (desire), A indicates users’ operations (i.e., behavior) corresponding to d at this moment, E represents

the environmental information, including terminal devices and location information, and ID represents users' identity information, including users' crowds and their roles played in the crowds.

Coincidence degree

Definition 6 Coincidence degree aims is to evaluate the consistency of the results returned by individual and overall users, as shown in Eq. (5), where R_{it} is the rating of task t by user i and \bar{R}_t is the average rating of task t by overall users. Therefore, the higher the coincidence degree is, the more consistent the rating of the user with overall users is.

$$C(U_{it}) = |R_{it} - \bar{R}_t| \quad (5)$$

User-task-suitability

Definition 7 User-task-suitability is how well a user performs a task. Suitability is calculated from the attributes and weights corresponding to the three task assign factors, namely, degree of completion, service response time, the degree of correlation, operation behaviors, and coincidence. Equation (6) shows that the greater the suitability, the more suitable the user is for this task. Accordingly, w_{ij} represents the j th attribute of user i , α_{it} represents the weight of the j th attributes of user i for task t , and p is the number of attributes. The specific weights of attributes are detailed in the next section.

$$S_{it} = \sum_{1 \leq j \leq p} \alpha_{ij}^t * w_{ij}^t \quad (6)$$

Establishment of task assign hierarchy

Hierarchy structure To determine the appropriate user for a task, users' various attributes should be considered. First, the issue should be organized and the corresponding hierarchy allocation system should be established, which includes the target, criterion and attribute layer. Thereafter, the attributes of each layer are further analyzed. The target layer aims to determine the crowd that is most suitable for a certain task. The criterion layer is the criterion for finding crowd users (i.e., factors that affect task assignment). The attribute layer is the attributes that make up each criterion layer factor. Figure 3 shows the hierarchical assignment structure.

Constructing judgment matrix at each layer Before using FAHP to determine the weight of each attribute, the importance of each layer's attributes is expressed by fuzzy triangular numbers $\tilde{a}_{ij} = (l_{ij}, m_{ij}, u_{ij})$ and a fuzzy reciprocal judgment matrix

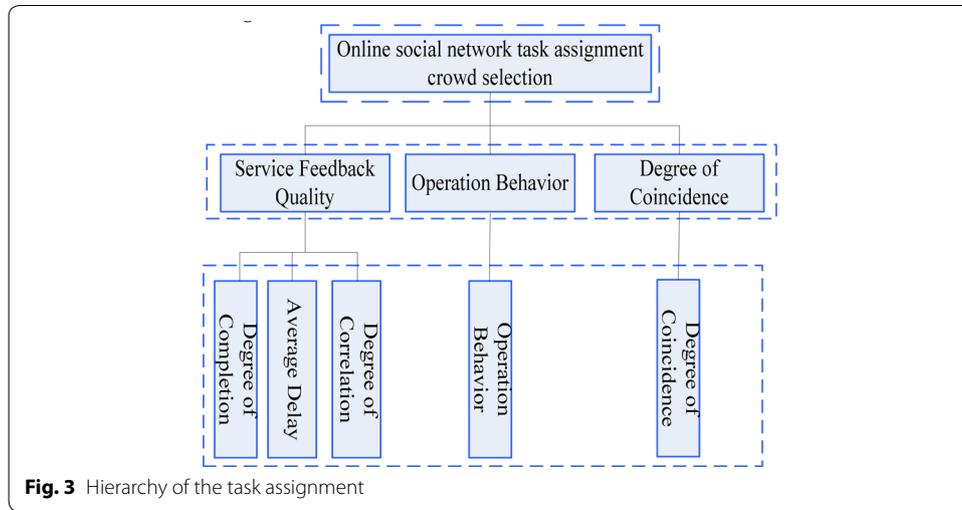


Table 1 Properties measurement table

Value range	Definition
$\tilde{a}_{ij} = 1$	Elements i and j are of the same importance to the previous factor
$\tilde{a}_{ij} = 3$	Element i is slightly more important than element j
$\tilde{a}_{ij} = 5$	Element i is more important than element j
$\tilde{a}_{ij} = 7$	Element i is considerably more important than element j
$\tilde{a}_{ij} = 9$	Element i is extremely more important than element j
$\tilde{a}_{ij} = 2n, n = 1, 2, 3, 4$	Importance of elements i and j is between $\tilde{a}_{ij} = 2n - 1$ and $\tilde{a}_{ij} = 2n + 1$
$\tilde{a}_{ji} = 1/\tilde{a}_{ij}$	Importance of elements i and j is opposite to that of elements \tilde{a}_{ij} , respectively

$$\tilde{A} = \begin{bmatrix} 1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \tilde{a}_{21} & 1 & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & \tilde{a}_{nn} \end{bmatrix}$$

can be constructed where $\tilde{a}_{ji} = 1/\tilde{a}_{ij}$. In order to facilitate operation, 1–9 and its reciprocal numbers are used as scales to determine the value of l_{ij}, m_{ij}, u_{ij} , where l_{ij}, m_{ij} and u_{ij} are the lower, the mean and the upper bounds of a triple (l_{ij}, m_{ij}, u_{ij}) , respectively. Hence, a relatively important standard seventeen meter is introduced when comparing the importance of each attribute using FAHP (Table 1).

Determination of weights and consistency check Given that the fuzzy judgment matrix A–B is reciprocal one, the non-linear programming modification of the Fuzzy Preference Programming (FPP) method that only rely on the elements of the upper right part for matrix A–B is used to estimate weights (Formula (7)). In formula (7), λ can be notated as the value of the consistency index and each w_i be expressed as weight of attribute.

maximize λ
 subject to

$$\begin{aligned}
 &(m_{ij} - l_{ij})\lambda w_j - w_i + l_{ij}w_j \leq 0, \\
 &(u_{ij} - m_{ij})\lambda w_j + w_i - u_{ij}w_j \leq 0, \\
 &\sum_{k=1}^n w_k = 1, w_k > 0, k = 1, 2, \dots, n \\
 &i = 1, 2, \dots, n - 1, j = 2, 3, \dots, n, j > i
 \end{aligned} \tag{7}$$

The detailed steps in obtaining weights of attributes are as follows.

1. The optimal solution (λ^*, w^*) of non-linear program problem for the formula (7) which includes one equality and six inequality constraints is solved.
2. The ratios of the acquired weights are calculated. If all solution ratios $\frac{w_i^*}{w_j^*}$ roughly satisfy the double-side inequalities, i.e. $l_{ij} \lesssim \frac{w_i^*}{w_j^*} \lesssim u_{ij}$, so the initial fuzzy judgements are agreeable. If not, the fuzzy judgment is unreasonable.
3. The fuzzy judgements consistency is validated. If the optimal value λ^* is positive, the consistency of fuzzy judgments is acceptable. Otherwise, these judgements are strongly inconsistent so that the fuzzy judgment matrix should be modified appropriately. Consistency detection prevents inaccurate weights because of subjective factors.
4. Lastly, by calculating the weight of each attribute in attribute layer C relative to target layer A, Suitability can be further calculated and the task can be assigned thereafter. Table 2 shows the weight values.

Task assignment algorithm of user task suitability based on SocialSitu

Data preprocessing

When a task publisher submits crowd tasks to crowd computing platforms, the quantity may be large and varied. Therefore, tasks submitted by the publisher should be processed into certain task types that can be handled by crowd computing platforms.

The daily data of user logs in online social networks are irregular and incomplete. First, data preprocessing should convert these irregular data into data formats that can be recognized by the proposed algorithm. The possible steps are as follows.

Table 2 Weights of attributes

Target layer A	Criterion layer B	Attribute layer C	C layer's weights relative to layer A
User crowd selection	B1 service feedback quality	C1 degree of completion	0.027
		C2 average delay	0.010
		C3 degree of correlation	0.067
	B2 operation behavior	C4 operation behavior	0.632
	B3 degree of coincidence	C5 degree of coincidence	0.260

1. Data cleaning: Useless and cluttered data are eliminated.
2. Data identification: This step comprises two parts: user identification and full session identification (i.e., to crowd data and select each user's completed session records).
3. Data conversion: Data are converted into formats that can be recognized by the proposed algorithm.

Specific steps of algorithm

The key issue in this study is how to accurately and efficiently assign suitable tasks to online social users. The main ideas are as follows: (1) use the *SocialSitu(t)* theory to obtain behavioral patterns under different intentions by analyzing the *SocialSitu(t)* sequence of users' behaviors, (2) find user crowd L_1 with relevant functional experience and calculate user task suitability, and (3) sort tasks by suitability based on the required number of users to form assignment collection L_2 . The assign algorithm is shown, where *SituBehaviorAnalytics* (DS , $Min_Support$, G) function refers to analysis of user behavior patterns under different intents, DS is user operation behavior data set, $Min_Support$ is minimum support, and G is users' intentions. Figure 3 shows the algorithm process. The specific steps of the algorithm are as follows.

1. Input task collection T and user collection U . Traverse unaccomplished task list $T = \{t_1, t_2, \dots, t_i\}$ and user list $U = \{u_1, u_2, \dots, u_j\}$, and obtain task t_i and information of user u_j .
2. The situation information λ of u_j is analyzed and processed and the behavior sequence patterns of user u_j are obtained through the user behavior pattern discovery algorithm *SituBehaviorAnalytics*(DS , $Min_Support$, G).
3. If the behavior sequence patterns of user u_j match the category ϕ of uncompleted task t_i , then the id of user u_j is stored in user collection L_1 of the unallocated tasks.
4. Suitability of user collection L_1 for task t_i is calculated $S_{L_m} t_i$ and is sorted from high to low.
5. The previous θ user is selected and stored in user collection L_2 .
6. Assign tasks to users in collection L_2 .
7. Repeat steps 1 to 6 until there are no newer unassigned tasks. And the detailed algorithm process is shown in Algorithm 1 and Fig. 4.

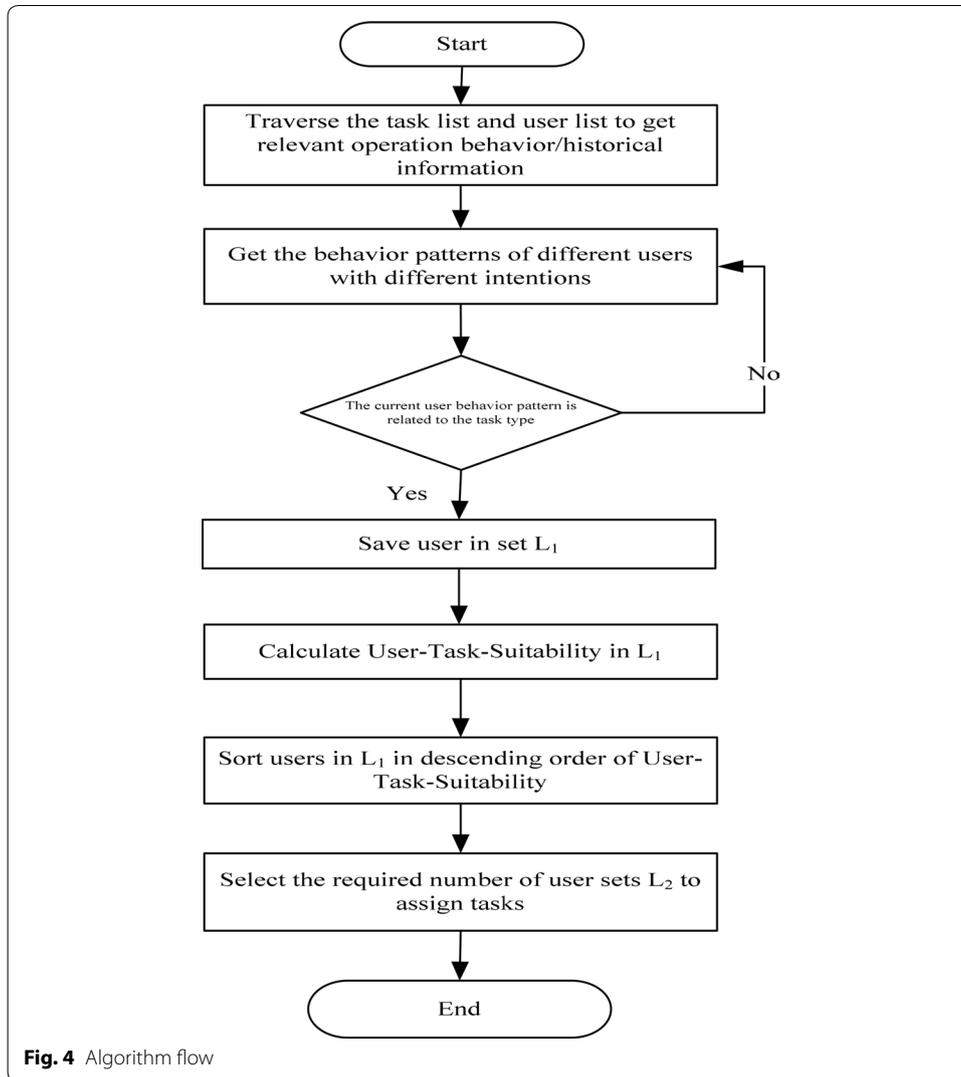


Fig. 4 Algorithm flow

Algorithm 1: Task assign algorithm

```

Input: T, U          // T represents task collections, U
                    // represents all user collections in the system
Output: L2         //Users to be assigned
1: Begin
2:for each ti ∈ T and i ← 1 to |T|
3: for each uj ∈ U and j ← 1 to |U|
4: ci.Behaviorpattern[intention]=SituBehaviorAnalytics(
DS, Min_Support, G) //Obtain behavior patterns
of different users with different intentions
5: if uj.Behaviorpattern ∈ ti. ϕ
6:   L1 = ∪ uj
7:   end if
8:   end for
9: for each Lm ∈ L1 and m ← 1 to |L1|
10: Compute SLmti //Calculate the suitability of
users Lm for tasks ti
11:endifor
12:Rank SLmti //Sort all users in collection L1 for
the suitability of tasks ti
13: Assign(ti, L2); //Task assign function
14:endifor
15:End

```

Experimental setup and analysis

The purpose of crowd task assignment is to obtain large-scale user computing resources in an online social environment. To verify the validity and correctness of the proposed algorithm, it is compared with two typical and popular algorithms, namely, random algorithm and algorithm based on user theme awareness. Random algorithm is currently the most popular allocation algorithm, which is easy to operate, has short time consumption, and

Table 3 The information of crowd assessment tasks

ID	Tasks types	Assessment tasks
1	Security and privacy	Whether platforms regularly repair security v-ulnerabilities and prom-pt them or not
2	Security and privacy	Whether platforms allow users that publish content to choose who has the authority to view or not
3	Security and privacy	Whether platforms provide and implement hierarchical digital copyright protection architecture or not
4	Security and privacy	Whether the content provided by the platform is confused with fake information or not
5	Quality assurance	Whether platforms have reliable feedback/repo-rting structure for offensive contents or not
6	Quality assurance	Whether platforms allow users to evaluate for the veracity of informat-ion that exists in the platform or not
7	Quality assurance	Whether platforms afford final state of successful or failed submission tips or not

is extensively used. Task assign algorithm based on user theme awareness detects users' themes by analyzing users' historical data to assign tasks. However, both algorithms have problems with inaccurate assignment.

Experimental environment

To analyze the correctness and effectiveness of the proposed algorithm, the self-developed technology social platform Shareteches (formerly CyVOD) [35] (<http://www.shareteches.com>) and its mobile applications are used as experimental platforms to conduct experiments and data analysis. The web server is used as task publisher and the client is used as mobile.

Experimental design

Social networks have gained widespread attention, and have been extensively applied as platforms for people to spread information on the Internet and conduct social exchange activities.

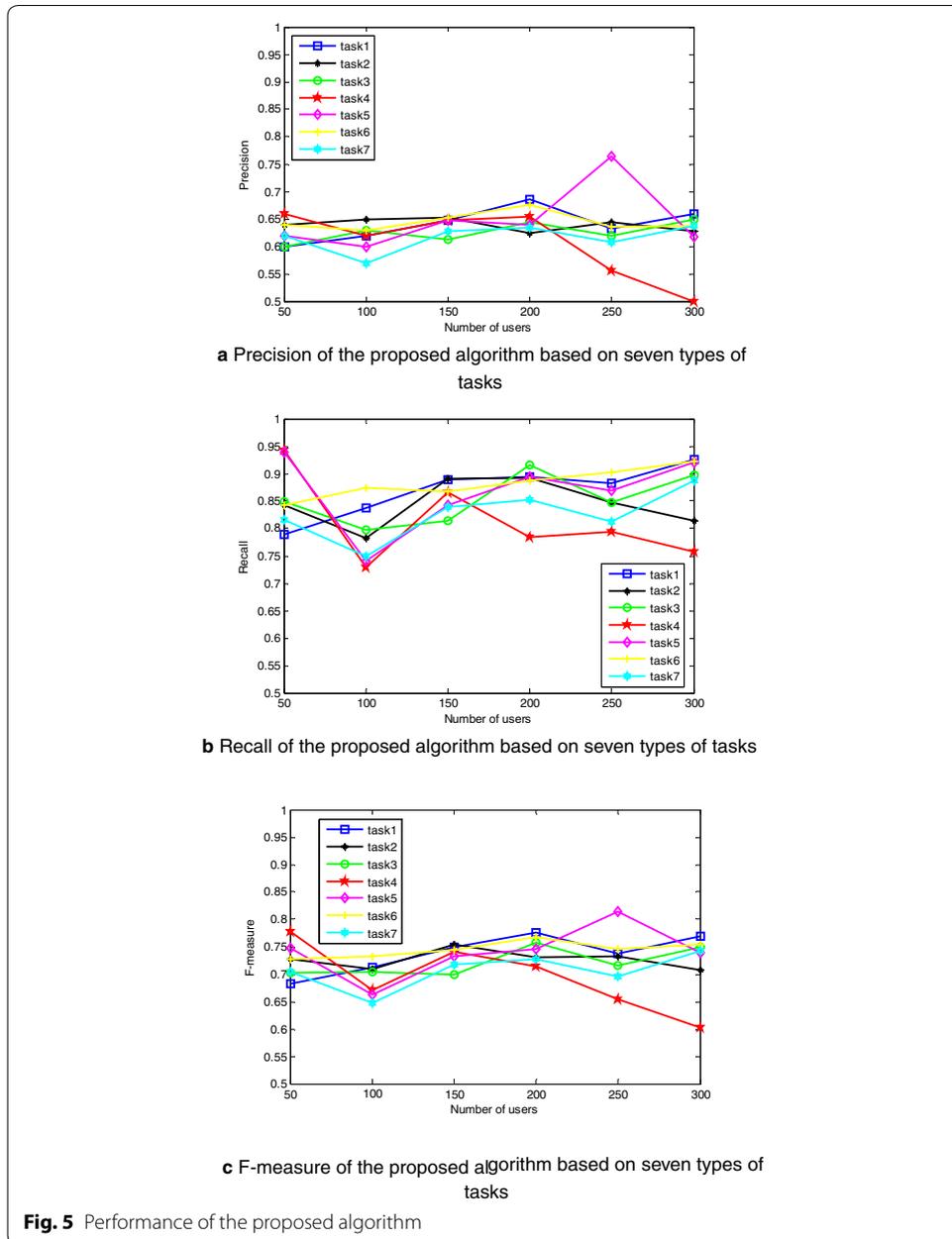
At present, security and privacy issues of social network platforms highlight the urgent need for social network users to assess platform functionality, security precautions, privacy protection, and other features [36]. According to the prevailing security and privacy issues in current social networks [37–39], this study designs seven assessment tasks of social networking platforms and evaluates security trust, functionality, and other aspects of social networking platforms. The details of crowd assessment information are showed in Table 3. In the course of the experiment, all the participants who come from Shareteches consist of ordinary users and expert users. Moreover, these users have experience of using this social networks platform and are able to perform better when they operate these assessment tasks.

Analysis of experimental results

Precision, recall, F-measure, and degree of competition are used as evaluation indicators of the algorithm. Precision and recall are two extensively used measures to assess the quality of returned results. Precision is the ratio of returned correct results to all returned results and measures the accuracy of task assignment. Recall refers to the ratio of the correct number of returned results to the total number of assignment and measures the coverage of task assignment. F-measure is the harmonic average of precision and recall. Degree of completion is the ratio of the number of returned results to the total number of assignment. We here use R_u to represent the total number of results for user u and T_u represents the total number of user u task assignments. Therefore, the formulas and notations for evaluation factors such as Precision, recall and F-measure are expressed as follows:

$$Precision = \frac{\sum_u |R_u \cap T_u|}{\sum_u |R_u|} \quad (8)$$

$$Recall = \frac{\sum_u |R_u \cap T_u|}{\sum_u |T_u|} \quad (9)$$



$$F - measure = \frac{2 * precision * recall}{precision + recall} \tag{10}$$

Algorithm correctness analysis

To reflect the correctness of proposed algorithm, it is used to assign 7 types of tasks about the security and privacy to 300 users of the social platform Shareteches. The precision and recall are shown in Fig. 5a, b. The figure shows that the precision is mainly stable between 0.60 and 0.70 and the recall is stable between 0.80 and 0.90. The F-measure is shown in Fig. 5c, which shows that the F-measure is mainly stable between 0.7 and 0.75.

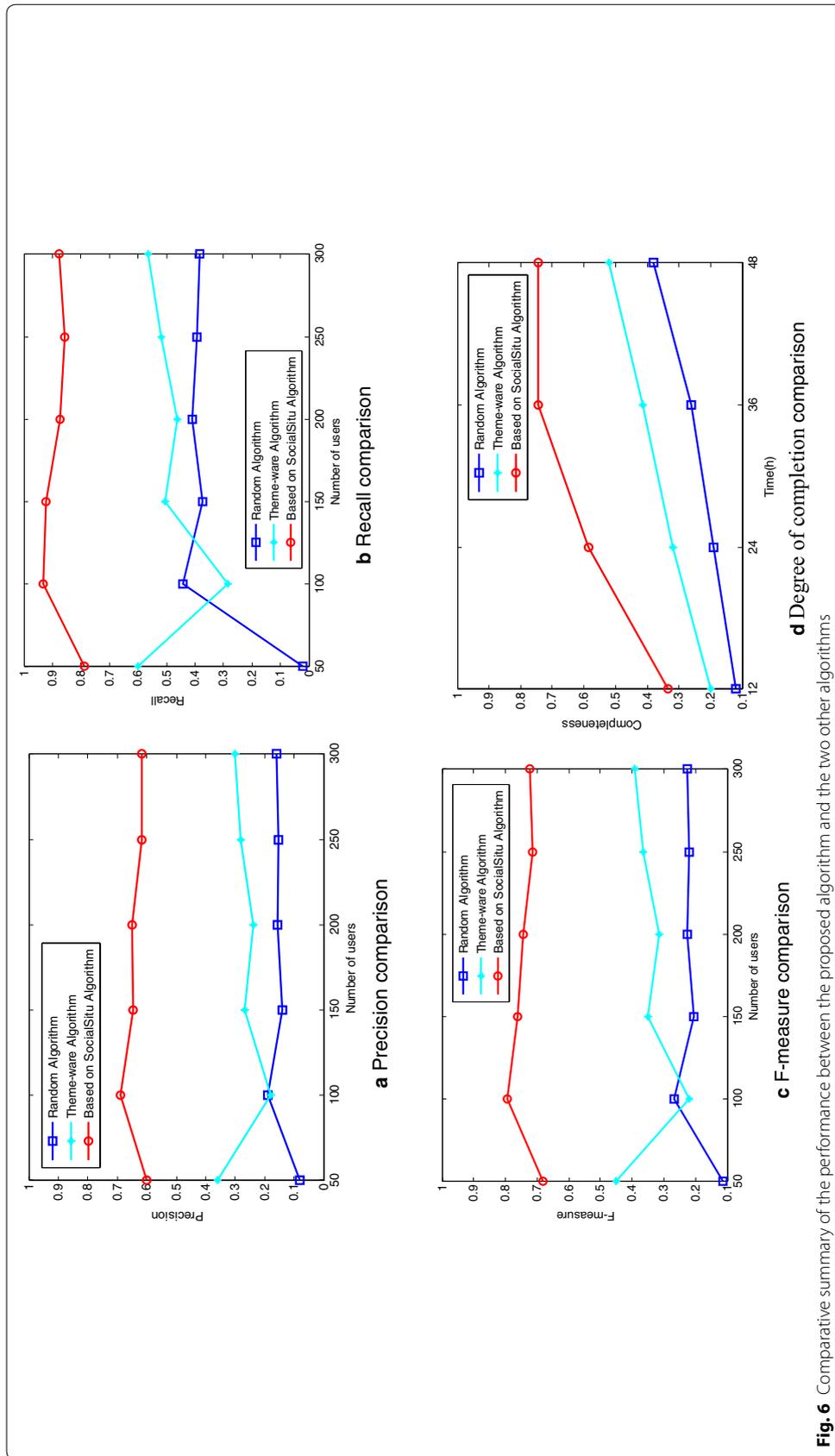


Fig. 6 Comparative summary of the performance between the proposed algorithm and the two other algorithms

Algorithm validity analysis

Task assignment was performed using random assignment, theme-aware assignment, and proposed algorithms. Precision, recall, and F-measure comparisons are shown in Fig. 6a–c. Given the increase of users, the precision and recall of the proposed and random assignment algorithms are gradually increasing and tend to be stable. Precision and recall of the theme-aware algorithm are slightly decreased and eventually tend to be stable. The precision of proposed algorithm is mainly stable at approximately 0.6, while recall is stable between 0.8 and 0.9. Both are higher than the other two algorithms. When the number of assigned users reaches 300, the F-measure of the proposed algorithm is approximately 0.3 and 0.5 higher than the other two algorithms, respectively. Moreover, the proposed algorithm has evident advantages in terms of timeliness. Figure 6d shows that the completeness of the three algorithms increases as the delay increases. The degree of completion of the proposed algorithm is higher than that of the other two algorithms with the same delay.

Conclusion

To accurately and efficiently assign assessment tasks about the security and trustworthiness of online social networks to social users, this study defined task assign factors and their attributes and user tasks suitability by analyzing users' situational information and historical records. Accordingly, this research designed a task assign method of crowd assessment for online social networks and proposed a task assign algorithm based on human-centric computational abstraction SocialSitu theory. Crowd assessment experiments were conducted on a real world social network Shareteches. The experimental results showed that the proposed method not only achieves both validity and effectiveness, but also further improves the security and trustworthiness for online social networks. In the further, we firstly mine more effective task allocation factors based on users' social behavior characteristics and content characteristics. Then, we further combine machine learning method with crowdsourcing theory to complete the security and trustworthiness assessment of social network platform.

Acknowledgements

We show gratitude to all members who have ever done researching and developing on Shareteches (formerly CyVOD) in Henan International Joint Laboratory of Cyberspace Security Applications, and would also like to thank the reviewers and editor for their valuable comments, questions and suggestions.

Authors' contributions

ZZ put forward the initial ideas and arguments. ZZ, JJ and WX contributed to writing the manuscript. ZZ, WX, KRC and BBG designed the methods and prepared figures. ZZ and JJ discussed and analyzed the results. All authors read and approved the final manuscript.

Authors' information

Zhiyong Zhang received his Master, Ph.D. degrees in Computer Science from Dalian University of Technology and Xidian University, P. R. China, respectively. He was ever post-doctoral fellowship at School of Management, Xi'an Jiaotong University, China. Nowadays, he is a full-time Henan Province Distinguished Professor and Dean with Department of Computer Science, College of Information Engineering, Henan University of Science & Technology, China. He is also a visiting professor of Computer Science Department of Iowa State University. His research interests include cyber security and computing, social big data, multimedia content security and Digital Rights Management. Recent years, he has published over 120 scientific papers and edited 6 books in the above research fields, and also holds 12 authorized patents. He is Chair of IEEE MMTTC DRMIG, IEEE Systems, Man, Cybernetics Society Technical Committee on Soft Computing, World Federation on Soft Computing Young Researchers Committee, Committeeman of China National Audio, Video, Multimedia System and Device Standardization Technologies Committee. And also, he is editorial board member and associate editor of *Multimedia Tools and Applications* (Springer), *Human-centric Computing and Information Sciences* (Springer), *IEEE Access* (IEEE), *Neural Network World*, *EURASIP Journal on Information Security* (Springer), leading guest editor or co-guest Editor of *Applied Soft Computing* (Elsevier), *Computer Journal* (Oxford) and *Future Generation Computer Systems* (Elsevier). And also, he is Chair/Co-Chair and TPC Member for numerous international conferences/ workshops on digital rights management and cloud computing security.

Junchang Jing received the B.E. degree and the M.S. degree from the College of Mathematics and Information Science, Henan Normal University, Xinxiang, China, in 2015 and 2018, respectively. He is currently pursuing the Ph.D. degree with College of Information Engineering, Henan University of Science & Technology, Luoyang, China. His research interests include social network security and computing, machine learning, and social big data.

Xiaoxue Wang received the B.E. degree and the M.S. degree from College of Information Engineering, Henan University of Science & Technology, China, in 2015 and 2018, respectively. Her research interests include crowd assessment, social network security and trustworthiness.

Kim-Kwang Raymond Choo received the Ph.D. in Information Security in 2006 from Queensland University of Technology, Australia. He currently holds the Cloud Technology Endowed Professorship at The University of Texas at San Antonio. In 2016, he was named the Cybersecurity Educator of the Year—APAC (Cybersecurity Excellence Awards are produced in cooperation with the Information Security Community on LinkedIn), and in 2015 he and his team won the Digital Forensics Research Challenge organized by Germany's University of Erlangen-Nuremberg. He is the recipient of the 2018 UTSA College of Business Col. Jean Piccione and Lt. Col. Philip Piccione Endowed Research Award for Tenured Faculty, IEEE TrustCom 2018 Best Paper Award, ESORICS 2015 Best Paper Award, 2014 Highly Commended Award by the Australia New Zealand Policing Advisory Agency, Fulbright Scholarship in 2009, 2008 Australia Day Achievement Medallion, and British Computer Society's Wilkes Award in 2008. He is a Fellow of the Australian Computer Society.

Brij B. Gupta received PhD degree from Indian Institute of Technology Roorkee, India in the area of Information and Cyber Security. In 2009, he was selected for Canadian Commonwealth Scholarship awarded by Government of Canada. He published more than 175 research papers in International Journals and Conferences of high repute including IEEE, Elsevier, ACM, Springer, Wiley, Taylor & Francis, Inderscience, etc. He has visited several countries, i.e. Canada, Japan, Malaysia, Australia, China, Hong-Kong, Italy, Spain etc to present his research work. His biography was selected and published in the 30th Edition of Marquis Who's Who in the World, 2012. Dr. Gupta also received Young Faculty research fellowship award from Ministry of Electronics and Information Technology, Government of India in 2017. He is also working as principal investigator of various R&D projects. He is serving as associate editor of IEEE Access, IEEE TII, and Executive editor of IJITCA, Inderscience, respectively. He is also serving as reviewer for Journals of IEEE, Springer, Wiley, Taylor & Francis, etc. He is also serving as guest editor of various reputed Journals. He was also visiting researcher/Professor with University of Murcia (UMU), Spain, Deakin University, Australia and Yamaguchi University, Japan in 2018, 2017 and 2015, respectively and many other universities. At present, Dr. Gupta is working as Assistant Professor in the Department of Computer Engineering, National Institute of Technology Kurukshetra India. His research interest includes Information security, Cyber Security, Cloud Computing, Web security, Intrusion detection and Phishing.

Funding

The work was sponsored by National Natural Science Foundation of China Grant No. 61972133 and 61772174, Plan For Scientific Innovation Talent of Henan Province Grant No. 174200510011. Project of Leading Talents in Science and Technology Innovation for Thousands of People Plan in Henan Province Grant No. 204200510021.

Availability of data and materials

The data that support the findings of this study are available from a real world social network Shareteches but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Shareteches.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Information Engineering College, Henan University of Science and Technology and Henan International Joint Laboratory of Cyberspace Security Applications, Luoyang 471023, People's Republic of China. ² Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA. ³ National Institute of Technology Kurukshetra, Kurukshetra 136119, India.

Received: 30 September 2019 Accepted: 25 April 2020

Published online: 02 June 2020

References

- Persia F, DAuria D (2017) A survey of online social networks: challenges and opportunities. In: IEEE international conference on information reuse and integration, pp 614–620
- Yan R, Li D, Wu W, Du D, Wang Y (2019) Minimizing influence of rumors by blockers on social networks: algorithms and analysis. *IEEE Trans Netw Sci Eng.* <https://doi.org/10.1109/TNSE.2019.2903272>
- Roy D, Lotan G, Zeng W (2015) The attention automaton: sensing collective user interests in social network communities. *IEEE Trans Netw Sci Eng* 2(1):40–52
- Howe J (2006) The rise of crowdsourcing. *Wired Mag* 14(14):1–5
- Parshotam K (2013) Crowd computing: a literature review and definition. In: South African Institute for Computer Scientists & Information Technologists Conference, pp 121–130
- Chang CK (2018) Situation analytics-at the dawn of a new software engineering paradigm. *Sci China Inform Sci* 61(5):050101:1–050101:14
- Xing Y, Wang L, Li Z (2019) Multi-Attribute crowdsourcing task assignment with stability and satisfactory. *IEEE Access* 7:133351–133361

8. Tran L, To H, Fan L (2018) A real-time framework for task assignment in hyperlocal spatial crowdsourcing. *ACM Trans Intell Syst Technol* 9(3):37:1–37:26
9. Alireza S, Shafiqheh H, Masoud RA (2018) Personality classification based on profiles of social networks' users and the five-factor model of personality. *Hum-Cent Comput Inform Sci* 8(1):24–38
10. Karger D R, Oh S, Shah D (2011) Iterative learning for reliable crowdsourcing systems. In: *International conference on neural information processing systems*, pp 1953–1961
11. Zhang X, Yang Z, Wu C, Sun W, Liu Y, Liu K (2014) Robust trajectory estimation for crowdsourcing-based mobile applications. *IEEE Trans Parallel Distrib Syst* 25(7):1876–1885
12. Zhang Z, Sun R, Wang X, Zhao C (2019) A situational analytic method for user behavior pattern in multimedia social networks. *IEEE Trans Big Data* 5(4):520–528
13. Srinivasan K, Agrawal P, Arya R, Akhtar N, Pengoria D, Gonsalves TA (2012) Context-aware, QoE-driven adaptation of multimedia services. In: *5th international conference on mobile wireless middleware, operating systems, and applications*, pp 236–249
14. Tekin C, Van Der Schaar M (2015) Contextual online learning for multimedia content aggregation. *IEEE Trans Multimed* 17(4):549–561
15. Bulterman D, Cesar P, Guimaraes R (2013) Socially-aware multimedia authoring: past, present, and future. *ACM Trans Multimed Comput Commun Appl* 9(1):35:1–35:23
16. Chang C, Jiang H, Ming H, Oyama K (2009) Situ: a situation-theoretic approach to context-aware service evolution. *IEEE Trans Serv Comput* 2(3):261–275
17. Chang C (2016) Situation analytics—a foundation for a new software engineering paradigm. *Computer* 49(1):24–33
18. An J, Gui L, He Q, Wu B (2015) Crowdsourcing assignment mechanism based on AHP in mobile crowd sensing. *J Beijing Univ Posts Telecommun* 38(5):37–41
19. Zhang X, Shang L, Yuan Y (2017) A crowd wisdom management framework for crowdsourcing systems. *IEEE Access* 4:9764–9774
20. Song T, Xu K, Li J et al (2019) Multi-skill aware task assignment in real-time spatial crowdsourcing. *Geoinformatica* 24:153–173
21. Guo B, Liu Y, Wang L (2018) Task allocation in spatial crowdsourcing: current state and future directions. *IEEE Internet Things J* 5(3):1749–1764
22. Sun Y, Tan W (2019) A trust-aware task allocation method using deep q-learning for uncertain mobile crowdsourcing. *Hum-Cent Comput Inform Sci* 9(1):1–27
23. Mao H (2016) OCC: Opportunistic crowd computing in mobile social networks. *Database systems for advanced applications*. Springer, Berlin
24. Zhang X, Li G, Feng J (2015) Theme-aware task assignment in crowd computing on big Data. *J Comput Res Dev* 52(2):309–317
25. Kim M, Gupta BB, Rho S (2018) Crowdsourcing based scientific issue tracking with topic analysis. *Appl Soft Comput*. <https://doi.org/10.1016/j.asoc.2017.09.028>
26. Yuan D, Li Q, Li G et al (2020) PriRadar: a privacy-preserving framework for spatial crowdsourcing. *IEEE Trans Inform Forensics Secur* 15:299–314
27. Li Y, Yi G, Shin B (2019) Spatial task management method for location privacy aware crowdsourcing. *Cluster Comput* 22:1797–1803
28. Sharma V, You I, Jayakody DNK et al (2019) Cooperative trust relaying and privacy preservation via edge-crowdsourcing in social Internet of Things. *Fut Gener Comput Syst* 92:758–776
29. Ma Y, Sun Y, Lei Y et al (2019) A survey of blockchain technology on security, privacy, and trust in crowdsourcing services. *World Wide Web*. <https://doi.org/10.1007/s11280-019-00735-4>
30. Wang Y, Cai Z, Tong X et al (2018) Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems. *Comput Netw* 135:32–43
31. Chi Z, Wang Y, Huang Y, Tong X (2018) The novel location privacy-preserving CKD for mobile crowdsourcing systems. *IEEE Access* 6:5678–5687
32. Jiang J, An B, Jiang Y, Lin D (2020) Context-aware reliable crowdsourcing in social networks. *IEEE Trans Syst Man Cybern Syst* 50(2):617–632
33. Huang Y, Chen M (2019) Improve reputation evaluation of crowdsourcing participants using multidimensional index and machine learning techniques. *IEEE Access* 7:118055–118067
34. Mikhailov L, Tsvetinov P (2004) Evaluation of services using a fuzzy analytic hierarchy process. *Appl Soft Comput* 5(1):23–33
35. Zhang Z, Sun R, Zhao C, Wang J, Chang C, Gupta B (2017) CyVOD: a novel trinity multimedia social network scheme. *Multimed Tools Appl* 76(18):18513–18529
36. Zhang Z, Wen J, Wang X, Zhao C (2018) A novel crowd evaluation method for security and trustworthiness of online social networks platforms based on signaling theory. *J Comput Sci* 26:468–477
37. Zhang Z, Gupta B (2018) Social media security and trustworthiness: overview and new direction. *Fut Gener Comput Syst* 86:914–925
38. Oukemeni S, Rifa-Pous H (2019) Privacy analysis on microblogging online social networks: a survey. *ACM Comput Surv* 52(3):1–20
39. Souri A, Hosseini R (2018) A state-of-the-art survey of malware detection approaches using data mining techniques. *Hum-Cent Comput Inform Sci* 8(1):3–25

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.