

Attribute based Weighted Mechanism for Community Detection in Social Internet of Things

Subash Rajendran (✉ subashrajendran.2021@gmail.com)

SRM Institute of Science and Technology

Jebakumar R

SRM Institute of Science and Technology

Research Article

Keywords: community detection, content spreading, credential, social attributes, Social Internet of Things, weighted mechanism

Posted Date: July 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-652736/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Attribute based Weighted Mechanism for Community Detection in Social Internet of Things

^{1*}Subash Rajendran, ²Jebakumar R

^{1*,2}Department of Computer Science and Engineering, College of Engineering and

Technology, SRM Institute of Science and Technology,

SRM Nagar, Kattankulathur, 603203, Kancheepuram, Chennai, India

^{1*}Corresponding Author Email: subashrajendran.2021@gmail.com

Abstract

A new paradigm of Internet of Things (IoT) is emerging rapidly by socializing the smarter physical devices called as Social Internet of Things (SIoT). Social relationships established between these objects make them autonomously connected for services, without any human intervention. Since SIoT is a large-scale network with huge data involved, the content spreading behaviour need to be exploited. In order to ensure the growth of the content spread, the large-scale SIoT network is divided into several communities based on the social attributes in this work. We first divided the SIoT network into high quality Sociality based Weighted Communities (SWC). Social attributes like user preferences, social similarities, and mutual friends' degrees are main metrics for achieving the best rate function. The weighted method based on these social attributes determine the nodes to be present in their respective communities. Also, the controlling of the local community augmentation using cluster concepts is done in our approach. Finally, a Credential Acclaimed Information Spreading (CAIS) mechanism is proposed which selects the best node with the maximum credential to surge the content spreading behaviour in the detected communities of SIoT network. The proposed social-driven attribute based weighted mechanism for community detection is validated using

three diverse real-world datasets such as CASAS, MIT and ARAS dataset containing 427 sensors. Investigational outcomes validate that the overall performance of the proposed method overwhelms the conventional community detection algorithms like Louvain, Girvan Newman, Bron Kerbosch, Infomax and the recent state-of-art-approaches interms of spreading outcomes, NMI, modularity, F-measure, precision, recall and computational time.

Keywords: community detection; content spreading; credential; social attributes; Social Internet of Things; weighted mechanism.

1. Introduction

With the massive growth in technology and by incorporating smartness in the physical devices, Internet of Things (IoT) marches the real-world communication. The ultimate aim of IoT is to deliver livelihood for mankind by performing similar services. Progression of IoT intends to be a boon to the broad variety of various application fields by enabling smarter device to device communication [1]. Implementing the social perceptions in the IoT paradigm, unveils a new era called Social Internet of Things (SIoT) [2]. The SIoT model indicates an environment which consents humans and smart devices to interact within a society through numerous varieties of relationships. SIoT uplifts the socialization of the smarter devices connected with one another without any human intervention [3]. SIoT devices are proficient of constructing any number of friends, which in turn to form several communities based on the interested attributes of the devices involved. Therefore, detecting and characterizing the large-scale SIoT network into various communities is imperative for the better service discovery. The nodes are classified and grouped as several communities which serve as the elementary component of SIoT networks. The leaders of these communities effectively influence the other nodes present.

The main two steps in the community detection is the identification of promising leaders in the SIoT network and examination of nodes likenesses to construct various communities.

Usually, community detection in networks is based on the several conventional community detection algorithms like Louvain, Girvan Newman and Bron Kerbosch. Louvain algorithm detects the disjoint communities in a directed social network using greedy optimization of modularity. Louvain algorithm tends to be one of the fastest community detection algorithms with a high modularity score. But this method is not suited for the smaller communities which limits the resolution [4]. Bron Kerbosch algorithm works goods for unweighted undirected graph for discovering the overlapping communities. This algorithm computes the utmost cliques by hunting for the perfectly linked nodes of a network. The major problem with Bron Kerbosch algorithm is that it does not hold good for output sensitive problems [5]. Girvan Newman algorithm eliminates the devices with the highest number of the shortest routes between the devices. The edges of the devices that are joining the other devices in a community is maintained to have maximum betweenness. The major problem with Girvan Newman algorithm is that, it is not suited for detecting the communities of huge and complex network structures [6]. Rosvall et al proposed the Infomap algorithm which detects the communities by engaging arbitrary strides to evaluate the content spreading behaviour in the networks. This algorithm does the encoding the content in the network as an encoded graph via a restricted channel. Finally, the original graph is decoded by constructing the set of probable participants. Lesser the number of participants, larger the content about the network, therefor, nor suited for several participants in a network [7].

Paper is planned as follows: section 2 precises the related works; the proposed SWC detection model is discoursed in section 3; CAIS mechanism is explained in section 4; experimental evaluation is discussed in section 5; the results and discussions are elaborated in section 6 and finally the conclusion and the proposed work is offered in section 7.

2. Related Works

Nowadays, these conventional community detection algorithms are used for detecting the communities of SIoT networks too. Khanfor et al explored the application of Louvain and Bron Kerbosch algorithms in the field of community detection in SIoT networks. The authors implemented these conventional algorithms by representing SIoT network as large-scale graph. These graphs are divided as overlapping and non-overlapping communities on the basis of various types of SIoT relationships such as Ownership Device Relationship (ODR), Social Device Relationship (SDR), Co-Location Device Relationship (CLDR) and Co-Work Device Relationship (CWDR) [8]. Khanfor et al extended his exploration of the conventional community detection algorithms by applying Order Statistics Local Method (OSLOM) algorithm for monitoring and reducing the difficulties in the traits of the service discovery. Natural Language Processing (NLP) is used for analyzing the text data contents of real world SIoT networks. Though this approach is well suited for monitoring the crowdsourcing applications, it involves the handling of only the textual information [9]. Shanbhaq implemented the variation of Louvain algorithm for the faster and accurate community detection in the large-scale networks with numerous granularity points. But the work handled on reducing the execution time with discussing the quality of partitions [10]. Zhang et al proposed a community detection technique to divide the SIoT network into many superiority communities and then the content spreading is maximized via two phases such as candidate and greedy phases to select the best candidate for the maximum content sharing. But the influence of node content on the information spread is ignored [11]. Incorporation of Louvain algorithm with the fuzzy network is proposed for finding the communities in a SIoT network. Shapely index is used as the primary degree for obtaining the fuzzy measures [12]. Liu et al presented the progresses in the community detection via deep learning networks. Deep learning

models learn the pattern of nodes, neighborhoods, and subgraphs present in their respective communities of the real-world scenario. Currently convolutional neural networks (CNN), autoencoders and generative adversarial networks (GAN) are mostly used for the community detection but the following are the gap between deep learning networks and community detection: detection and recognition of the spatial variations among various communities are not done, combination of temporal based information and spatial content-based information are yet to be learned in these deep networks [13]

Though these algorithms are suited for social networks, but really not appropriate for SIoT networks. In this work, we aim to divide the SIoT network into several smaller communities and maximize the content spread among these communities. The major contributions presented in our work are the following:

1. A Sociality based Weighted Community detection (SWC) algorithm for dividing the SIoT network into high quality smaller communities is developed.
2. An effective mechanism for maximizing the content spreading behaviour among the detected communities is proposed via a Credential Acclaimed Information Spreading (CAIS) strategy.
3. The suggested model is estimated on three different datasets like ARAS, MIT and CASAS datasets.
4. To end, the performance of the proposed attribute-based community detection is compared with various available approaches.

3. Social Attribute based Community Detection

From the research studies, a SIoT network is a random large-scale network containing several nodes of diverse relationships. Generally, SIoT networks are represented in terms of weighted graphs by including the social belongings of the links between nodes. In our work, we used a weighted method which is based on the social attributes such as user preferences,

social similarities, degree of the mutual friends are for achieving the best rate function. If two nodes are connected to a node with a lesser degree, then those nodes behave with the linked characteristics. This is same like, when few persons discuss on an uncommon theme, then those persons are with the similar interests. Therefore, it is clear that in a SIoT network, content spreading behaviour can be maximized only if it is divided into several smaller communities. Our aim is to divide a large SIoT network into several small communities based on the communication relationships between the nodes present in each community.

3.1 Sociality based Weighted Community (SWC) Division Mechanism

Let us consider that our SIoT network possesses only local structures with a sub graph S containing few nodes in it. Hence, we initially choose local clusters and then the size of these clusters is increased consistently by choosing the nodes with the best rate function. Let φ regulates the number of communities to be formed and it is always a positive integer. When $\varphi = 0 \geq 0.5$, it forms only one community and when $\varphi > 2$, it forms several communities. Since, we pursue to form several smaller communities, we chose $\varphi = 1$.

Our algorithm uses the social attributes such as user preferences, social similarities, highest degree and maximum mutual friends are for achieving the best rate function. Here, the social similarities which obtained from the user preferences are considered to be Direct Intimacy (DI) and degree of the mutual friends are considered to be Indirect Intimacy (II). Let us consider two nodes, a and b . The DI is measured as the summation of the weights associated to node a multiplied by the summation of the weights associated to node b . When there is no social similarity between the nodes a and b , then $DI = 1$. The II is measured as the reciprocal of the mutual friends' degrees. This DI and II contribute to the weights of the edge between the respective nodes, $w = \delta(DI) + [1 - \delta(II)]$. Here δ is balancing constraint between DI and II . Usually, $0 > \delta > 1$ and we selected $\delta = 0.8$.

Consider a node N to sub graph S , the difference in the best rate function to S with and without N is equal to the best rate function of S with node N added with the best rate functions of S without node N . If the difference in the best rate function plus the summation of the weights to S with and without N is greater than zero, this specifies that node N improved the rate function of its community fitness and therefore it is then added to sub graph S . If the difference in the best rate function plus the summation of the weights to S with and without N is lesser than zero, this specifies that node N reduced the rate function of its community fitness and therefore it is not joined to sub graph S . Thus, only if a node makes some improvement to the community fitness function, it joins that respective community. Thus, this method is suitable for exploring the overlapping community detection by means of controlling the number of communities via adjusting the value of.

3.2 Local Community Augmentation Control

SIoT network contains huge amount of triangular assemblies, which emerges several smaller communities. appear in many social networks. So, it is important to control the local community expansions to maintain the reliability of the community structures. This is achieved via clustering coefficient concept. Usually, clustering coefficient lies in the range of $[0,1]$ and the value of the clustering coefficient of a community keep increasing when the size of the network starts increasing. We aim to find several smaller and articulate clusters from the entire SIoT network. The clustering coefficient augments the estimation of local network features, and computes the number of triangles in its respective community. The reliability of the community structures is estimated from these clustering coefficients. The reliability of node N to the community S is equal to the sum of stability of S with node N added with the stability of S without node N . The reliability of node N is used to estimate the influence of N to the reliability of the community S . Arbitrarily a node is deleted from the entire SIoT network and then the Divergence (D) in the reliability of the community S for the respective scenario is

noted down. If the value of $D > (-0.05)$, then the respective node positively influences the community S and that node is added to S . If the value of $D < (-0.05)$, then the respective node negatively influences the community S and that node is not added to S . Therefore, if the community can be more reliable, the node is added to S . Else, even if the node is with the best fitness function, it is not added to S .

3.3 SWC Algorithm

Our proposed Sociality based Weighted Community (SWC) division algorithm discover smaller and evocative clusters from a random large-scale SIoT network containing several nodes of diverse relationships. In our work, a weighted method which is based on the social attributes obtained from the user preferences (DI) and degree of the mutual friends (II) are used as weights to the links in SIoT network for achieving the best rate function. Then, the regularization of the local cluster formation is done using a reliability parameter. This parameter helps to form several smaller and meaningful communities. The entire SWC algorithm is described in the following steps of Algorithm 1

Algorithm 1: Sociality based Weighted Community (SWC) division algorithm

Begin

Consider a SIoT network possessing only local structures with a sub graph S and choose the value for φ .
--

Select a random node N , which does not belong to any of the local communities.

Use the Direct Intimacy (DI) and Indirect Intimacy (II) functions to compute the rate function of the neighboring.
--

Choose a node a with the best rate function. If the best rate function is nonnegative, estimate the reliability of node a comparative to community S for sub graph S .

If the best rate function is negative, then repeat from step 3.

If the reliability of node a comparative to the local community S is superior than -0.05 , then a will be added in S , creating a larger local community; else, repeat from step 3.

Recompute the best rate function and reliability of each node.

If any of a node possess a negative rate function and its reliability of the same does not satisfy the constraint, then discard that node from the larger local community, which in turn generates new sub graph.

If step 9 occurs, repeat from step 8 again.

If all nodes fulfil the constraint of rate function and its reliability, then return to step 4.

End

4. Credential Acclaim based Information Spreading (CAIS) Mechanism

After discovering several communities and regularizing it from a SIoT environment, next our aim is to increase the information spreading quality of then communities formed. It can be achieved only by electing a leader, who is capable of spreading the maximum information among the nodes of their respective community. In this work, we used Credential Acclaim based Information Spreading (CAIS) mechanism which is solely based on allotting credentials for the best performing nodes with high popularity. The best performing with high popularity nodes are chosen using a recent and most efficient metaheuristic algorithm called Sail Fish Optimizer (SFO). SFO is a nature inspired algorithm, molded by the inspiration on the attacking strategy of the group hunting predators (sail fishes) on their prey (sardines). The algorithm is mathematically modeled by extracting the natural mannerisms of sail fish and

sardines. It contains four phases such as: initialization and estimation of fitness, elitism, strategy of attack, stalking and catching the prey.

The population of sail fish and sardines are initialized randomly. The current position of sail fishes is marked using $sf_{current}$ and sardines as $s_{current}$, as shown in the equations (1) and (2), where v indicates the number of variables, i indicates the number of sail fishes and j indicates the number of sardines.

$$sf_{current} = \begin{bmatrix} sf_{1,1} & sf_{1,2} & \cdots & sf_{1,v} \\ sf_{2,1} & sf_{2,2} & \cdots & sf_{2,v} \\ \vdots & \vdots & \cdots & \vdots \\ sf_{i,1} & sf_{i,2} & \cdots & sf_{i,v} \end{bmatrix} \quad (1)$$

$$s_{current} = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,v} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,v} \\ \vdots & \vdots & \cdots & \vdots \\ s_{j,1} & s_{j,2} & \cdots & s_{j,v} \end{bmatrix} \quad (2)$$

Then the fitness of the sail fishes and sardines are estimated using the equations (3) and (4), f indicates the objective fitness function. sf_{fit} saves the value of the sail fish fitness and s_{fit} saves the value of the sardine fitness functions respectively

$$sf_{fit} = \begin{bmatrix} f(sf_{1,1} & sf_{1,2} & \cdots & sf_{1,v}) \\ f(sf_{2,1} & sf_{2,2} & \cdots & sf_{2,v}) \\ \vdots & \vdots & \cdots & \vdots \\ f(sf_{i,1} & sf_{i,2} & \cdots & sf_{i,v}) \end{bmatrix} = \begin{bmatrix} F_{sf_1} \\ F_{sf_2} \\ \vdots \\ F_{sf_i} \end{bmatrix} \quad (3)$$

$$s_{fit} = \begin{bmatrix} f(s_{1,1} & s_{1,2} & \cdots & s_{1,v}) \\ f(s_{2,1} & s_{2,2} & \cdots & s_{2,v}) \\ \vdots & \vdots & \cdots & \vdots \\ f(s_{j,1} & s_{j,2} & \cdots & s_{j,v}) \end{bmatrix} = \begin{bmatrix} F_{s_1} \\ F_{s_2} \\ \vdots \\ F_{s_j} \end{bmatrix} \quad (4)$$

The position of the best sail fish and injured sardine with the best fitness value is saved in each iteration and considered as the elite, the position of sail fishes and sardines are then updated towards the best solution as given in the equations (5) and (6).

$$Y_{sf_{new}}^i = Y_{sf_{elite}}^i - \alpha_i \left(rand(0,1) \left(Y_{sf_{elite}}^i + Y_{s_{injured}}^i / 2 \right) - Y_{sf_{old}}^i \right) \quad (5)$$

$$Y_{s_{new}}^i = k \left(Y_{sf_{elite}}^i - Y_{s_{old}}^i + ap \right) \quad (6)$$

In the equations (5) and (6), $Y_{sf_{elite}}^i$ indicates the best position of the elite sail fish which has the best fitness value and $Y_{s_{injured}}^i$ indicates the best position of the injured sardine fish. $Y_{sf_{old}}^i$ and $Y_{s_{old}}^i$ are the old positions of the sail fish and sardine fish respectively. $\alpha_i = 2 \times rand(0,1) \times DP - DP$, here DP is the density of the sardine prey and it is given as $DP = 1 - (N_{sf}/N_{sf} + N_s)$. Here, N_{sf} and N_s are the number of sail fishes and sardine fishes respectively. In the equation (6), k represents the numbers between 0 and 1 and ap indicates the attack power, which is given as $ap = \sigma \times (1 - (2 \times Itr \times \tau))$. Here $\sigma = 4$ and $\tau = 0.001$. The number of sail fish which updates their position is $\mu = N_s \times ap$ and the number of sardine fish which updates their position is $\omega = n_i \times ap$. Here n_i indicates the number of variables.

Then estimate the fitness of the sardine fishes, if there is a better fitness solution in sardine, then that injured sardine is replaced with the elite sail fish as given in the equation (7)

$$Y_{sf}^i = Y_s^i ; \text{ if } f(s) > f(sf) \quad (7)$$

Such best fitted sail fishes are given higher credentials and are considered as the leader of the other nodes in its respective community. The content of the information is spread using the elected leader via SFO algorithm. The entire CAIS algorithm is described in the following steps of Algorithm 2

Algorithm 2: Credential Acclaim based Information Spreading (CAIS) mechanism

Inputs: Nodes N present in a community of SIoT network. N_{sf} and N_s are the number of sail fishes and sardine fishes respectively

Output: Elite sail fish and injured sardine as the leader node

Initialization: Consider a random population of sail fish and sardine as given in the equations 1 and 2, $\sigma = 4$ and $\tau = 0.001$

Begin

Calculate the fitness of the sail fishes and sardines using the equations (3) and (4)

Calculate $DP = 1 - (N_{sf}/N_{sf} + N_s)$

Calculate $\alpha_i = 2 \times rand(0,1) \times DP - DP$

Update the position of sail fish using the equation (5)

Estimate the attack power, $ap = \sigma \times (1 - (2 \times Itr \times \tau))$

Update the position of all the sardine fishes using the equation (6), if $ap > 0.5$

Calculate $\mu = N_s \times ap$ and $\omega = n_i \times ap$ if $ap < 0.5$

Update the position of selected sail fish using the equation (5)

Calculate the fitness of all the sardines using the equation (4)

Replace elite sail fish by injured sardine using the equation (7)

Remove the hunted sardine fish from the population

Update the best sail fish and sardine

Give high credentials to such best sail fish and sardine fishes.

Choose the highest credentialed fish/node as the leader for its own community

Spread the content of the information via such an elected leader

End

5. Experimental Evaluation

To prove the versatility of our proposed model, we validated it using three different real world datasets are Center for Advanced Studies in Adaptive Systems (CASAS) [15], Massachusetts Institute of Technology (MIT) [16] and Activity Recognition with Ambient Sensing (ARAS) [17] for recognising the actions using machine learning. We tested our model on 16 subjects. Totally 15791 number of actions are collected from 427 sensors which are preinstalled in 11 flats. The details of the datasets used in our work is shown in the Table 1.

Table 1: Three Datasets

Dataset	Flat No.	No. of Subjects	No. of Sensors	No. of Actions
CASAS	1	1	32	5714
	2	1	32	4320
	3	1	32	3361
	4	2	27	600
	5	2	51	37
	6	1	32	1513
	7	2	20	166
MIT	1	1	77	13
	2	1	84	13
ARAS	1	2	20	27
	2	2	20	27

Total	11	16	427	15791
-------	----	----	-----	-------

5.1 Experimental Setup

SWC and CAIS algorithm is implemented in C++ platform and for simulation NS3 tool is used. Every plot is earmarked as an average of 10 runs for each experiment. Sensors are deployed in each portion of the flat for tracking the actions of the contributors. Obtained sensor data confined the following items such as sensor id, information, start and end time of every actions. CASAS dataset comprises the information obtained from 7 smart flats, containing a total of 10 occupants staying for about 14 days in it. All the flats are preinstalled with the 226 sensors on the doors, windows, furniture and appliances for acquiring 15711 actions. MIT dataset comprises the information obtained from smart flats, each containing a single occupant for a period of one month. Both the flats are deployed with the 161 sensors on the doors, windows, furniture and appliances for acquiring 35 actions. ARAS dataset comprehends the data attained from 2 smart communities each comprising 2 inhabitants. Both the communities are preinstalled with 40 types of sensors for gaining about 54 actions. Entirely 427 sensors are considered as one single SIoT network, which are aimed to be divided into number of communities based on the preferences and social similarities.

6. Results and Discussions

The most commonly used metrics in the evaluation of community detection algorithms are influence speed, Normalized Mutual Information (NMI), modularity, F-measure, precision, recall and computation time [17].

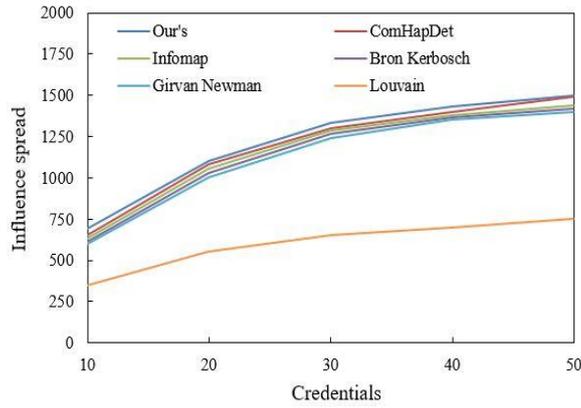


Fig. 1. Influence Spread plot for different community detection algorithms

Three datasets such as CASAS, MIT and ARAS for the assessment of attribute based weighted model. For the validation purpose, we compared the performance of the proposed model with five different community detection algorithms such as Infomap, Girvan Newman, ComHapDet, Bron Kerbosch and Louvain.

Influence spread is maximized based on the credentials acquired by the nodes which are elected as the leader via SFO algorithm. Fig. 1. depicts the influence spread plot for different community detection algorithms. It is evident that the information spread is maximum for our proposed attribute-based method with the increase in the credentials. NMI is estimated via a confusion matrix, each row corresponds to the number of originally existing community and each column corresponds to the number of detected communities in a SIoT network.

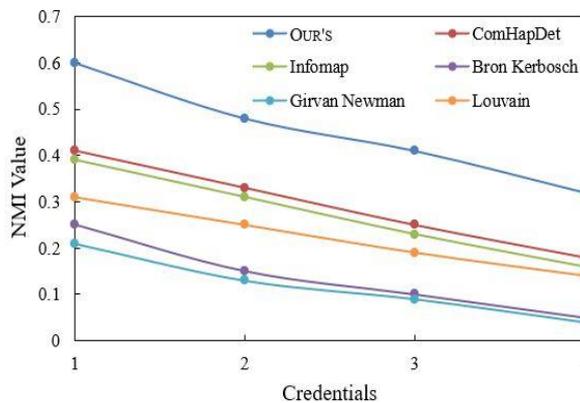


Fig. 2. Normalized Mutual Information plot for various community discovery algorithms

Fig. 2. depicts the NMI plot for various community discovery algorithms. The proposed method achieved a NMI value of about 0.32 for the highest credential node whereas

ComHapDet, Infomap, Bron Kerbosch, Girvan Newman and Louvain achieved about 0.18, 0.16, 0.05, 0.04 and 0.14 respectively.

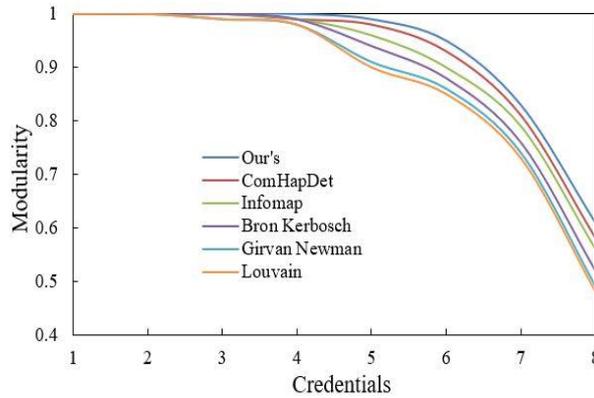


Fig. 3. Modularity plot for different community detection algorithms

The quality of the divided community is validated using the modularity index. Modularity is the difference between the edge counting of the original community and the edge counting of the community divided. Fig. 3. depicts the modularity plot for different community detection algorithms. The proposed method achieved a modularity value of about 0.61 for the highest credential node whereas ComHapDet, Infomap, Bron Kerbosch, Girvan Newman and Louvain achieved about 0.58, 0.56, 0.52, 0.49 and 0.48 respectively.

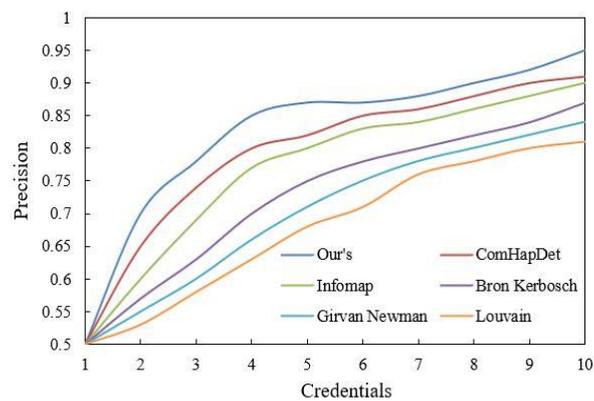


Fig. 4. Precision plot for different community detection algorithms

Precision is defined by the ratio of number of nodes in that are correctly divided after the community detection to the sum of number of nodes in that are correctly divided and number of nodes in that are erroneously divided after the community detection. Fig. 4. depicts the precision plot for different community detection algorithms. The proposed method achieved a

precision value of about 0.95 for the highest credential node whereas ComHapDet, Infomap, Bron Kerbosch, Girvan Newman and Louvain achieved about 0.91, 0.90, 0.87, 0.84 and 0.81 respectively.

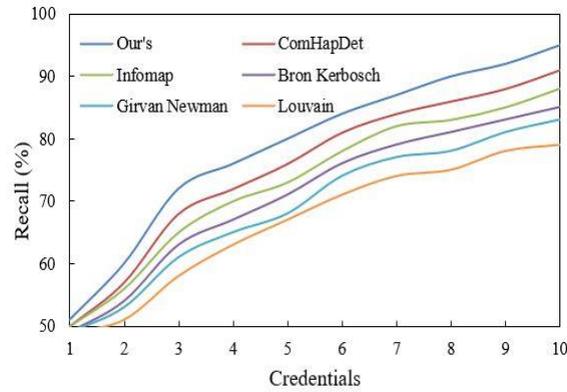


Fig. 5. Recall plot for different community detection algorithms

Recall is defined by the ratio of number of nodes in that are correctly divided after the community detection to the sum of number of nodes in that are correctly divided and number of nodes in that are erroneously divided after the community detection. Fig. 5. depicts the precision plot for different community detection algorithms. The proposed method achieved a precision value of about 0.95 for the highest credential node whereas ComHapDet, Infomap, Bron Kerbosch, Girvan Newman and Louvain achieved about 0.91, 0.88, 0.85, 0.83 and 0.79 respectively.

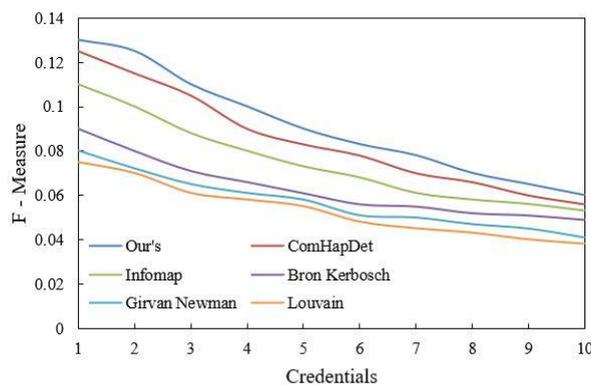


Fig. 6. F-Measure plot for different community detection algorithms

F-measure is the harmonic mean of precision and recall indexes. Fig. 6. Depicts the F-measure plot for different community detection algorithms. The proposed method achieved a

F-measure value of about 0.95 for the highest credential node whereas ComHapDet, Infomap, Bron Kerbosch, Girvan Newman and Louvain achieved about 0.91, 0.88, 0.85, 0.83 and 0.79 respectively.

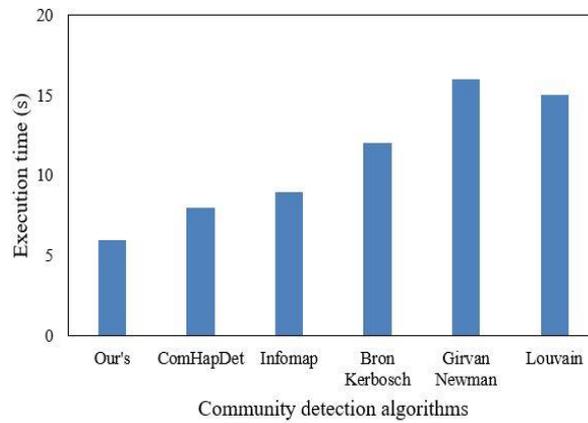


Fig. 4. Comparison of computational time for different community detection algorithms

Fig. 7. depicts the comparison of computational time for different community detection algorithms. ComHapDet needed an execution time of about 8 s while Infomap, Bron Kerbosch, Girvan Newman and Louvain required 9, 12, 16, 15 s respectively. Whereas, our proposed attribute based weighted mechanism consumed only 6 s.

7. Conclusion

Proposed an attribute-based weighted mechanism for the detection of communities in a SIoT environment. We focussed our work in dividing the SIoT network into high quality Sociality based Weighted Communities (SWC). We exploited the important social attributes like user preferences, social similarities, mutual friends' degrees are main metrics for achieving the best rate function. The weighted method based on these social attributes determine the nodes to be present in their respective communities. We also presented a mechanism for controlling the local community augmentation. A Credential Acclaimed Information Spreading (CAIS) mechanism is implemented for selecting the best node with the maximum credential to surge the content spreading behaviour in the detected communities of SIoT network. Experimental results prove that the proposed social-driven attribute based weighted mechanism for

community detection beats the conventional community detection algorithms like Louvain, Girvan Newman, Bron Kerbosch, Infomax and the recent state-of-art-approaches in terms of spreading outcomes, NMI, modularity, F-measure, precision, recall and computational time.

Declaration: The authors declare that they do not have any conflict of interest

References

- 1 Jie Lin, Wei Yu, Nan Zhang, Xinyu Yang, Hanlin Zhang, Wei Zhao, “A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications”, in IEEE Internet of Things Journal (Volume: 4, Issue: 5, Oct 2017), March 2017.
- 2 Atzori, Iera, Morabito, Nitti, The Social Internet of Things (SIoT) – When social networks meet the Internet of Things: Concept, architecture and network characterization. Computer Networks. 56 (16) (2012), 3594-3608. <https://doi.org/10.1016/j.comnet.2012.07.010>.
- 3 Atzori, Iera, Morabito, From "Smart Objects" to "Social Objects": The Next Evolutionary Step of the Internet of Things. IEEE Communications Magazine. 52(1) (2014), 97-105
[10.1109/MCOM.2014.6710070](https://doi.org/10.1109/MCOM.2014.6710070)
- 4 V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” Journal of statistical mechanics: theory and experiment, vol. 2008, no. 10, Oct. 2008.
- 5 C. Bron and J. Kerbosch, “Algorithm 457: finding all cliques of an undirected graph,” Communications of the ACM, vol. 16, no. 9, pp. 575– 577, Sept. 1973.
- 6 M. Girvan, M. E. J. Newman. Community structure in social and biological networks, PNAS June 11, 2002 99 (12) 7821-7826. <https://doi.org/10.1073/pnas.122653799>
- 7 Rosvall, M., Axelsson, D. & Bergstrom, C. T. The map equation. The European Physical Journal Special Topics 178, 13–23 (2010)

- 8 Abdullah Khanfor, Hakim Ghazzai, Ye Yang, Yehia Massoud, Application of Community Detection Algorithms on Social Internet-of-things Networks, 2019
- 9 Abdullah Khanfor, Hakim Ghazzai, Ye Yang, Mohammad Rafiqul Haider, Yehia Massoud, Automated Service Discovery for Social Internet-of-Things Systems, IEEE International Symposium on Circuits and Systems (ISCAS'20), Seville, Spain, Oct. 2020.
- 10 Shanbhaq, Sunanda Vivek, "A faster version of Louvain method for community detection for efficient modeling and analytics of cyber systems" (2016). Open Access Theses. 814. https://docs.lib.purdue.edu/open_access_theses/814
- 11 Zufan Zhang, Xieliang Li, Chenquan Gan, Identifying influential nodes in social networks via community structure and influence distribution difference, Digital Communications and Networks, 2020, <https://doi.org/10.1016/j.dcan.2020.04.011>
- 12 I. Gutierrez, D. Gomez, J. Castro, R. Espinola, A New Community Detection Algorithm Based on Fuzzy Measures, In book: Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making, 2020. Advances in Intelligent Systems and Computing, vol 1029. Springer, Cham. https://doi.org/10.1007/978-3-030-23756-1_18
- 13 Fanzhen Liu, Shan Xue, Jia Wu¹, Chuan Zhou, Wenbin Hu, Cecile Paris, Surya Nepal, Jian Yang, Philip S. Yu, Deep Learning for Community Detection: Progress, Challenges and Opportunities, IJCAI-2020, doi:10.24963/ijcai.2020/693
- 14 Cook, Learning Setting-Generalized Activity Models for Smart Spaces, Smart Environment. IEEE Intelligent Systems. 99(1), (2012). <https://doi.org/10.1109/MIS.2010.112>
- 15 E. M. Tapia, S. S. Intille, K. Larson, Activity recognition in the home using simple and ubiquitous sensors. In International conference on pervasive computing. (2004), pp 158–175
- 16 Hande Ozgur Alemdar, Halil Ertan, Ozlem Durmaz Incel, Cem Ersoy, ARAS Human Activity Datasets in Multiple Homes with Multiple Residents, 2013.

- 17 Ranjan Kumar Behera, Santanu Kumar Rath, Sanjay Misra, Robertas Damasevicius, Rytis Maskeliunas, Large Scale Community Detection Using a Small World Model, Applied Sciences 2017, 7, 1173; doi:10.3390/app7111173
- 18 Guangxia Xu, Xinkai Wu, Jun Liu, and Yanbing Liu, A Community Detection Method Based on Local Optimization in Social Networks, IEEE Access, Big Data Intelligent Networking, 2020.
- 19 Peng He., Tong Tang. Community-Oriented Multimedia Content Maximization Mechanism in Social Internet of Things, IEEE Access. Special Section on Mobile Multimedia: Methodology and Applications (8) (2020). <http://dx.doi.org/10.1109/ACCESS.2020.2970453>