IoT Data Collection Based on Social Network Models

Nataly Zhukova
St. Petersburg Institute for Informatics and Automation of
the Russian Academy of Sciences
St. Petersburg, Russia
nazhukova@mail.ru

Aung Myo Thaw
Department of Informatics and Applied Mathematics,
ITMO university
St. Petersburg, Russia
aungmyothaw52660@gmail.com

Man Tianxing
Department of Informatics and Applied Mathematics,
ITMO University
St. Petersburg, Russia
mantx626@gmail.com

Mustafin Nikolay
Petersburg Electrotechnical University "LETI" (ETU)
St. Petersburg, Russia
ngmustafin@etu.ru

Abstract— **Nowadays, the heterogeneous network and devices of IoT become more complex and encountered several challenges. The investigations of social networks show that social network analysis methods are useful for the evolution and development of communication and social network models allow solving Internet-related problems. And they can provide more exact answers for complex problems. This article analyzes some methods of SNA and investigated compatible methods to apply in IoT.**

I. Introduction

In recent years, many researches are devoted to the domain of "Internet of Things". According to [1][2], Internet of Things (IoT) is already widely used in different areas that surround us such as smart home, smart mobility and transport, smart city, smart health, smart industry, e-society and many other. According to Cisco's prediction, 500 billion IoT devices will be connected to the Internet by 2030. The widespread diffusion of the IoT has led to the integration of IoT with many other technologies. As a result, new ecosystems are emerging in different kinds of areas.

Nowadays data collection has become a highly complicated task of the different IoT ecosystems. The dynamic changes in device location are the main challenge of such systems. Comparing to the static devices, it makes the task of data collection much more complex. First, all produced data cannot be gathered and processed in time. Second, multiple new opportunities for network attacks appear. For example, malicious nodes can be added.

Fortunately, several researches investigated social network concepts to solve Internet-related problems. The social network model is considered for the improvement of IoT. It allows objects to autonomously establish relationships by using social networking elements in the IoT. By applying methods of social relationships, the provision of a level of trustworthiness can be established among different IoT devices as "friends". Therefore, smart things build social relationships with other objects they might come into contact with, to create an overlay social network to be exploited for information search and collection for required information. Moreover, the big data problem of IoT can be solved by applying social data classification. Therefore, a social-oriented approach is expected to improve the discovery, selection and services collaboration and the quality of the data provided by distributed devices and, thus, the accuracy of the solved tasks.

In this paper, we apply some social network methods to the proposed data collection model for improving security and trust in the dynamic IoT network.

This paper is organized as follows. Section I is the introduction. Section II describes social network analysis (SNA) models that allow performing the requirements of IoT. Section III presents the data collection model for the physical layer based on existing IoT models and social networks models. The last section is the conclusion.

II. Background

To date, a considerable number of techniques and models for data collection have been developed in different ecosystems of IoT.

A. Cyber–Physical Social Systems

Before 2010, Internet of Things was mainly focused on Cyber–Physical Systems (CPS), and many applications emerged in different experimental areas such as smart cities, smart transportation, Industry 4.0 and smart grids, machine-to-machine, smart healthcare, smart environments, and so on [3][4]. After the emergence of these applications, many researchers were focused on the convergence of CPS and social system which is denoted as Cyber–Physical Social Systems (CPSS) [5][6]. CPSS is based on the side effect of social interactions and relationships that take place in the physical world and social interactions observed in Online Social Network sites such as Facebook, Twitter, LinkIn and many others.

The concept of CPS and CPSS may overlap in the case of Social Internet of Things systems (SIoT) [7][8], where IoT network devices are capable of establishing social relationships with other things in an autonomous way with respect to the rules set by the owner. According to [9], applying social networking concepts to the IoT can lead to several advantages:

- Guarantee the Network navigability.
- Establish the levels of trustworthiness.
- Designed models can be reused to address IoT related

issues.

- Prediction of the actual behavior with time.

Moreover, in the research of [9], authors investigate different types of social interactions among humans that can be considered in relation to the interaction between mobility and static devices in IoT. But some security and trust algorithms need to be improved due to dynamic changes in the IoT network.

Luigi Atzori introduced the novel concept of the Social Internet of Things, based on analyzed types of social relationships among objects [10]. Over dynamic change of network topologies, user privacy data is to be secured according to proposed architecture.

In the work of [11], authors analyzed possible solutions to reduce the distance between requester and provider of a service in the IoT network by using two social strategies: a caching algorithm and a friendship selection mechanism. The best fact of this research is: their results improve the ability of these nodes to quickly reach more central nodes. But sending information between notes, the concepts of trust and security among objects need to be considered.

Zhiting Lin and Liang Dong introduced a comprehensive trust model, which proposed to use tailored nodes in IoT. This model performs the evaluation of the trustworthiness between nodes. Trustor evaluates the potential trustees base on some aspects such as success rate, gain, damage, and cost. If a node behaves maliciously in a task, the trustor's decision will be changed [12]. But their evaluation of the trust model needs improvement for mobile devices that changes the network topology overtimes.

As described above, many other researchers developed their solutions on the convergence of IoT and social networks [13][14]. The main purpose of these works is to construct the most secure and trusted IoT systems by applying social relation methods. Moreover, there are some security and trust models in the IoT system different from social concepts.

B. Security and trust model in existing IoT System

In [15], Alkhamisi proposed a cross-layer model for the data collection model. It chooses a cluster member with the highest energy as a cluster head. It is based on aggregating node identification. By sending a query message with id from a cluster head to a child node, the intermediate nodes are marked with the specific query id of the child node. This way, data collection routes are scheduled. This mechanism can reduce traffic load and energy consumption, but it suffers from high latency.

In [16], the edge centric model is proposed. It develops fog computing techniques between end devices and cloud service. This model consists of devices and network elements (fog, Specific Gateway, and edge). The model performs the tasks of data collection and processing of the traffic produced by IoT devices. Calculation of data transmission metrics for a large number of real-time, low latency applications shows that the service latency and power consumption in the models oriented on fog computing are significantly lower than in the models that use cloud computing. But fog computing oriented models suffer from a low level of security.

In [17], the authors preserve the authenticity of the data at stake, as well as the privacy of the participants by using the PAgIoT model. The basic principle of this model is the distribution of the encryption key that is shared between the sink, cluster head, and member nodes. This mechanism enables data collection of several attributes of each entity in a single operation ensuring data authenticity and privacy so that it is appropriate to be used in a large-scale IoT network. But this model suffers from high latency and power consumption.

By evaluating the trust value of the sensor node and mobile sink, trustworthy data collection model [18] can protect the security of sensor cloud systems. Except for resisting malicious attacks, this model also considers network performance parameters, such as consumption of energy, transmission distance, and network throughput.

Above mentioned researches cannot fulfill the requirements to security in the dynamic IoT network. The goal of the research was to define the methods of social network that meets the requirements of trust and security in dynamic IoT. The proposed baseline data collection model is a combination of fog techniques and clustering techniques. In this model social methods to improve trust and security between different IoT devices are used. Below we will investigate some social network models that support and improve the security of dynamic changes in IoT network topology.

C. Social Network Methods

1) Graph

Graph is a general-purpose structure for connecting data between devices. Hence anything that is somehow connected can be modeled and represented as a graph. Therefore, the expectation is that applying a graph model shall improve effectiveness in discovering risks and defects in IoT resources of the physical layer.

2) Degree distribution

Generally, sink node queries and collect data from other nodes in the network. In this case, the degree distribution of the node can apply as an indicator of the number of packets received by this node. And also, the reception of packets is related to the energy consumption of this node. Furthermore, the node with a very high degree can cause a security attack.

For modeling cluster head-based networks, the most suitable graph model is the scale-free network model. The directed graph can express the distinction between nodes. In this case, the out-degree $d+1$ of a node is the number of edges, i.e. the number of nodes that can transmit data from node i. The in-degree of node $d-1$ indicates the number of the nodes that can receive data to node i. Based on this concept, we can set a maximum limit for the number of connections that a cluster head could have in order to avoid the invalid nodes and to estimate the behavior of member nodes. The probability of the new node i links to the cluster head c can be defined as:

$$P_s = \left(1 - \frac{k_c}{k_{max}}\right) \frac{E_c k_c}{\sum_{i=1}^{N(t)} E_i k_i} \quad (1)$$

Where $N(t)$ is the number of cluster heads at the time t.

In this case, the height of value of $E_c k_c$ defines possibility of a new node connection. If the degree of cluster head reaches the maximum limit for the number of connections k_{max}, the probability of the new node i to connect cluster head becomes zero value. This means the cluster head cannot connect any new node.

3) Local Clustering Coefficient

Watts and Strogatz [19] presented the ability for a node to quickly reach the network controller with its requests that is organized by the network cluster where nodes are interlinked. This characteristic can be estimated with the hight of the value of the local clustering coefficient. It measures the status of the nodes in the cluster as follows:

$$C_{local}(n) = \frac{2 * e_n}{k_n(t) * (k_n(t) - 1)} \quad (2)$$

Where k_n (t) indicates the number of neighbors of the node at time (t) and e_n is the number of edges among the neighbors.

III. Data Collection Model at Physical layer

The scheme of data collection in the proposed model assumes that the network has cluster-based topology which comprises nodes of three main types namely:

- Cluster Heads
- Normal Nodes
- Gateway and Fog nodes

Some clustering techniques as LEACH-Mobile, LEACH-ME can be used for the data collection of mobility devices. They are based on two-layer distributed clustering and autonomously selected cluster head. In our model, the changes in mobile devices' location are calculated with the help of random waypoint mobility. Due to the movement of nodes, in every next round the new cluster head is selected and re-clustering with fuzzy C-mean clustering is performed.

In our model, we assume the fog node is the first access point of the IoT devices. Each IoT device has the ability to connect to the Fog node. The architecture is similar to the social virtual object [20], which collects data regarding the user's need, implements the social behavior and relations with other IoT devices. We assume that when the fog node receives the service request from the IoT devices, based on the available resources offered by the other devices, fog assists them in matching their available resources to the received requests.

Fig. 1. represents a network, where node 5 wants to get access to the data owned by node 3 and neighbors. The optimum path may lead through node 6 or other nodes because of autonomously selected cluster head. After creating the friendship link between nodes through social relation management, the cluster-based data collection strategies will perform.

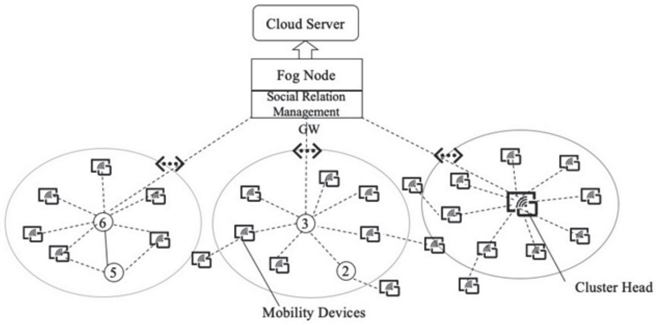

Fig. 1. The Data Collection Model

A. Selection of Network Friendship Links

According to the research [9], objects can create several types of friendships with the other nodes based on the rules set by the owners. The following possible types of social relationships can be defined:

- the parental object relationship (POR) between devices of the same model (example: same production or manufacture batch);
- the social object relationship (SOR) between devices that reflect social activities such as consequence meeting with other devices through the corresponding application;
- the ownership object relationship (OOR) between devices owned by the same user or organization;
- the co-location object relationship (COR) between devices when user together uses the devices at the workplace, like the laptop and printer in the office.

With the development of the IoT networks, other types of friendships can be added in the future. Currently, by defining the cluster head considering related relationships, the cluster-based data collection can be performed that mmets the requirements of the IoT system.

The main idea of this research is making the service search process more efficient and scalable. The goal is to help the mobility nodes in selecting the best set of friends and optimum path. All nodes can accept the friendship requests of the maximum number of friends k_{max}. This parameter is intended to limit the consumption of energy, computational resources of the nodes. And also improve the security by avoiding malicious node. The following scheme is applied to manage any requests.

1) A node needs to reject any new request of friendships after reaching k_{max}.
2) Each node needs to sort its friends list in their decreasing order of degree. To maximize the number of friends, the node needs to accept the first k_{max} friends.
3) Each node needs to sort its friends list in their increasing order of degree. To minimize the number of friends, the node need to reject the first k_{max} friends.
4) In order to maximize its own local cluster coefficient, the node sorts its friends list in decreasing order in term of their common friends and rejects the lowest k_{max} friends.

5) In order to minimize its own local cluster coefficient, the node sorts its friends in increasing order of the common friends. Then rejects the nodes with the highest value of k_{max}.

As an example shown in Fig. 2, the maximum friendship of nodes is set to k_{max} = 7. Suppose node 5 sends a friendship request to node 3. If node 3 has already reached the maximum friendship, node 3 needs to manage this request with one of above-mentioned schemes.

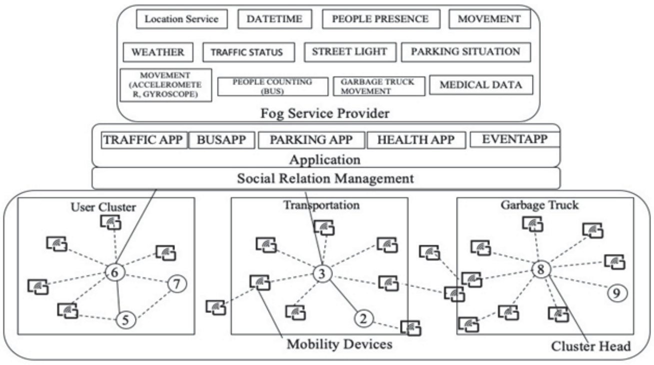

Fig. 2. Data Collection Function

If node 3 has not reached the maximum friendship list, it can accept the node 5 requests. Then, node 3 senses data from member clusters and send it to node 5.

According to scheme 2), node 3 checks the degree of all its friendship and node 5's friendship, and then it rejects the friendship with node 2 in order to accept the friendship request from node 5.

In scheme 3), node 3 terminates the relation with node 2, which has more friendship links.

With scheme 4), node 3 compares the common friends with node 5 and rejects the node 2 with which it has least common friends.

In scheme 5, node 3 rejects the relation with node 5 with which it has the highest number of common friends.

B. Friendship Links in Mobility Devices

After creating the friendship links, each IoT devices can sense the movement or behavior of other devices. In this paper, we assume that the devices have alternate movements with time, so whenever they move from one location to another, they inform the fog node with an evaluation of their movement time in that location. In this case, the data collection of their movement information is performed with clustering techniques in order to be delivered to the device requesting them.

For the proposed scenario, fog node needs to define the arrival and departure time from a given location for each device. Moreover, the fog node has to match the received requests with the available resources, it has to take into account the energy, trust level of each device. Above mentioned mobility devices management is to be performed by the service provider fog node.

When users want to know the current or estimated state of traffic congestion, they need to request the fog service provider through the corresponding application. The fog service is the mediator to construct the friendship link between user and transportation sensors. In this case, if the corresponding transportation sensor has already reached the maximum friendship, this sensor needs to perform the request with a friendship management scheme and consider the degree of requested node at current time. In fact, friendship creation is the basis for constructing trust between IoT devices.

The degree of trusted friendship links keeps stable when the next round of collecting data from mobility devices is performed. Before receiving the next new friendship request, the sensor can process the movement or behavior with their best set of friends k_{max}.

Similarly, when the garbage truck wants to know the situation of traffic, the friendship link with the transportation sensors need to be constructed. By constructing friendships, the private devices of users and the public devices are getting access to the required data through fog service. The number of the requests to fog service is shown in Fig. 3.

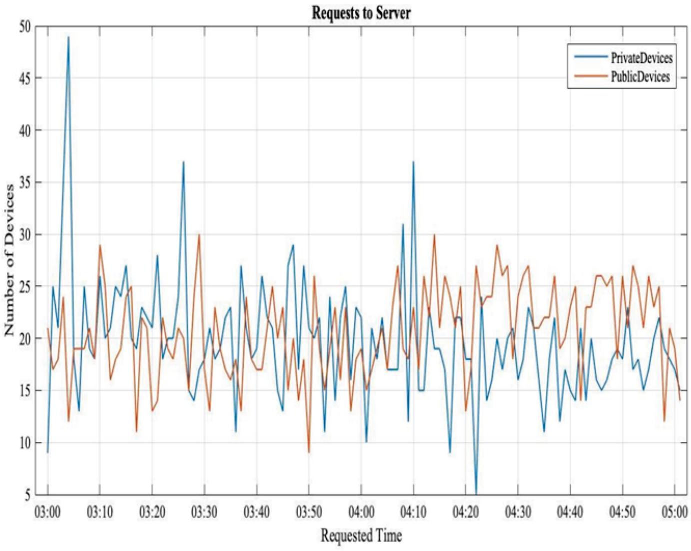

Fig. 3. Number of Requests to Server

C. Simulation

To analyze the navigability of the IoT network by using social network friendship, we need information about the requests for establishing new relationships.

We use the dataset of city of Santander [9] to evaluate our model. This dataset contains information regarding the objects, such as typology and the owner, their profiles, mobility and devices positions, as well as the social relations that each node can create with the others based on the rules set by the user. The dataset is based on real IoT objects available in the city of Santander. The total number of objects is 16216, of which 14600 from private devices and 1616 from public devices. Private devices are divided into two categories which are mobile and static. The mobile devices are smartphone, car, tablet and smart watch etc. Static devices are PC, Printer and Home sensors.

Especially, the simulation defines SOR and OOR relationships based on the above dataset. SOR relations depend on the random meeting among the objects, hence they are stochastic in nature and depend on the mobility pattern of the objects. OOR relations depend on the devices owned by the same user or organization. As shown in Fig. 3, OOR relations are established in short time and the number of SOR relations is increases with time evolution. It is observed due to the SOR create frequent reciprocal meetings. This social relation management is performed by fog service. After creating social relations, the fog service needs to collect the required data from related nodes.

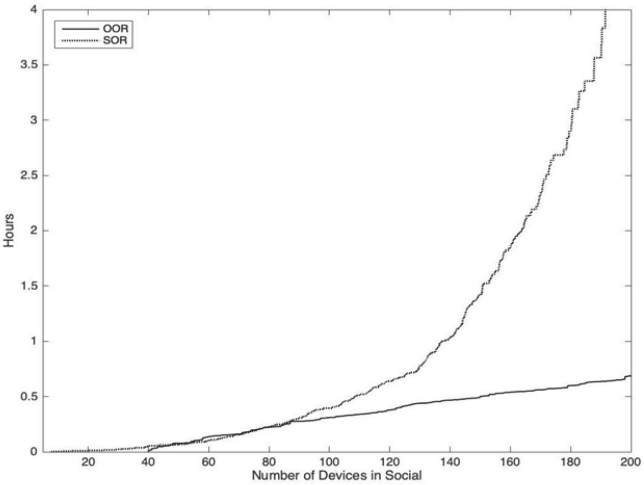

Fig. 4. Time Evolution of Social Connected Devices

In above dataset, there is no information regarding their power sources, capacity and function of IoT devices. Therefore, we assume that the 300 mobile devices with initial energy 2J are randomly disposed and move in the 300x300 m^2 area of network. The average node speed is 1 m/s and the bandwidth is 1 Mbps in our simulation. Moreover, we consider that static objects are usually plugged to the grid and then have no energy consumption issues. The average number of devices owned by user is between 3 and 5.

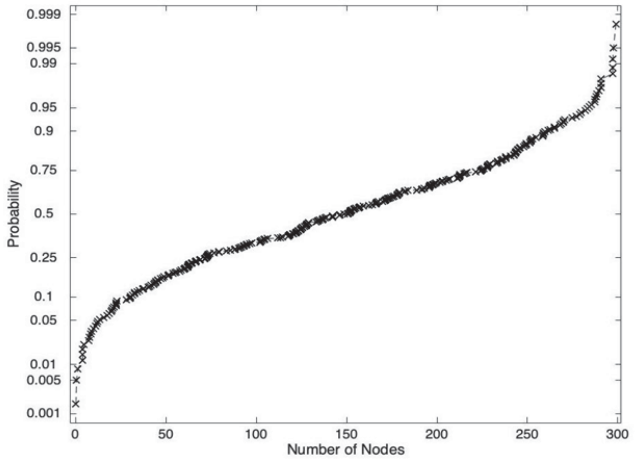

Fig. 5. The distribution of five CHs Collected Data from the Friendship 300 nodes in SOR Network

After receiving the request from the user, the fog service needs to sense the required data form corresponded node that completed friendship link creation with the requestor node. In this case, the data collection of fog service is performed with the help of cluster-based data collection. Fig. 5 shows the data collection of cluster heads from the friendship members of the cluster. In Fig. 6, the friendship of cluster head defined only maximum friendship link "50" nodes.

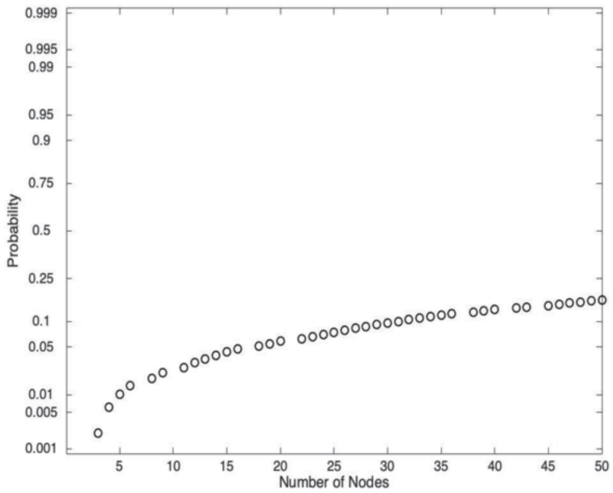

Fig. 6. The Max distribution of CH Collected Data from the defined 50 Friendship nodes in SOR Network

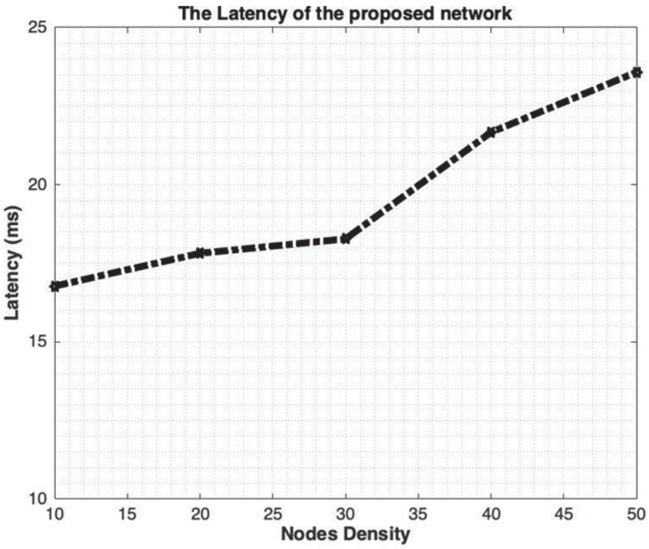

Fig. 7. The Latency of the Data Collection from the defined Friendship 50 Node.

The defining of cluster head needs to consider some security test. Because the CHs are the intermediary nodes that play the main role in collecting data from the end nodes and sending them to the fog. The fog or controller node calculates the intrusion ratio based on the received and transmitted data. Fig. 7. shows the latency of the data collection from the friendship 50 node. Fig. 8. shows average avoidance rate of the malicious node as cluster head in the network.

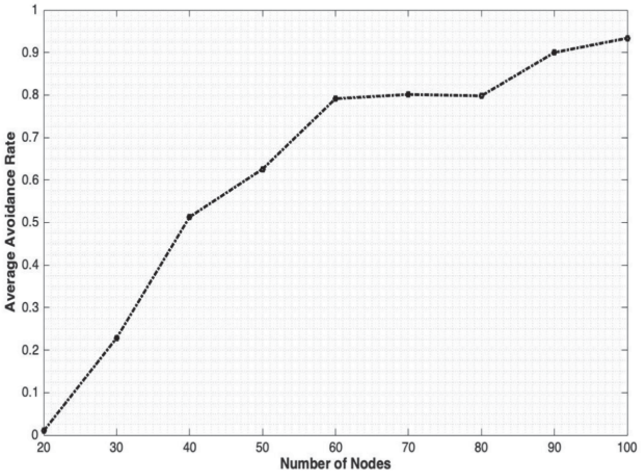

Fig. 8. Average Avoidance Rate of Malicious Node as CHs in Social Object Relationship Network

IV. Conclusion

In the paper, a new data collection model is combined with the social network friendship model. By constructing friendship between IoT devices, the model allows reach high level of trust and security.

The proposed model is being developed towards using social network models and methods in order to increase the efficiency of the management of data collection in IoT. It is expected that further extension of number of models and methods of social networks used in IoT networks can further significantly improve data collection in mobile IoT networks.

Reference

[1] Medagliani, Paolo et al. “Internet of Things Applications - From Research and Innovation to Market Deployment.” (2014).

[2] Atzori, Luigi et al. “The Internet of Things: A survey.” Computer Networks 54 (2010): 2787-2805.

[3] Ilić, Marija D. et al. “Modeling of Future Cyber–Physical Energy Systems for Distributed Sensing and Control.” IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 40 (2010): 825-838.

[4] Ray, Partha Pratim. “Creating Values out of Internet of Things: An Industrial Perspective.” Journal Comp. Netw. and Communic. 2016 (2016): 1579460:1-1579460:11.

[5] Xiong, Gang et al. “Cyber-physical-social system in intelligent transportation.” IEEE/CAA Journal of Automatica Sinica 2 (2015): 320-333.

[6] Cassandras, Christos G. “Smart Cities as Cyber-Physical Social Systems.” (2016).

[7] Atzori, Luigi et al. “SIoT: Giving a Social Structure to the Internet of Things.” IEEE Communications Letters 15 (2011): 1193-1195.

[8] Pruthvi, M. et al. ““SMART COLLEGE”- Study of Social Network and IoT Convergence.” 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on (2018): 100-103.

[9] Atzori, Luigi et al. “The Social Internet of Things (SIoT) - When social networks meet the Internet of Things: Concept, architecture and network characterization.” Computer Networks 56 (2012): 3594-3608.

[10] Atzori, Luigi et al. “SIoT: Giving a Social Structure to the Internet of Things.” IEEE Communications Letters 15 (2011): 1193-1195.

[11] Nitti, Michele and Luigi Atzori. “What the SIoT needs: A new caching system or new friendship selection mechanism?” 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT) (2015): 424-429.

[12] Lin, Zhiting and Liang Dong. “Clarifying Trust in Social Internet of Things.” IEEE Transactions on Knowledge and Data Engineering 30 (2018): 234-248.

[13] Nitti, Michele et al. “Friendship Selection in the Social Internet of Things: Challenges and Possible Strategies.” IEEE Internet of Things Journal 2 (2015): 240-247.

[14] P. Chen and K. Chen, “Intentional attack and fusion-based defense strategy in complex networks,” in IEEE Global Telecommunications Conference (GLOBECOM), Dec. 2011, pp. 1–5.

[15] Alkhamisi, Abrar Omar et al. “A cross-layer framework for sensor data aggregation for IoT applications in smart cities.” 2016 IEEE International Smart Cities Conference (ISC2) (2016): 1-6.

[16] Maiti, Prasenjit et al. “Efficient Data Collection for IoT Services in Edge Computing Environment.” 2017 International Conference on Information Technology (ICIT) (2017): 101-106.

[17] González-Manzano, Lorena et al. “PAgIoT - Privacy-preserving Aggregation protocol for Internet of Things.” J. Netw. Comput. Appl. 71 (2016): 59-71.

[18] Wang, Tian et al. “A Comprehensive Trustworthy Data Collection Approach in Sensor-Cloud System.” (2019).

[19] D. J. Watts and S. H. Strogatz, “Collective dynamics of small- world networks,” nature, vol. 393, no. 6684, pp. 440–442, 1998.

[20] Farris, Ivan et al. “Social Virtual Objects in the Edge Cloud.” IEEE Cloud Computing 2 (2015): 20-28.