This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

Privacy Preserved Ranking of Industrial Sensing Services Using Topological Information

Aroosa Hameed, Muhammad Usman, and Onaiza Maqbool

***Abstract*— Sensing-as-a-Service paradigm has realized a number of contemporary Industrial Internet of Things (IIoT) applications. In this paper, we have considered an order driven scenario of a smart factory, where wearable devices provide various services and communicate with the service providers. The discovery process generates numerous smart factory services. The selection of an appropriate (or a set of) service(s) remains a challenge while preserving the service data privacy. It is required to involve an anonymous communication mechanism to design a privacy preserved ranking model. The prevalent techniques for the prioritization of semantically equivalent sensor-provided services rely on Quality-of-Service (QoS) information. However, the QoS information is not always readily available at the node level. Moreover, the existing topological information-based (i.e., node importance and energy) solutions do not consider imperative features such as degree. The objective of this study is to design a privacy preserved ranking model, based on the onion routing technique and features along with the stochastic shortest route. Onion routing enables anonymous communication and prevents unauthorized entities from accessing ranking results. The weighted valuation is then derived to compute the cost of the homogeneous and dynamic sensor-provided services. Finally, the ranking is computed based on each service cost. The proposed model is extensively evaluated in two different network configurations of varying sizes. The evaluation results show that the proposed method performs 10% better in terms of ranking quality and 32% in terms of energy efficiency across different network configurations as compared to the existing cost-based method while offering a desired level of privacy.**

.

***Index Terms*—Internet of Things, privacy preserved ranking, sensing-as-a-service, service valuation, wireless sensor network.**

I. Introduction

INTERNET of Things (IoT) technology permits multiple addressable devices to be connected over the Internet to trade the sensed data. IoT is foreseen to reach 500 billion devices that are expected to be connected to the Internet by 2030 [1]. The sensed data, produced by the IoT solutions, can be traded between entities, namely, IoT data owners and consumers using a business model, namely, Sensing-as-a-Service (S^2aaS) [2]. S^2aaS allows consumers to request data from owners which is fulfilled in the form of services stored at the S^2aaS cloud. This highlights the fact that there exist many sensor-provided services that are exposed by the sensing objects of a smart environment. The Industrial Internet of Things (IIoT) makes manufacturing a flexible, cost effective, and responsive process. The IIoT includes various applications of manufacturing area. An order-driven smart factory scenario can be used to explain the need of IIoT ranking process in the S^2aaS environment.

The smart factories comprise wearable devices and sensors in an IIoT environment. These wearable gadgets and sensors are source of offering functionalities as services. These devices detect Wi-Fi availability in the factory and communicate information to the service providers. We assume that the registration of each factory with a service provider in the S^2aaS model is already performed. The service providers communicate with each factory owner in order to get permission regarding their data publishing. Each smart factory publishes its services at S^2aaS cloud after a legal agreement describing what service to be published, the allowed bidding, and the type of return.

There is a company ‘A’ that requires production order to be delivered by the most relevant service of the smart factory according to the specified requirements. Several factory services are discovered as a result of the discovery process. There is a need to have a ranking mechanism which ranks the most relevant smart factory service at the top of the list. Thus, each factory service is ranked by utilizing a ranking technique. Once the process of ranking is completed, the company ‘A’ can send an order request to the smart production factory ‘F_1’. The manager of the factory ‘F_1’ then checks its production status and adapts the flexible production process to meet the order, reconstructing a production line, such as by replacing or extending assets. The decision of whether to publish their factory services or not is taken by the owners. Thus, the model offers a control of the owner privacy in their own hands.

The existing sensor-provided service ranking techniques utilize the Quality of Service (QoS) features, content information, contextual information or the state-of-art methods such as machine learning and Multi-criteria Decision-making (MCDM). However, the computation of QoS information remains a challenge in sensor networks due to resource constraints, limited bandwidth, heterogeneous, and dynamic network topology [7]. Furthermore, the content and context-based techniques suffer from the scalability issue and the state-of-the-art techniques require a high computational complexity leading towards polynomial time.

[1]A. Hameed is with the Department of Computer Science, Quaid-i-Azam University, Pakistan. (e-mail: aroosa.hameed@cs.qau.edu.pk).
M. Usman is with Faculty of Computing, Engineering and Science, University of South Wales, Treforest, CF37 1DL, United Kingdom. (e-mail: muhammad.usman@southwales.ac.uk).
O. Maqbool is with the Department of Computer Science, Quaid-i-Azam University, Pakistan. (E-mail: onaiza@qau.edu.pk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

A related work employs the deterministic Dijkstra algorithm to find a shortest path [15]. However, the wireless sensor networks are typically highly dynamic in nature. The nodes are subject to failures, making the shortest path problem a stochastic problem. Therefore, we have proposed a probabilistic method to find the shortest paths among various competing paths in order to access a service and provide anonymous communication while preserving the data privacy. Furthermore, the authors of the previous work did not directly consider the degree of each node involved in the service access cycle [15]. The high degree of a node is an indicator of the importance of the node, causing the cost of accessing a service through such nodes to be higher. Another limitation is its efficiency in worst cases, which leads to polynomial time. Furthermore, they have not computed ranking considering the security and privacy concerns. This highlights the need of additional factors, namely, time and privacy in the calculation of the access cost of services.

This study has focused on designing a privacy preserved ranking method using anonymous communication by considering the features: (i) influence (importance) value (ii) energy level, and (iii) degree of each IIoT node (service). The major contributions of this study are following: (i) in order to find the shortest route, we have suggested a probabilistic technique to analyze the likelihood of each competing path. The proposed probabilistic technique has considered two important factors, namely, the distance among nodes and the count of the number of nodes that exist on the path (route). This provides probability of each alternative path for the selection of an appropriate shortest route, which is a key building block of our method (Section IV. A); (ii) an anonymous communication mechanism is provided that adapts the onion routing technique in order to provide the privacy of data (Section IV.A.1); (iii) a weighted valuation function, that adapts the topological service (corresponding to a sensor node) features to achieve the access cost of a sensor-provided service, is proposed (Section IV.B); (iv) proposed an extended method to find the influence of each service with respect to the shortest route (Section IV.B.1). To determine a service influence, we have computed transition probabilities and importance values of every node in the shortest path. These two factors specify whether a service is highly influential or not within the shortest route.

The rest of the paper is structured as follows: Section II presents highlights of some of the previous works in the IIoT service ranking domain. It also elucidates challenges in the ranking process. Section III provides the network model and preliminaries necessary for comprehending the ranking problem. Section IV elucidates the proposed ranking model. Section V explains our proposed privacy preserved ranking algorithm and its asymptotic analysis. Section VI provides the details of the experiment setup, the performance evaluation results, and their detailed analyses. Finally, Section VII concludes this paper and provides future directions.

II. RELATED WORK

The S^2aaS paradigm transforms everything into services while allowing sharing and reusing of data among multiple consumers. These IoT services, provided by sensors, are of high demand by different business organizations and users. Therefore, there is a need to design a mechanism that can deliver best match services according to the user defined criteria. In this context, two challenges exist in the IoT domain, i.e. (i) discovery of services and (ii) ranking of services. The discovery process generates many candidate services. However, selection of the relevant service, from those large number of services, which meets the customer needs remains a challenge while preventing unauthorized entities to access ranking results. A multi-objective decision-making process that decides the ranking of service set based on multiple criteria is known as IoT service ranking [16].

The existing studies mostly rely on QoS information for the ranking of services. Guinard et al. designed a ranking method that incorporates different types of data (e.g., temperature, noise), QoS information, i.e., latency, and attributes such as location coordinates [3]. Each criterion has a weight, assigned by the consumers in the query. Chatterjee et al. proposed a ranking algorithm using the QoS parameters as service link capacity, types of delays (i.e. transmission, hop and processing), and accuracy [4]. Neha et al., proposed a ranking method using the user preference [5]. Firstly, consumers provide point-based requirements, secondly, Preference-Based Weighted Index (PBWI) is assigned to each service by aggregating user requirements and contextual values. Finally, the services are ranked using PBWI values. Another ranking approach utilized Primitive Cognitive Network Process (P-CNP) to map ranking of the IoT services as an MCDM problem [6]. The above-cited works, [3], [4], [5] and [6], employed QoS information which may not be readily available at sensor node level. Furthermore, it may not always be possible to collect such information about every single sensor node in a large-scale network because of resource constraints, bandwidth limitations, heterogeneous and dynamic nature of the network. Moreover, dynamic nature causes connections or disconnections between nodes because of the energy dissipation of sensor nodes. The assessment of the performance of these methods is difficult, because it may vary from time to time.

There are other existing works that employed contextual information [8-9] and content information [10-12] to address the ranking problem. A ranking model, namely, Context Aware Sensor Search, Selection and Ranking Model for Internet of Things Middleware (CASSARAM), proposed by Perera et al., is based on the contextual information such as reliability, availability, battery life, and quantitative reasoning [8]. CASSARAM is further extended in a way that services are described using the ontology theory [9]. A method by Niu et al, aggregates the User QoS Assessments (UQA) and Context QoS Assessment (CQA) using the fuzzy logic for ranking [10]. Furthermore, this work ranked services by utilizing the timeliness feature. Truong et al. proposed a method that computes a ranking score for each service based on the fuzzy sets and a range given in the query [11]. What type of sensor data is required, how it is sensed, and when to sense are the factors that are employed by Babu et al. to prioritize services [12]. However, these works [8-12] suffer from the scalability

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

issue, as these techniques require a number of sensor nodes to be minimum within the network to perform efficiently in terms of time. This is because such methods need to calculate contextual or content information for each of the services involved within the WSN. Therefore, the above-cited approaches are not adaptable at larger scale.

Some related works employ predictive modelling to perform the ranking of services. Cassar et al. applied a machine learning approach, i.e., Latent Dirichlet Allocation (LDA) technique that extracts the latent factors from each sensor service description to rank the services [13]. Zhang et al. estimated the state of the services based on a prediction model. It is observed that the communication overhead is reduced by utilizing the matching predictive probabilities of services for ranking [14]. However, these works [13-14] require a simple environment and queries.

The work carried out by Wang et al. relies on the topological features to rank the sensor-provided services [15]. Thus, we have compared our proposed work with [15]. Wang et al. computes the service access cost based on the shortest path and topological information, i.e., energy levels of sensor nodes and their importance values. The shortest path is computed based on the deterministic Dijkstra algorithm. However, determining the shortest path in sensor networks is a stochastic problem due to the dynamic nature of sensor networks (limitation 1). The other limitation of Wang et al. method is the exponential growth of running time in the worst cases as the overall access cost depends on the shortest path calculations. Thus, making the ranking method an inefficient method (limitation 2). Furthermore, there are dependencies between importance and degree; energy and degree; and influence and energy that need to be studied with respect to the ranking computation process (limitation 3). Moreover, the above-cited research works lack in terms of privacy preserved ranking (limitation 4).

This work aims at providing a privacy preserved ranking mechanism to address the above-cited limitations in such a way: (1) providing a probabilistic method that computes the shortest route, based on its likelihood; (2) to overcome the limitation 2, we provide the solution that takes logarithmic time even in the worst scenarios for ranking of the sensor-provided services; (3) we examine and analyze the relations among proposed topological features with respect to our proposed ranking strategy in order to address the limitation 3; (4) Incorporated an onion routing based communication technique [26], in order to provide anonymous communication among nodes while preserving the privacy.

III. NETWORK MODEL AND PRELIMINARIES

A. Network Model

A sensor-provided service network, SN, is an undirected graph, G, and represented as a tuple $< S, L, W >$, where S is the set of services provided by the service network, such that, $S = \{s_1, s_2, s_3, \dots., s_m\}$ as $S \in SN$ and $m < n$, where m is the total number of services and n is the total number of nodes in the network, L is the set of links, representing relations between services (nodes), such that, $L = \{l_1, l_2, \dots., l_k\}$, where $1 \leq k \leq (t_n - 1)$, with t_n representing total number of nodes, and W denotes the set of weights (distance among links) on each link within the SN, such that, $W = \{w_1, w_2, \dots\dots, w_z\}$, where, $1 < z < k$. Each service, s_m in set S is assigned a public key, pu_m and a private key pr_m which are used in the encryption process.

Definition 1 (Sensor-Provided Service): A sensor-provided service, s_i, is a tuple $<nm, dsp, l, r, i, e, dg>$, where nm is the name of the service, s_i, (ii) dsp is the description of the service, s_i, (iii) l is the location of the service, s_i, in terms of x and y coordinates, (iv) r is the communication range of node v_i which offers the service s_i, (v) i is the influence value of the node v_i, (vi) e is the current energy of the node v_i, and (vii) dg is the degree of the node v_i.

Definition 2 (Feature Set): A topological feature set, FS, is a set consisting of sensor node characteristics and represented as: $FS = \{fs_1, fs_2, fs_3\}$ where $fs_1 = I_i$ is the influence of sensor node i, $fs_2 = E_i$ is energy level of the sensor node i, $fs_3 = dg_i$ represent the outgoing degree of the sensor node i indicating the connections of a node.

B. Sensing-as-a-service ranking problem

Given the two sets $R = \{r_1, r_2, r_3, \dots, r_{nn}\}$ and $S = \{s_1, s_2, s_3, \dots., s_{mm}\}$, in a search space, V, such that, $S \subset V$, where R be the set of nn number of requirements describe by the consumer and S be the set of mm number of sensor-provided services that match the requirements of consumer, then the ranking problem is to sort the set S such that, $S = \{s_1 > s_2 > s_3 > \dots > s_{mm}\}$, where a highly relevant element in S appears first followed by the second and so on. The ranking is performed according to some criteria, i.e., described in terms of relevancy of the services with respect to the needs expressed in user requirements while preserving the privacy.

C. Requirements for sensing-as-a-service ranking

The requirements for the Sensing-as-a-Service ranking are listed below:

- The remaining energy of the sensor nodes is represented as $RE = \{e_1, e_2, \dots, e_n\}$, where n denotes the energy of nth sensor node. It is imperative to design a ranking scheme that helps to preserve the energy of sensor nodes [15].
- The anonymous communication is provided by using onion routing approach, thus each service acts as an onion router (OR) in WSN. The requested service is returned to the consumer through anonymous communication.
- The services, within S^2aaS paradigm, represent ‘things’ in the IoT domain. These services differ from traditional web services in terms of heterogeneity [15], as each service may possess varying sensing characteristics (i.e. sampling rates, error distributions, and spatial resolution) and other varying characteristics (e.g., manufacturer, battery life, transmission range). The ranking system needs to be independent of such complexity and heterogeneity.

IV. PROPOSED RANKING METHOD

The proposed ranking method consists of three different stages, as shown in Figure 1. The first stage estimates the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

shortest route by employing the proposed probabilistic method and provides anonymous communication. The second stage calculates the access cost value of each service within the valuation unit and the third stage ranks the services based on their access cost values within the sorting unit.

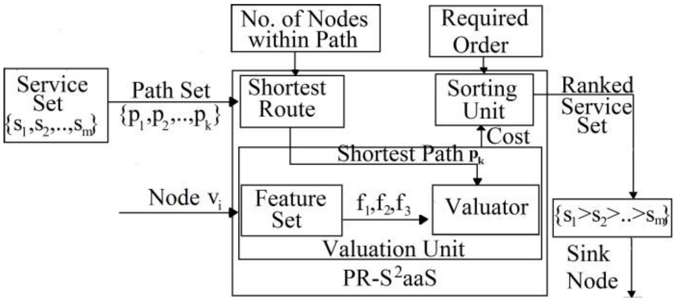

Figure 1: Ranking of Sensing-as-a-Service

A. Estimating the shortest route

The service access process includes communication among different sensor nodes within the path (route). The access cycle generates some cost, which is dependent on the selected path. However, the number of nodes involved within the route are also important as the distance between nodes is a stochastic value. This value can vary due to the failure of nodes, which can result in disconnections and changes in the network topology over a period of time. This result in change in number of nodes within the alternate routes. Therefore, we propose a probabilistic approach, which locates the shortest route towards the service, say s_1, based on the distance and number of nodes within the route. The route, with the minimum probability, is the shortest route from the sink node to the required service node.

$$p(i) = \frac{n*ds}{\sum_{j=1}^{m} ds_j} \qquad \forall\ n \in \mathbb{N},\ m \in \mathbb{Z}+ \tag{1}$$

where $p(i)$ denotes the probability of the possible path i towards the smart factory service s_1; n represents the total number of nodes in the path; ds represents the sum of the distances among all nodes n in the path i and it is calculated using (2); m denotes the total number of alternative paths towards the sensor service, s_1. As the value of $n * ds$ increases, the probability of the path i reach towards the highest probability value, i.e., 1.

$$ds = \sum_{j=1}^{n} d_j \tag{2}$$

where d represents the distance between the two sensor nodes, say, v_k and v_l in the path i. The distance, d between two sensor nodes, say, v_k and v_l may be calculated using the Euclidean method. It is pertinent to observe that the Euclidean distance is the straight-line distance, which may not be feasible with the geographical coordinates, as they do not consider the curvature of the earth and the geographical barriers. When calculating the distance over the projected geographical coordinates, spherical calculations must be done along the curved surface of earth. Thus, there is a need to perform a projection to compute the circular distance among the nodes [17]. The Lambert Transformation method is applied for this transformation because it is relatively simple as it involves trigonometric functions that can be solved in $O(\log(n))$. Further circular distance between two points can easily be computed using Euclidean distance [18].

1) Lambert projection method

Since the earth has a spherical shape, the service coordinates (spherical) set, i.e., SCS needs to be transformed into Cartesian coordinates, i.e. the set, CCS using Lambert conformal conic projection technique [21]. The Cartesian coordinates set, CCS from the service coordinates (spherical) set, SCS, are calculated using (3) and (4).

$$x = p \sin\theta \tag{3}$$

$$y = p_0 - p \cos\theta \tag{4}$$

where p and p_0 denote constants and θ represents an angle which depends on the longitude values of the sensor-provided services in LG.

2) Anonymous Communication

We obtain anonymity using onion routing approach [25], where each sensor within the network acts as an onion router. Each node in the shortest route provides its own onion layer for a privacy preserved ranking. Information among nodes is communicated anonymously. Thus, infected (compromised) nodes are unaware of the communicated readings. We have utilized the data collection phase of work [25] to perform anonymous communication among nodes. Each node receives a $< OR_i, R_i >$ where OR_i represents the layer and R_i denotes an actual reading. At each node, one layer of the onion is removed using decryption and data is encrypted within the innermost layer by the sensor. Each node is unaware of the position in WSN. This ensures anonymity, thus preserves privacy.

B. Sensing service valuation

The valuation phase estimates the cost of each service based on the features. Such cost of each service is then fed into the sorting phase. The service with the lowest cost value is assigned the high rank, whereas, the service with the highest cost value is assigned the least rank.

1) Weight calculation

We consider each sensor node having a feature set as service influence, service energy level, and service degree. The reason for the selection of these features is already explained in Section I. The range of the service influence is between 0 and 1; the energy level is between 0 and 100. However, range of the service degree depends on the number of nodes in the network configurations. For the configuration 1 it is between 0 and 49; for the configuration 2 it is between 0 and 99 (explained in Section IV.A). The maximum degree of any node is 49 in configuration 1 and 99 in the configuration 2. We normalize the ranges of each feature between 0 and 1.

It is observed that the influence of features, on the computed ranking values, is not equal. For example, high energy of a sensor node, low influence value of a node, and low degree value of a node lead towards the high-ranking value of a service. By high energy, we mean the value of energy is greater than a minimum threshold. It is also noticed that some features, such as energy with higher values, are best for ranking. On the other hand, other features such as influence and degree with low values are best for the computation of ranking. Therefore, we capture this effect by assigning numerical weights to the features as (5):

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

$$w_i = \begin{cases} (1+\log(fs_i)) * \log\left(1+\frac{N}{\sum_{i=1}^{N} fs_i}\right), & if\ fs_i > 0 \\ 0, & Otherwise \end{cases} \quad (5)$$

where, $fs_i \in F$ represents the features in feature set F; N denotes the total number of nodes in the communication path; $\sum_{i=1}^{N} fs_i$ represents the sum of feature value for all N nodes in the communication path. Adding 1 to the log values and fraction values avoids the odd behavior of the extreme values of fs_i and $\sum_{i=1}^{N} fs_i$.

2) Service access cost calculation

The cost of each participatory node is independently valuated and then the cost values of all participatory nodes within the route are summed up to formulate the service access cost value, as shown below.

$$s_i = \sum_{i=1}^{k} c_{pn_i} \quad (6)$$

where s_i is the i_{th} sensor-provided service that is to be valuated; k is the total number of participating nodes in the shortest route R and c_{pn_i} denotes the cost of the i_{th} participating node. The proposed method to find the cost of each participatory node in the path P is computed using (7) depending on the feature set of the node,

$$c_{pn_i} = \frac{w_1(I_i)+w_2(E_i)+w_3(dg_i)}{\sqrt{(w_1)^2+(w_2)^2+(w_3)^2}} \quad \forall\ I_i, E_i, dg_i\ \geq\ 0 \quad (7)$$

where $w_i \in W$ represents the weights of the features and W={w_1, w_2, w_3}, I_i denotes the influence of the i_{th} sensor node; E_i represents the energy level of the i_{th} sensor node; and dg_i shows the degree of the i_{th} sensor node. If the energy level is less than the specified minimum energy threshold value, then its influence and degree may be high. Therefore, the service access cost through such nodes will be high.

a) Influence of a node

The influence of sensor node quantifies the importance of a node within the shortest route, R, as illustrated in Figure 2. There are 11 sensor nodes within the network, where the node n_{11} provides the service, accessed via a route, $sink \rightarrow n_1 \rightarrow n_7 \rightarrow n_9 \rightarrow n_{11}$. It can be observed that the node, n_7, is a highly influential node in SN, as it has 5 outgoing links. The highly influential node causes high cost to access a service because it may have several outgoing connections, performing sensing and relaying functions, and consuming most of the energy. To compute the influence value, the WSN is modeled as an undirected graph with n number of nodes. The gateway updates the adjacency matrix A of WSN either with 0 or 1 in order to get an overview of all communication links.

The computation of the node influence depends on the transition probability and importance matrices. The transition matrix represents the probability of transiting from one node to another node and the importance matrix denotes the importance value of the adjacency list of each node. We define transition probability P_{ij} as a probability to transit from the node i to j, as shown in (8). It should be noticed that P_{ij} is different from the transition probability P_{ji}.

$$P_{ij} = \frac{1}{K_{ij}} \qquad 0 \leq P_{ij} \leq 1 \quad (8)$$

where K_{ij} denotes the total number of outgoing edges of the node i to node j. It can be observed that as the outgoing degree of a node i increase, the transitional probability decreases. After estimating the transitional probabilities of each sensor node, it is possible to compute the influential metric. For this, we extend the original influence metric, discussed in [20] as (9), because the transitional probability of a node is increased with respect to the importance of the connected node in the shortest route.

$$I_i^h = e^{(-(l_j\times P_{ij})^h log(l_j\times P_{ij})^h)} \quad (9)$$

where e denotes a constant whose value is approximately 2.71828; I_i is the influence value of the node i; h are the number of nodes in the shortest route R; I_j denotes the importance of the node j to which node i is connected and P_{ij} denotes the transition probability from the node i to node j.

The negative sign in the expression $\left(-(l_j\times P_{ij})^h log(l_j\times P_{ij})^h\right)$ indicate that the probability values must be between 0 and 1. The upper and lower limits in (9) are $e^{(-(l_j\times P_{ij})^h log(l_j\times P_{ij})^h)} < 1$ as the value of $\left((l_j\times P_{ij})^h log(l_j\times P_{ij})^h\right) < 0$. The influence of a node is modeled as the exponential function that exhibits the exponential decay behavior as:

If $\left((l_j\times P_{ij})^h log(l_j\times P_{ij})^h\right) \rightarrow \infty$

Then $e^{(-(l_j\times P_{ij})^h log(l_j\times P_{ij})^h)} \rightarrow 0$

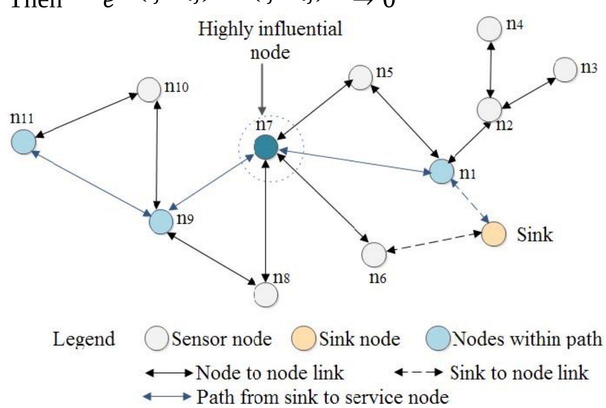

Figure 2: Influential node within the sensor network

b) Energy of a node

The energy of a sensor node is a key factor in the process of service access cost valuation because it indicates the status of the sensor node within the network. The cost not only depends upon the energy of sensors that provide services, but also on the energy of participatory nodes. We assume that the gateway of the sensor network has a permanent energy source. The energy consumption of each node, within the shortest path includes the sensing energy, receiving energy, transmitting energy, and processing energy. The transmission energy of nodes is highly dependent on the distance to the next node or the sink node. The energy of the i_{th} sensor node for packet transmission can be computed as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

$$E_i = l(e_t * d_i^{\alpha} + e_o) \quad (10)$$

where l denotes the length of a packet in bits to be transmitted; e_t represents the loss coefficient of bit transmission; d_i is the distance between node i and next hop; α represents the path loss exponent; and e_o denotes the overhead energy required to transmit packets including sensing, receiving, and processing energy. The path loss has values from 2 to 4.

c) Degree of a node

The change in the degree of a highly influential sensor node could result in disconnections of the sensor nodes. The sensor node degree dg_i consists of the outgoing degree of each participatory node pn_i within the shortest route R.

C. Service Ranking

The final step involves the ranking of the smart factory services, based on the cost value produced for each sensor service in (6). The ranking can be performed by sorting the sensor services in the ascending order according to their respective cost values, such that sensor service with the lowest cost value is assumed to be highly desired rank as compared to service with higher access cost values.

V. THE RANKING ALGORITHM

A. The Algorithm

The proposed Privacy Preserved Ranking of Sensing-as-a-Service (PR-S^{2}aaS) algorithm is of recursive nature, consisting of two major procedures as: RANKING and VALUATION. The RANKING procedure sets the initial cost of each service to zero and then checks the total number of services. If the discovered service set consists of more than one service, then the algorithm recursively divides the service set into two halves. The recursion follows until each sub array contains one service. The two major operations, which are performed within the RANKING procedure on each service in sub arrays, are (i) computation of the shortest route (phase 1) by employing the function, $shortestroute()$ and (ii) valuation is performed by calling VALUATION procedure that generates access cost of each service (phase 2).

In the VALUATION procedure influence, degree, and energy for each of the services are extracted and their weights are calculated using $weight()$. Finally, $val()$ computes the cost of each node in the shortest route. These cost values are then aggregated to additive form to compute the overall service cost. Finally, the RANKING procedure sort services using $rank()$ function, according to their cost values (phase 3). The services having a low-cost value is ranked high. The low-cost value of a service is an indicator of the following: (i) access to this service requires low energy consumption and (ii) the participatory nodes tend to be less influential. The algorithmic description of the ranking process is given in Algorithm 1.

Algorithm 1: PR-S^{2}aaS Algorithm

RANKING $(\boldsymbol{S, p, r, n, m, d, I, E, deg})$

// S is service set array; p denotes the starting index of S; r is the last index of S; d is distance between nodes; n is number of nodes in paths; m is number of alternative paths; I is the influence matrix; deg is the degree matrix; E is energy matrix.

1. set $Service_{value} \leftarrow 0$
2. ***if*** $p < r$
3. ***then*** $q \leftarrow \left\lfloor \frac{p+r}{2} \right\rfloor$
4. RANKING $(S, p, q, n, m, d, I, E, deg)$
5. RANKING $(S, q+1, r, n, m, d, I, E, deg)$
6. $SR \leftarrow shortest_{route}\ (S, d, n, m)$ ▻ *Phase* 1
7. $Service_{value} \leftarrow VALUATION\ (SR, I, E, deg)$ ▻ *Phase* 2
8. $rank\ (S, Service_{value})$ ▻ *Phase* 3
Output $s_1 > s_2 > \cdots > s_m$

VALUATION Procedure

VALUATION $(\boldsymbol{SR, I, E, deg})$

//SR is the shortest route to access a service

1. **for** $\boldsymbol{i} \leftarrow$ **0** ***to*** SR
2. **do** $ind \leftarrow SR[i].nodeID$
3. $inf[i] \leftarrow inf[i] + I\ [ind]$
4. $energy\ [i] \leftarrow energy\ [i] + E\ [ind]$
5. $\text{degree}[i] \leftarrow \text{degree}[i] + deg[ind]$
6. $w1 \leftarrow weight\ (inf, \text{inf}[ind])$
7. $w2 \leftarrow weight\ (energy, energy\ [ind])$
8. $w3 \leftarrow weight\ (degree, degree\ [ind])$
9. $pn \leftarrow pn + val(w1, w2, w3, I[ind], E[ind])$
10. return pn

B. Asymptotic Analysis

Proposition 1. The computational complexity of VALUATION procedure is $O(n)$.

Proof of Proposition 1. The running time of VALUATION procedure depends on the size of the shortest route. We have two cases as: (i) in the best scenario, the shortest route consists of one sensor node, thus, constant time is required to compute cost value i.e., $O(1)$. Because the for loop in VALUATION executes only one time. Further, all statements (step 2 to step 9) in the loop body of the VALUATION procedure are the assignment statements; thus, execute in $O(1)$. (ii) In the worst scenario, the shortest route may consist of n number of sensor nodes, then the time complexity of the VALUATION procedure is $O(n)$. This is because the loop executes n number of times and each statement within the loop body takes constant time i.e., $O(1)$. Thus, the overall complexity is $O(n) * O(1) = O(n)$.

Proposition 2. The computational complexity of PR-S^{2}aaS is $\theta(nlgn)$ in the worst case.

Proof of Proposition 2. As the algorithm consists of two procedures, namely, RANKING and VALUATION. The RANKING procedure divides the problem set into the sub problems, each of size $\frac{n}{2}$. If the running time of the RANKING procedure is $T(n)$, then the step 4 and step 5 executes in $T\left(\frac{n}{2}\right)$ times. Step 7 takes $O(n)$ time and other steps are taking constant time, i.e., $O(1)$. Thus, the running time of the RANKING procedure forms recurrence equation of the form as:

$$T(n) = \begin{cases} O(1) & if\ n = 1 \\ 2T\left(\frac{n}{2}\right) + O(n) & if\ n > 1 \end{cases} \quad (11)$$

The above recurrence (11) is solved using case 2 of the master theorem [21], we have $a = 2$; $b = 2$; $k = 1$ and $p = 0$. As $a = b^k$ i.e., $2 = 2^1$ and $p = 0$, the solution of recurrence is (12):

$$T(n) = \theta\left(n^{\log_b a} \log n\right) \quad (12)$$

By inserting the values of a and b in (12), we have

$$T(n) = \theta\left(n^{\log_2 2} \log n\right) = \theta(nlogn) \quad (13)$$

Thus, the time complexity is $\theta(nlogn)$. As the logarithmic function grows slower than the linear function, our proposed algorithm performs better than the existing linear time algorithm [15] whose running time is linear: $O\left(D * (n + m)\right)$ where D is a service set size and exponential: $O(n^2)$ in the worst cases.

VI. Performance Evaluation

A. Experiment Setup

We considered specifications of IRIS nodes, manufactured by Memsic, for the realistic simulation scenarios [22]. Two network configurations, consisting of 50 and 100 sensors, are deployed uniformly in the 2D plane of 100×100 meter with x and y coordinates generated between 0 and 100. The underlying topology of the network depends on the distance among sensor nodes and radio range, i.e., 30 m. The initial connection value is set to be 1 in the adjacency matrix, if the distance value is less than the 30 m; otherwise, 0. The 1 Joule initial energy value is assigned to each node and energy consumption of the network is simulated using an energy model with e_t and e_o set as 0.0013 pJ/bit/m4 and 50 nJ/bit, respectively [23]. The maximum energy of nodes, mx, is set to be 1 and the minimum energy, mn, is set to be 0.3. The packet length l is taken as 1000 bits and the path loss coefficient is set as $\alpha = 4$.

The energy consumption is calculated and updated in the matrix throughout the experiments. An event-driven traffic is simulated within the network, according to the Poisson distribution with 1 packet generated per second, thus, $\lambda = 1$. The traffic within WSN is generated anonymously without showing the current positions of sensor nodes. Two different datasets are generated for each of the network configurations listed in Table I. The dataset consists of several services over the network with a set of 100 queries. We assumed that the discovery is already performed, resulting 100 sets matching 100 queries. For the simulation purposes, we have taken a varying number of competing paths. Two experiments were performed for a time period of 100 seconds where 1 query is processed at each second, and the ranking of service set is carried out. In order to assess the performance of the proposed algorithm, two metrics are considered: ranking quality and energy preservation.

TABLE I
Network Configuration Parameters

N/w settings	Sensor nodes	Sink nodes location	Sensor services	Number of paths
1	50	(50,50)	49	3000
2	100	(100,100)	99	4000

1) Ranking Quality

The ranking quality is measured through Normalized Discounted Cumulative Gain (NDCG) [24]. As all services are not of equal relevance in terms of features, NDCG evaluates and assigns some relevance grade to the services as (14) [27],

$$NDCG_a = \frac{DCG_a}{IDCG_a} \quad (14)$$

where DCG_a is the Discounted Cumulative Gain of the sensor service, s_a, calculated using (15) and $IDCG_a$ is the Ideal Discounted Cumulative Gain of the service s_a.

$$DCG_a = \sum_{b=1}^{N} \frac{rel_b}{\log(b+1)} \quad (15)$$

where b represents the position of a service s_a within the ranked set; N represents the total number of services in the set; and rel_b is the relevance grade of the service s_a at the b^{th} position.

2) Energy preservation

We compare our proposed algorithm with the existing cost-based method [15] in order to identify the ranking method which preserves more energy of the sensor network. For this, we have maintained an energy matrix that is updated at regular intervals. The energy consumption of the network in ranking of the service sets is estimated after every 50 and 100 seconds in all settings.

B. Results and Discussions

1) Ranking Quality Results

The ranking evaluations for two different configurations of the first 5 services are shown in Table II across 100 queries. The results indicate that the proposed algorithm, PR-S^{2}aaS has highest values of NDCG for some services as compared to the

TABLE II
Comparison of values of NDCG@5 of both methods for Configuration 1 and Configuration 2

N/w settings	Nodes	1	2	3	4	5
1	PR-S^{2}aaS	0.660	0.677	0.810	0.740	0.876
	Cost	0.648	0.544	0.736	0.617	0.905
2	PR-S^{2}aaS	0	0.886	0	0.791	0.858
	Cost	0	0.666	0	0.816	0.769

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

cost-based method in both configurations across 100 queries.

For the configuration 1, PR-S^{2}aaS improves the NDCG values for service 1 by 1%, service 2 by 13%, service 3 by 7%, and service 4 by 12%. However, in the configuration 2, the PR-S^{2}aaS generates improvement of NDCG as: 22% for service 2 and 8% for service 5. The reason is that for the cost-based method, the shortest route may consist of highly important nodes which consume high energy. This considerably reduces the energy of nodes. Therefore, those services were not preferred by the ranking mechanism. However, there exist NDCG@5 in the configuration 1 and NDCG@4 in the configuration 2 with the highest NDCG values for the cost-based method, because some participating nodes along the shortest route are highly influential. Those nodes have high outgoing degree. Thus, causing the energy level to be low. However, for all other services in the both configurations, the NDCG values for the cost-based method are low as compared to the proposed method, because of the above described findings. Furthermore, NDCG@1 and NDCG@3 in the configuration 2 are zero for both methods, because these services may not match any query and not result in any of the discovery sets.

For the complete overview of the performance evaluation, the NDCG values of the services are averaged across queries for the both configurations and resultant graphs are plotted in Figure 3 and Figure 4. The performance of PR-S^{2}aaS is consistent in terms of the NDCG as compared to the cost-based method. The PR-S^{2}aaS is approximately 3% and 7% better than the cost-based method for the configuration 1 and the configuration 2. There is an aggregated improvement of 10% in the ranking quality by the PR-S^{2}aaS method as compared to the cost-based method. It should be noted that we have no NDCG values for some services because during experiments those nodes may have not satisfied any query.

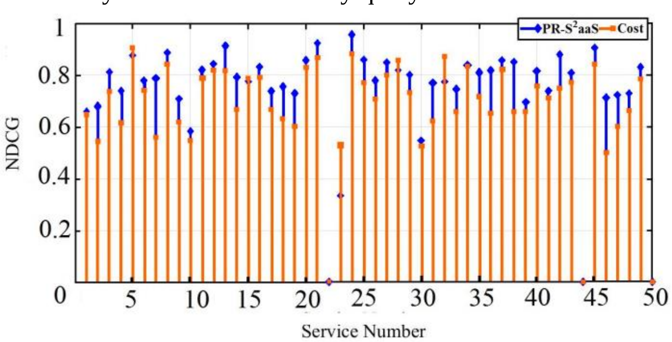

Figure 3: Averaged NDCG values for Configuration 1

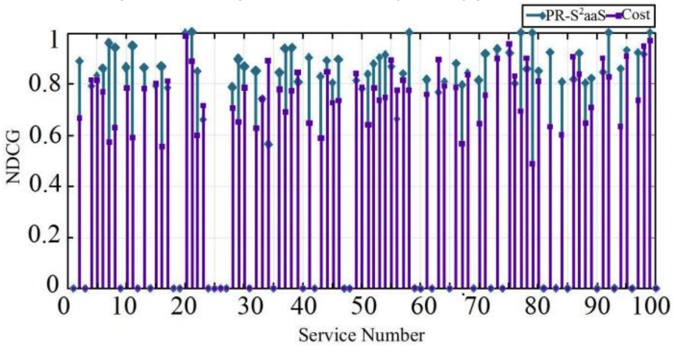

Figure 4: Averaged NDCG values for Configuration 2

Furthermore, the cost-based method improves the NDCG value of some services as compared to the PR-S^{2}aaS method. These services are: 5, 15, 23, 28, and 32 in the configuration 1 and 4, 15, 23, 34, 39, 49, 56, 63, 75, 76, 78, 86, 91, and 99 in the configuration 2. This is because of the above-cited reasons.

2) *Statistical Analysis*

Another interesting observation from Figure 3 and Figure 4 is the quality of both methods increase and decrease showing some random behavior which we have studied by calculating the statistical measures as shown in Table III. The standard deviation and mean values for PR-S^{2}aaS are low as compared to the cost-based method. We can see from Figures 3 and 4 that values of the ranking quality for PR-S^{2}aaS have less variation around mean, i.e., standard deviation of 0.20 and 0.38 close to the mean values (0. 69 and 0.84 in both configurations). In the configuration 1, the deviation of the ranking quality from the mean value for the proposed method is about 20%. However, the deviation of the ranking quality from the mean for the cost-based method is high, i.e., 0.3139. In the configuration 2, the deviation of the ranking quality from the mean value is about 30% for PR-S^{2}aaS and about 40% for the cost-based method.

TABLE III
STATISTICAL ANALYSIS OF BOTH METHODS FOR CONFIGURATIONS 1 AND 2

Measures	Mean		Standard deviation		Range	
Configuration	1	2	1	2	1	2
PR-S^{2}aaS	0.69	0.84	0.20	0.38	0.91	0.89
Cost	0.73	0.79	0.31	0.41	0.96	0.92

The range represents the dispersion of the ranking quality of both methods. The maximum ranking quality value and the minimum ranking quality value for the PR-S^{2}aaS method are 0.9069 and 0.0013 in the configuration 1 and 0.9183 and 0.0331 in the configuration 2. The value of the range for the proposed method is less than the range value for the cost-based method. Therefore, the ranking quality values deviated less for the proposed method and deviated higher for the cost-based method from their respective mean values in both configurations.

3) *Energy Results*

The evaluation of the energy consumption of the PR-S^{2}aaS method against the cost-based method in collection of reading values and onion forwarding is discussed in this section. The initial energy assigned to each sensor node was 1 Joule. We assumed sensor node energy level in percentage for the plotting purpose, where 10J is equivalent to 0.1% and 100J is equal to 1.0%. It can be seen from Figure 5 and Figure 6 that the number of dead nodes for PR-S^{2}aaS were 3 and 5 in the configuration 1 and configuration 2. However, the total number of nodes died completely for the cost-based method are 12 and 42 in the configuration 1 and configuration 2. We also analyze the number of nodes having the remaining energy below than 20%. These nodes were 2 and 4 for the PR-S^{2}aaS method and 14 and 22 for the cost-based method in the configuration 1 and configuration 2, respectively. The performance of the PR-S^{2}aaS method in terms of energy consumption was 11% and 21% better than the cost-based method in the configuration 1 and configuration 2. Thus, there is an aggregated improvement of 32% in energy efficiency by the PR-S^{2}aaS method as compared to the cost-based method.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

Further, Table IV provides a summary of the dead nodes and nodes in the low energy level zone in both network configurations after processing 50 and 100 queries. The number within a bracket shows completely dead nodes and the number outside the bracket indicates the number of nodes having energy less than the minimum energy threshold level.

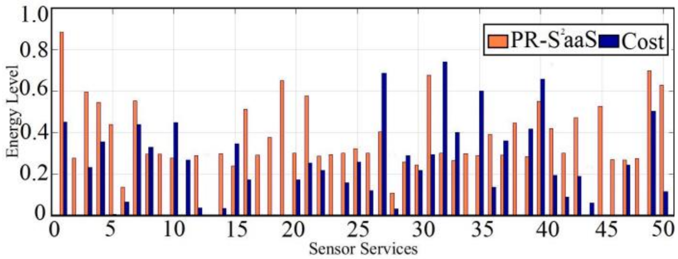

Figure 5: Remaining energy of all nodes after processing 100 queries for Configuration 1

TABLE IV
SUMMARY OF NODES ENERGY LEVELS FOR BOTH CONFIGURATIONS

Queries	50		100	
Configurations	1	2	1	2
PR-S^2aaS	0(0)	3(0)	2(3)	4(5)
Cost	6(4)	18(40)	14(12)	22(42)

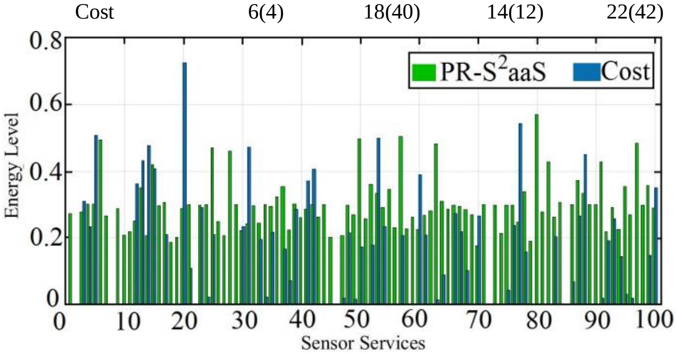

Figure 6: Remaining energy level of all nodes after processing 100 queries for Configuration 2

4) Privacy Results

We present the privacy related results illustrating the ranking quality, i.e., average value of NDCG and percentage of infected (compromised) nodes for the proposed work within each network configuration in Figure 7. The x-axis represents the network configurations and y-axis represents the percentage value for NDCG and infected nodes in each configuration. In the configuration 1, the number of infected nodes (blue bar) is 18%, i.e., 0.18 and in the configuration 2, the number of infected nodes (blue bar) is 22%, i.e., 0.22.

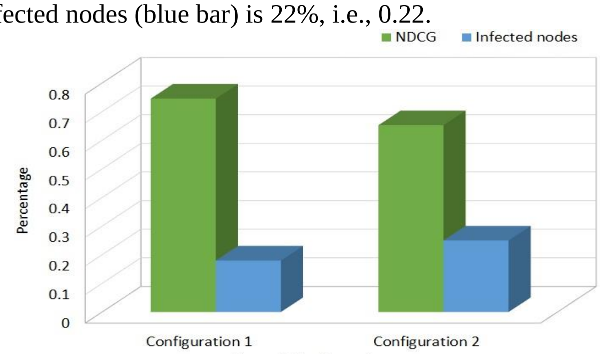

Figure 7: Ranking Quality and Number of Infected Nodes

However, the ranking quality for each of configuration is not much affected, i.e., NDCG (green bar) in Configuration 1 is 74% i.e., 0.74 and NDCG (green bar) in Configuration 2 is 65%, i.e.,0.65.

5) **Topological Features**

We have plotted the topological features against the query 1 consisting of 12 services in Figure 8 for the configuration 1 and Figure 9 for the configuration 2. It can be observed from Figure 8 that in terms of the influence (feature 1), the service 9 is considered as highly influential. However, the service 11 is best as it has low value of influence, i.e., it is less important and located at less dense area. In terms of the degree (feature 2), the service 9, the service 12, the service 2, and the service 4 have the highest values, followed by the service 1, the service 3, the service 6, the service 5, the service 7, the service 8, the service 10 and the service 11. However, the service 11 and the service 8 are better because of the low degree values, i.e., 0.19 and 0.22, respectively. The service 11 has high energy (feature 3) as compared to the other services, because it is less influential, having minimum degree value.

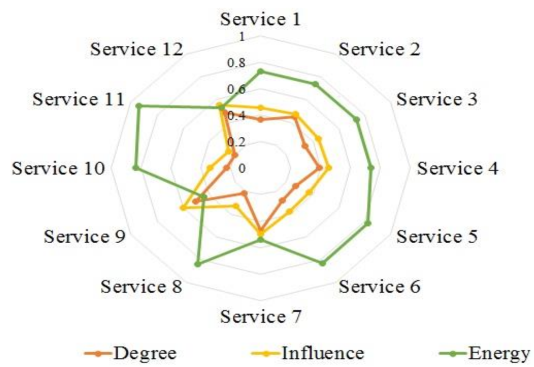

Figure 8. Radar Chart for services for query 1 in Configuration 1

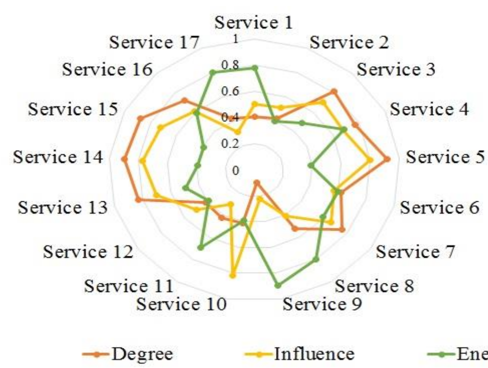

Figure 9. Radar Chart for services for query 1 in Configuration 2

Figure 9 shows the service 9 is better as it has a less degree value, i.e., 0.10 as well as it is a less influential service. However, the service 5 is highly influential. Therefore, the outgoing connections of the Service 5 are high. The service 8, the service 9, and the service 17 have high-energy levels as compared to the other services. Furthermore, if the influence of a node is high, then the energy level of the node is low, because several tasks are performed by the highly influential node. On the other hand, if the node energy is high, then the node has the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2020.2981563, IEEE Transactions on Industry Applications

low influence. Moreover, if a node has less value of the influence, then it has low value for the degree feature. The direct relation between the influence and degree against the query 1 is shown in Figure 10 and Figure 11.

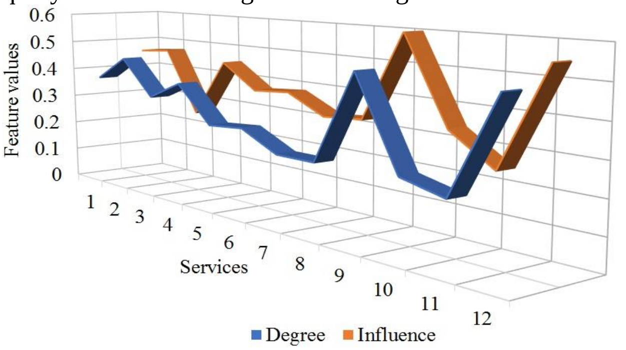

Figure 10. Relation between degree and influence in Configuration 1

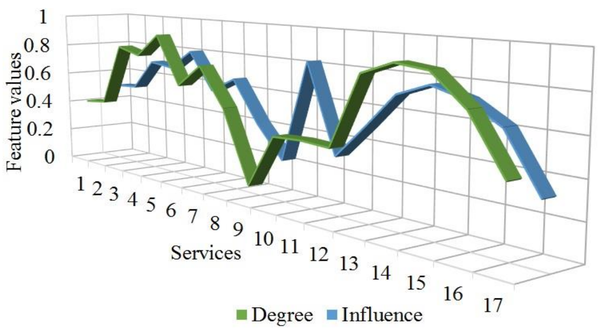

Figure 11. Relation between degree and influence in Configuration 2

VII. Conclusion

The QoS information-based ranking strategies suffer from the resource constraint property of sensor nodes. We have presented a novel energy and time efficient privacy preserved ranking method, PR-S^{2}aaS, for the ranking of services. The ranking is computed based on the information available at the sensor node level. The PR-S^{2}aaS employs a feature set which is computed using topological information such as influence value, energy level, and degree of the sensor nodes. The proposed method computes a stochastic shortest route of each service and performs anonymous communication. The valuation technique, to find the value (cost) of each service using a topological information-based feature set, is also presented. Finally, the services are ranked according to the computed values. The proposed work is compared with an existing cost-based ranking method in the different realistic network configurations. The results indicate a significant improvement in the ranking quality (up to 10%). The proposed scheme also has the potential to preserve energy consumption (up to 32%) in different network settings.

The future direction of this work includes the incorporation of the PR-S^{2}aaS method into IoT middleware where it can rank heterogeneous nature of services around a given spatial region.

References

[1] Cisco.com. (2019). [online] Available at: https://www.cisco.com/c/dam/en/us/products/collateral/se/internet-of-things/at-a-glance-c45-731471.pdf [Accessed 24 Dec. 2019].

[2] X. Sheng, J. Tang, X. Xiao and G. Xue. (2013). Sensing as a Service: Challenges, Solutions and Future Directions. *IEEE Sensors*. [Online]. *13(10), pp.* 3733-3741.

[3] D. Guinard, V. Trifa, S. Karnouskos, P. Spiess and D. Savio. (2010). Interacting with the SOA-Based Internet of Things: Discovery, Query, Selection, and On-Demand Provisioning of Web Services. *IEEE Transactions on Services Computing*. [online]. 3(3), 223-235.

[4] S. Chatterjee and S. Misra, "QoS estimation and selection of CSP in oligopoly environment for Internet of Things," in *Proc. IEEE WCNC*, Doha, 2016, pp. 1-6.

[5] Neha and S. Saxena, "Vector method for ranking of sensors in IoT," in *Proc. ICICT*, Coimbatore, 2016, pp. 1-5.

[6] K. K. F. Yuen and W. Wang, "Towards a ranking approach for sensor services using primitive cognitive network process," in *Proc. The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*, Hong Kong, 2014, pp. 344-348.

[7] C. Basaran and K. Kang, Quality of Service in Wireless Sensor Networks. in Computer Communications and Networks, ed. by S Misra, I Woungang ,S Misra (Springer, London, 2009), p. 305-321.

[8] C. Perera, A. Zaslavsky, P. Christen, M. Compton, and D. Georgakopoulos, "Context-Aware Sensor Search, Selection and Ranking Model for Internet of Things Middleware," in *Proc. MDM,* Washington, 2013.

[9] C. Perera, A. Zaslavsky, C. H. Liu, M. Compton, P. Christen and D. Georgakopoulos. (2014, Feb.). Sensor Search Techniques for Sensing as a Service Architecture for the Internet of Things. *IEEE Sensors Journal*. [Online]. 14(2), pp. 406-420.

[10] W. Niu, J. Lei, E. Tong, G. Li, L. Chang, Z. Shi and S. Ci. (2013, Jan.). Context- Aware Service Ranking in Wireless Sensor Networks. J. Network and Systems Management [online]. 22(1), pp. 50–74.

[11] C. Truong and K. Römer, "Content-based sensor search for the Web of Things," in Proc. *GLOBECOM*, Atlanta, GA, 2013, pp. 2654-2660.

[12] K. R. R. Babu, S. J. George and P. Samuel, "Optimal sensor selection from sensor pool in IoT environment," in *Proc. iCATccT*, Bangalore, 2016, pp. 697-702.

[13] G Cassar, P Barnaghi, W Wang and K Moessner, "A Hybrid Semantic Matchmaker for IoT Services," in *Proc. IEEE CGCC*, Besancon, France, 2012, pp. 20-23.

[14] P. Zhang, Y. Liu, F. Wu, S. Liu and B. Tang. (2016, April.). Low-Overhead and High-Precision Prediction Model for Content-Based Sensor Search in the Internet of Things. *IEEE Communications Letters*. [online]. 20(4), pp. 720-723.

[15] W. Wang, F. Yao, S. De, K. Moessner and Z. Sun. (2015). A Ranking Method for Sensor Services Based on Estimation of Service Access Cost. *J. Information Sciences.* [online]. 319, 1-17.

[16] Y. Fathy, P. Barnaghi and R. Tafazolli. (2018). Large-Scale Indexing, Discovery, and Ranking for the Internet of Things (IoT). *ACM Computing Surveys*. [online]. 51(2), pp. 1-53.

[17] I. Hussain, J. Pilz and G. Spoeck. (2011). Homogeneous climate regions in Pakistan. *International Journal of Global Warming*. [online]. 3(12), pp. 55.

[18] K. T. Chang. (2015.) *Introduction to Geographic Information Systems. Coordinate Systems.* (8th ed.) [online].

[19] F Spitzer. (2013). The Classification of Random Walk. *Principles of Random Walk*. [online]. 34, pp. 1-35.

[20] B. Travençolo and L. da F. Costa. (2008). Accessibility in complex networks. *J. Physics Letters* 373, pp. 89-95.

[21] T. Cormen, C. Leiserson, R. Rivest and C. Stein, *Introduction to algorithms*.

[22] *Memsic.com*, 2019. [Online]. Available: http://www.memsic.com/userfiles/files/Datasheets/WSN/IRIS_Datasheet.pdf. [Accessed: 20- Apr- 2019].

[23] M Noori, M Ardakani, "A probabilistic lifetime analysis for clustered wireless sensor networks," in *Proc. WCNC,* Las Vegas, NV, USA, 2008.

[24] K. Järvelin and J. Kekäläinen. (2002). Cumulative gain-based evaluation of IR techniques. ACM Transactions on Information Systems. [online]. 20, pp. 422-446.

[25] E. De Cristofaro, X. Ding and G. Tsudik, "Privacy-Preserving Querying in Sensor Networks," *2009 Proceedings of 18th International Conference on Computer Communications and Networks*, San Francisco, CA, 2009, pp. 1-6.

[26] J. Sengupta, S. Ruj, and S. Das Bit, "End to end secure anonymous communication for secure directed diffusion in IoT," 2019 *Proceedings of the 20th International Conference on Distributed Computing and Networking* (ICDCN '19), New York, NY, USA, pp. 445-450.

BIOGRAPHIES

Aroosa Hameed received her MPhil degree in Computer Science from the Quaid-i-Azam University, Pakistan in 2018. She is currently pursuing her Ph.D. research at ETS Montreal (University of Quebec), Canada. Her current research interests include IoT services, IoT security, IoT analytics, wireless sensor networks, and machine learning.

Dr Muhammad Usman received the MS (Computer Science) from the PMAS-AAUR, Pakistan with first class and obtained Ph.D. from School of Information and Communication Technology, Griffith University, Australia. He severed as a Postdoc Research Fellow in Cyber Security and Machine Learning at University of Surrey, UK. He possesses 17 years of experience during which he has held several academic and industrial positions in different parts of the globe including Australia, Asia and Europe. His current research interests include design and analysis of security and privacy techniques for complex cyber-physical systems such as aviation and maritime transport systems; security and privacy issues of IoT-driven and smart systems; security, trust and privacy of cross-discipline domains; formal and statistical modelling; and applied machine learning and data analytics in several domains such as healthcare, social science, etc. He has published more than thirty-five research papers in international journals and conferences including prestigious journals such as IEEE Transactions on Consumer Electronics and IEEE Transactions on Industry Applications. He has also published a book focused on Mobile Agent-Based Anomaly Detection and Verification System for Smart Home Sensor Networks. He has successfully supervised over 25 postgraduate research and undergraduate project students and supervising PhD students. He has been recipient of several research and travel grants. He has led, and acted as an associate guest editor, in special issues in IEEE transactions. He has served in different capacities such as organising committee member, TCP member, focal person and/ or publication chair of several international IEEE conferences. He is a member of Network Security Research Group at the Institute for Integrated and Intelligent Systems at Griffith University, Australia. He is also a member of Computer Science Teachers Association, endorsed by ACM, USA. He is a Juniper certified networking and security specialist. His research paper has received the best paper award in IEEE ComTech 2017.

Onaiza Maqbool received her Ph.D. degree in Computer Science from the Lahore University of Management Sciences in 2006. She is a Tenured Associate Professor at the Department of Computer Science, Quaid-i-Azam University, Islamabad, Pakistan. Prior to joining Quaid-i-Azam University, she worked in the software industry for some years. Her research interests lie in exploring machine learning techniques to solve problems in various areas including software engineering, social networks, and natural language processing. She has published more than thirty papers in well reputed journals and conferences.