Qamar, A, Asim, M, Maamar, Z, Saeed, S and Baker, T

A Quality-of-Things Model for Assessing the Internet-of-Thing's Non-Functional Properties

Article

Qamar, A, Asim, M, Maamar, Z, Saeed, S and Baker, T (2019) A Quality-of-Things Model for Assessing the Internet-of-Thing's Non-Functional Properties. Transactions on Emerging Telecommunications Technologies. ISSN 2161-3915

A Quality-of-Things Model for Assessing the Internet-of-Thing's Non-Functional Properties

Ayesha Qamar, Muhammad Asim, Zakaria Maamar, Saad Saeed, and Thar Baker

***Abstract*—The Internet of Things (IoT) is in a "desperate" need for a practical model that would help in differentiating things according to their non-functional properties. Unfortunately, despite IoT growth, such properties either lack or ill-defined resulting into ad-hoc ways of selecting similar functional things. This paper discusses how things' non-functional properties are combined into a Quality-of-Things (QoT) model. This model includes properties that define the performance of things' duties related to sensing, actuating, and communicating. Since the values of QoT properties might not always be available or confirmed, providers of things can tentatively define these values and submit them to an Independent Regulatory Authority (IRA) whose role is to ensure fair competition among all providers. The IRA assesses the values of non-functional properties of things prior to recommending those that could satisfy users' needs. To evaluate the technical doability of the QoT model, a set of comprehensive experiments are conducted using real datasets. The results depict an acceptable level of the QoT estimation accuracy.**

***Index Terms*—Competition, Internet-of-Things, Quality-of-Things, Thing Selection.**

I. Introduction

In an open environment like the Internet, coming up with the right price for a product or service is always "tricky" [1], [2]. Indeed, many factors contribute to pricing such as availability of alternative products/services, potential customers' profiles, and existing legislations. In the ICT field, an example of online services are Web services (WS) that usually come with specific functionalities (e.g., book train ticket) and a set of non-functional properties (e.g., response time and availability rate) defining what the R&D community refers to as Quality-of-Service (QoS) [3], [4]. Upon a WS selection among other similarly-functional peers, a Service Level Agreement (SLA) is usually established between this WS's provider/owner and the user/requestor. A SLA is a set of clauses that define the modalities of using a WS (or any other service or product) along with the penalties that this WS's provider would be subject to, should the announced QoS values be not satisfied at run-time [5]. Over time, a provider has the opportunity of adjusting the QoS of its WSs to minimize and/or avoid such penalties (e.g., financial and reputation).

The Internet of Things (IoT) is another ICT discipline that would highly benefit from criteria for differentiating similarly-functional IoT-compliant things (things, for short). According to Gartner[1], 6.4 billion connected things were in use in 2016, up 3% from 2015, and will reach 20.8 billion by 2020. So, how to select the "right" and "best" things? To the best of our knowledge, selection criteria with an IoT flavor do not exist, yet, and, hence, developing a Quality-of-Things' (QoT) model (like QoS) would be necessary when specialized marketplaces for things will become available soon according to Perera et al. [6]. A QoT model would consist of a set of non-functional properties that would be specifically geared towards the peculiarities of things in terms of what they do, with whom they interact, and how they interact. In this paper, we refer to what-things-do as duties and specialize them into *sensing*, *actuating*, and *communicating* [7].

Since the QoT model is something relatively new, coming-up with values for things' non-functional properties is not straightforward (e.g., there is not any benchmark, yet). To this end, we adopt a 2-step strategy in which the values of QoT properties are tentatively assigned and then, adjusted over time. The strategy also ensures fair competition among all things' providers. This happens through an Independent Regulatory Authority (IRA) that checks if the tentative values of the non-functional properties are neither under-estimated nor over-estimated. Under-estimated could lead to monopoly. And, over-estimated could lead to limited business. The IRA benchmarks the different things' non-functional properties so, that, a scale of performance would be established over time. IRA is analogous to a cloud/service broker who is an intermediate entity between users and providers, and that helps the users choose services tailored to their requirements [8].

In line with Issarny et al.'s statement that revisiting service-oriented architecture for the IoT context is a must [9], we embrace the same when it comes to adjusting existing selection techniques like those used in the context of WSs, and/or developing new selection techniques[2]. According to Issarny et al., many features make IoT context unique including scalability, deep heterogeneity, high dynamics, and uncertainty. On top of these features we include diversity and multiplicity of things' development and communication technologies [1], users' re-

Ayesha Qamar is with National University of Computer and Emerging Sciences, Islamabad, Pakistan. (e-mail: i161069@nu.edu.pk)

Muhammad Asim is with National University of Computer and Emerging Sciences, Islamabad, Pakistan. (e-mail: muhammad.asim@nu.edu.pk)

Zakaria Maamar is with Zayed University, Dubai, U.A.E. (email:Zakaria.Maamar@zu.ac.ae)

Saad Saeed is with National University of Computer and Emerging Sciences, Islamabad, Pakistan. (e-mail: saadsaeed01@live.com)

Thar Baker is with Liverpool John Moores University, Liverpool, UK. (email: t.baker@ljmu.ac.uk)

[1] www.gartner.com/newsroom/id/3165317.

[2] Thing discovery does not fall into this work's scope.

luctance and sometimes rejection because of privacy invasion caused by things [2], and, passive nature of things being mainly data suppliers (with some *actuating* capabilities) [10]. In this paper, we present the design and development of an approach for defining and assessing the non-functional properties of things. These properties contribute towards defining a QoT model. The main contributions of our work are as follows:

- A QoT model that captures things' non-functional properties. These properties refer to the performance of things' duties in the form of sensing, actuating, and communicating.
- Different from other existing QoS prediction approaches, the proposed approach deals with a problem in which no real usage experience is available for a thing-to-evaluate.
- A novel approach for thing selection and recommendation which enhances Spearman's rank correlation by considering information of resources (energy) linked to things' duties during the correlation.
- Several experiments are performed to evaluate the estimation accuracy of the proposed approach.

The rest of this paper is organized as follows. Section II presents some related work and a case study. Section III describes the QoT model for IoT. Section IV discusses the details of the thing selection and recommendation. Experiments and evaluation details are given in Section V. Conclusions and future directions are drawn and listed in Section VI, respectively.

II. Background

This section first, discusses related work, and, then presents a case study of smart city where different things are in operation.

A. Related work

The plethora of techniques (e.g., selection and recommendation) applied to WSs are deemed inappropriate for IoT [9]. Things operate in a cyber-physical surrounding, while WSs operate in a cyber-surrounding, only. To address this inappropriateness and lack of IoT driven references, we discuss some WS works that helped in defining our QoT model as well as some works on QoS for IoT applications [11]. Today's IoT user-application stakeholders are expected to engage in collaborative scenarios requiring the coordination of millions of heterogeneous devices that have reduced size, restricted connectivity, continuous mobility, limited energy, and constrained storage. Moreover, these applications are expected to combine sensing, actuating, and communicating into composite scenarios that would run over networks of different sizes with unknown network topologies as well as uncertain features of things [1].

The existing literature on QoS refers to traditional WSs in a non-IoT context. Wang et al. [12] propose a Collaborative Filtering (CF)-based approach for QoS prediction for recommending services in mobile edge computing. They consider mobility of users and volatility of historical QoS data. Location-aware personalized CF approaches [13], [14] have been proposed in which both locations of users and locations of WSs are considered when selecting similar neighbors for the target user or service. Yu et al. [15] propose a QoS prediction approach by taking both time and location-aware CF into account. But multiple time slots are used to predict missing QoS values, which makes the approach slow. A time-aware and data sparsity tolerant WSs recommendation approach [16] integrates time information and transitive similarity. Time information of temporally close QoS experience from 2 users on a same service and more recent QoS experience from 2 users on a same service is integrated into the similarity measurement. A trustworthy 2-phase WSs discovery mechanism based on CF and QoS has been proposed by Lin et al. [17]. Observer agents collect records about users' behaviors, including querying and invoking WSs to monitor actual QoS. Reputation-aware QoS value prediction approaches [18], [19] for WSs have been presented to identify untrustworthy users. However, the scope is limited in term of user-based similarity and similar services are not identified.

White et al. [20] quantitatively evaluate already existing different matrix factorization approaches (e.g., CloudPred, Extended Matrix Factorization, and Latent Factor Models) to make QoS predictions by investigating the past usage of similar users and IoT services. In [21], White et al. present a matrix factorization-based IoT Predict approach for collaborative QoS predictions for IoT services. No additional invocations of IoT services are required which lowers the load on the network.

Unfortunately, the aforementioned works do not discuss a comprehensive strategy of how things should be described, announced, selected, invoked, rewarded, etc. So, there is a pressing need for such a strategy that would sustain IoT growth.

B. Case study

We refer to a smart city in which multiple systems backed by different things (e.g., humidity sensors, road cameras, and traffic signs) are in operation. Examples of systems include Transportation ($\mathcal{T}_{sys}$) that controls road traffic and Environment ($\mathcal{E}_{sys}$) that monitors air pollution so, that, it advises $\mathcal{T}_{sys}$ to divert traffic on some roads.

Following a car accident and, then, tunnel closure, the city's cameras and passing-by vehicles' embedded cameras form an ad-hoc collaborative environment in a way that permits to provision live streaming of the accident scene to rescue teams. However, including 3rd-party unknown cameras (or things, in general) in this environment would require assessing these cameras' QoT properties with respect to the requirements of the current situation that is rescuing passengers. Since the values of these QoT properties might not always be available or confirmed, we proceed by allowing providers/owners of things to tentatively estimate the QoT properties of their things. The providers "hope" that these properties' values are fair; neither under-estimated nor over-estimated. To achieve this fairness, providers of things (e.g., cameras) resort to the IRA that would use recommendation techniques to assess this fairness.In the context of smart cities, city halls could act as an IRA.

III. Quality-of-Things model in IoT

Defining a QoT model for things is in line with the trend of defining similar models in other ICT fields for instance, cloud services and WSs [22]. Eisa et al. define a quality model as a degree to which a set of attributes/properties of a service fulfils stated requirements [3].

To enable a competitive selection of similarly-functional things with respect to situations' non-functional requirements, we develop properties that would constitute a thing's QoT model [7]. This model would revolve around 3 duties (Fig. 1): sensing (in the sense of collecting/capturing data), actuating (in the sense of processing/acting upon data), and communicating (in the sense of sharing/distributing data). According to Fig. 1, a thing senses the cyber-physical surrounding, so, that, it generates some outcomes; a thing actuates outcomes with focus on the outcomes that result from sensing; and a thing communicates with the cyber-physical surrounding the outcomes that result from both sensing and actuating.

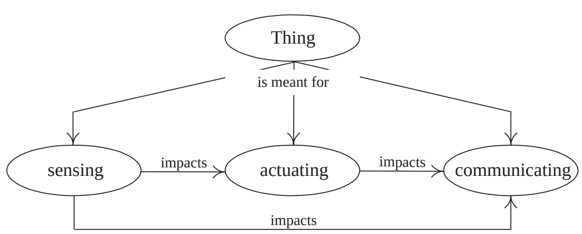

Fig. 1: Duties upon which a thing's QoT model is built

Below we propose some QoT properties (but not limited to) per type of duty:

1) QoT properties for assessing sensing include:
   - *Frequency of sensing* (e.g., continuous *versus* intermittent).
   - *Quality of sensed outcome* determines for instance, the accuracy and validity of the outcome (e.g., high *versus* low accuracy.
   - *Resource* (e.g., energy, CPU, and storage) consumption during sensing (e.g., high *versus* low energy).
2) QoT properties for assessing actuating include:
   - *Quality of actuated outcome* determines for instance, the accuracy and validity of the outcome.
   - *Resource* (e.g., energy, CPU, and storage) consumption during actuating (e.g., high *versus* low).
3) QoT properties for assessing communicating include:
   - *Reception rate of sensed and/or actuated outcome* (incoming flow) determines for instance, data loss, and data volume with respect to a bandwidth.
     – *Acceptance rate of received outcome* is about the outcome that has been accepted for distribution; some received outcome could be rejected.
   - *Delivery rate of sensed and/or actuated outcome* (outgoing flow) determines data loss, and data volume with respect to a bandwidth, etc.
     – *Acceptance rate of delivered outcome* is about the outcome that has been accepted after distribution at the recipient end; some delivered outcome could be rejected.
   - *Resource* (e.g., energy and bandwidth) consumption during communicating (e.g., high *versus* low bandwidth).

IV. Thing selection and recommendation approach

This section provides details about our thing selection and recommendation approach backed up by its architecture. It further explains this architecture's 3 modules namely *thing clustering*, *thing similarity assessment*, and *thing QoT estimation*.

A. Architecture

Fig. 2 represents the modules and chronology of interactions in our approach for defining and assessing the QoT properties of things. It all starts when providers assign tentative values for their "new" things' QoT properties and submit these values to the IRA. We assume that details about other things' QoT properties (values and locations like IP addresses) are available to the IRA by consulting the relevant repository (Fig. 2). The IRA clusters all things on the basis of their locations through the *thing clustering module* (Subsection IV-C). Then, by identifying those existing things that fall into the same cluster as the "new" thing, the IRA evaluates their similarities using the *thing similarity assessment module*. This consists of checking the things' respective duties and users' past experiences with invoking these things (Subsections IV-D and IV-E). The *thing similarity assessment module* helps the IRA evaluate if the tentative QoT values of the "new" thing are fair (neither under- or over-estimated according to specific thresholds) through the *thing QoT estimation module* (Subsection IV-F). Finally, the evaluation's feedback is sent back to the "new" thing's provider for further actions like adjustment, if necessary.

B. Formal definitions

To support the definition and assessment of things' QoT properties, we provide the following formal definitions:

- $\mathsf{T} = \{\mathsf{t}_1, \mathsf{t}_2, ..., \mathsf{t}_s\}$ is a set of all things that could be running in an ecosystem of IoT (e.g., smart cameras, ambient sensors, and vehicles) that users can invoke their duties.
- Thing-to-Evaluate (e.g., a new smart camera) represented as $\mathsf{t}_i^? \in \mathsf{T}$ has its QoT properties' values tentatively defined and hence, these values need to be confirmed by the IRA before $\mathsf{t}_i^?$ takes part in any situation (e.g., passenger rescue).
- Existing-Thing (e.g., an existing smart camera) represented as $\mathsf{t}_j^{ex} \in \mathsf{T}$ has its QoT properties' values already known and could constitute a competitor to the thing-to-evaluate, should both have the same duty among other details.

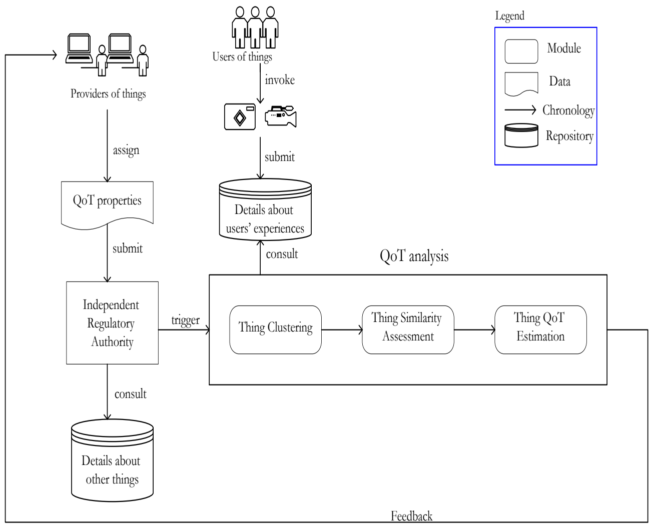

Fig. 2: Architecture of the approach for defining and assessing things' QoT properties

- Cl = $\{\mathsf{cl}_1, \mathsf{cl}_2, ..., \mathsf{cl}_k\}$ is a set of all things' location-based clusters where $\mathsf{cl_j}^{\mathsf{t}_i^?}$ is the j^{th} cluster to which $\mathsf{t}_i^?$ belongs.
- Pr = $\{\mathsf{pr}_1, \mathsf{pr}_2, ..., \mathsf{pr}_k\}$ is a set of all providers of things where $\mathsf{pr_i}^{\mathsf{t}_i^?}$ is a provider of $\mathsf{t}_i^?$ and $\mathsf{pr_j}^{\mathsf{t}_j^{ex}}$ is a provider of t_j^{ex}; it happens that $\mathsf{pr_i}$ and $\mathsf{pr_j}$ are same.
- D = $\{\mathsf{d}_{\mathsf{t}_1^?}, ..., \mathsf{d}_{\mathsf{t}_j^{ex}}\}$ is a set of all things' duties where $\mathsf{d}_{\mathsf{t}_1^?}$ is any specific duty ($\mathsf{sensing}$ or $\mathsf{actuating}$ or $\mathsf{communicating}$) that $\mathsf{t}_1^?$ performs and $\mathsf{d}_{\mathsf{t}_j^{ex}}$ is any specific duty that t_j^{ex} performs, too. $\mathsf{d}_{\mathsf{t}_1^?}$ and $\mathsf{d}_{\mathsf{t}_j^{ex}}$ could be either different or same (i.e., competition).
- QP = $\{\mathsf{t}_1^{ex}\mathsf{v}_1^d, ..., \mathsf{t}_s^{ex}\mathsf{v}_s^d\}$ is a set of all values of QoT properties for all existing things' duties where $\mathsf{t}_j^{ex}\mathsf{v}_j^d$ is a value of a particular QoT property for a particular duty d that t_j^{ex} offers to users.
- TV = $\{\mathsf{t}_i^?\mathsf{tv}_1^d, ..., \mathsf{t}_i^?\mathsf{tv}_s^d\}$ is a set of all $\mathsf{tentative}$ values of QoT properties for all thing-to-evaluate duties where $\mathsf{t}_i^?\mathsf{tv}_i^d$ is a $\mathsf{tentative}$ value of a particular QoT property for a particular duty d, declared by $\mathsf{pr_i}^{\mathsf{t}_i^?}$ for $\mathsf{t}_i^?$.

C. Clustering of things

Physical location of things play an important role in assessing their similarity [23]. Those things which are geographically located close to each other are more likely to exhibit similar QoT values e.g., response time [14]. Since the IRA has access to the location details about $\mathsf{t}_i^?$ and $\{\mathsf{t}_j^{ex}\}$ in the repository (Fig. 2), it proceeds with processing the QoT records of existing things $\{\mathsf{t}_j^{ex}\}$ that are closely located to $\mathsf{t}_i^?$. Moreover, this location-based processing saves time and computation as similarity targets those things that are in the "close vicinity" of $\mathsf{t}_i^?$, instead of considering all available things (T). We use K-means to cluster things on the basis of Latitude-Longitude information. We refer readers to [13] about the benefits of using Latitude-Longitude when assessing WSs (example of software things) similarity. The identifiers of the geographically selected $\{\mathsf{t}_j^{ex}\}' \subseteq \{\mathsf{t}_j^{ex}\}$ and $\mathsf{t}_i^?$ are sent forward to the *thing similarity assessment module* for processing.

D. Assessment of similarity of things

Based on the things' identifiers that the *thing clustering module* sends the *similarity assessment module*, this latter begins by narrowing down the most similar existing things $\{\mathsf{t}_j^{ex}\}'$ that have similar historical QoT records (in terms of duties and QoT properties) as that of $\mathsf{t}_i^?$. For this purpose, we use Spearman's rank correlation coefficient (ρ) [24], which gives the similarity between things (not necessarily IoT-compliant things) based on the ranked values of 2 variables. This is done separately for each QoT property as there can be multiple QoT properties for which we have to find the similarity of $\mathsf{t}_i^?$ with $\{\mathsf{t}_j^{ex}\}'$ that fall into $\mathsf{cl_j}^{\mathsf{t}_i^?}$ (Equation 1):

$$\rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})} = 1 - \frac{6\sum(diff)^2}{m(m^2-1)} \tag{1}$$

where $\mathsf{t}_j^{ex'} \in \{\mathsf{t}_j^{ex}\}'$, $\rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})}$ represents the similarity of $\mathsf{t}_i^?$ with $\mathsf{t}_j^{ex'} \in \mathsf{cl_j}^{\mathsf{t}_i^?}$, "*diff*" is the difference of rankings of $\mathsf{t}_i^?\mathsf{tv}_i^d$ and $\mathsf{t}_j^{ex'}\mathsf{v}_j^d$ for a particular QoT property for a duty d, and m=$|\mathsf{QP}|$ is the number of elements in the set QP. 6 comes from the sum of squares of integers in the denominator[3]. The similarity value computed from Equation (1) is in the continuous range of [-1, 1]. The greater the computed value, the more similar things are. Since things offer duties that are either atomic (e.g., $\mathsf{sensing}$, only) or combined (e.g., $\mathsf{sensing}$ and $\mathsf{actuating}$), we revise Equation (1) to consider this difference in duties. In Section 3, our proposed QoT model shows a hierarchical structure of duty specialization into $\mathsf{sensing}$, $\mathsf{actuating}$, and $\mathsf{communicating}$, and how each duty is further linked to a set of QoT properties. IoT is a constrained entity that depends on resources. This is depicted in our QoT model through resource property that is common to all duties. Therefore, we link resources to things' duties for instance, energy, to demonstrate the impact of resources on similarity assessment of things' duties. We consider intersection of duties performed by things for this purpose.

- $\mathsf{Atomic\ duties}$: When things perform single duties, i.e., either $\mathsf{sensing}$ or $\mathsf{actuating}$ or $\mathsf{communicating}$, we use the following enhanced Spearman's rank correlation coefficient (Equation 2) to add the effect of a thing's resource/energy on similarity assessment of things:

$$SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP = \frac{1}{\mid E(\mathsf{d}_{\mathsf{t}_1^?}) - E(\mathsf{d}_{\mathsf{t}_j^{ex'}}) \mid} * \rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})} \tag{2}$$

where $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})|QP$ represents the similarity measure after adding the effect of resources linked to things' atomic duties on thing similarity assessment, and $E(\mathsf{d}_{\mathsf{t}_1^?})/E(\mathsf{d}_{\mathsf{t}_j^{ex'}})$ represents energy consumption during any specific duty d that $\mathsf{t}_i^?/\mathsf{t}_j^{ex'}$ performs. When energy

[3] http://mathforum.org/library/drmath/view/52774.html.

consumption for a duty d for both $\mathsf{t}_i^?$ and $\mathsf{t}_j^{ex'}$ is same, we take $|E(\mathsf{d}_{\mathsf{t}_1^?}) - E(\mathsf{d}_{\mathsf{t}_j^{ex'}})| = 1$.

- Combined duties: When things perform multiple duties, e.g., sensing and actuating, we revise Equation (2) as follows (Equation 3):

$$SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP = \frac{1}{\sum_{x=y=1}^{N=(n^? \cap m^{ex'})} |E(\mathsf{d}_{\mathsf{t}_1^?}(x)) - E(\mathsf{d}_{\mathsf{t}_j^{ex'}}(y))|} * \rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})} \tag{3}$$

where $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})|QP$ represents the similarity measure after adding the corresponding collaborative effect of things' duties on similarity assessment of things. $n^?/m^{ex'}$ represents the total number of separate duties that $\mathsf{t}_i^?/\mathsf{t}_j^{ex'}$ perform, and N represents the total number of common combined duties that $\mathsf{t}_i^?$ and $\mathsf{t}_j^{ex'}$ perform. We add the difference of energy consumption among all the common combined duties of $\mathsf{t}_i^?$ and $\mathsf{t}_j^{ex'}$ to see their overall combined impact on thing similarity. If the energy consumption among all the common combined duties of $\mathsf{t}_i^?$ and $\mathsf{t}_j^{ex'}$ is same, we take $\sum_{x=y=1}^{N=(n^? \cap m^{ex'})} \mid E(\mathsf{d}_{\mathsf{t}_1^?}(x)) - E(\mathsf{d}_{\mathsf{t}_j^{ex'}}(y)) \mid = 1$.

E. Joint similarity computation

After assessing the similarity of things for each QoT property included in the QoT model, we now need to combine all the individual assessments of all these QoT properties to evaluate against an output label ($\mathsf{L_m}$), e.g., cost of thing, where $\mathsf{L_m} \in \mathsf{L}$ (L is a set of all output labels and L is a subset of QoT properties). This step has to be done as $\mathsf{L_m}$ can be dependent on multiple QoT properties, e.g., it is common that $\mathsf{pr_i}^{\mathsf{t}_i^?}$ announces the tentative values of multiple QoT properties for $\mathsf{t}_i^?$ using one particular cost. To calculate the joint similarity between $\mathsf{t}_i^?$ and $\mathsf{t}_j^{ex'}$, we use Equation (4):

$$Joint_{Sim}(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) = SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP_1 * SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP_2 * ...SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP_{\mathrm{x}} \tag{4}$$

where $Joint_{Sim}(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})$ is a joint product of the similarities based on all the QoT properties, i.e., $\mathrm{QP_1}$, $\mathrm{QP_2}$, $\mathrm{QP_3}$,......,$\mathrm{QP_x}$ computed by Equations (1), (2), and (3).

Next step is to reduce the search space on the basis of similarity assessment of things' duties. If $\mathsf{L_m}$ is to be estimated against tentative values of dependable QoT properties then, after performing the joint similarity computation of $\mathsf{t}_i^?$ with each $\mathsf{t}_j^{ex'}$ (Equation 4), we narrow down those Top-K $\mathsf{t}_j^{ex'}$ from $\{\mathsf{t}_j^{ex}\}'$ that have higher computed similarity value and create a set of these Top-K most similar existing things, i.e., $\{\mathsf{t}_j^{ex}\}'' \subseteq \{\mathsf{t}_j^{ex}\}'$. Let us suppose that K is 5, i.e., top most 5 similar existing things, i.e., $\mathsf{t}_j^{ex''} \in \{\mathsf{t}_j^{ex}\}''$ are selected. If $\mathsf{t}_i^?\mathsf{tv}_i^d$ for an individual QoT property has to be estimated then, after assessment of similarity of things (Equation 1, 2, and 3), we directly do this step (no joint similarity computation is done for this case). Information about $\{\mathsf{t}_j^{ex}\}''$ and $\mathsf{t}_i^?$ is sent forward to the *thing QoT estimation module* for processing.

F. Thing QoT estimation

To estimate the QoT properties, we apply Decision Tree Regression (DTR) technique on data of Top-K most similar selected things, i.e., $\{\mathsf{t}_j^{ex}\}''$. DTR is a promising and efficient technique to address classification and regression problems [25]. It has a hierarchical decision pattern which assists interpret the decision at every level within the decision tree building process. In our case, we deal with a regression problem as our inputs and outputs both consist of continuous values, so it is appropriate to apply DTR to estimate values of QoT properties.

In a DTR-based model, our central choice is to select which QoT property corresponding to $\mathsf{t}_j^{ex''}$ is most useful for splitting the QoT data into multiple points. The QoT property corresponding to $\mathsf{t}_j^{ex''}$ having the lowest Sum of Squared Errors (SSE, Equation 5) among all QoT properties is selected as the root node or split point. This process continues in a recursive manner during the decision tree building.

$$SSE = \sum_{k=1}^{n} (\mathsf{t}_j^{ex''} \mathsf{v}_j^d - avg(QP_i(\mathsf{t}_j^{ex''})))^2 \tag{5}$$

where n=$|QP_i(\mathsf{t}_j^{ex''})|$ is the total number of values of a QoT property (QP_i) for $\mathsf{t}_j^{ex''}$. $\mathsf{t}_j^{ex''}\mathsf{v}_j^d$ represents a particular value of QP_i for $\mathsf{t}_j^{ex''}$, and $avg(QP_i(\mathsf{t}_j^{ex''}))$ represents the mean value of QP_i for $\mathsf{t}_j^{ex''}$.

Fig. 3 shows the regression model that takes ($\mathsf{QP}\{\mathsf{t}_j^{ex}\}''$) and $\mathsf{t}_j^{ex''}\mathsf{v}_j^d(\mathsf{L_m})$ as inputs. ($\mathsf{QP}\{\mathsf{t}_j^{ex}\}''$) represents the QoT properties' values for $\{\mathsf{t}_j^{ex}\}''$ and $\mathsf{t}_j^{ex''}\mathsf{v}_j^d(\mathsf{L_m})$ represents the QoT values of $\mathsf{L_m}$ for $\mathsf{t}_j^{ex''}$. The regressor is trained to build a model that estimates the output, i.e., $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m})$ based on the test data, i.e., $\mathsf{ins}(\mathsf{QP}[\mathsf{t}_i^?])$ that is an instance of QoT properties' values of $\mathsf{t}_i^?$ ("null" values for real usage experience of $\mathsf{t}_i^?$ and $\mathsf{t}_i^?\mathsf{tv}_i^d$ for tentative QoT value of $\mathsf{t}_i^?$). DTR-based model is not trained on $\mathsf{ins}(\mathsf{QP}[\mathsf{t}_i^?])$. $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m})$ is the estimated QoT value of $\mathsf{L_m}$ for $\mathsf{t}_i^?$.

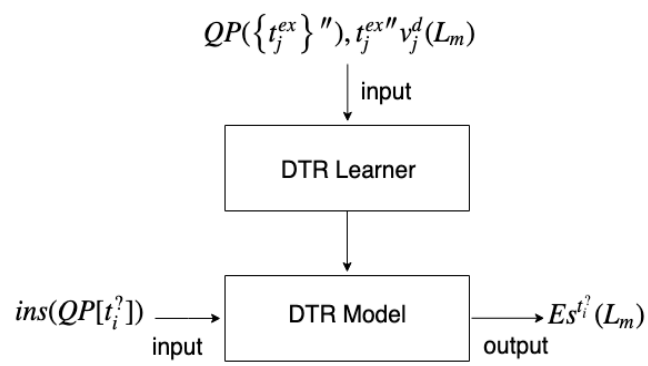

Fig. 3: A DTR-based model for thing QoT estimation

After training the regressor with inputs (Equation 6), the output, i.e., $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m})$ w.r.t. $\mathsf{ins}(\mathsf{QP}[\mathsf{t}_i^?])$, is given by an estimation model (Equation 7).

$$model \leftarrow regressor.train((\mathsf{QP}\{\mathsf{t}_j^{ex}\}''), \mathsf{t}_j^{ex''}\mathsf{v}_j^d(\mathsf{L_m})) \tag{6}$$

$$\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m}) \leftarrow model.estimate(\mathsf{ins}(\mathsf{QP}[\mathsf{t}_i^?])) \quad (7)$$

Algorithm 1: Tentative QoT values estimation

1 **INPUT:** T; Pr; D; QP; TV; L
2 **OUTPUT:** Estimated QoT value
3 **BEGIN:**
4 load QoT properties' data provided by $\mathsf{pr_i}^{\mathsf{t}_i^?}$ and providers of $\{\mathsf{t}_j^{ex}\}$
5 *similar_thingsset* $\leftarrow$ *null;*
6 Initialize centroids $\mu_1, \mu_2, ..., \mu_\mathrm{k}$ for location-based k clusters
7 **Repeat until convergence of 'k' clusters:** {
8 **for** *all things* $\in T$ **do**
9 $\mathsf{cl_j} \leftarrow$ index (1 to k) of cluster centroid closest to $\mathsf{t}_i^?$ and $\{\mathsf{t}_j^{ex}\}$
10 **end for**
11 **for** *cl* $\leftarrow$ *1 to k* **do**
12 $\mu_{\mathsf{cl}} \leftarrow$ average of points assigned to cluster cl
13 **end for**
14 }
15 **return** $\{\mathsf{t}_j^{ex}\}'$ from $\mathsf{cl_j}^{\mathsf{t}_i^?}$;
16 **for** $t_i^?$*, each* $t_j^{ex'}$*, and each QoT property* **do**
17 compute $\rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})}$ by Equation (1)
18 **return** $\rho_{(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})}$;
19 Case1: **if** *atomic duty is performed by things* **then**
20 compute $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP$ for atomic duties by Equation (2)
21 **return** $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP$;
22 **end if**
23 Case2: **if** *common combined duties are performed by things* **then**
24 compute $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP$ for combined duties by Equation (3)
25 **return** $SIM(\mathsf{t}_i^?, \mathsf{t}_j^{ex'}) \mid QP$;
26 **end if**
27 **if** *tentative values of dependable QoT properties are announced* **then**
28 compute $Joint_{Sim}(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})$ by Equation (4)
29 **return** $Joint_{Sim}(\mathsf{t}_i^?, \mathsf{t}_j^{ex'})$;
30 **end if**
31 **for** $K \leftarrow 1$ **to** i **do**
32 *similar_thingsset*.add($\mathsf{t}_j^{ex''}$);
33 *K*++;
34 **end for**
35 **return** Top_*K(similar_thingsset)*; // $\{\mathsf{t}_j^{ex}\}''$
36 **end for**
37 train_X $\leftarrow$ ($\mathsf{QP}\{\mathsf{t}_j^{ex}\}''$);
38 train_Y $\leftarrow$ $\mathsf{t}_j^{ex''}\, \mathsf{v}_j^d(\mathsf{L_m})$;
39 model $\leftarrow$ decision tree regressor(train_X, train_Y)
40 $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m}) \leftarrow model.estimate(\mathsf{ins}(\mathsf{QP}[\mathsf{t}_i^?]))$;
41 **return** $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m})$;
42 **END**

Algorithm 1 describes tentative QoT values estimation. For thing selection and recommendation, the IRA runs Algorithm 1. It takes as input T, Pr, D, QP, TV, and L (line 1) and produces an estimated QoT value (line 2). It groups all the things on the basis of their locations by using K-means Clustering (lines 6 to 15). It assesses the similarity of things through lines 17 and 18. For atomic duties, it does assessment of similarity of things through lines 19 to 22. For combined duties, it does assessment of similarity of things through lines 23 to 26. It computes the joint similarity for dependable QoT properties (lines 27 to 30). Top-*K* similar things set i.e., $\{\mathsf{t}_j^{ex}\}''$ is returned by computation (lines 31 to 35). DTR is trained with inputs (lines 37, 38, and 39). Finally, the DTR-based model estimates $\mathsf{Es}^{\mathsf{t}_i^?}(\mathsf{L_m})$ (lines 40 and 41).

V. Experiments and Evaluation

For evaluation purposes, we carried out several experiments after developing 4 in-house Python programs on a Dell notebook with the following technical specification: Intel(R) Core(TM) i5-2540M CPU @ 2.60GHz, 4GB RAM with Windows 10 Enterprise. These programs test the thing selection and recommendation approach for evaluating thing QoT estimation. Specifically, we raised the following questions: *(i)* what is the impact of duties of things on similarity assessment and QoT estimation, *(ii)* does different number of Top-*K* similar things impact the accuracy of QoT estimation, and *(iii)* what is the accuracy of our approach compared to existing approaches.

A. Dataset

The experiments were carried out using 3 real datasets: sensors dataset [26], COMBED energy dataset [27], and WS-Dream dataset1 [28]. Specifically, we enriched the existing IoT sensors dataset with details from WSDream and COMBED in order to map things' location and energy consumption during things' duties onto QoT data. Through this mapping, we made a dataset suitable for thing selection and recommendation. The enriched sensors dataset consists of QoT data (i.e., response-time (RT) and throughput (TP) values) for low-power devices. We also used location information (Latitude-Longitude) from the WSDream and energy consumption information from COMBED against the QoT values of sensors dataset. As a result, we formed a new hybrid IoT dataset for RT Matrix and TP Matrix.

B. Impact of duties with respect to top-K on thing QoT estimation

We first choose 'k' for clustering of things. We plotted Within-Cluster the Sum of Squares (WCSS)/Inertia as a function of k (number of clusters) to determine the elbow point for the optimal k. Fig. 4 depicts the elbow point, that is k = 4.

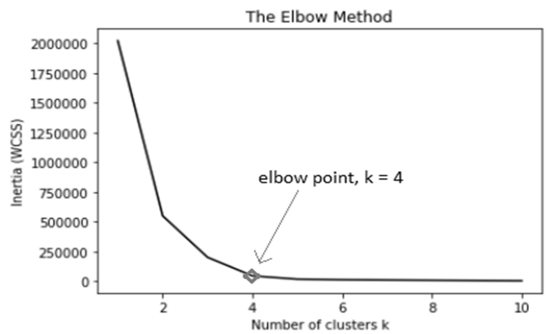

Fig. 4: Elbow plot for optimal k to be used

Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalized Mean Absolute Error (NMAE) are widely used to determine how accurate a recommendation is [13], [14], [15], [21]. A smaller value of MAE, RMSE, and NMAE represents an excellent accuracy. We also used these metrics for thing selection and recommendation approach to calculate deviation through error measurement. Based on these metrics, we can conclude the fairness of tentative QoT values. Individual differences have equal weight in MAE while extreme errors have large weight due to the squaring term in RMSE. NMAE is used because various things' QoT properties have different ranges of values, so we need to normalize MAE. MAE, RMSE, and NMAE are defined as follows (Equation 8, 9, and 10):

$$MAE = \frac{\sum_{t_i^? \in T} \mid t_i^? tv_i^d(L_m) \; - \; Es^{t_i^?}(L_m) \mid}{N} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{t_i^? \in T} (\, t_i^? tv_i^d(L_m) \; - \; Es^{t_i^?}(L_m)\,)^2}{N}} \tag{9}$$

$$NMAE = \frac{(N * MAE)}{\sum_{t_i^? \in T} t_i^? tv_i^d(L_m)} \tag{10}$$

where N is the total number of estimated values per QoT property and $\sum_{t_i^? \in T}$ denotes all $t_i^?$ for which QoT estimations are made.

To determine the impact of things' duties on QoT estimation, we conducted multiple experiments using different configurations of Top-K i.e., K = 5, 10, 15, 20, and 25, respectively. In an IoT context, the matrix density for QoS properties is expected to be quite low [20], so it is important to identify how our thing selection and recommendation approach performs with low matrix density. Initially, we selected a set of 200 different t_j^{ex} (existing smart cameras as stated in case study), i.e., $\{t_j^{ex}\} \in T$. Then, we selected one $t_i^?$ (new smart camera as stated in case study) with separate $t_i^? tv_i^d$ for RT and TP. We obtained clustered $\{t_j^{ex}\}'$ on the basis of things' locations. Then, we narrowed down $\{t_j^{ex}\}''$ corresponding to the value of Top-K. The more lower the NMAE, the more fair is $t_i^? tv_i^d$. The experiments targeted atomic and combined duties:

- Atomic duties: We considered different cases for atomic duties and evaluated them on different values of K. For RT estimation, when both $t_i^?$ and $t_j^{ex''}$ perform sensing, $t_i^? tv_i^d$ evaluated to be fair for K = 15 and 20, but over-estimated for K = 5 and 25, and under-estimated for K = 10. That is why error on K = 5, 10, and 25 is high for sensing. When both $t_i^?$ and $t_j^{ex''}$ perform actuating, $t_i^? tv_i^d$ evaluated to be fair for K = 10, but over-estimated for K = 5, and under-estimated for K = 15, 20, and 25. When both $t_i^?$ and $t_j^{ex''}$ perform communicating, $t_i^? tv_i^d$ evaluated to be fair for K = 10 and 20, but over-estimated for K = 5, 15, and 25 (Fig. 5).
For TP estimation, when both $t_i^?$ and $t_j^{ex''}$ perform sensing, $t_i^? tv_i^d$ evaluated to be fair for K = 10 and 15, but over-estimated for K = 20, and under-estimated for K = 5 and 25. When both $t_i^?$ and $t_j^{ex''}$ perform actuating, $t_i^? tv_i^d$ evaluated to be fair for K = 10, but over-estimated for K = 5, 15, 20, and 25. When both $t_i^?$ and $t_j^{ex''}$ perform communicating, $t_i^? tv_i^d$ evaluated to be fair for K = 10 and 25, but under-estimated for K = 5, 15, and 20 (Fig. 6).

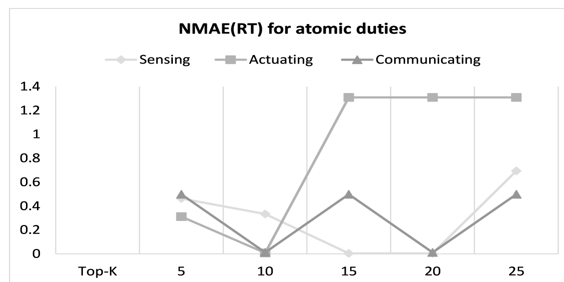

Fig. 5: Impact of atomic duties with respect to top-K on RT estimation of $t_i^?$

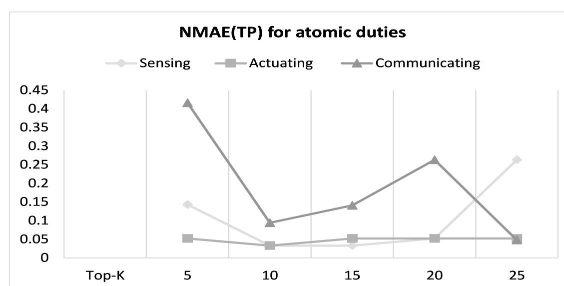

Fig. 6: Impact of atomic duties with respect to top-K on TP estimation of $t_i^?$

- Combined duties: We considered different cases for combined duties and evaluated them on different values of K. For RT estimation, when both $t_i^?$ and $t_j^{ex''}$ perform sensing and actuating, $t_i^? tv_i^d$ evaluated to be fair for K = 10 and 15, but over-estimated for K = 5 and 20, and under-estimated for K = 25. When both $t_i^?$ and $t_j^{ex''}$ perform sensing and communicating, $t_i^? tv_i^d$ evaluated to be fair for K = 5, 10, and 15, but over-estimated for K = 20, and under-estimated for K = 25. When both $t_i^?$ and $t_j^{ex''}$ perform actuating and communicating, $t_i^? tv_i^d$ evaluated to be fair for K = 15 and 20, and over-estimated for K = 5, 10, and 25. When both $t_i^?$ and $t_j^{ex''}$ perform sensing, actuating and communicating, $t_i^{?}tv_i^{d}$ evaluated to be fair for K = 15, and over-estimated for K = 5, 10, 20, and 25 (Fig. 7).

For TP estimation, when both $t_i^{?}$ and $t_j^{ex''}$ perform sensing and actuating, $t_i^{?}tv_i^{d}$ evaluated to be fair for K = 10, 15, 20, and 25 but under-estimated for K = 5. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing and communicating, $t_i^{?}tv_i^{d}$ evaluated to be fair for K = 10 and 25, but under-estimated for K = 5, 15, and 20. When both $t_i^{?}$ and $t_j^{ex''}$ perform actuating and communicating, $t_i^{?}tv_i^{d}$ evaluated to be fair for K = 15, but over-estimated for K = 20 and 25, and under-estimated for K = 5 and 10. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing, actuating and communicating, $t_i^{?}tv_i^{d}$ evaluated to be fair for K = 5, 15, 20, and 25, but over-estimated for K = 10 (Fig. 8).

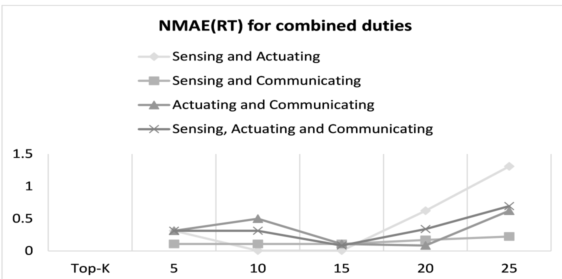

Fig. 7: Impact of combined duties with respect to top-K on RT estimation of $t_i^{?}$

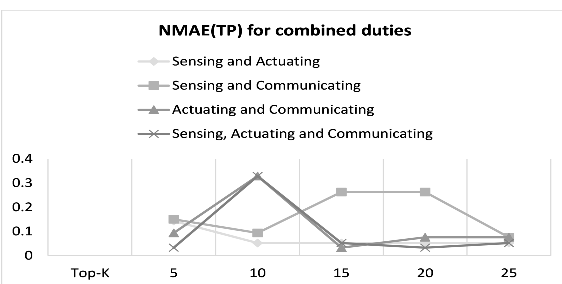

Fig. 8: Impact of combined duties with respect to top-K on TP estimation of $t_i^{?}$

C. Cost estimation for QoT properties

In addition to estimating individual QoT properties, we also estimate dependable QoT properties (dependable because cost is tentatively set against tentative values for RT and TP). For experimentation, we considered the case of $pr_i^{t_i^{?}}$ announcing a tentative cost value for tentative values of multiple QoT properties of $t_i^{?}$. We took L_m as cost, $t_i^{?}tv_1^{d}$ as tentative value for RT, $t_i^{?}tv_2^{d}$ as tentative value for TP, and $t_i^{?}tv_3^{d}$ as tentative value for cost. We evaluated that $t_i^{?}tv_3^{d}$ is fair or justified for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$ or not. The lowest NMAE is, the more fair $t_i^{?}tv_3^{d}$ is for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. We used K = 5, 10, 15, 20, 25, and 30 respectively, and conducted experiments for atomic and combined duties:

- Atomic duties: We considered different cases for atomic duties to estimate a thing's cost for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$ and evaluated them on different values of K. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 25, and 30, but under-estimated at K = 5, 10, 15, and 20, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. When both $t_i^{?}$ and $t_j^{ex''}$ perform actuating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 20, 25, and 30, but under-estimated at K = 5, 10, and 15, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. When both $t_i^{?}$ and $t_j^{ex''}$ perform communicating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 5, 10, 15, and 20, but over-estimated at K = 25, and 30, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$ (Fig. 9).

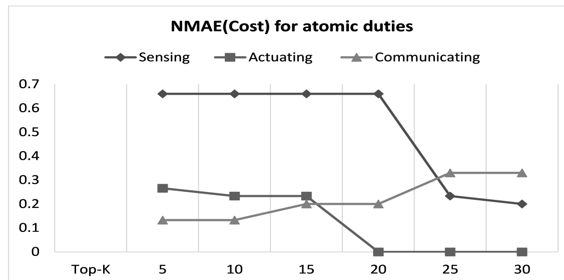

Fig. 9: Impact of atomic duties with respect to top-K on cost estimation of $t_i^{?}$

- Combined duties: We considered different cases for combined duties to estimate a thing's cost for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$ and evaluated them on different values of K. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing and actuating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 10, 15, and 20, but over-estimated at K = 25 and 30, and under-estimated at K = 5, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing and communicating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 5 and 10, but under-estimated at K = 15, 20, 25, and 30, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. When both $t_i^{?}$ and $t_j^{ex''}$ perform actuating and communicating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 15, 20, and 25, but over-estimated at K = 5, 10, and 30, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$. When both $t_i^{?}$ and $t_j^{ex''}$ perform sensing, actuating and communicating, $t_i^{?}tv_3^{d}$ evaluated to be fair at K = 5, 10, and 15, but under-estimated at K = 20, 25, and 30, for $t_i^{?}tv_1^{d}$ and $t_i^{?}tv_2^{d}$ (Fig. 10).

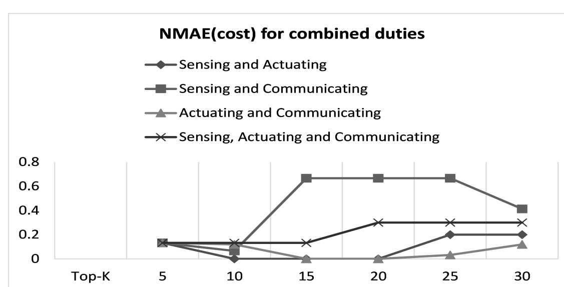

Fig. 10: Impact of combined duties with respect to top-K on cost estimation of $t_i^{?}$

Based on the above findings, we conclude that we need to narrow down the optimal number of similar things to achieve accurate QoT estimations. On the one hand, when K is very low, there are chances that more similar things are not selected for QoT estimation. On the other hand, when K is increased enormously, there are chances that more dissimilar things are selected for QoT estimation which affect the accuracy of their selection and recommendation. Moreover, each set of duty has different behavior for QoT estimation with respect to K. We have considered $K = 15$ as optimal parameter to further evaluate thing selection and recommendation approach. Fig. 11 and Fig. 12 show the behavior of cost estimation for $t_i^?$ at $K = 15$ for atomic and combined duties, respectively. In cost estimation for atomic duties, the deviation of $t_i^? tv_3^d$ is lowest for communicating and highest for sensing. In cost estimation for combined duties, the deviation of $t_i^? tv_3^d$ is lowest for sensing and actuating, actuating and communicating, and sensing, actuating and communicating, but highest for sensing and communicating. NMAE(cost) is zero for sensing and actuating as well as actuating and communicating. Table. I lists the estimations of RT, TP, and cost for $t_i^?$ at $K = 15$ for atomic and combined duties, respectively where s, a, and c are for sensing, actuating, and communicating, respectively.

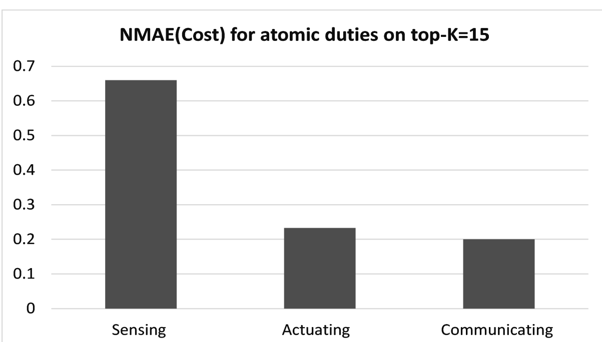

Fig. 11: Behavior of cost estimation of $t_i^?$ at $\boldsymbol{K}$ = **15** for atomic duties

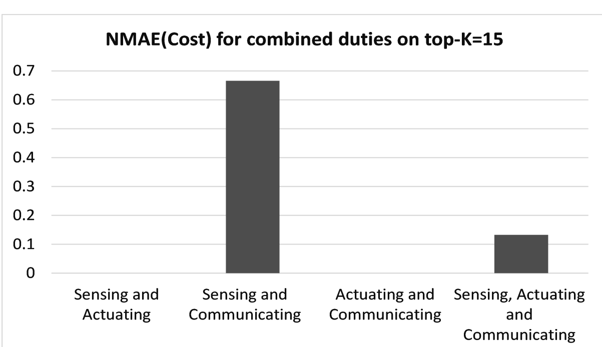

Fig. 12: Behavior of cost estimation of $t_i^?$ at $\boldsymbol{K}$ = **15** for combined duties

TABLE I
Estimations of RT, TP, and cost of $t_i^?$ at $\boldsymbol{K}$ = **15** for atomic and combined duties

Thing's duties	RT	TP	Cost
s	Fair	Fair	Under-estimated
a	Under-estimated	Over-estimated	Under-estimated
c	Over-estimated	Under-estimated	Fair
s a	Fair	Fair	Fair
s c	Fair	Under-estimated	Under-estimated
a c	Fair	Fair	Fair
s a c	Fair	Fair	Fair

D. Performance evaluation of thing selection and recommendation approach

To evaluate our approach, we further experimented on optimal cases and parameters of atomic and combined duties extracted from the above experiment (Subsection V-B). We considered RT, TP and cost estimation cases while setting the threshold of NMAE ≤ 0.1 to declare $t_i^? tv_i^d$ as fair for RT estimation, NMAE ≤ 0.05 to declare $t_i^? tv_i^d$ as fair for TP estimation, and NMAE ≤ 0.1 to declare $t_i^? tv_i^d$ as fair for cost estimation. But, defining thresholds is highly dependent upon the type of IoT application i.e., best effort (no QoS), differentiated services (soft QoS), and guaranteed services (hard QoS) [29]. We also performed the experiments for RT and TP estimations on existing approaches using the same IoT dataset and same parameters for Top-K.

Benchmarks for performance comparison: We compared the performance of thing selection and recommendation approach on the basis of accuracy metrics with other CF-based QoS prediction approaches. The benchmarked approaches are as follows:

1) IPCC: It is an item-based CF approach in which Pearson Correlation Coefficient (PCC) is used as similarity measure to employ similar services for the purpose of service recommendation [30];
2) IMEAN: Average QoS performance of a service evaluated by other users is used for service recommendation for an active user [31];
3) ILACF: It is an item-based and location-aware CF approach for service recommendation proposed by Liu et al. [13];
4) IoTPredict: It is an approach for collaborative QoS prediction IoT which is user-based as well as item-based, and proposed by White et al. [21]. It is based on matrix factorization.

- Atomic duties: We picked the optimal case when both $t_i^?$ and $t_j^{ex''}$ perform sensing and set $K = 15$. We considered 5 different $t_i^?$ i.e., $t_1^?$, $t_2^?$, $t_3^?$, $t_4^?$, $t_5^?$ with 5 different $t_i^? tv_i^d$ for RT, TP, and cost separately. In each iteration, we estimated $Es^{t_i^?}(L_m)$ for RT, TP, and cost for one $t_i^?$. In this way, we performed QoT estimations for 5 different $t_i^?$. Table. II exhibits the deviations of RT, TP, and cost estimations for each thing-to-evaluate when they performed sensing. These results depict that $t_2^? tv_i^d$ is fair for RT, TP, and cost estimations. So, $t_2^?$ is the "right" and "best" thing to select with respect to RT, TP, and cost estimations for dependable QoT properties. For RT estimation, $t_1^?tv_i^d$, $t_2^?tv_i^d$, and $t_4^?tv_i^d$ are fair. So, $t_1^?$, $t_2^?$, and $t_4^?$ are the "right" and "best" things to select with respect to RT estimation. For TP estimation, $t_1^?tv_i^d$, $t_2^?tv_i^d$, $t_4^?tv_i^d$, and $t_5^?tv_i^d$ are fair. So, $t_1^?$, $t_2^?$, $t_4^?$, and $t_5^?$ are the "right" and "best" things to select with respect to TP estimation. $t_3^?tv_i^d$ is slightly over-estimated for RT as well as TP estimation. $t_4^?tv_i^d$ is slightly over-estimated for TP estimation and $t_5^?tv_i^d$ is slightly under-estimated for RT estimation. Fig. 13 shows the behavior of RT estimation of each thing-to-evaluate at $K = 15$ for sensing and Fig. 14 shows the behavior of TP estimation of each thing-to-evaluate at $K = 15$ for sensing. Fig. 15 shows the behavior of cost estimation of each thing-to-evaluate at $K = 15$.

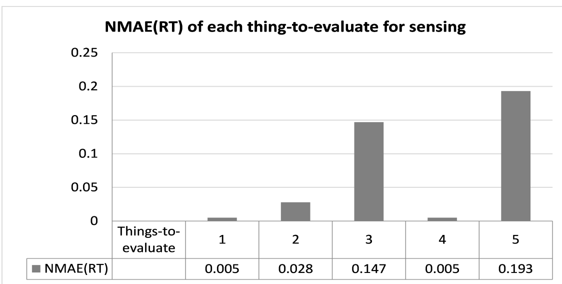

Fig. 13: Behavior of RT estimation of each thing-to-evaluate at $\boldsymbol{K} = \mathbf{15}$ for s

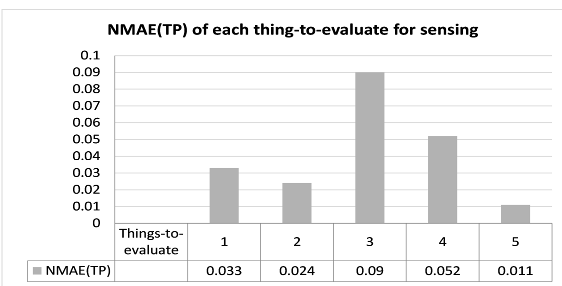

Fig. 14: Behavior of TP estimation of each thing-to-evaluate at $\boldsymbol{K} = \mathbf{15}$ for s

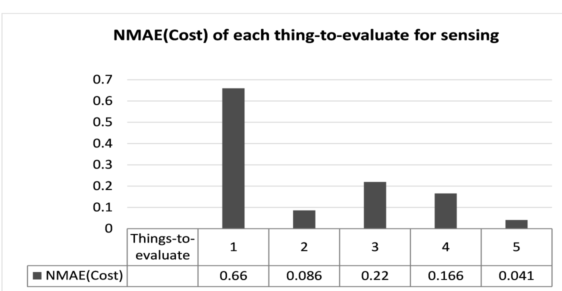

Fig. 15: Behavior of cost estimation of each thing-to-evaluate at $\boldsymbol{K} = \mathbf{15}$ for s

Moreover, we also experimented on our IoT dataset for $K = 15$ by using existing benchmark approaches. Table. III displays the results for MAE, RMSE, and NMAE of RT and TP estimations for atomic duties. These results depict that QoT estimation accuracy of thing selection and recommendation approach with respect to atomic duties is higher as compared to other benchmark approaches. Fig. 16 shows the performance comparison of RT estimation for atomic duties and Fig. 17 shows the performance comparison of TP estimation for atomic duties with different approaches.

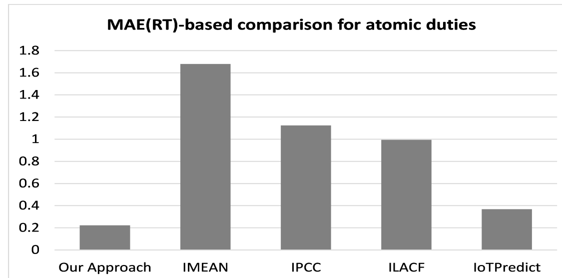

Fig. 16: Performance comparison of RT estimation for atomic duties

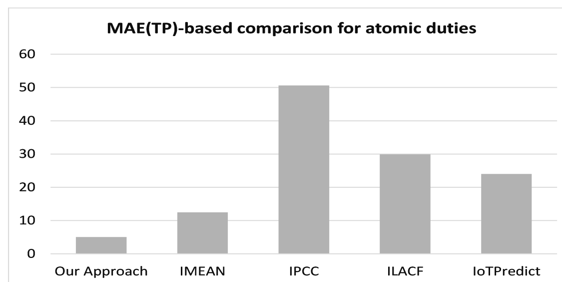

Fig. 17: Performance comparison of TP estimation for atomic duties

- Combined duties: We picked the optimal case when both $t_i^?$ and $t_j^{ex''}$ perform all the three duties i.e., sensing, actuating and communicating combinely, and set $K = 15$. We repeated the same process as we did for atomic duties for 5 different $t_i^?$ i.e., $t_1^?$, $t_2^?$, $t_3^?$, $t_4^?$, $t_5^?$ with 5 different $t_i^?tv_i^d$ for RT, TP, and cost separately. Table IV exhibits the deviations of RT, TP, and cost estimations for each thing-to-evaluate when they performed sensing, actuating, and communicating. These results depict that $t_1^?tv_i^d$ and $t_5^?tv_i^d$ are fair for RT, TP, and cost estimations. So, $t_1^?$ and $t_5^?$ are the "right" and "best" things to select with respect to RT, TP, and cost estimations for dependable QoT properties. For RT estimation, $t_1^?tv_i^d$, $t_2^?tv_i^d$, $t_4^?tv_i^d$, and $t_5^?tv_i^d$ are fair. So, $t_1^?$, $t_2^?$, $t_4^?$, and $t_5^?$ are the "right" and "best" things to select with respect to RT estimation. For TP estimation, $t_1^?tv_i^d$, $t_2^?tv_i^d$, and $t_5^?tv_i^d$ are fair. So, $t_1^?$, $t_2^?$, and $t_5^?$ are the "right" and "best" things to select with respect to TP estimation. $t_3^?tv_i^d$ is slightly over-estimated for RT as well as TP estimation. $t_4^?tv_i^d$ is slightly over-estimated for TP estimation. Fig. 18 shows the behavior of RT estimation of each thing-to-evaluate at $K = 15$ for combined duties and Fig. 19 shows the behavior of TP estimation of each thing-to-evaluate at

TABLE II
Individual performance of different things-to-evaluate for RT, TP, and cost estimations at $\boldsymbol{K}$ = **15** for atomic duties

Sr. No.	Things-to-Evaluate	Duty	RT	TP	Cost
1	$t_1^?$ provided by $pr_1^{t_1^?}$	s	Fair	Fair	Under-estimated
2	$t_2^?$ provided by $pr_2^{t_2^?}$	s	Fair	Fair	Fair
3	$t_3^?$ provided by $pr_3^{t_3^?}$	s	Over-estimated	Over-estimated	Under-estimated
4	$t_4^?$ provided by $pr_4^{t_4^?}$	s	Fair	Fair	Over-estimated
5	$t_5^?$ provided by $pr_5^{t_5^?}$	s	Under-estimated	Fair	Fair

TABLE III
Performance comparison of RT and TP estimations for atomic duties

Approaches	MAE(RT)	RMSE(RT)	NMAE(RT)	MAE(TP)	RMSE(TP)	NMAE(TP)
IMEAN	1.6791	1.9732	0.5645	12.5041	13.9752	0.1076
IPCC	1.1245	1.2851	0.3781	50.6448	51.0916	0.4358
ILACF	0.9942	1.1270	0.3343	29.8943	30.2893	0.2572
IoTPredict	0.3687	0.4232	0.6448	24.0191	24.0211	10.0389
Our Approach	**0.2227**	**0.3187**	**0.0749**	**5.0569**	**6.1591**	**0.0435**

K = 15 for combined duties. Fig. 20 shows the behavior of cost estimation of each thing-to-evaluate at K = 15 for combined duties.

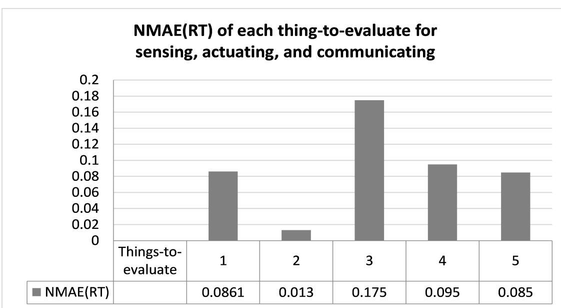

Fig. 18: Behavior of RT estimation of each thing-to-evaluate at $\boldsymbol{K}$ = **15** for s a c

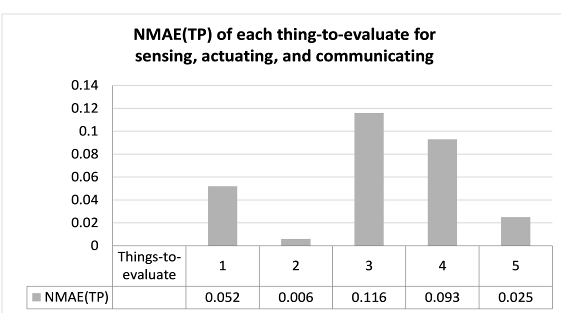

Fig. 19: Behavior of TP estimation of each thing-to-evaluate at $\boldsymbol{K}$ = **15** for s a c

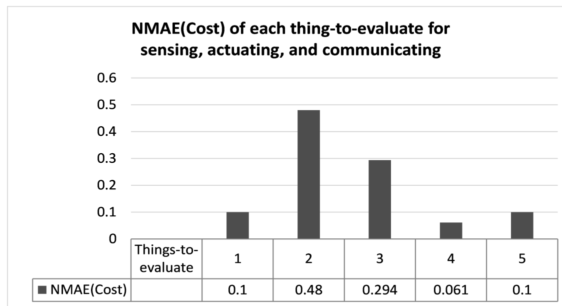

Fig. 20: Behavior of cost estimation of each thing-to-evaluate at $\boldsymbol{K}$ = **15** for s a c

We compared our results with existing approaches on the same parameters. Table. V displays the results for MAE, RMSE, and NMAE of RT and TP estimations for combined duties. These results depict that QoT estimation accuracy of thing selection and recommendation approach with respect to combined duties is higher as compared to other benchmark approaches. Fig. 21 shows the performance comparison of RT estimation for combined duties and Fig. 22 shows the performance comparison of TP estimation for combined duties with different approaches.

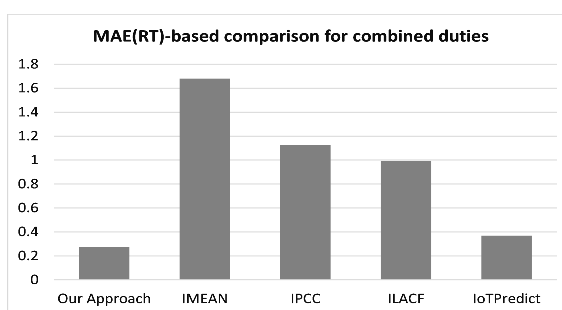

Fig. 21: Performance comparison of RT estimation for combined duties

TABLE IV
Individual performance of different things-to-evaluate for RT, TP, and cost estimations at K = **15** for combined duties

Sr. No.	Things-to-Evaluate	Duties	RT	TP	Cost
1	$t_1^?$ provided by $pr_1^{t_1^?}$	s a c	Fair	Fair	Fair
2	$t_2^?$ provided by $pr_2^{t_2^?}$	s a c	Fair	Fair	Over-estimated
3	$t_3^?$ provided by $pr_3^{t_3^?}$	s a c	Over-estimated	Over-estimated	Over-estimated
4	$t_4^?$ provided by $pr_4^{t_4^?}$	s a c	Fair	Over-estimated	Fair
5	$t_5^?$ provided by $pr_5^{t_5^?}$	s a c	Fair	Fair	Fair

TABLE V
Performance comparison of RT and TP estimations for combined duties

Approaches	MAE(RT)	RMSE(RT)	NMAE(RT)	MAE(TP)	RMSE(TP)	NMAE(TP)
IMEAN	1.6791	1.9732	0.5645	12.5041	13.9752	0.1076
IPCC	1.1245	1.2851	0.3781	50.6448	51.0916	0.4358
ILACF	0.9942	1.1270	0.3343	29.8943	30.2893	0.2572
IoTPredict	0.3687	0.4232	0.6448	24.0191	24.0211	10.0389
Our Approach	**0.2736**	**0.3244**	**0.0942**	**6.7289**	**8.2993**	**0.0606**

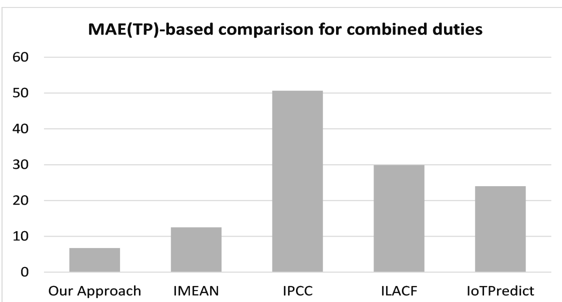

Fig. 22: Performance comparison of TP estimation for combined duties

VI. Conclusion

We presented a novel approach for thing selection and recommendation in the context of IoT. It first discusses how things' non-functional properties are captured into a QoT model based on the performance of things' duties in the form of sensing, actuating, and communicating. It, then, evaluates how fair the QoT properties' values of things are, so that fair competition between things' providers occurs. Different from other existing QoS prediction approaches, we address the problem of lack of data usage of things-to-evaluate. Our things selection and recommendation approach starts with clustering things on the basis of their locations. By identifying those existing things that fall into the same cluster as the "new" thing whose tentative QoT properties' values are provided for assessment, we do their similarity assessment. We enhanced Spearman's rank correlation coefficient (ρ) with things' duties information and energy as a resource. Top-K most similar existing things are narrowed down to perform thing QoT estimation using Decision Tree Regression technique. This enables us to evaluate if the tentative QoT values of the "new" thing are fair by setting acceptable ranges (neither under- or over-estimated according to specific thresholds). Deviation of QoT values is calculated through evaluation metrics and finally, the evaluation's feedback is sent back to the "new" thing's provider for further actions like adjustment, if necessary. In addition, we also discussed a specific case of cost estimation in which we evaluated that a tentative cost value against tentative QoT values of dependable QoT properties announced by a provider is justified fair or not. Comprehensive experiments are conducted using real datasets to evaluate the technical doability of the proposed approach. Our future work will include estimating the initial reputation and trust values of newcomer things and the impact they create on the assessment of things' similarity and QoT estimation.

References

[1] D. Andročec, B. Tomaš, and T. Kišasondi, "Interoperability and lightweight security for simple iot devices," in *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017 40th International Convention on*, pp. 1285–1291, IEEE, 2017.

[2] A. Martínez-Ballesté, P. A. Pérez-Martínez, and A. Solanas, "The pursuit of citizens' privacy: a privacy-aware smart city is possible," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 136–141, 2013.

[3] M. Eisa, M. Younas, and K. Basu, "Analysis and representation of qos attributes in cloud service selection," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 960–967, IEEE, 2018.

[4] J. Rittinghouse and J. Ransome, "Cloud computing: Implementation, management, and security," 2009.

[5] R. B. Uriarte, F. Tiezzi, and R. D. Nicola, "Slac: A formal service-level-agreement language for cloud computing," in *Proceedings of the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing*, pp. 419–426, IEEE Computer Society, 2014.

[6] C. Perera, C. H. Liu, S. Jayawardena, and M. Chen, "A survey on internet of things from industrial market perspective," *IEEE Access*, vol. 2, pp. 1660–1679, 2014.

[7] Z. Maamar, T. Baker, M. Sellami, M. Asim, E. Ugljanin, and N. Faci, "Cloud vs edge: Who serves the internet-of-things better?," *Internet Technology Letters*, vol. 1, no. 5, p. e66, 2018.

[8] S. Sundareswaran, A. Squicciarini and D. Lin, "A Brokerage-Based Approach for Cloud Service Selection," in *IEEE Fifth International Conference on Cloud Computing, Honolulu, HI*, pp. 558-565, 2012.

[9] V. Issarny, G. Bouloukakis, N. Georgantas, and B. Billet, "Revisiting service-oriented architecture for the iot: a middleware perspective," in *International Conference on Service-Oriented Computing*, pp. 3–17, Springer, 2016.

[10] A. M. Mzahm, M. S. Ahmad, and A. Y. Tang, "Agents of things (aot): An intelligent operational concept of the internet of things (iot)," in *Intelligent Systems Design and Applications (ISDA), 2013 13th International Conference on*, pp. 159–164, IEEE, 2013.

[11] I. Awan, M. Younas, and W. Naveed, "Modelling qos in iot applications," in *Network-Based Information Systems (NBiS), 2014 17th International Conference on*, pp. 99–105, IEEE, 2014.

[12] S. Wang, Y. Zhao, L. Huang, J. Xu, and C.-H. Hsu, "Qos prediction for service recommendations in mobile edge computing," *Journal of Parallel and Distributed Computing*, 2017.

[13] L. Jianxun, M. Tang, Z. Zheng, X. F. Liu, and S. Lyu, "Location-aware and personalized collaborative filtering for web service recommendation," *IEEE Transactions on Services Computing*, vol. 9, no. 5, pp. 686–699, 2016.

[14] X. Chen, Z. Zheng, Q. Yu, and M. R. Lyu, "Web service recommendation via exploiting location and qos information," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1913–1924, 2014.

[15] C. Yu and L. Huang, "A web service qos prediction approach based on time-and location-aware collaborative filtering," *Service Oriented Computing and Applications*, vol. 10, no. 2, pp. 135–149, 2016.

[16] Y. Hu, Q. Peng, and X. Hu, "A time-aware and data sparsity tolerant approach for web service recommendation," in *2014 IEEE International Conference on Web Services (ICWS)*, pp. 33–40, IEEE, 2014.

[17] S.-Y. Lin, C.-H. Lai, C.-H. Wu, and C.-C. Lo, "A trustworthy qos-based collaborative filtering approach for web service discovery," *Journal of Systems and Software*, vol. 93, pp. 217–228, 2014.

[18] W. Qiu, Z. Zheng, X. Wang, X. Yang, and M. R. Lyu, "Reputation-aware qos value prediction of web services," in *Services Computing (SCC), 2013 IEEE International Conference on*, pp. 41–48, IEEE, 2013.

[19] C. Wu, W. Qiu, Z. Zheng, X. Wang, and X. Yang, "Qos prediction of web services based on two-phase k-means clustering," in *2015 IEEE International Conference on Web Services (ICWS)*, pp. 161–168, IEEE, 2015.

[20] G. White, A. Palade, C. Cabrera, and S. Clarke, "Quantitative evaluation of qos prediction in iot," in *Dependable Systems and Networks Workshop (DSN-W), 2017 47th Annual IEEE/IFIP International Conference on*, pp. 61–66, IEEE, 2017.

[21] G. White, A. Palade, C. Cabrera, and S. Clarke, "Iotpredict: collaborative qos prediction in iot," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)(PerCom 2018), Athens, Greece*, 2018.

[22] D. A. Menascé, "Qos issues in web services," *IEEE internet computing*, vol. 6, no. 6, pp. 72–75, 2002.

[23] M. Tang, T. Zhang, J. Liu, and J. Chen, "Cloud service qos prediction via exploiting collaborative filtering and location-based data smoothing," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 18, pp. 5826–5839, 2015.

[24] X. Zheng, L. Da Xu, and S. Chai, "Qos recommendation in cloud services," *IEEE Access*, vol. 5, pp. 5171–5177, 2017.

[25] M. Xu, P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, "Decision tree regression for soft classification of remote sensing data," *Remote Sensing of Environment*, vol. 97, no. 3, pp. 322–336, 2005.

[26] G. White, A. Palade, and S. Clarke, "Forecasting qos attributes using lstm networks," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018.

[27] N. Batra, O. Parson, M. Berges, A. Singh, and A. Rogers, "A comparison of non-intrusive load monitoring methods for commercial and residential buildings," *arXiv:1408.6595*, 2014.

[28] Z. Zheng, Y. Zhang, and M. R. Lyu, "Investigating qos of real-world web services," *IEEE transactions on services computing*, vol. 7, no. 1, pp. 32–39, 2014.

[29] G. White, V. Nallur, and S. Clarke, "Quality of service approaches in iot: A systematic mapping," *Journal of Systems and Software*, vol. 132, pp. 186–203, 2017.

[30] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, pp. 285–295, ACM, 2001.

[31] Z. Zheng, H. Ma, M. R. Lyu, and I. King, "Qos-aware web service recommendation by collaborative filtering," *IEEE Transactions on services computing*, vol. 4, no. 2, pp. 140–152, 2011.

Ayesha Qamar is a master student in the Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad, Pakistan. She received her B.S. degree in Computer Science from the same university. Her research interests include Internet of Things, recommendation and selection, and machine learning.

Dr. Muhammad Asim is an Assistant Professor at the Department of Computer Science, National University of Computer and Emerging Sciences, Pakistan. Having attained a Ph.D. in 2010 from Liverpool John Moores University, he researches in the fields of Cloud Computing, Computer Networks, Network Security, Internet of Things and Wireless Sensor Networks. Muhammad serves as an editorial board member and potential reviewer for several scientific journals and conferences.

Dr. Zakaria Maamar is a Professor in the College of Technological Innovation at Zayed University, Dubai, United Arab Emirates. His research interests include Cloud/Fog Computing and Internet of Things. Zakaria has a Ph.D. in Computer Science from Laval University, Quebec City, Canada.

Saad Saeed is a researcher at the Department of Computer Science, National University of Computer and Emerging Sciences, Pakistan. His research interests include Internet of Things, Cyber Security and Machine Learning.

Dr. Thar Baker is a Senior Lecturer in Network Systems and Head of Applied Computing Research Group (ACRG) in Liverpool John Moores University, UK. Thar has published numerous referred research papers in multidisciplinary research areas including: Cloud Computing, Distributed Software Systems, Algorithm Design, and Network Protocols and Algorithms. He has been actively involved as a member of editorial board and review committees for a number of international journals.