Age of Information as a QoS Metric in a Relay-Based IoT Mobility Solution

Anderson Augusto Simiscuka  
School of Electronic Engineering  
Dublin City University  
anderson.simiscuka2@mail.dcu.ie

Gabriel-Miro Muntean  
School of Electronic Engineering  
Dublin City University  
gabriel.muntean@dcu.ie

Abstract—Internet of Things (IoT) networks handle multiple data types generated by numerous types of devices, with the additional challenge of delivering this data at high QoS levels. A great number of IoT applications require real-time updates and fresh data, in areas such as health, vehicular, UAVs and sensor monitoring (temperature, acceleration, motion, etc.), therefore, a metric such as the Age of Information (AoI) is useful to measure how recent the information is, considering the difference between the time it was generated and the time it is successfully delivered. In this paper, we propose an architecture consisting of IoT objects, which provide diverse services including video applications and CCTV monitoring, smart gateways, which support network connectivity, and a cloud-deployed platform utilised for resource management. The performance, quality and mobility challenges of IoT services are improved through an algorithm which uses Quality of Service (QoS), AoI, objects location and service relevance metrics in order to cluster the networked IoT objects, attach them to the most suitable smart gateway or relay device and improve the performance of their services.

Keywords—Internet of Things (IoT), Age of Information (AoI), smart gateways, QoS, performance, mobility.

I. INTRODUCTION

The Internet of Things (IoT) is expected to comprise 500 billion devices by 2030 [1]. An IEEE report [2] identified areas of research focusing on the challenges brought by the new innovation paradigms of IoT, especially regarding mobility, which is crucial for devices such as robots, drones, cars, buses or trains, wearable devices (such as smart watches and wrist bands) and smartphones. These research areas include performance, network design with real-time coordination, concurrency, mobility patterns, data storage, etc. Quality of Service (QoS) attributes such as packet loss, delay, throughput, energy consumption, along with others, have been used for performance measurements and decision-making in IoT [3]–[5], resulting in adjustments in quality of offered services, routing and clustering techniques and connection control mechanisms [6]–[14].

The real-time requirements of many IoT services, however, demand an additional metric that can be used to measure the freshness of the information being delivered to the users, the Age of Information (AoI). AoI is the time difference between the generation and successful delivery of information to a destination [15]. AoI has already been implemented in Wireless Sensor Networks (WSN) [16], mobile devices [17] and vehicular networks (VANET) [18].

In this paper, we propose the use of AoI along with other QoS performance and mobility metrics, to fully support IoT requirements. Our Internet of Things Multimedia-Enabled Mobility Solution (ITMEMS), as seen in figure 1, considers mobile objects, AoI, performance, and objects’ distances to WiFi access points and LTE base stations. The main goal is to provide the best connectivity to objects by assigning them to the optimal IoT smart gateway in the IoT network or using certain devices as relay nodes. ITMEMS’ architecture contains IoT devices, smart gateways and a cloud-based IoT Integration Platform (ITINP). IoT devices provide and use services from other devices, such as multimedia streams and sensorial data, the smart gateways provide network connectivity to the IoT objects and ITINP is responsible for the management of the smart gateways. Simulation-based testing intends to show how ITMEMS is able to provide improvements in connectivity to devices, increasing their performance and QoS. We test a scenario where a smart device is connected to a gateway in a WLAN (fig. 1 ⊙) and moves towards another WLAN (fig. 1 ⊙). However, while moving, the device is out of the reach of the available WLANs and is connected by a relay node with LTE access (fig. 1 ⊙).

This paper is organised as follows. In section II related works are presented and section III introduces the architecture for the proposed ITMEMS solution. Section IV details ITMEMS algorithm, and section V presents simulations-based testing and results. Conclusions and future work directions finalise this paper.

II. RELATED WORKS

The proposed solution is built upon three areas of research with a number of state-of-the-art works being...
evaluated in each area: a) **QoS evaluation and clustering in IoT**, b) **AoI in IoT** and c) **Mobility in IoT**.

### A. QoS Evaluation and Clustering in IoT

IoT devices are usually clustered in order to provide meaningful information to users and to improve QoS of the network by decreasing the amount of messages exchanged.

An access WSN gateway was proposed in [6] to provide end-to-end (E2E) connectivity for Machine Type Communication (MTC) traffic flow within the license spectrum of LTE-APro media. The solution addressed heterogeneity of applications, services and terminal devices and the related QoS issues among them. Simulation results (in services such as voice, video conferencing, FTP and HTTP) showed significant improvements in QoS performance, regarding Packet Loss Ratio delay, and Packet Delay Variation (PDV), while retaining good Quality of Experience (QoE). In Voice over LTE, however, some degradation was encountered in Network QoS.

In [19], the authors described an approach to build low-rate overlay networks in LoRaWAN, ultimately finding correlation between certain QoS metrics (Received Signal Strength Indicator (RSSI), packet delay, Signal-to-Noise Ratio (SNR) and jitter) and the distance between a node and a gateway. The authors have yet to decrease packet delay times, required by certain type of overlay applications.

Quality scores are useful to simplify the complexity and diversity of applications QoSs, such as the scoring system proposed in [20] along with a mapping mechanism for multiple QoS attributes in order to find the optimal deployment decision with the highest overall application level quality score. The problem was modelled as a maximum weighted bipartite problem, and solved by the use of integer linear programming (ILP). Simulations showed that the ILP algorithm outperforms the greedy matching algorithm in terms of QoS matching for applications. ILP, however, takes more time to compute results, which may not be suitable for certain IoT applications.

An approach introduced in [21] aimed to enhance energy efficiency while ensuring a desired QoS threshold by employing a game theoretical solution concept: the satisfaction equilibrium. Authors reach an equilibrium in slow and fast fading channels by using distributed schemes.

The authors in [22] proposed an algorithm for energy-centred and QoS-aware services selection. Services were preselected offering the QoS level required to satisfy the user's requirements in terms of cost, response time, reputation, reliability, and availability, then, the relative dominance relation is used to select among the preselected services the best candidate services for the composition process. The algorithm has yet to be able to handle the dynamic re-selection of services by predicting QoS and energy parameters changes, with respect to the different composition operators used in a given plan.

A multi-objective optimisation problem of the tradeoff between QoS provisioning and energy efficiency was addressed with a non-dominated sorting genetic algorithm in [23], with the authors also introducing a method to select optimum cooperative coalition for cluster heads and cooperative nodes, by using exhaustive search combined with quantum-inspired particle swarm optimisation.

In [24], authors investigated the delivery of uniform connectivity and service experience to converged multi-radio heterogeneous deployments, where distributed unlicensed-band networks (e.g., WiFi) may take advantage of centralised control functions in cellular networks (e.g., LTE). Results showed that the scheme could improve the performance of a WiFi-preferred scheme based on a minimum SNR threshold. Further work includes the investigation of user equipment-based algorithms with load variation in the network.

The Equalised Cluster Head Election Routing Protocol (ECHERP) presented in [25], focused on energy conservation through balanced clustering. Network is modelled as a linear system and with a Gaussian elimination algorithm, combinations of nodes that can be cluster heads are calculated to extend network lifetime. Performance evaluation evidenced the effectiveness of the protocol in terms of network energy efficiency when compared against other protocols. ECHERP has yet to be enhanced by aggregating other QoS metrics and time constraints.

Authors in [26] proposed an adaptive method for determining the number of clusters that can be found in a stream based on the data distribution. An online clustering mechanism was used to cluster the incoming data from the streams. A set of experiments in a traffic use-case scenario, compared against a non-adaptive stream cluster algorithm, showed significant improvements in the cluster quality metric of silhouette coefficient.

### B. Age of Information in the Internet of Things

The AoI has been studied in several works in areas closely related to IoT, such as mobile networks, ad hoc networks and WSNs.

Authors in [16] proposed a solution with framing and scheduling techniques for Cognitive Networks to optimise energy efficiency of communication systems under strict constraints on the expected AoI. Packet lengths are regularised based on sensing and communication parameters, and channel quality factors. Investigations showed that in light traffic, shorter packets are preferred to reduce waiting time for packet formation, while in heavy traffic, longer packets experience higher discard rates, increasing the service and waiting time. Packet transmission times, then, should be much smaller than the mean available interval to avoid interruptions.

A scheduling rule that can optimise the tradeoff between the AoI and energy efficiency was identified in [17]. The tradeoff was determined by the weight coefficient in the cost function. Results found helped in the identification of suitable operation parameters for the stopping rule.
An analytical model was proposed in [18] in order to evaluate the Age of Information of a VANET. Results show that the model is able to capture and approximate how the average AoI changes with respect to the beacon sending frequency, which is related to vehicle density. Higher message sending periods are beneficial for vehicles moving in sparser zones, but degrading AoI in crowded areas. Limitations in the model included inaccuracy as the mean vehicle density grows.

Investigations of the problem of age of information from a network perspective were performed in [27] and [28]. Authors proposed a new technique of optimising the scheduling in the network in order to enhance the freshness of information and also developed an algorithm offers better scalability than ILP. The study has yet to be extended to other scenarios, multi-hop networks and a setting that incorporates the freshness of information, the quality of service, and the fairness among the links.

In [29], authors studied optimal ways of managing the freshness of information in communications, using AoI. An age penalty function characterised the level of dissatisfaction on data staleness and algorithms find the optimal update policy of the zero-wait policy. Zero-wait policies are not optimal if the age penalty function grows quickly with respect to the age; the packet transmission times over the channel are positively correlated over time; and the packet transmission times are highly random. The jointly optimal control of signal sampling at the source node and remote estimation at the destination was not investigated.

An experimental evaluation of AoI was presented in [30] verifying how well the queue models approximate a network for the purpose of age. Results showed that the D/D/1 queue model models the average age better than the D/M/1, specifically in low loss cases with distribution of latencies being smooth and similar. Mixing both models can provide a good prediction when latencies are less smooth.

A model for mobile caching that considers popularity and freshness of the information was proposed in [31]. Authors proved that by caching packets that have the highest request values, the number of missed requests are minimised. The evolution of the request rate was analysed in two versions of the proposed model, providing sufficient conditions for the settling time of request rate. A mechanism in which arriving packets are updated to cached content has yet to be studied.

C. Mobility in the Internet of Things

A number of solutions have been proposed in the field of mobility for IoT. A study on an architecture with routing and mobility features for IoT devices was presented in [32]. The scalable routing architecture used Bloom Filters, a space-efficient randomised data structure for membership queries with false positive errors. The solution for IoT applications was implemented in an environment wherein 500,000 IoT devices move in a square field independently.

Authors introduced an aggregation scheme that make the size of routing information small, however, the scheme sacrifices accuracy and the solution did not suppress overheads to update routing information.

In [33], the authors presented a protocol for inter-WSN mobility. This protocol among other tasks, carry out the moving signalling, decreasing the number of interchanged messages among the mobile IoT nodes at a hospital environment.

MOSDEN, Mobile Sensor Data Processing Engine, was introduced in [34]. MOSDEN collects data from several different sensors and does a unified data processing. It is also compatible with the cloud-based Global Sensor Network Middleware. The authors have yet to develop a model suitable for different types of devices. 

III. ITMEMS ARCHITECTURE

ITMEMS is deployed on an architecture consisting of clusters of IoT devices, smart gateways, and the cloud-based IoT Integration Platform (ITINP), as presented in figure 2. These components are described next as well as features such as Reputation and Utility Functions and a Re-Clustering Algorithm.

A. ITMEMS Main Components

IoT objects (e.g. wireless devices, sensor based devices, appliances) provide one or more services to other objects. If an IoT object is more relevant to another object cluster in a distant area (i.e. is associated with high number of requests from that area), it will be attached to a smart gateway serving the new area using long range wireless connectivity (e.g. LTE). Objects can also change from LTE to WiFi and vice versa according to their position and the availability, movement patterns, performance and position of the access points or base stations. Objects attached to a smart gateway (SG) or relay node (RN) are identified via an ID (Oi) and provide one or more services Sijkl where i identifies the gateway or relay node, j indicates the object and k, the service. Objects are initially attached to the closest smart gateway, in a physical cluster, defined by distance.
However, certain objects’ offered services are accessed more often by objects from other clusters, and in this case, the current gateway is only being used as a bridge for communications. This scenario, which negatively influences the performance of communications, can be prevented by reassigning each object regularly to the most relevant smart gateway. If the associated smart gateway is inaccessible by WiFi, for example, a long range wireless solution such as LTE will be used for object-smart gateway communications, forming a *logical cluster*. When objects are moving, an algorithm will find the best balance between location and performance, and consider relay nodes (e.g. a smartphone or other nodes with access to the internet) for networking these moving devices.

A *smart gateway* provides connectivity to IoT objects, manages admission control, aggregates data and stores performance-related information in form of reputation scores. Objects communicate to each other via the smart gateway they are currently attached to. When an object tries to communicate with another object in a different smart gateway, ITINP will be used for the communication, as it is accessible from all smart gateways. Smart gateways also are aware of objects’ locations and uses it for admission control and re-clustering of objects. In figure 3, the device $O_2$ moves into an area with signal from 2 access points and 1 base station. The utility function that will decide which gateway will reply to the connection request first is calculated by ITINP. Device $O_4$, for example, is closer to BS$_i$, but could be connected to AP$_2$ depending on its reputation score.

ITINP is a cloud-based server platform that creates and manages optimal clusters of objects and assigning them to the best available smart gateway. This is done by an algorithm that considers objects’ reputations, calculated by an utility function that computes QoS, AoI, location and relevance scores (all provided by the gateways). ITINP receives smart gateway locations, therefore, when an object is unable to access the most suitable gateway using a short range wireless connection such as WiFi, the object will either 1) be requested to switch to a long range wireless technology to connect to the selected gateway, or 2) be requested to use an available relay node that is able to connect itself and other objects to ITINP.

### B. Reputation Score and Utility Functions

The reputation score is a combination of the utility functions of ITMEMS. The score is used for admission and performance control (i.e. re-clustering and re-assigning objects to other Smart Gateways/Relay Nodes) and identification of objects with low performance.

The QoS score ($P_i$), given in [3] in eq. (1), associated with an object is obtained by normalising the sum of the normalised values of each metric $M_x$ measured per service provided by the object. Each gateway keeps this score for objects that have been connected to it. The QoS score of a smart gateway ($P$) is also obtained by using this formula.

$$P_{jk} = \sum \left( \frac{M_x}{\max(M_x)} \right)$$

The service scores are, then, averaged in eq. (2) to compose the QoS score of an object.

$$P_{ij} = \text{avg}(P_{ijk})$$

The Relevance score ($R_{ij}$) of objects [3] is measured in an array of tuples composed by each gateway that has communicated with an object and the percentage of total packet communication exchanged between the gateway and the object. For example, an IoT object that has exchanged messages with two different smart gateways, 200 packets were sent and received from devices in the same gateway ($G_1$) and 400 packets from objects associated with a different gateway ($G_2$). Therefore: $R_{ij} = \{[[G_1, 33.3\%], [G_2, 66.6\%]]\}$.

The average Age of Information ($A_{ij}$) per object is calculated by a function for the M/M/1 queue model described in eq. (3) and [15]. The arrival rate (the generation and submission of packets) is determined by a Poisson process $\lambda$ and average service time is given by $1/\mu$. The utilization $\rho = \lambda/\mu$ varies according to each service provided by objects.

$$A_{ij} = \frac{1}{\mu} \left(1 + \frac{1}{\rho} + \frac{\rho^2}{2(1-\rho)}\right)$$

The Location score ($L_{ij}$) of an object is determined by its RSSI or Reference Signal Received Power (RSRP) in relation to the reachable gateways. The utility function – eq. (4) – of ITMEMS calculates the reputation of an object in relation to a gateway trying to find a balance between performance, relevance, AoI and location of devices, resulting on the normalised $U_{ij}$ score per object. The sum of the weights $= 1$, and gateways might prioritise one score over another.

$$U_{ij} = \text{wp} \cdot \frac{P_{ij}}{\max(P_j)} + \text{wr} \cdot \frac{R_{ij}}{\max(R_j)} + \text{wa} \cdot \frac{A_{ij}}{\max(A_j)} + \text{wl} \cdot \frac{|L_{ij}|}{\max(|L_i|)}$$

The $U_{ij}$ score is used in the re-clustering of low-performing objects (see Algorithm 1). For objects that are moving, the $L_{ij}$ score is used (see Algorithm 2).
C. ITMEMS Algorithms

When objects move, a handover algorithm adapted from [35] is used (see Algorithm 1). It reduces objects involvement, with gateways handling movement detection and attachment prediction. Gateways have a threshold of how low the average RSSI or RSRP \( L_0 \) of objects can be (default: -90 dBm), and the algorithm is triggered once the threshold is reached, in order to find another gateway or relay node.

ITMEMS implements a simple re-clustering algorithm (see Algorithm 2) in the cloud-based ITINP. If an object is not in association with the closest gateway (strongest signal in \( L_0 \)), then it will be reattached. The relevance, AoI and performance of objects in relation to the gateways (computed in \( U_q \)) are also considered for re-clustering.

IV. PERFORMANCE ANALYSIS

The ITMEMS solution was modelled and evaluated using Network Simulator 3 (NS-3) [36]. Table I presents the simulation setup. The scenario considered contains a WiFi-only smart watch in a smart house with several other devices connected to a gateway using a WiFi 802.11ac access point. Two gateways, one connected at a WiFi access point and one at an LTE base station, contain devices with bitrates of 20, 10 and 4 Mbps (one device at each speed) and seven devices at 1 Mbps or lower, representing diverse types of mobile and non-mobile IoT devices. The smart watch, which has a bit rate of 2 Mbps, moves towards another WLAN. However, while moving, the device is out of the reach of the available WLANs and is connected by a relay node with LTE access. ITMEMS regularly checks performance and RSSI of devices, and proposes moving devices affecting performance to other gateways connected to ITINP to improve quality of data delivery for all devices.

Results show that after the execution of ITMEMS, QoS metrics have improvements in terms of throughput, packet loss and delay, the QoS measured in the study (fig. 4). The smart watch and the devices in the LTE base station benefited from ITMEMS while the devices in the WiFi access point did not run the algorithms. The smart watch moved from one gateway to the other after around 6.5 seconds. Especially on the downlink, main improvements are a decrease of 11% in average packet loss, almost eliminating it and average delay decreased from 133ms to 5ms. The other devices in the first WLAN also had improvements in terms of downlink performance. Average delay decreased from 132ms to 89ms, an improvement of 32.5%, average packet loss dropped to 7%, and throughput increased in most devices. Regarding the uplink of the devices, the smart watch throughput had a small increase of \( \sim 2\% \), while average packet loss had a slight increase of 0.2% and the average uplink delay was reduced to 3ms. The other devices were not impacted much in terms of their average packet loss, delay and throughput, which remained roughly the same. The devices associated with bitrates of 20 Mbps, 10 Mbps and 4 Mbps do not have any uplink traffic.

V. CONCLUSIONS AND FUTURE DIRECTIONS

This paper introduced ITMEMS, its algorithms and scores. Our assessment of the solution through NS-3 demonstrate improvements in terms of throughput, packet loss and delay.

Future work includes an optimisation of the relevance score by prioritising different types of services and content.

ACKNOWLEDGMENT

This work was supported by the Irish Research Council grant number EPSPG/2015/178 and in part by the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement no. 688503 for the NEWTON project (http://newtonproject.eu).

---

Algorithm 1 – Mobility Support for Objects

<table>
<thead>
<tr>
<th>Line</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>if ( L_0 \leq \text{threshold} )</td>
</tr>
<tr>
<td>2</td>
<td>( SG ), broadcasts message with Object ID ( O_i ) and ( L_0 )</td>
</tr>
<tr>
<td>3</td>
<td>( SG ), starts timer</td>
</tr>
<tr>
<td>4</td>
<td>( SG ), waits other ( SGs ) and ( RNs ) (that can detect ( O_i )) reply</td>
</tr>
</tbody>
</table>
| 5    | \( O_i \), receives message broadcasted by \( SG \).
| 6    | \( O_i \), broadcasts message to \( SGs \) and \( RNs \) |
| 7    | \( SGs \) and \( RNs \) compute \( L_0 \) received from \( O_i \)
| 8    | \( SGs \) and \( RNs \) send message with \( L_0 \) and \( U_q \) to \( SG \),
| 9    | \( SG \), stops timer |
| 10   | \( SG \), selects best \( L_0 \) |
| 11   | if \( SG \), has best \( L_0 \) |
| 12   | end |
| 13   | else |
| 14   | \( SG \), notifies \( SG/RN \) with best \( L_0 \)
| 15   | \( SG/RN \), attach \( O_i \)
| 16   | \( SG \), detache \( O_i \)

Algorithm 2 – Re-Clustering Low-Performing IoT Objects

| Require: \( P_{o_i}, R_{o_i}, A_s \) and \( L_0 \) per \( O_i \) in \( SG \) |

<table>
<thead>
<tr>
<th>Line</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>for each ( O_i )</td>
</tr>
<tr>
<td>2</td>
<td>ITINP computes ( U_q )</td>
</tr>
<tr>
<td>3</td>
<td>if ( SG ), is not max ( L_0 )</td>
</tr>
<tr>
<td>4</td>
<td>or if ( SG ), is not most relevant ( SG )</td>
</tr>
<tr>
<td>5</td>
<td>and most relevant ( SG, P_i )</td>
</tr>
<tr>
<td>6</td>
<td>Most Relevant ( SG ), attach ( O_i )</td>
</tr>
<tr>
<td>7</td>
<td>( SG ), detach ( O_i )</td>
</tr>
<tr>
<td>8</td>
<td>if ( O_i ), has low ( L_0 )</td>
</tr>
<tr>
<td>9</td>
<td>and most relevant ( SG = SG_i )</td>
</tr>
<tr>
<td>10</td>
<td>Next most relevant ( SG ), attach ( O_i )</td>
</tr>
<tr>
<td>11</td>
<td>( SG ), detach ( O_i )</td>
</tr>
</tbody>
</table>

---

Table I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulator</td>
<td>NS-3.24.1</td>
</tr>
<tr>
<td>Duration of the Simulation</td>
<td>14+10s before and after sim.</td>
</tr>
<tr>
<td>Initial dist. betw. nodes and antennas</td>
<td>3m</td>
</tr>
<tr>
<td>WiFi Data Rate</td>
<td>40 Mbps</td>
</tr>
<tr>
<td>WiFi Standard</td>
<td>802.11ac (40MHz, MCS 9)</td>
</tr>
<tr>
<td>LTE eNB Antenna Model Type</td>
<td>Isotropic Antenna Model</td>
</tr>
<tr>
<td>LTE Data Rate</td>
<td>100 Mbps</td>
</tr>
<tr>
<td>Remote Station Manager</td>
<td>ConstantRateWifiManager</td>
</tr>
<tr>
<td>Mobility Model</td>
<td>ConstantVelocityMobilityModel</td>
</tr>
<tr>
<td>Speed of Smart Watch user</td>
<td>2 metres per second</td>
</tr>
<tr>
<td>Smart Watch Bit Rate</td>
<td>2 Mbps</td>
</tr>
<tr>
<td>Other Devices Bit Rate</td>
<td>20, 10, 4 Mbps and &lt;=1 Mbps</td>
</tr>
</tbody>
</table>
REFERENCES


