An Intelligent Data Uploading Selection Mechanism for Offloading Uplink Traffic of Cellular Networks

Qian Wang [1], Juan Fang [1], Bei Gong [1,*], Xiaojiang Du [2] and Mohsen Guizani [3]

[1] College of Computer Science, Beijing University of Technology, Beijing 100124, China; wangqian2020@bjut.edu.cn (Q.W.); fangjuan@bjut.edu.cn (J.F.)

[2] Department of Computer and Information Science, Temple University, Philadelphia, PA 19122, USA; dxj@ieee.org

[3] Department of Computer Science and Engineering, Qatar University, Doha 2713, Qatar; mguizani@ieee.org

* Correspondence: gongbei@bjut.edu.cn

Abstract: Wi-Fi uploading is considered an effective method for offloading the traffic of cellular networks generated by the data uploading process of mobile crowd sensing applications. However, previously proposed Wi-Fi uploading schemes mainly focus on optimizing one performance objective: the offloaded cellular traffic or the reduced uploading cost. In this paper, we propose an Intelligent Data Uploading Selection Mechanism (IDUSM) to realize a trade-off between the offloaded traffic of cellular networks and participants' uploading cost considering the differences among participants' data plans and direct and indirect opportunistic transmissions. The mechanism first helps the source participant choose an appropriate data uploading manner based on the proposed probability prediction model, and then optimizes its performance objective for the chosen data uploading manner. In IDUSM, our proposed probability prediction model precisely predicts a participant's mobility from spatial and temporal aspects, and we decrease data redundancy produced in the Wi-Fi offloading process to reduce waste of participants' limited resources (e.g., storage, battery). Simulation results show that the offloading efficiency of our proposed IDUSM is (56.54×10^{-7}), and the value is the highest among the other three Wi-Fi offloading mechanisms. Meanwhile, the offloading ratio and uploading cost of IDUSM are respectively 52.1% and (6.79×10^3). Compared with other three Wi-Fi offloading mechanisms, it realized a trade-off between the offloading ratio and the uploading cost.

Keywords: uplink traffic offloading; opportunistic communications; mobility prediction; mobile crowd sensing applications

1. Introduction

With the proliferation of smart devices with various sensors (e.g., smart phones, iPads, intelligent vehicles etc.), Mobile Crowd Sensing (MCS) has become an appealing paradigm by empowering smart devices as participants to contribute data sensed or generated from them, and then aggregating and fusing these data in the cloud platform for crowd intelligence extraction and human-centric service delivery [1]. It is gradually being used in more and more large-scale sensing applications such as air condition monitoring, health monitoring, and so forth. It can also support Intelligent Transportation System (ITS) to provide innovation services, improve cost-effectiveness and efficiency of transportation and traffic management systems [2,3], for example, a bus arrival time prediction is realized relying on the participatory sensing of bus passengers [4], and a bus route optimization planning is designed based on the sensing data of participants [5].

In MCS, after participants finished assigned tasks the data should be uploaded to the backend server located in the cloud for further processing. Nowadays most existing work assumes that

the data is uploaded by cellular networks as soon as the data is produced. However, according to Cisco forecast, cellular networks are very likely to become overloaded due to the dramatically increased traffic [6], even though it upgrades to the 5th Generation (5G) mobile networks with higher capacity. 5G networks are expected to handle a 1000-fold increase in capacity [7,8]. However, in the long term the emerging Internet of Things (IoT), massive Machine Type Communication (mMTC), always-connected devices such as smartphones, tablets, video-game consoles, virtual/augmented reality devices, wearable electronics introducing a huge number of connections may still cause the network to be congested [9–12].

With the increase of the number of MCS applications and participants, if all data is uploaded through cellular networks, the overload problem of cellular networks may be worsened. Therefore, there is an urgent demand to offload the traffic of cellular networks generated by the data uploading process of MCS. One promising solution for this issue is Wi-Fi uploading that source participants transmit collected data to Wi-Fi Access Points (APs) by direct or indirect short-range radios (e.g., bluetooth, Wi-Fi) with the help of encountered participants, and then the data is uploaded to the cloud through wired links of Wi-Fi APs. Through the method, part traffic of cellular networks is offloaded to Wi-Fi networks. However, the Wi-Fi uploading method might lead to uncontrollable uploading delay, since short-range radios between participants or between participants and APs are opportunistic, and the data transmission method is called opportunistic transmissions. For the data with time constraint, such as traffic condition and air quality, if the data is not uploaded to the server before it expires, the source participant will reupload the data through cellular networks to ensure the quality of the application. Hence, if the collected data fails to be uploaded through Wi-Fi uploading, not only the traffic of cellular networks will not be reduced, but also limited resources (storage, battery, etc.) of participants participating in Wi-Fi uploading process will be wasted. In addition, it may produce additional transmission cost because of participants' data plans.

Pay As You Go (PAYG) and Unlimited Data Plan (UnDP) are two data plans widely used by most telecom operators [13,14]. A participant with a PAYG data plan means he/she needs to pay the operator according to the amount of data transferred via cellular networks. An UnDP participant can transfer unlimited amounts of data via cellular networks during a time period (usually for a month) without additional cost, because the payment for UnDP is fixed. In the Wi-Fi uploading failure scenario, if the source participant's data plan is PAYG, he/she needs to pay additional uploading cost. However, if from the start the source participant, who cannot successfully upload the data through Wi-Fi networks, directly chooses cellular uploading, in which he/she forwards data to an UnDP participant within valid time of the data by opportunistic transmissions and the UnDP participant uploads the data by cellular networks. The additional transmission cost may be saved and part participants' resources will not be wasted.

Therefore, in this paper we study how to design an appropriate data uploading mechanism to make a balance between the offloaded traffic of cellular networks and participants' uploading cost considering the differences among participants' data plans and direct and indirect opportunistic transmissions. The uploading cost contains accessing Wi-Fi APs and the transmission of cellular networks. We proposed an Intelligent Data Uploading Selection Mechanism (IDUSM). The mechanism first helps the source participant choose an appropriate data uploading manner from Wi-Fi uploading and cellular uploading based on our proposed probability prediction model, and then continues optimizing the performance objective for the chosen data uploading manner, for example, the performance objective of Wi-Fi uploading is increasing the offloaded cellular traffic and the performance objective of cellular uploading is decreasing participants' uploading cost. The contributions of the paper can be summarized as follows:

- Our proposed IDUSM is distributed. It as a background program runs in every participant helping the participant make data uploading decision. It is feasible in large scale scenarios. Meanwhile, the proposed IDUSM realized a trade-off between the offloaded cellular networks traffic and

participants' uploading cost considering the differences among participants' data plans and probabilities of opportunistic transmissions.

- The decision of data uploading manner in IDUSM is based on our proposed probability prediction model. The model precisely predicts a participant's probability of successfully transferring data to a Wi-Fi AP considering direct and indirect opportunistic transmission and time variation of opportunistic contact patterns. It increases the accuracy of data uploading decision.

- The proposed IDUSM considers the waste of participants' limited resources by the redundant data produced in the Wi-Fi uploading process. It reduces the data redundancy by only allowing participants carrying data to replicate data to participants with higher probability of contacting the Wi-Fi AP.

The rest of the paper is organized as follows—Section 2 gives a brief overview of related work. In Section 3, we describe the proposed IDUSM. Section 4 presents the participant's mobility prediction model. Section 5 evaluates the performance of our proposed approach by simulation, and Section 6 concludes the paper.

2. Related Work

Nowadays, there are a lot of work studying the mobile data offloading problem from the aspect of offloading the data from the overloaded cellular networks to other networks (such as the Wi-Fi network) by opportunistic communications [15]. These work can be divided into two categories, offloading downlink traffic and uplink traffic according to the flow direction.

Offloading downlink traffic: according to the number of data requesters the downlink offloading works can also be divided into two categories—the data requested by many users [16–25] or by one user [26–29]. References [16–25] proposed to offload the data by selecting a subset of nodes downloaded data from cellular networks and other unselected nodes get the data by opportunistic transmissions. References [26–29] proposed methods to optimize the probability of offloading data to a specific destination.

All the methods mentioned above study how to offload downlink traffic of cellular networks. However, they are not suitable for offloading uplink traffic. Firstly, each mobile user's data is unique in the uploading process unlike downlink offloading where many mobile users request the same data. Secondly, there is not an appointed destination for the upload offloading. All contacted Wi-Fi APs are useful for offloading uplink traffic.

Offloading uplink traffic: for increasing the offloaded uplink traffic of cellular networks, Reference [30] proposed a weighted proportionally fair bandwidth allocation algorithm aiming to improve the energy efficiency of the participants and increase the offloaded data volume under the concurrent use of Wi-Fi APs and cellular networks; Reference [31] proposed a spatiotemporal opportunistic transmission algorithm by analyzing the spatiotemporal visiting probabilities and similarities of nodes to achieve high offloading success rate; Reference [32] developed an energy-efficient multi-hop data forwarding method based on Dynamic Source Routing (DSR) to deliver the sensory data to participants, and then upload the data to the backend server through cellular networks. These three methods ignore the cost of data uploading.

Part of the work is proposed to reduce the cost of uploading data by opportunistic communications. The Cost-Effective Multi-Mode Offloading (CEMMO) mechanism [33] selected the most effective upload offloading mode from cellular delivery, delay tolerance delivery and peer-assisted delivery through the prediction of user mobility and connectivity with Wi-Fi APs to reduce the overall cost in terms of financial settlement, energy consumption, and user satisfaction. ecoSense [34] attempted to minimize the organizer's data refund budget considering the two most common real-life 3G price plans UnDP and PAYG. Reference [35] proposed two algorithms including progress-balanced algorithm and social-aware forwarding algorithm to minimize total uploading cost of all users including data plan costs and extra costs for uploading data outside the plans. Reference [36]

proposed the Prediction-based User Recruitment for mobile crowdsEnsing-Delegation Forwarding (PURE-DF) method by delivering data to UnDP user to minimize data uploading cost.

　　Some other work optimizes other metrics in addition to uploading cost. Reference [37] proposed the Cost-aware Energy efficient Offloading (CEO) policy to minimize the energy consumption within a given deadline and a monetary cost constraint. An energy-efficient and cost-effective data uploading framework effSense [13] reduced the data cost by maximally offloading data to Wi-Fi APs or encountered data-plan users. Reference [38] proposed a general Delay-Aware Wi-Fi Offloading and Network Selection (DAWN) algorithm aiming to achieve a good tradeoff between the participant's payment and Quality of Service (QoS) characterized by the data's valid time. Reference [39] proposed a personalized data offloading scheme to provide maximum throughput within the cellular budget through employing an adaptive model to predict the throughput of Wi-Fi APs and the network usage of participants. Reference [40] proposed a preference-oriented offloading strategy for mobile node to make a proper cost-delay tradeoff.

　　However, none of the above-mentioned methods simultaneously optimizes the offloaded traffic of cellular networks and participants' cost of uploading data. In the paper, we propose the distributed IDUSM to select the appropriate data uploading manner based on the proposed probability prediction model realizing a trade-off between the offloaded traffic and the uploading cost, and meanwhile reducing unnecessary data copies in opportunistic offloading process to reduce the waste of participants' limited resources.

3. The Intelligent Data Uploading Selection Mechanism

　　The paper defines two data uploading manners: **Wi-Fi uploading** and **cellular uploading**, as illustrated in Figure 1, where the black dotted arrow is opportunistic transmitting and solid arrow represents non-delayed data link including cellular transmission and wired link transmission. **Wi-Fi uploading** utilizes opportunistic transmissions to offload the traffic of cellular networks to Wi-Fi networks. It contains source participants directly offload the data to Wi-Fi APs such as path 1 or indirectly offload the data to Wi-Fi APs with the help of other encountered participants such as path 2, and then the Wi-Fi networks upload the data to the cloud side by wired links. **Cellular uploading** utilizes cellular networks to upload data. It contains the source participant directly uploads data by cellular networks like path 3, or indirectly uploads data by cellular networks with the help of opportunistically encountered participants like path 4.
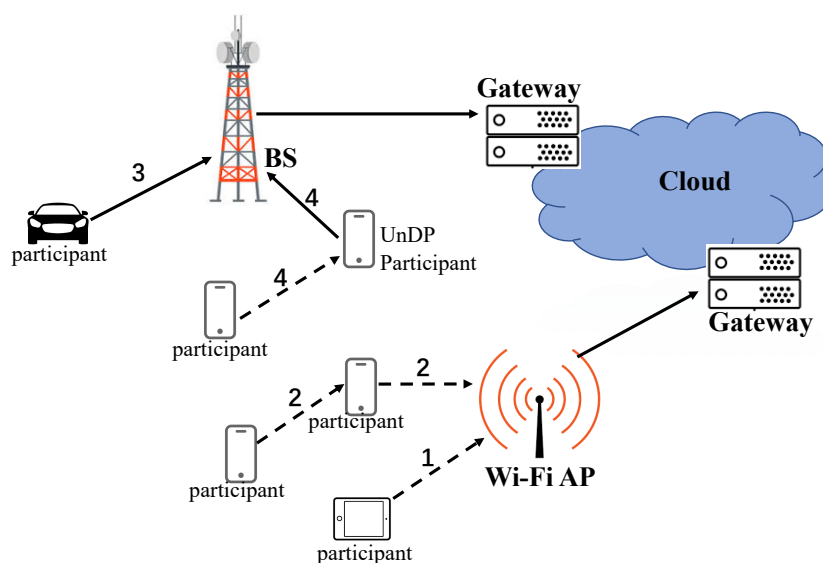

Figure 1. Data uploading manners.

Our proposed IDUSM as a background program runs in every participant. Its process is illustrated in Figure 2. It firstly helps the participant $u_i(i \in 1, 2, ..., n)$ choose an appropriate data uploading manner from Wi-Fi and cellular uploading. Wi-Fi uploading aims to offload the traffic of cellular networks, but cellular uploading aims to reduce transmission cost. For finding a balance between them, we set u_i chooses Wi-Fi uploading when his/her probability P_i is larger than α that means u_i has high possibility to offload data to Wi-Fi networks, or u_i chooses cellular uploading. P_i represents the probability that u_i transmits the data to a Wi-Fi AP within the valid time of the data by direct and indirect opportunistic transmissions, and it is calculated by our proposed probability prediction model (Section 4). The value of α is set by the server located in the cloud in advance. It influences the performance of proposed mechanism, and we will discuss it in Section 5.2.
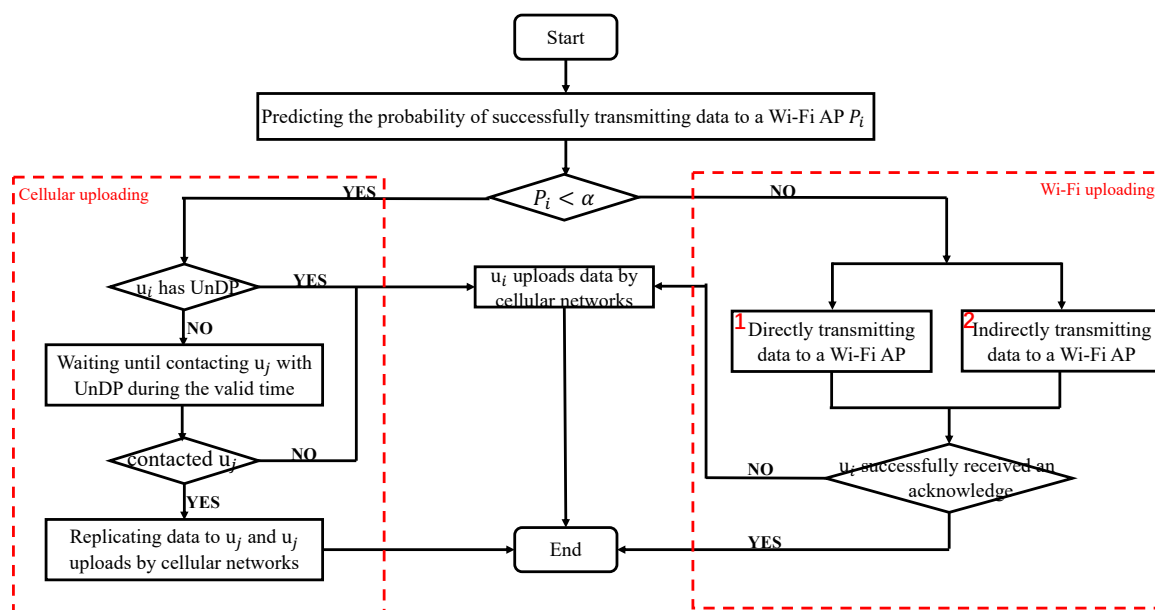
Figure 2. The proposed Intelligent Data Uploading Selection Mechanism (IDUSM).

If cellular uploading is chosen, u_i will execute the following processes to reduce the cost of uploading data by cellular networks. If the data plan of u_i is UnDP, he will directly upload the data by cellular networks, or he will wait until encountering a participant with UnDP $u_j(j \in 1, 2, ..., n)$ or at the end of the valid time of the data. If u_i encounters u_j, he will replicate data to u_j and u_j uploads the data by cellular networks, or u_i will directly upload the data by cellular networks. Through this way the transmission cost can be reduced.

If Wi-Fi uploading is chosen, u_i will execute the following processes to increase the offloaded cellular traffic while reducing the waste of participants' limited resources produced by unnecessary data replication. As shown in Figure 2, u_i executes sub-process 1 and 2 in parallel, and the sub-processes end when u_i receives an acknowledgement message or the current time exceeds the valid time of the data. If u_i does not receive an acknowledge message within valid time of the data, he/she will directly upload the data by cellular networks. Sub-process 1 represents u_i transmits data to directly encountered Wi-Fi APs. Sub-process 2 represents u_i indirectly transmits the data to Wi-Fi APs with the help of other encountered participants. Since transmitting data more than 2 hops does not improve the efficiency of opportunistic transmissions, even dramatically increases data redundancy [41,42], the path of opportunistically transmitting data to a Wi-Fi AP is at most two hops.

Redundant data produced by opportunistic transmissions in sub-process 2 consumes participants' limited resources, even data offloading ratio may be decreased. For example, there is a situation that the encountered participants do not have enough space to carry the data because of carrying other participants' data, specifically, the encountered participant can successfully offload the data to the Wi-Fi AP but some of its carried data cannot be successfully offloaded to the Wi-Fi APs. Therefore,

for reducing redundant data in sub-process 2, we propose u_i only replicates data to the encountered participant u_j whose probability p_j^D is larger than the probability p_i^D. p_j^D is the probability that u_j directly transmits the data to a Wi-Fi AP within valid time of the data, and p_i^D represents the probability that u_i directly transmits the data to a Wi-Fi AP within valid time of the data calculated in Section 4.1. The detailed procedures of sub-process 2 are illustrated in Figure 3.
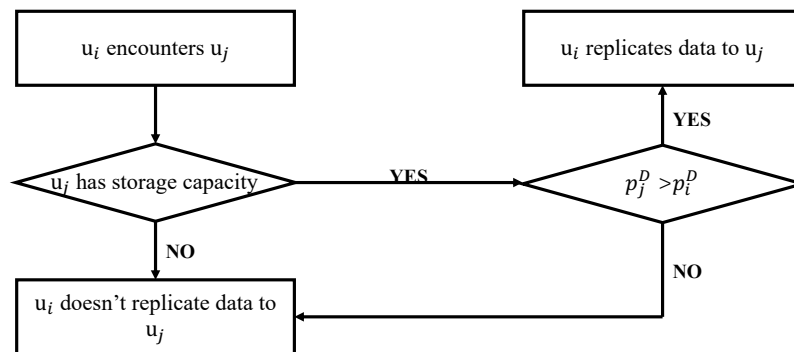

Figure 3. Data forwarding selection in sub-process 2.

4. The Probability of Transmitting Data to a Wi-Fi AP

In the paper, we focus on offloading traffic data, which can be completely transferred within one contact, such as texts or numerical data. Hence, the probability of a participant of transmitting data to a Wi-Fi AP within the valid time of the data is same with the probability that the participant directly or indirectly contacts the Wi-Fi AP within the valid time of the data. In the section, predicting the probability contains two parts: the probability of directly contacting a Wi-Fi AP within the valid time of the data (Section 4.1) and the probability of indirectly contacting a Wi-Fi AP within the valid time of the data (Section 4.2).

We consider a participant moves within the coverage of cellular networks such that the cellular connection is always available to the participant. Occasionally, the participant may be able to access Wi-Fi APs at some regions. This means the Wi-Fi connection is location dependent and may not be available to the participant at all time. Therefore, in the paper, we propose the Small Base Station (SBS) in cellular networks divides its covered area into small regions with the same size. The length of the diagonal of a small region should be smaller than the coverage radius of a Wi-Fi AP that ensures that a participant can connect to a Wi-Fi AP in a small region if the region has one or multiple Wi-Fi APs. A participant can get the location information of regions by its connected SBS when he/she accesses the covered area of the SBS for the first time. Figure 4 shows a participant's sample moving scenario. The area covered by a SBS is divided into small regions with unique identifiers and a participant can contact the Wi-Fi AP only in regions r_5, r_7, r_9. Since References [43–45] observe that the movement of an individual exhibits a high degree of regularity and an individual regularly visits a small set of locations/regions and moves between those locations, in this paper we use participants' historical contact records to predict the participants' probabilities of directly contacting a Wi-Fi AP or indirectly contacting a Wi-Fi AP with the help of other peer participants within a specific time constraint. The specific time constraint represents the valid time of the data. A participant records his/her transition and direct contacting records based on the location information of the divided regions, and the records of indirectly contacting Wi-Fi APs with the help of contacted peer participants are got by short-range wireless communications (e.g., bluetooth) when two participants contact each other.
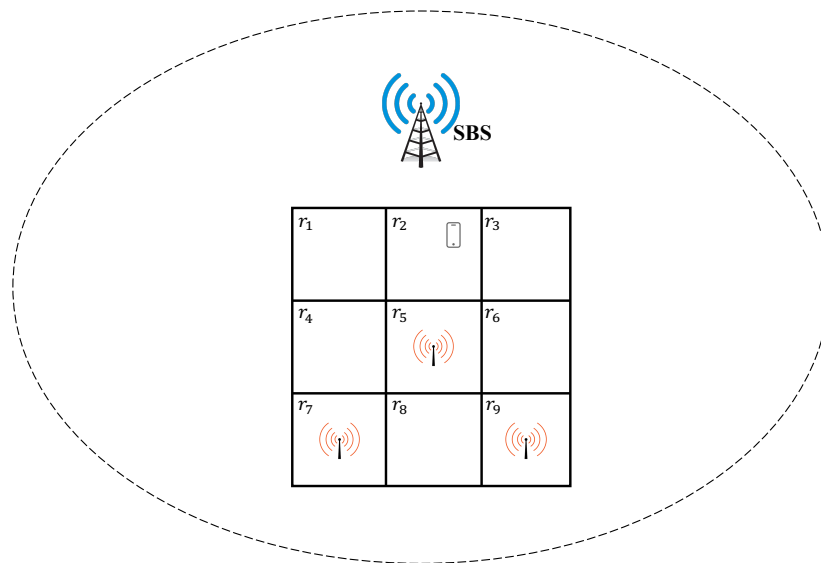
Figure 4. A sample scenario.

4.1. The Probability of Directly Contacting a Wi-Fi AP

Since contact patterns are time-varying, we predict the probability from spatial and temporal aspects. The spatial aspect predicts the probability that the participant moves between regions, and the temporal aspect predicts the duration of the participant of staying in a region. In this section, we integrate spatial and temporal aspects to predict the probability that a participant directly contacts a Wi-Fi AP within a specific time constraint.

4.1.1. Spatial Prediction

The movement of a participant is regular. Therefore, we set the time period to express regularity, and at every time period the participant's mobility is similar, for example, participants visit certain types of locations such as home and workplace at the same time everyday, and in this case a day is a time period. What's more, to accurately acquire the time-dependent transition probability we split the time period into a series of small time intervals, whose durations are the same.

Based on the historical transition information, the participant uses a matrix $M(t)$ to record his/her transition probability $p_{ij}(t)$ from a region $r_i \in \{\mathcal{R}_p - \mathcal{R}_w\}$ without a Wi-Fi AP to another region $r_j \in \mathcal{R}_p$ during the time interval t, where $\mathcal{R}_p(|\mathcal{R}_p| = m)$ represents regions that are visited at least x time periods aiming to eliminate regions visited randomly, and $\mathcal{R}_w(|\mathcal{R}_w| = m')$ is the Wi-Fi available regions. The parameter x is set by the participant based on real data sets in advance. The transition probability $p_{ij}(t)$ is calculated as follows:

$$p_{ij}(t) = N(r_i r_j, t)/N(r_i, t), \tag{1}$$

where $N(r_i r_j, t)$ and $N(r_i, t)$ are statistics based on historical records of the participant; $N(r_i, t)$ is the number of the participant visiting region r_i during t; $N(r_i r_j, t)$ is the number of the participant moving from r_i to r_j during t; when $r_j = r_i$, $N(r_i r_j, t)$ represents the number of the participant staying in r_i during t. Since a participant only stays in its original region or moves to adjacent regions during a time interval, for example, the participant can only stays in r_2 or move to its adjacent regions r_1, r_3, r_5 at the current time interval in Figure 4, the probabilities of moving to non-adjacent regions are 0. The probability calculation method refers to Reference [33].

4.1.2. Temporal Prediction

Temporal prediction gets the duration that a participant stays in a region before making a transition to some other regions at a time interval. Using historical records of previous durations, we estimate probabilities of the durations in a region at a time interval. We act the duration, which has maximum probability, as the participant's staying time in the region at the time interval. The detailed calculation method is described as follows:

$\mathcal{H}_i(t) = \{d|d \in \{1, 2, 3, ...\}\}$ represents historical records of previous duration of the participant in the region r_i at t, where d is the number of time intervals which is divided by duration. $p_i^t(d)$ calculated by (2) represents the probability that at t the participant stays in the region r_i d time intervals, where $N(r_i t, d)$ is the number of the participant staying in r_i for d time intervals at t. We choose d whose probability is the largest from the set $\{p_i^t(d)|d \in \mathcal{H}_i(t)\}$ as the staying duration of u_i in region r_i at t.

$$p_i^t(d) = N(r_i t, d) / \sum_{d \in \mathcal{H}_i(t)} N(r_i t, d). \tag{2}$$

4.1.3. The Combination of Spatial and Temporal Prediction

A participant may need to move to other regions to contact a Wi-Fi AP, for example in Figure 4, although the participant cannot contact a Wi-Fi AP at the current time interval, he can contact a Wi-Fi AP through multiple movements such as $r_2 \to r_5$ or $r_2 \to r_1 \to r_4 \to r_5$ or $r_2 \to r_1 \to r_4 \to r_7$, and so forth. Thus, a participant's probability of contacting a Wi-Fi AP within a specific time constraint should be the sum of probabilities of multiple movements.

The movement of a participant in a set of possible regions following a Markovian mobility model, which means the next visiting region of a participant only depends on the previous region. The model is widely used in the literature [33,38,46], hence we also use a first-order Markov model as the spatial prediction model to predict the probability of the participant of moving to the next region. Moreover, since the probability of moving from a region to another depends on time, our proposed probability prediction model combines spatial and temporal prediction. We describe one step transition probability of a participant in the following.

We use r_s to record the starting region of the participant; t to record the current time instance located time interval; t_e to record the time interval that the deadline of the time constraint located in. We get the number of the participant's staying time intervals d based on the temporal prediction method (Section 4.1.2). Thus, the time interval that the participant moves to the next region is $(t + d)$ and the transition probability matrix is $M(t + d)$. If $((t + d) < t_e)$, the participant will move to the next possible adjacent region r_a, and the probability $p_a(t + d)$ that the participant reaches the region r_a is calculated by (3),

$$p_a(t + d) = p_s(t) \cdot p_s^t(d) \cdot p_{sa}(t + d), \tag{3}$$

where $p_s(t)$ is the probability that the participant reaches the region r_s at the time interval t; $p_s^t(d)$ represents the probability of staying in the region r_s for d time intervals; $p_{sa}(t + d)$ gotten from $M(t + d)$ is the probability that the participant moves from r_s to r_a at the time interval $(t + d)$.

Since a participant's next region has some possibilities, he/she might contact a Wi-Fi AP by multiple movements, for example, in Figure 4 the participant can contact a Wi-Fi AP by directly moving to the region r_5 or multiple movements like $r_1 \to r_4 \to r_7$ or $r_3 \to r_6 \to r_9$. For calculating the probability that a participant directly contacts a Wi-Fi AP within a time constraint, we need to traverse all possible movements during the time constraint. Hence we recursively call our proposed one step transition probability calculation method, where every possible next region, which do not contain a Wi-Fi AP as the starting region, to calculate the participant's probability of directly contacting the Wi-Fi AP within a time constraint. An iteration ends when the time interval exceeds the time interval that the deadline of the time constraint located in or the participant reaches a region contained a Wi-Fi AP, and the probability will be added to p_i^d if the participant arrives the region with a Wi-Fi AP within

the specific time constraint. The above-mentioned procedures end when all possible next regions are traversed, and we get p_i^d ultimately. The whole process is described in Algorithm 1.

Algorithm 1 Predicting the probability of u_i directly contacting a Wi-Fi Access Point (AP) within a time constraint

---

Input: the starting region r_s, the starting time interval t, the stop time interval t_e.

Output: P_i^D

Initialize: $P_i^D = 0, p_s(t) = 1$

1: **OneStepTransition()** {
2: **if** (r_s contains a Wi-Fi AP) **then**
3:    $P_i^D += p_s(t)$;
4: **else**
5:    Getting the staying time intervals d and $p_s^t(d)$ using (2);
6:    $t += d$;
7:    **if** ($t < t_e$) **then**
8:        **for** every possible next region r_a **do**
9:            Calculating $p_{sa}(t)$ using (1);
10:            Calculating $p_a(t)$ using (3);
11:            $r_s = r_a$;
12:            $p_s(t) = p_a(t)$;
13:            Calling **OneStepTransition()**
14:        **end for**
15:    **end if**
16: **end if**}
17: **return** P_i^D;

We use an example to illustrate the Algorithm 1. The moving scenario of the participant u_1 is shown in Figure 4. His start region is r_2. The current time interval is 1 and the stop time interval is 9. The whole processes of the Algorithm 1 is shown in Figure 5. Since $p_2^1(2)$ is the largest among other possible staying durations, u_1 stays in r_2 for 2 time intervals, and then u_1 moves to the next adjacent region at the 3rd time interval . u_1 finds he has three choices: r_1, r_3, r_5, whose probabilities are not zero based on his transition matrix. We found the summary of probabilities of arriving all possible regions (r_1, r_3, r_5) is not 1, because the participant also randomly moves to other regions besides the periodic visiting regions. If u_1 reaches r_1, he will continue moving until reaching r_5 or r_7, because r_1 and r_4 do not contain a Wi-Fi AP and the transition time interval is smaller than the stop time interval. The probabilities of reaching r_5 and r_7 are added to P_1^D, and this branch ends. If u_1 arrives r_3, he will move to the next region since r_3 does not contain a Wi-Fi AP and the transition time interval is smaller than the stop time interval. Based on the transition matrix u_1's next step has only one choice: r_6. However, when u_1 reaches r_6, the branch will end since the transition time interval exceeds the stop time interval. What's more, since r_5 contains a Wi-Fi AP, u_i's probability of arriving r_5 is directly added to P_1^D and this branch ends. Until now, all possible arriving regions are traversed and the whole process ends. P_1^D is equal to 0.31.
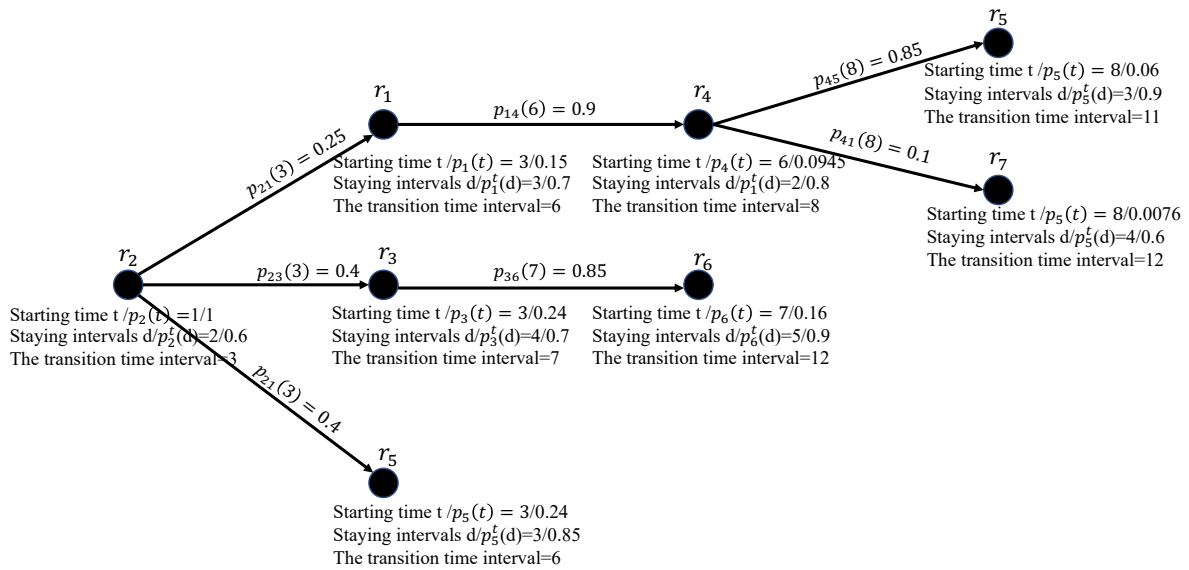
Figure 5. An example of Algorithm 1.

4.2. The Probability of Indirectly Contacting a Wi-Fi AP

Calculating the probability p_i^I that a participant $u_i(i \in \{1, 2, ..., n\})$ indirectly contacts a Wi-Fi AP with the help of other contacted peer participants within a specific time constraint includes two parts. The first is calculating the probability of contacting a peer participant u_j at the time interval t_j, which is smaller than the stop time interval t_e, as illustrated in Figure 6 where t is the starting time interval. Since transmitting data more than 2 hops does not improve the forwarding efficiency, even dramatically increases data redundancy, the second part is to calculate u_j's probability of directly contacting a Wi-Fi AP and the starting time interval is t_j.
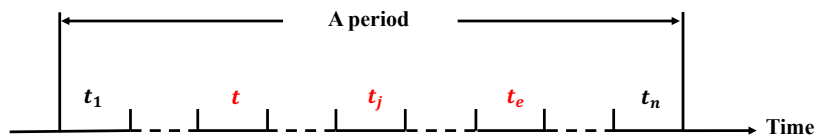

Figure 6. The position of t_j.

u_i calculates the probability p_i^I based on his historical contact records $< u_j, n_j, t_j, r_j >$, which represents the participant contacts u_j n_j times at the time interval t_j in the region r_j. For every historical contact record, the chosen contacted participant u_j should satisfy $(n_j > x)$ to eliminate the influence of randomly contacted participants and $(t_j < t_e)$ to limit the contacted time. The parameter x is defined in Section 4.1.1. The probability $p_j^j(t_j)$ that u_i encounters u_j at the time interval t_j in the region r_j is calculated by (4), where n is the sum of all contact times in historical contact records. After calculating the probability $p_j^j(t_j)$, we need to calculate the probability p_j^D that u_j directly contacts a Wi-Fi AP within the rest of time. p_j^D is calculated through calling the Algorithm 1, and the input parameters including the starting region, the starting time interval and the stop time interval are r_j, t_j and t_e respectively. p_i^I is calculated using (5). The whole process is described in Algorithm 2.

$$p_j^j(t_j) = n_j/n; \tag{4}$$

$$p_i^I = p_i^I + p_j^j(t_j) \cdot p_j^D. \tag{5}$$

We integrate the above-mentioned two methods of calculating the probability of directly contacting a Wi-Fi AP p_i^D and the probability of indirectly contacting a Wi-Fi AP p_i^I to get the

probability P_i that the participant u_i contacts a Wi-Fi AP within a specific time constraint shown in (6). In conclusion, the participant executes our proposed probability calculation algorithm when he/she has data to be uploaded, and the computation complexity of the algorithm is $O(\frac{T}{D_s} \cdot (N_e + 1))$, which is decided by the number of time intervals T split by the valid duration of the data, the number of encountered participants N_e during the valid time of the data, and the average number of time intervals D_s staying in a region. Moreover, the probability is calculated based on the participants' historical contact records. The participant directly records contacted Wi-Fi APs and peer participants, and the records of indirectly contacting Wi-Fi APs are got by short-range wireless communications when the participant contacts peer participants. The methods of getting information do not occupy communication resources of cellular networks.

$$P_i = p_i^D + p_i^I. \tag{6}$$

Algorithm 2 Predicting the probability of u_i indirectly contacting a Wi-Fi AP within a time constraint

---

Input: the starting region r_s, the starting time interval t, the stop time interval t_e.

Output: p_i^I

Initialize: $p_i^I = 0, p_s(t) = 1$

1: **for** every historical records of u_i **do**
2: 　**if** $((t < t_j < t_e)\&(n_j > x))$ **then**
3: 　　Calculating $p_j^j(t_j)$ using (4);
4: 　　Calling Algorithm 1(r_j, t_j, t_e) to get p_j^D;
5: 　　$p_i^I + = p_j^j(t_j) \cdot p_j^D$;
6: 　**end if**
7: **end for**
8: **return** p_i^I;

---

5. Performance Evaluation

In this section, we first describe the setup of simulation environment and performance evaluation metrics, then discuss the value of α in our proposed IDUSM, and finally we compare performance of our proposed IDUSM with three related offloading mechanisms under different simulation scenarios.

5.1. Environment Setup

The simulation is executed on the ONE simulation platform [47]. In the simulation, there are 200 participants with UnDP and PAYG data plan, and the exact number of each kind is varied based on different experiments. The participants move within a 4.5 km $\times$ 3.5 km section of Helsinki, Finland, and they follow the Working Day Movement model, which is a realistic mobility model that simulates the daily mobility of people. A cellular BS locates at the center of the map and covers all the users within the field. Wi-Fi APs are randomly located with the sites in the map, and the number of Wi-Fi APs is varied based on different experiments.

Every participant executes a Poisson process to fire an event for data generation, where the interval between two sequential events follows the Poisson distribution with parameter $\lambda = 1800$ s. The size of data is 2 MB. We also set the bandwidth of BS and Wi-Fi AP are unlimited that ensure the data can be completely transferred during a contact. Considering participants' storage capacities, we set their caches are 50 MB. For evaluating the cost of transferring data, we assume the usage price of cellular networks of the participant with PAYG data plan is 1 RMB/MB and the participant with UnDP is free. The cost of transferring data through the Wi-Fi AP is 0.05 RMB/MB. The values of uploading cost of

PAYG participant and Wi-Fi AP aims to express their ratio referred to References [33,48]. They are not the real cost.

In all experiments, we set the region size is equal to 300 m × 340 m, and time interval duration is equal to 10 min. It corresponds to 150 regions and 144 distinct time intervals. The total duration of our experiment is 20 days. We use 10 days to train, which aims to build the prediction model, and the rest 10 days to evaluate. We evaluate the performance from the following three metrics:

1. Offloading ratio: is used to measure the offloaded cellular network traffic. It is expressed by the fraction of the total amount of generated data that is offloaded through Wi-Fi networks shown as follow:

offloading ratio = (Total data offloaded through Wi-Fi networks)/(Total data generated)

2. Offloading efficiency: Offloading ratio is improved by increasing data redundancy in the network, but the transmission of redundant data wastes participants' limited resources such as storage and battery. In the paper, each generated data or replicated data is referred to as a data segment. We use offloading efficiency, which is represented by the fraction of offloading ratio to the number of data segments in the network, to measure the efficiency of an offloading mechanism shown as follow:

offloading efficiency = (Offloading ratio)/(The number of data segments in the network)

3. Uploading cost: is the summary of the data uploading cost of all participants. A participant's uploading cost contains the cost of using cellular networks and Wi-Fi APs to upload the data shown as follows:

$$\text{uploading cost} = \sum_n (C_c \cdot D + C_w \cdot D'),$$

where n is the number of participants; C_c and C_w represent transmission cost of cellular networks and the Wi-Fi APs respectively; D and D' is the size of data uploaded through cellular networks and Wi-Fi APs.

5.2. Performance of the Proposed IDUSM

For our proposed IDUSM, we discuss the value of α, which is used to decide the selection of Wi-Fi or cellular uploading manner realizing a trade-off between the offloading ratio and the uploading cost. We set the valid time of the data is 2 h. The number of participants with UnDP and PAYG is 100 respectively. The number of Wi-Fi APs is 20. As shown in Figure 7a, we find that when ($\alpha < 0.6$), with the increase of α the offloading ratio decreases slowly, but the uploading cost decreases rapidly. The reason is the increase of α makes some participants, who cannot successfully upload data through Wi-Fi uploading, directly choose cellular uploading, which opportunistically forwards data to contacted UnDP participants and utilizes the UnDP participants upload the data by cellular networks. Through this way the uploading cost produced by the failure of Wi-Fi uploading is reduced, but its influence on the offloading ratio is slight.

When ($\alpha \geq 0.6$), with the increase of α, more and more participants, who can upload data through Wi-Fi uploading, are forced to choose cellular uploading. However, some of the participants with PAYG data plan cannot contact UnDP participants during the valid time of the data. They have to upload the data using their network traffic, which leads to a slow decline in uploading cost. Therefore, when ($\alpha \geq 0.6$), the offloading ratio sharply decrease and the uploading cost decreases slowly with the increase of α.

In conclusion, the value of α influences the performance of the proposed IDUSM, and its value should be set based on different experiment scenarios. We set ($\alpha = 0.6$) to achieve the optimal performance in this experiment environment.
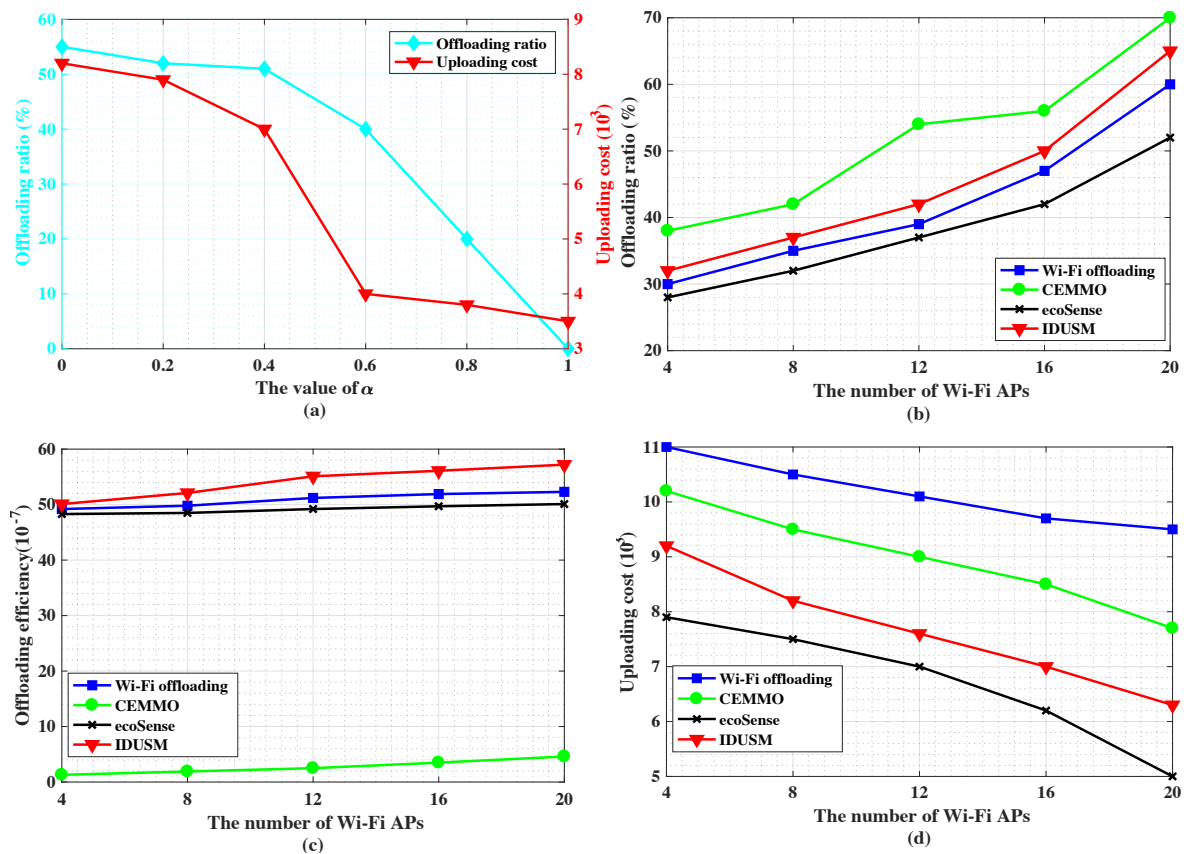
Figure 7. (**a**) The influence of the value of α on the performance of IDUSM. (**b**) The impact of the number of Wi-Fi APs on offloading ratio. (**c**) The impact of the number of Wi-Fi APs on offloading efficiency. (**d**) The impact of the number of Wi-Fi APs on uploading cost.

5.3. Performance Comparison with Other Mechanisms

In this paper, we compare the performance of our proposed IDUSM with three related offloading mechanisms: Wi-Fi Offloading [49], CEMMO [33] and ecoSense [34] from the aspects of offloading ratio, offloading efficiency and uploading cost.

Wi-Fi Offloading: in addition to directly transfer data, participants also indirectly transfer data to Wi-Fi APs with the help of contacted peer participants. The data can be transferred to Wi-Fi APs through multi-hops. If the data is not transferred to the destination within its valid time, it will be uploaded through cellular networks.

CEMMO: is a centralized mechanism and does not distinguish participants' uploading cost. It contains three modes of uploading: cellular delivery, direct Wi-Fi offloading and peer-assisted Wi-Fi offloading. Direct Wi-Fi offloading means the participant carries the data without forwarding until contacting a Wi-Fi AP. In peer-assisted Wi-Fi offloading mode, the participant will flood data to participants in a passing region where the summary of participants' probabilities is the maximum among other passing regions, and the participants offload data to Wi-Fi APs. If the cost of direct offloading is smaller than peer-assisted offloading, direct offloading will be chosen, or peer-assisted offloading mode will be chosen. If the data does not be uploaded to the cloud within its valid time through direct and peer-assisted offloading, it will be uploaded through cellular delivery.

ecoSense: aims to minimize the uploading cost. If a participant has UnDP, he/she directly uploads data by cellular networks, or he/she uploads data by the OneHopFlooding strategy that the PAYG participant would relay data to another directly contacted PAYG participant until he/she meets either one of the two following stopping criteria: (1) he/she directly encountered an UnDP participant or (2) the server notifies that he/she could stop flooding.

The performance of mechanisms is evaluated under different simulation environment. Parameters used to control simulation environment are the density of Wi-Fi APs, the valid time of data, the ratio of UnDP participants' numbers to PAYG participants' numbers represented by β.

As shown in Figures 7b,c, 8a,b,d and 9a, firstly, the offloading ratio of CEMMO outperforms other three Wi-Fi offloading mechanisms no matter how parameters change and its average value is 58.9%, since it makes source participants indirectly transmit data to Wi-Fi APs by flooding the data to participants in a region besides directly transmitting data, but other three Wi-Fi offloading mechanisms only replicates data to encountered parts of participants, therefore, the offloading efficiency of CEMMO is the lowest among all offloading mechanisms, and its average value is (1.75×10^{-7}). Secondly, the average values of offloading ratio and offload efficiency of our proposed IDUSM are respectively 3.4% and (4.9×10^{-7}) larger than Wi-Fi Offloading mechanism in all simulation scenarios, since in the process of opportunistic offloading besides accurate mobility prediction our proposed IDUSM only replicates data to encountered participants whose probabilities are higher than carried data participants, that reduces the situation that encountered participants do not have enough space to carry the data because of carrying other undelivered data. Thirdly, the average values of offloading ratio and offloading efficiency of Wi-Fi Offloading are respectively 5% and (2.99×10^{-7}) larger than ecoSense mechanism, since ecoSense aims to reduce uploading cost but Wi-Fi Offloading focuses on traffic offloading. It is worth noting that although the values of offloading efficiency of mechanisms are almost same, the tiny gap reflects the huge gap in the number of redundant data generated by mechanisms.

The total uploading cost is related with offloading ratio and uploading participants' data plans. As shown in Figures 7d, 8c and 9b, the uploading cost sorting from cheap to expensive is ecoSense, IDUSM, CEMMO and Wi-Fi Offloading mechanism in all simulation scenarios, and the average values are respectively (5.68×10^{3}), (6.79×10^{3}), (8.18×10^{3}), (9.78×10^{3}). The reasons are illustrated as follows. Firstly, in ecoSens mechanism participants with UnDP directly upload data by cellular networks aiming to minimize uploading without considering offloading the traffic of cellular network, but in our proposed IDUSM, participants with high probability of offloading data to the Wi-Fi AP will choose Wi-Fi offloading. Not only the cellular traffic is offloaded, but also the uploading cost is reduced because the uploading cost of Wi-Fi APs is low in IDUSM. However, the cost of uploading data by participants with UnDP is free, therefore, uploading cost of ecoSense is lower than IDUSM. Secondly, the uploading cost of the proposed IDUSM is lower than CEMMO mechanism, since CEMMO does not consider differences among participants' data plans. Finally, the uploading cost of CEMMO is lower than Wi-Fi offloading mechanism, since the offloading ratio of CEMMO is bigger than Wi-Fi offloading mechanism and the transmission cost of Wi-Fi APs is much cheaper than cellular networks.

In the following, we illustrate changes of performance of mechanisms and reasons of the changes under different simulation environment in detail.

5.3.1. Impact of the Density of Wi-Fi APs

In the experiment, we evaluate the impact of the density of Wi-Fi APs on the performance of Wi-Fi Offloading, CEMMO, ecoSense and our proposed IDUSM. We set the valid time of data is 2 h. The number of UnDP and PAYG participants is 100, respectively. The number of Wi-Fi APs ranges from 4 to 20. According to Figure 7b, the offloading ratios of all mechanisms increase with the increase of the number of Wi-Fi APs, since participants' possibilities of offloading data to Wi-Fi APs are increased. At the same time, the total cost decreases with the increase of the number of Wi-Fi APs shown in Figure 7d, because more data is uploaded through Wi-Fi APs, whose uploading cost is cheap.
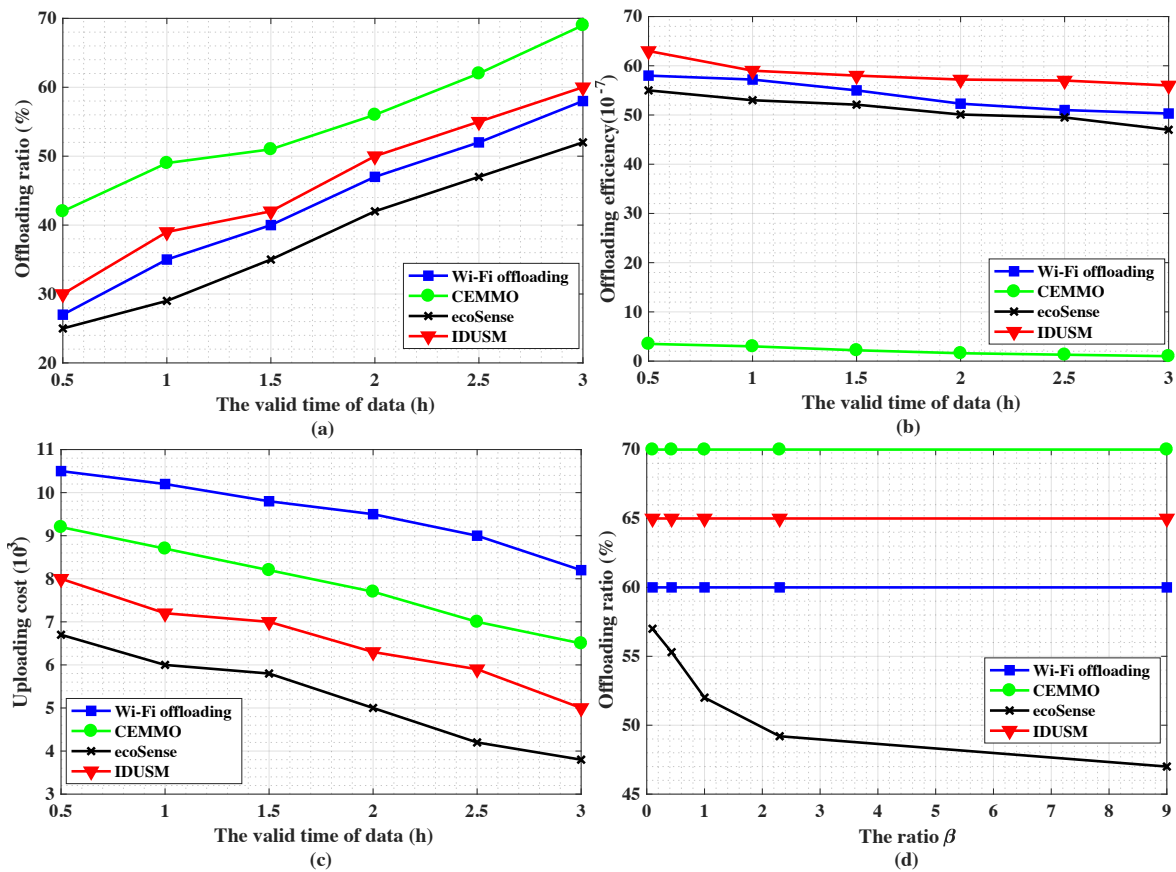
Figure 8. (**a**) The impact of the valid time of data on offloading ratio. (**b**) The impact of the valid time of data on offloading efficiency. (**c**) The impact of the valid time of data on uploading cost. (**d**) The impact of the ratio β on offloading ratio.
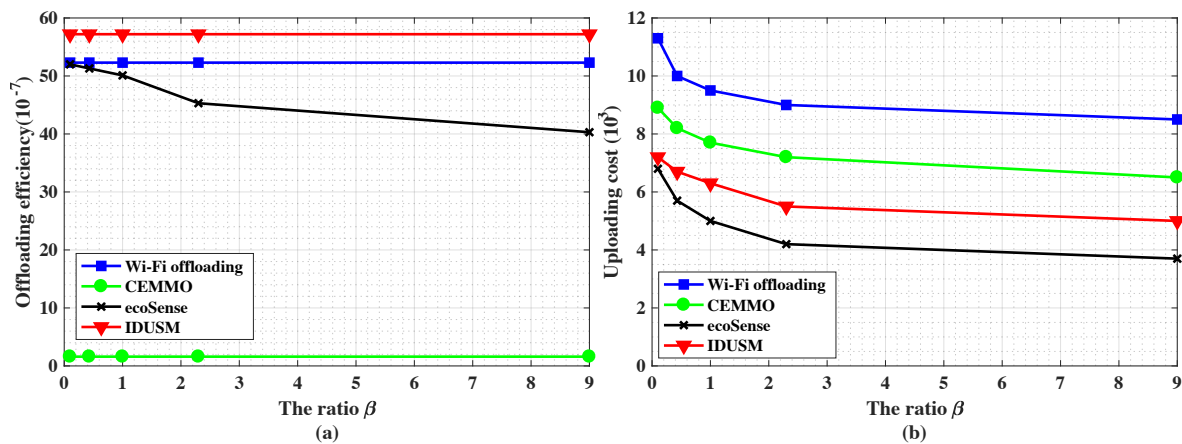
Figure 9. (**a**) The impact of the ratio β on offloading efficiency. (**b**) The impact of the ratio β on uploading cost.

For Wi-Fi Offloading, ecoSense, IDUSM and CEMMO, the increase of the density of Wi-Fi APs makes participants more quickly encounter Wi-Fi APs so that the duration of replicating data to encountered peer participants becomes short, therefore the number of data in the network decreases as the increase of Wi-Fi APs. Therefore, with the increase of the density of Wi-Fi APs, the offloading efficiency of Wi-Fi Offloading, CEMMO, ecoSense and IDUSM increase illustrated in Figure 7c.

5.3.2. Impact of the Valid Time of Data

The valid time of data plays an importance role in the process of data offloading. In the experiment, we set the number of Wi-Fi APs is 20. The number of UnDP user and PAYG user is equal to 100 respectively. Data's valid time ranges from 0.5 to 3 h. Figure 8a shows that the offloading ratio of all mechanisms increases with the increase of the delay tolerant time of the data, since participants' probabilities of encountering Wi-Fi APs increase, and meanwhile, the uploading cost of all mechanisms decreases because the transmission cost of Wi-Fi APs is cheap shown in Figure 8c.

What is more, Figure 8b shows the change of offloading efficiency with the increase of the valid time of data. For all offloading mechanisms, the number of data in the network increases as the increase of the data's valid time, because the duration of replicating data to participants increases and the participants carried data encounters more peer participants in this process. The increased number of data in the network is much larger than the increased offloading ratio, thus the offloading efficiency of Wi-Fi Offloading, ecoSense, CEMMO and IDUSM decreases with the increase of the valid time of data.

5.3.3. Impact of the Ratio β

In this experiment, we examine the influence of the ratio of UnDP participants' number to PAYG participants' number on the performance of all mechanisms. We set the number of Wi-Fi APs is 20. Data's valid time is 2 h. The number of all participants is 200, and the number of UnDP participants varies from 20 to 180.

In Figures 8d and 9a, for Wi-Fi Offloading and our proposed IDUSM, the selection of data uploading method and replicating data to the encountered participant only consider the probability of encountering Wi-Fi APs. The cost of transferring data would not influence these processes. For CEMMO, although the source participant selects the data uploading method based on transmission cost, it acts all participants' transmission cost through cellular networks as the same without considering their differences. Hence, for Wi-Fi Offloading, CEMMO and our proposed IDUSM, the offloading ratio and offloading efficiency will not change with the increase of β. For ecoSense, with the increase of β, more and more participants with UnDP upload data by cellular uploading instead of Wi-Fi offloading, therefore, the offloading ratio decreases. However, in the situation the number of data in the network does not change, therefore, the offloading efficiency also decreases.

As shown in Figure 9b, all mechanisms' actual uploading cost decreases with the increase of β, since data has greater probability of uploading through UnDP participants beside Wi-Fi uploading, that does not need to pay for the use of cellular networks.

In conclusion, the offloading efficiency of our proposed IDUSM is the highest among all offloading mechanisms, and its average value is (56.54×10^{-7}). The offloading ratio and uploading cost of IDUSM are not optimal in all experiments. The average value of offloading ratio of CEMMO is 6.53% larger than IDUSM, and the average uploading cost of ecoSense is (1.11×10^3) higher than IDUSM. However, at the same time, the average uploading cost of CEMMO is also (1.39×10^3) higher than IDUSM and the average value of offloading ratio of ecoSense is 9.27% lower than IDUSM. Therefore, the experimental results prove our proposed IDUSM realized a trade-off between the offloading ratio and the uploading cost.

6. Conclusions and Future Work

In this paper, we studied how to design an appropriate data uploading mechanism to offload cellular traffic produced by the uploading process of mobile crowd sensing applications, and the proposed the IDUSM executed in every participant's background to make a balance between the offloaded cellular traffic and the uploading cost. Simulation results proved that the offloading efficiency of our proposed IDUSM outperforms other three offloading mechanisms. Meanwhile, although the offloading ratio and uploading cost of IDUSM are not optimal in all experiments, it realized a trade-off between the offloading ratio and the uploading cost. However, our proposed IDUSM only considers

that the data can be completely transferred during one contact. It is not suitable for scenarios in which participants need to upload the data with large size, such as video. Therefore, in the future we will consider offloading data with large size, and focus on maximizing the probability of offloading data to Wi-Fi APs by fragmenting the data and assigning the fragments to different direct and indirect paths generated by opportunistic contacts.

There are still many interesting open research areas about traffic offloading needed to be investigated. One focus is to offload downlink traffic requested by many users. Some interesting approaches to solve the issue are user selection [21,50,51] and edge caching [52–55]. User selection means the selected users download the data by cellular networks, and other requested users get the data from the selected users by short-range wireless communications. Edge caching offloads the downlink traffic by caching popular contents on edge nodes, and the requested users get the data by wireless communications. In addition, with the popularity of Artificial Intelligence (AI), utilizing AI technology to solve the problem of traffic offloading is also a meaningful topic to research [56,57].

Author Contributions: Q.W. designed and wrote the manuscript. J.F. revised the manuscript. B.G., X.D. and M.G. gave great suggestions to the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guo, B.; Yu, Z.; Zhou, X.; Zhang, D. From participatory sensing to mobile crowd sensing. In Proceedings of the 2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS), Budapest, Hungary, 24–28 March 2014; pp. 593–598.
2. Capponi, A.; Fiandrino, C.; Kantarci, B.; Foschini, L.; Kliazovich, D.; Bouvry, P. A survey on mobile crowdsensing systems: Challenges, solutions, and opportunities. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2419–2465. [CrossRef]
3. Sun, G.; Song, L.; Yu, H.; Du, X.; Guizani, M. A Two-tier Collection and Processing Scheme for Fog-based Mobile Crowd Sensing in the Internet of Vehicles. *IEEE Int. Things J.* **2020**. [CrossRef]
4. Zhou, P.; Zheng, Y.; Li, M. How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. In Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services, Manchester, UK, 22–25 June 2012; pp. 379–392.
5. Kong, X.; Xia, F.; Li, J.; Hou, M.; Li, M.; Xiang, Y. A shared bus profiling scheme for smart cities based on heterogeneous mobile crowdsourced data. *IEEE Trans. Ind. Inform.* **2019**, *16*, 1436–1444. [CrossRef]
6. Cisco, C. Global mobile data traffic forecast update, 2017–2022. *White Paper*, 28 February 2020.
7. Romanous, B.; Bitar, N.; Imran, A.; Refai, H. Network densification: Challenges and opportunities in enabling 5G. In Proceedings of the 2015 IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD), Guildford, UK, 7–9 September 2015; pp. 129–134.
8. Bader, A.; ElSawy, H.; Gharbieh, M.; Alouini, M.S.; Adinoyi, A.; Alshaalan, F. First mile challenges for large-scale IoT. *IEEE Commun. Mag.* **2017**, *55*, 138–144. [CrossRef]
9. Gharbieh, M.; Bader, A.; ElSawy, H.; Yang, H.C.; Alouini, M.S.; Adinoyi, A. Self-organized scheduling request for uplink 5G networks: A D2D clustering approach. *IEEE Trans. Commun.* **2018**, *67*, 1197–1209. [CrossRef]
10. Akpakwu, G.A.; Silva, B.J.; Hancke, G.P.; Abu-Mahfouz, A.M. A survey on 5G networks for the Internet of Things: Communication technologies and challenges. *IEEE Access* **2017**, *6*, 3619–3647. [CrossRef]
11. Al-Falahy, N.; Alani, O.Y. Technologies for 5G networks: Challenges and opportunities. *IT Prof.* **2017**, *19*, 12–20. [CrossRef]
12. Feng, B.; Zhang, C.; Liu, J.; Fang, Y. D2D communications-assisted traffic offloading in integrated cellular-WiFi networks. *IEEE Int. Things J.* **2019**, *6*, 8670–8680. [CrossRef]

13. Wang, L.; Zhang, D.; Yan, Z.; Xiong, H.; Xie, B. effSense: A novel mobile crowd-sensing framework for energy-efficient and data uploading. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 1549–1563. [CrossRef]

14. Yu, H.; Cheung, M.H.; Huang, L.; Huang, J. Predictive delay-aware network selection in data offloading. In Proceedings of the 2014 IEEE Global Communications Conference, Austin, TX, USA, 8–12 December 2014; pp. 1376–1381.

15. Xu, D.; Li, Y.; Chen, X.; Li, J.; Hui, P.; Chen, S.; Crowcroft, J. A survey of opportunistic offloading. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2198–2236. [CrossRef]

16. Zhou, H.; Wang, H.; Zhu, C.; Leung, V.C. Freshness-Aware Initial Seed Selection for Traffic Offloading Through Opportunistic Mobile Networks (WCNC). In Proceedings of the 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 15–18 April 2018; pp. 1–6.

17. Barua, B.; Khan, Z.; Han, Z.; Abouzeid, A.A.; Latva-aho, M. Incentivizing selected devices to perform cooperative content delivery: A carrier aggregation-based approach. *IEEE TWC* **2016**, *15*, 5030–5045. [CrossRef]

18. Rebecchi, F.; de Amorim, M.D.; Conan, V. Should I Seed or Should I Not: On the Remuneration of Seeders in D2D Offloading. In Proceedings of the 2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), Coimbra, Portugal, 21–24 June 2016; pp. 1–9.

19. Chang, Z.; Gong, J.; Zhou, Z.; Ristaniemi, T.; Niu, Z. Resource Allocation and Data Offloading for Energy Efficiency in Wireless Power Transfer Enabled Collaborative Mobile Clouds. In Proceedings of the 2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Hong Kong, China, 26 April–1 May 2015; pp. 336–341.

20. Wang, W.; Wu, X.; Xie, L.; Lu, S. Joint storage assignment for D2D offloading systems. *Comput. Commun.* **2016**, *83*, 45–55. [CrossRef]

21. Zhou, H.; Chen, X.; He, S.; Zhu, C.; Leung, V.C. Freshness-aware seed selection for offloading cellular traffic through opportunistic mobile networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 2658–2669. [CrossRef]

22. Balaji, C.G.; Anu Monisha, A.; Murugan, K. A Novel Hybrid UE Selection Scheme for Efficient Data Offloading Using D2D Communication. *Comput. J.* **2020**, *63*, 1513–1523.

23. Dash, S.K.; Dash, S.; Mishra, J.; Mishra, S. Opportunistic mobile data offloading using machine learning approach. *Wirel. Pers. Commun.* **2020**, *110*, 125–139. [CrossRef]

24. Zhang, Y.; Li, J.; Li, Y.; Xu, D.; Ahmed, M.; Li, Y. Cellular Traffic Offloading via Link Prediction in Opportunistic Networks. *IEEE Access* **2019**, *7*, 39244–39252. [CrossRef]

25. Du, X.; Lin, F. Improving sensor network performance by deploying mobile sensors. In Proceedings of the PCCC 2005 24th IEEE International Performance, Computing, and Communications Conference, Phoenix, Arizona, 7–9 April 2005; pp. 67–71.

26. Wang, N.; Wu, J. Opportunistic WiFi Offloading in a Vehicular Environment: Waiting Or Downloading Now? In Proceedings of the IEEE INFOCOM 2016—The 35th Annual IEEE International Conference on Computer Communications (INFOCOM), San Francisco, CA, USA, 10–14 April 2016; pp. 1–9.

27. Wang, N. Optimal Cellular Traffic Offloading Through Opportunistic Mobile Networks by Data Partitioning. In Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas, MO, USA, 20–24 May 2018; pp. 1–6.

28. Wang, N.; Wu, J. Optimal Data Partitioning and Forwarding in Opportunistic Mobile Networks. In Proceedings of the 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 15–18 April 2018; pp. 1–6.

29. Du, X.; Wu, D. Adaptive cell relay routing protocol for mobile ad hoc networks. *IEEE Trans. Veh. Technol.* **2006**, *55*, 278–285. [CrossRef]

30. Miliotis, V.; Alonso, L.; Verikoukis, C. Weighted proportional fairness and pricing based resource allocation for uplink offloading using IP flow mobility. *Ad Hoc Netw.* **2016**, *49*, 17–28. [CrossRef]

31. He, X.; Liu, M.; Yang, G. Spatiotemporal opportunistic transmission for mobile crowd sensing networks. *Pers Ubiquit Comput*. **2020**, *4*, 1–11. [CrossRef]

32. Jung, Y.; Baek, Y. Multi-hop data forwarding method for crowd sensing networks. *Peer-to-Peer Netw. Appl.* **2016**, *9*, 628–639. [CrossRef]

33. Komnios, I.; Tsapeli, F.; Gorinsky, S. Cost-effective multi-mode offloading with peer-assisted communications. *Ad Hoc Netw.* **2015**, *25*, 370–382. [CrossRef]

34. Wang, L.; Zhang, D.; Xiong, H.; Gibson, J.P.; Chen, C.; Xie, B. ecoSense: Minimize participants' total 3G data cost in mobile crowdsensing using opportunistic relays. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *47*, 965–978. [CrossRef]

35. Gong, W.; Huang, X.; Huang, G.; Zhang, B.; Li, C. Data Offloading for Mobile Crowdsensing in Opportunistic Social Networks. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; pp. 1–6.

36. Wang, E.; Yang, Y.; Wu, J.; Liu, W.; Wang, X. An efficient prediction-based user recruitment for mobile crowdsensing. *IEEE Trans. Mob. Comput.* **2017**, *17*, 16–28. [CrossRef]

37. Yang, C.; Stoleru, R. CEO: Cost-Aware Energy Efficient Mobile Data Offloading via Opportunistic Communication. In Proceedings of the 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 17–20 February 2020; pp. 548–554.

38. Cheung, M.H.; Huang, J. DAWN: Delay-aware Wi-Fi offloading and network selection. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 1214–1223. [CrossRef]

39. Kim, S.; Chon, Y.; Lee, S.; Cha, H. Prediction-based personalized offloading of cellular traffic through WiFi networks. In Proceedings of the 2016 IEEE international conference on pervasive computing and communications (PerCom), Sydney, Australia, 14–16 March 2016; pp. 1–9.

40. Zhou, X.; Ye, T.; Lee, T.T. Designing and Analysis of A Wi-Fi Data Offloading Strategy Catering for the Preference of Mobile Users. *arXiv* **2020**, arXiv:2005.13195.

41. Phe-Neau, T.; De Amorim, M.D.; Conan, V. The Strength of Vicinity Annexation in Opportunistic Networking. In Proceedings of the 2013 Proceedings IEEE INFOCOM, Turin, Italy, 14–19 April 2013; pp. 3369–3374.

42. Cheng, N.; Lu, N.; Zhang, N.; Yang, T.; Shen, X.; Mark, J.W. Vehicle-assisted device-to-device data delivery for smart grid. *IEEE Trans. Veh. Technol.* **2015**, *65*, 2325–2340. [CrossRef]

43. Wang, H.; Zeng, S.; Li, Y.; Jin, D. Predictability and prediction of human mobility based on application-collected location data. *IEEE Trans. Mob. Comput.* **2020**. [CrossRef]

44. Shi, H.; Li, Y.; Cao, H.; Zhou, X.; Zhang, C.; Kostakos, V. Semantics-aware hidden Markov model for human mobility. *IEEE Trans. Knowl. Data Eng.* **2019**. [CrossRef]

45. Tu, L.; Wang, S.; Zhang, D.; Zhang, F.; He, T. ViFi-MobiScanner: Observe Human Mobility via Vehicular Internet Service. *IEEE Trans. Intell. Transp. Syst.* **2019**. [CrossRef]

46. Masdari, M. Markov chain-based evaluation of the certificate status validations in hybrid MANETs. *J. Netw. Comput. Appl.* **2017**, *80*, 79–89. [CrossRef]

47. Keränen, A.; Ott, J.; Kärkkäinen, T. The ONE simulator for DTN protocol evaluation. In Proceedings of the 2nd International Conference on Simulation Tools and Techniques, Rome, Italy, 2–6 March 2009; pp. 1–10.

48. Gao, G.; Xiao, M.; Wu, J.; Han, K.; Huang, L. Deadline-sensitive mobile data offloading via opportunistic communications. In Proceedings of the 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), London, UK, 27–30 June 2016; pp. 1–9.

49. Dimatteo, S.; Hui, P.; Han, B.; Li, V.O. Cellular traffic offloading through WiFi networks. In Proceedings of the 2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems, Valencia, Spain, 17–22 October 2011; pp. 192–201.

50. Zhang, B.; Wang, L.; Han, Z. Contracts for joint downlink and uplink traffic offloading with asymmetric information. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 723–735. [CrossRef]

51. Xu, D.; Li, Y.; Xia, T.; Li, J.; Tarkoma, S.; Hui, P. Portfolio Optimization in Traffic Offloading: Concept, Model, and Algorithms. *IEEE Trans. Mob. Comput.* **2019**. [CrossRef]

52. Ning, Z.; Huang, J.; Wang, X.; Rodrigues, J.J.; Guo, L. Mobile edge computing-enabled Internet of vehicles: Toward energy-efficient scheduling. *IEEE Netw.* **2019**, *33*, 198–205. [CrossRef]

53. Zheng, J.; Gao, L.; Wang, H.; Li, X.; Xu, P.; Wang, L.; Jiang, B.; Yang, X. Joint downlink and uplink edge computing offloading in ultra-dense HetNets. *Mob. Netw. Appl.* **2019**, *24*, 1452–1460. [CrossRef]

54. Aloqaily, M.; Al Ridhawi, I.; Salameh, H.B.; Jararweh, Y. Data and service management in densely crowded environments: Challenges, opportunities, and recent developments. *IEEE Commun. Mag.* **2019**, *57*, 81–87. [CrossRef]

55. Berdik, D.; Otoum, S.; Schmidt, N.; Porter, D.; Jararweh, Y. A Survey on Blockchain for Information Systems Management and Security. *Inform. Proc. Manag.* **2021**, *58*, 102397. [CrossRef]

56. Wang, X.; Han, Y.; Wang, C.; Zhao, Q.; Chen, X.; Chen, M. In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Netw.* **2019**, *33*, 156–165. [CrossRef]

57. Sun, W.; Liu, J.; Yue, Y. AI-enhanced offloading in edge computing: When machine learning meets industrial IoT. *IEEE Netw.* **2019**, *33*, 68–74. [CrossRef]