The An Binh Nguyen, Pratyush Agnihotri, Christian Meurisch, Manisha Luthra, Rahul Dwarakanath, Jeremias Blendin, Doreen Böhnstedt, Michael Zink, Ralf Steinmetz. *Efficient Crowd Sensing Task Distribution Through Context-aware NDN-based Geocast*. In Proc. 42nd Conference on Local Computer Networks (LCN), p.52 - 60, IEEE 2017. ISBN 978-1-5090-6523-3

Efficient Crowd Sensing Task Distribution Through Context-aware NDN-based Geocast

The An Binh Nguyen*, Pratyush Agnihotri*, Christian Meurisch*, Manisha Luthra*, Rahul Dwarakanath*,
Jeremias Blendin*, Doreen Böhnstedt*, Michael Zink†, Ralf Steinmetz*
*TU Darmstadt, Germany
Emails: {firstname.lastname}@kom.tu-darmstadt.de, pratyushagnihotri03@gmail.com, meurisch@tk.tu-darmstadt.de
†University of Massachusetts, Amherst, USA
Email: zink@ecs.umass.edu

***Abstract*—Crowd sensing exploits users' smart devices and human mobility to collect information on a large scale. To realize a crowd sensing campaign, sensing tasks with spatio-temporal requirements are distributed to the devices that can provide the requested information. Typically, the distribution of sensing tasks relies on a centralized communication infrastructure such as cloud servers. However, such an approach will be unsuitable if access to communication infrastructure is restricted, for example in disaster relief scenarios. To fill this gap, we propose a distributed context-aware framework for disseminating sensing tasks, based on the Named Data Networking (NDN) paradigm. By adding context attributes to Interest packets, we allow a device to utilize this information to make forwarding decisions autonomously, thus guiding the sensing tasks towards the suitable sensing devices. Through intensive evaluation, we show that our framework achieves a timely delivery of sensing tasks, while keeping the communication overhead to a minimum compared to pure geo forwarding and flooding approaches.**

***Index Terms*—Crowd Sensing, Opportunistic Forwarding, Named Data Networking.**

I. INTRODUCTION

Crowd sensing is a sensing paradigm that utilizes mobile devices to collect information [1]. Comparing to the Wireless Sensor Network (WSN) paradigm, crowd sensing allows for more flexibility and enables monitoring of large scale phenomena through the mobility of the devices' owners. One central problem of crowd sensing is to find and to assign the sensing tasks to appropriate devices satisfying the requirements of a sensing campaign [2]. To solve this problem, most of the research work relies on a centralized platform to track potential participants [3]; using this, the task allocation can then be optimized. Nevertheless, the centralized approach will not be suitable in case of impaired communication infrastructure or platform failures, for example in disaster relief scenarios, in which not all devices can be connected to the platform. Communication in disaster relief scenario is often based on opportunistic ad-hoc network techniques for disseminating information. Crowd sensing research, like the work by Zhao et al. [4], is also based on opportunistic ad-hoc networks to (re-)distribute sensing tasks among devices directly, aiming at solving the coverage problem. Still, the coverage problem is quite specific and focuses only on finding a sufficient number of devices, thus assuming homogeneity. A distributed framework to disseminate sensing tasks without relying on a centralized platform is still missing.

Recently, Named Data Networking (NDN) [5] has been proposed and recognized as a promising networking architecture for information retrieval solely based on content name. The usage and performance of NDN on wireless ad-hoc networks have been investigated and confirmed by several works [6], [7]. Furthermore, Moreira et al. [8] also discuss and show the suitability of the NDN approach for crowd sensing in a decentralized opportunistic network. A device, called consumer, interested in receiving particular information can initiate and propagate an Interest packet through the network; and a device, called producer possessing the information matched with the request from the received Interest can reply with Data packets. Special characteristics of NDN such as in-network caching, flexibility in naming content, and provider agnostic requests are desired features for crowd sensing. However, the mobility of producers still remains an unsolved challenge in NDN [9]. This problem is equivalent to the problem of disseminating the sensing task to appropriate mobile devices in crowd sensing, by forwarding tasks through opportunistic contacts of devices. Based on the NDN paradigm, our goal is to design a framework to ensure decentralized dissemination of sensing tasks to suitable producers, considering mobility and spatio-temporal requirements of the tasks.

The main contributions of this paper are threefold:

- We design a system model and a corresponding NDN-based framework for distributed dissemination of sensing tasks, targeting opportunistic ad-hoc networks.
- We integrate context information, such as distance attributes and packet-sent counters, into Interest packets which allows individual devices to autonomously make forwarding decisions.
- We design a context-aware forwarding concept to guide the sensing tasks towards the requested location and to search for capable producers. Our forwarding concept is inspired by virtual gradient fields for self-organizing network patterns.

The rest of this paper is organized as follows. In Section II, we provide an overview of the related work. We discuss an application scenario and the system model in Section III.

Our context-aware NDN forwarding concept is described in Section IV. Section V presents the results of the evaluation. Conclusion and several possible research directions for future work are presented in Section VI.

II. RELATED WORK

In this section, we review related work from two relevant research disciplines—opportunistic networks and NDN wireless ad-hoc networks.

A. *Opportunistic Networks*

With the proliferation of mobile devices and location based services, research on methods to disseminate information to geographical area—geocast for opportunistic networks has gained attention. GSAF [10] uses a set of coordinates to define a cast. Messages contained by nodes moving far away from cast region will be discarded if the buffer of such nodes are full or when the message life time expires. Tuncay et al. [11] use profile-cast to recruit mobile nodes with a target profile for opportunistic sensing. Sensing tasks are distributed to nodes, that satisfy part of the profile, i.e, nodes do not necessarily have visited all requested locations. Geoopp [12] leverages the regularity in mobility pattern of humans to choose the nodes with best chance to move closer to the destination. These are nodes with their predicted future locations near the intended region. Geospray [13] is designed for vehicular delay-tolerant networks. Geospray utilizes both multiple-copies and single-copy opportunistic forwarding. It first injects several copies of the message in order to spread the message over several paths; these messages will then be forwarded with only one copy towards the destination. Another scheme called approach and roam is proposed by Cao et al. [14]. Based on information of historical locations, this scheme estimates the ranges of mobile nodes and replicates messages only within this range. However compare to our approach, none of these aforementioned works deals directly with the unavailability of capable mobile nodes at requested geographical locations.

B. *Named Data Networks*

Recently, NDN has been studied on multihop wireless ad hoc networks, such as MANET, VANET. For information retrieval in such contexts, Interest packets are broadcast, which results in *broadcast storm* problem. Amadeo et al. [7], [15] propose a set of defer timers to minimize congestion. Geographical forwarding is not considered in these works. Deng et al. [16] consider two types of Interest, i.e., location-dependent and location-independent for forwarding in VANET. Nevertheless, the authors only consider a static producer in case of location-dependent forwarding, such as gas stations; while in our work, we specifically target mobile producers. NDN-Q [17] is an NDN based query dissemination for vehicular networks, that exploits static Road Size Units to disseminate Interest to cars. Even though, NDN-Q architecture is comparable to our system model, NDN-Q relies only on pure broadcast. On the contrary, our forwarding approach is designed for geographical forwarding and is able reduce the overhead generated by pure broadcast remarkably. Similar to our approach, Kuai et al. [18] also aim to utilize opportunistic forwarding techniques in NDN for VANET. Their approach relies on a special beacon broadcast by neighbors to make forwarding decision, whereas in our approach we leverage only context attributes extracted from Interest packets. CODIE [19] focuses on data delivery for mobile consumer in vehicular NDN. The authors include hop count in Interest packet and a *data dissemination limit* in Data packet, that allows to control flooding of Data packet. Since our main target is to ensure the forwarding of Interest packets as request to mobile producers and the main target of CODIE is to deliver Data packets as reply to mobile consumers, CODIE is complementary to our work. Rehman et al. [20] include distance information from Consumer to an Interest packet and check on an energy threshold on each node to decide on forwarding. In comparison, our work provides a more holistic view by utilizing both local and distributed context attributes for forwarding decision.

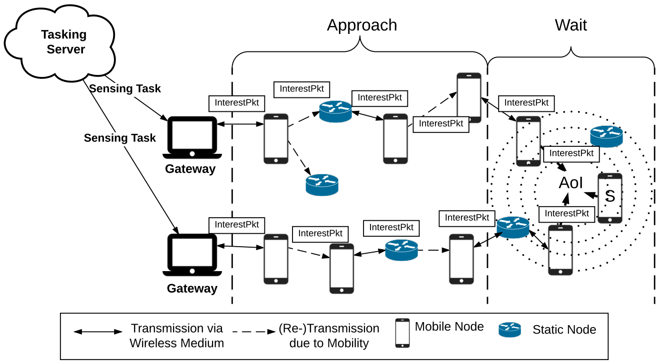


Fig. 1: NDN-based System Model and Two Phases Forwarding Approach

III. APPLICATION SCENARIO AND SYSTEM MODEL

In this section, we first discuss the application for crowd sensing using a disaster relief scenario; this application scenario is representative for the challenges and the requirements of a distributed crowd sensing framework. Second, we elaborate on the assumptions we make for designing NDN-based crowd sensing, by targeting decentralized ad-hoc networks. Third, we introduce the system model designed according to the assumptions and the application scenario.

A. *Application Scenario*

A disaster relief scenario is a typical example for crowd sensing application. Sensing tasks, such as heat map generation, are distributed to participants to collect situational information. The situational information is useful for planning relief operations, such as to identify and to verify hot spots for planning and delegating relief workers, efficiently. Still, access to communication infrastructure might be restricted in a disaster situation. Consequently, the distribution of sensing tasks cannot be completed centrally. Devices in isolated areas can still exchange information through opportunistic ad-hoc networks. Moreover, based on *store, carry and forward* via opportunistic contacts, mobile devices can disseminate information to distant areas even though an end-to-end connection might not exist. As shown in [21], several devices belonging to relief workers may still have a connection to a head-quarter and can act as a gateway to inject sensing tasks into an isolated area. Thus, the first objective is to exploit devices' mobility to bring the sensing tasks injected through the gateway towards the *area of interest* (henceforth, AoI). Furthermore, due to heterogeneity and mobility of the devices, the availability of a capable device for a sensing task cannot be ensured. As a consequence, the second objective is to search for the capable devices that can perform the sensing tasks. In doing so, the requirements of the sensing tasks need to be satisfied.

B. Assumptions

Based on the discussion of the above application scenario, we make the following assumptions:

- Decentralized ad-hoc networks: In this paper, we focus on the distribution of the sensing tasks through a decentralized ad-hoc network. The tasks can be injected into the ad-hoc network through gateway devices. We assume, that the sensing tasks are not injected to the area of interest directly. Furthermore, we assume that the gateway devices are static, since the problem of mobile consumers can be addressed by reissuing the same request [9].
- Heterogeneity: Due to various devices, there are different types of sensors available on participating nodes. Not all participating devices in the network are capable of performing the requested sensing tasks. We assume that only a subset of nodes act as information producers, which can provide Data to the requested Interest.
- Collaboration: Selfish behaviour of participating devices is beyond the scope of this paper. We assume that the nodes that fulfill the requirements of the sensing tasks are willing to participate in forwarding and to provide information.
- Location-aware: We assume that each device can determine its current location, e.g., through GPS.

C. System Model

The overview of our proposed system model is illustrated in Figure 1. Our system model consists of a tasking server, gateway devices, and NDN based ad-hoc networks. The tasking server first distributes the sensing tasks with specific requirements to gateway devices nearest to the AoI. The gateway devices have two functions—they act as client to receive tasks from the tasking server and as NDN information consumers to initiate Interest requests. Due to this reason, these devices take care of transforming the sensing task into an Interest packet and propagating this Interest packet into NDN based ad-hoc networks. The forwarder devices of the NDN ad-hoc networks can be either static or mobile, such that the mobility of the devices can be exploited to bring forward the Interest packet nearer to the AoI. Wireless Adhoc WiFi is used for communication between devices. In the illustration, we show two phases for the system model, i.e., the approach phase and the wait phase. The approach phase corresponds to the objective of forwarding Interest packets closer to the requested AoI. The wait phase is designed to combat the mobility problem of NDN producers. The idea of the wait phase is to *bind* Interest packets geographically, waiting for capable producers that can provide the requested information.

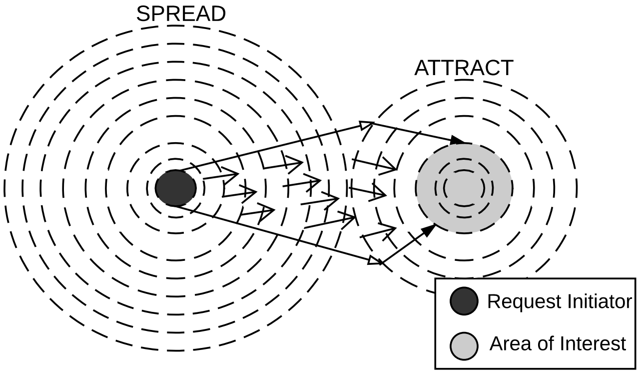


Fig. 2: Combination of Spread and Attract Gradient Fields for Sensing Tasks Dissemination

IV. NDN BASED CROWD SENSING

In this section, we give a detailed overview of our tasking framework for crowd sensing. First, we introduce the structure of the application naming scheme and the Interest packets. Second, we describe our approach for context-aware Interest forwarding. Third, we elaborate on the overall workflow to retrieve information by our framework.

A. Structure of Application Naming Scheme and Interest Packet

As introduced in Section III, our system model relies on gateways to transform the sensing tasks with their specific requirements into a NDN Interest packet. Wang et al. [2] summarize and discuss different dimensions for crowd sensing task's requirements. According to them, a sensing task is characterized by 5 main dimensions—*what to measure, where to measure, when to measure, who to measure, and how to measure*. Based on this observation, we decide on the naming scheme for our NDN based crowd sensing tasking framework as follows: */CrowdSensing/<geographical information>/<sensor type>/<time>*. <*Geographical information*> contains information about the AoI, (i.e., longitude, latitude and its corresponding radius R_{AoI}), in which the sensing tasks have to be performed. We use the representation as introduced by Pesavento et al. [22] to encode the <*geographical information*> in our naming scheme. The authors propose to use Cantor pairing function [23] to encode a pair of coordinates (longitude, latitude) into a sequence of natural numbers. Adapting this representation into our naming scheme allows a sensing task to be matched efficiently against in-network cached data with regard to the geographical requirement. <*Sensor type*> indicates which type of information is being requested. Requested

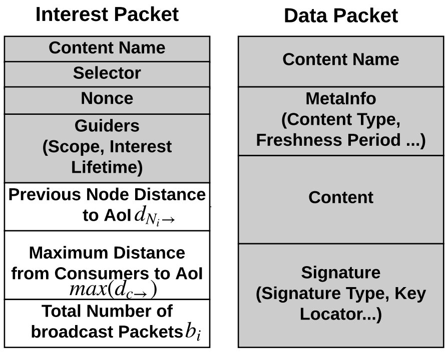


Fig. 3: Default Data packet and modified Interest packet with 3 additional fields, illustrated in white (default fields are adapted from [5])

information might require the availability of a special sensor type on the sensing devices. <*Time*> indicates time related requirements, e.g., when to perform sensing (scheduling) or how often to perform sensing (frequency). The proposed naming scheme covers all dimensions required for crowd sensing and allows the NDN capable nodes to act according to the requirements.

In general, an NDN network facilitates information retrieval by propagating an Interest packet through Faces defined in the Forwarding Information Base (FIB) table [5]. However, in decentralized opportunistic ad-hoc networks, the topology of the network is unstable due to the mobility of nodes. An end-to-end path among nodes may not exist. Therefore, propagating Interest packets based on the FIB table is not possible for this type of network. Consequently, each Interest packet has to be (re-)broadcast [18]. This results in an uncoordinated forwarding at each intermediate node. To alleviate this problem, we include context attributes of the current node into each Interest packet before broadcasting. Upon receiving an Interest packet, each node extracts the embedded context information, which is used to adapt current forwarding behavior. For our forwarding concept, we include distance of the current node to the requested AoI, maximum distance to the requested AoI from the gateway devices (as static consumers), and the number of Interest packets broadcast by the current node (cf. Figure 3). These three attributes will be used together with local attributes of the node (i.e., residual energy, speed and moving direction) to make forwarding decisions autonomously. Details on how these attributes are used, will be discussed in the next subsection.

B. Context-aware two-phases Forwarding

Our forwarding concept is inspired by self-organizing design patterns for distributed coordination among autonomous agents [24]. The basic idea is to utilize a virtual gradient field that will guide the Interest packet to capable mobile producers within the AoI. The gradient field used in this paper is illustrated in Figure 2. We combine two gradient fields—a *spread* field which aims to spread and push Interest packet first far away from the request initiator; and an *attract* field which aims to attract the Interest that comes nearer to the AoI. Such combination results in a direction flow towards the AoI that the Interest packets will follow. Accordingly, our forwarding approach comprises two phases, i.e., the *approach phase* and the *wait phase*. The *spread* field is applied for the *approach phase*, while the *attract* field is applied for the *wait phase*.

Algorithm 1 Interest Processing at Intermediate Node N_i during Approach Phase with Buffer Zone Radius R_{BZ}

```
Receive InterestPkt
$(d_{N_p\rightarrow}, d_{c\rightarrow}, b_p) \leftarrow readContext(InterestPkt)$
$d_{\overrightarrow{N_i}\rightarrow} \leftarrow d(N_i, AoI)$
$\overrightarrow{md} \leftarrow N_i's$ current moving direction
$s_i \leftarrow N_i's$ current speed
$e_i \leftarrow N_i's$ current energy
$b_i \leftarrow N_i's$ total broadcast packets
$d_{max} \leftarrow max(d_{c\rightarrow})$
if (matched Data found in Content Store) then
    $DT_{data} \leftarrow Time_{DeferSlot} * (\frac{d_{N_i\rightarrow}}{d_{max}} + T_{Random})$
    Schedule to broadcast Data after $DT_{data}$
else
    if $(d_{N_i\rightarrow} < d_{N_p\rightarrow})$ and $(\angle(\overrightarrow{md}, \overrightarrow{N_i}) < \angle_{threshold})$ and $(e_i > e_{threshold})$ and $(s_i = 0$ or $s_i > s_{threshold})$ then
        $isForwarder \leftarrow TRUE$
    end if
    if (¬ isForwarder) then
        Drop Interest
    end if
    if ( $d_{N_i\rightarrow} > R_{BZ}$) then
        if (Interest ← Find(PIT)) then
            Discard Interest
            Increase Interest's lifetime and update PIT
        else
            Add $(d_{N_i\rightarrow}, d_{max}, b_i)$ to Interest
            Insert Interest to PIT
            $DT_{Int} \leftarrow Time_{DeferSlot} * (T_d + T_e + T_s + T_{md} + T_{Random})$ {cf. Equation 2 for details}
            if $b_i < median(b_p)$ then
                $DT_{Int} \leftarrow DT_{Int} * \frac{b_i}{median(b_p)}$
            end if
            Schedule to broadcast Interest after $DT_{Int}$
        end if
    else
        Process Interest according to Wait Phase
    end if
end if
```

The wait phase is characterized by a buffer zone with the requested AoI being the center and a buffer zone radius R_{BZ}, where $R_{BZ} \geq R_{AoI}$ since we want to increase the chance of reaching a mobile producer within the AoI. (In Figure 2, the buffer zone corresponds to the outermost circle of the *Attract* field). Respectively, a node located outside the buffer zone is in *approach phase*. As soon as a node enters the buffer zone, its phase will be switched to the *wait phase*. Finding the optimal size of R_{BZ} for our forwarding concept proves to

Algorithm 2 Interest Processing at Intermediate Node N_i during Wait Phase with Buffer Zone Radius R_{BZ}

```
Receive InterestPkt
$(d_{N_p\rightarrow}, d_{max}, b_p) \leftarrow readContext(InterestPkt)$
$d_{N_i\rightarrow} \leftarrow d(N_i, AoI)$
$b_i \leftarrow N_i^{\prime}s$ total broadcast packets
if ($d_{N_i\rightarrow} < R_{BZ}$)) then
  if (matched Data found in Content Store) then
    $DT_{data} \leftarrow Time_{DeferSlot} * (\frac{d_{N_i\rightarrow}}{d_{max}} + T_{Random})$
    Schedule to broadcast Data after $DT_{data}$
  else
    if (Interest ← Find(PIT)) then
      Discard Interest
      Increase Interest lifetime and update PIT
    else
      if ($d_{N_i\rightarrow} < R_{AoI}$) then
        $d_{max} \leftarrow R_{AoI}$
      else
        $d_{max} \leftarrow R_{BZ}$
      end if
      Add $(d_{max}, b_i)$ to Interest
      Insert Interest to PIT
      $DT_{Int} \leftarrow Time_{DeferSlot} * (T_d + T_e + T_s + T_{md} + T_{Random})$ {cf. Equation 2 for details}
      if $b_i < median(b_p)$ then
        $DT_{Int} \leftarrow DT_{Int} * \frac{b_i}{median(b_p)}$
      end if
      $n_{REP} \leftarrow n_{max} * \frac{R_{BZ} - d_{N_i\rightarrow}}{R_{BZ}}$
      for i ← 1 to $n_{REP}$ do
        Schedule to broadcast Interest after $DT_{Int}$
      end for
    end if
  end if
end if
```

be challenging, since the topology of the network may change constantly and thus is not known to either tasking server or gateway devices, beforehand. In the evaluation, we empirically assess the effect of the buffer zone radius to study the trade-off between performance and overhead for our concept.

The sensing task in form of an Interest packet is first initiated by the gateway devices acting as NDN consumers. In each Interest packet, the consumer includes its own distance to the AoI ($d_{c\rightarrow}$) and starts broadcasting this Interest packet. Every node receiving Interest packets will check which phases they are currently in, by comparing their distance to the AoI ($d_{N_i\rightarrow}$) with R_{BZ}, and execute the forwarding algorithms accordingly. The wait phase is activated only when $d_{N_i\rightarrow} < R_{BZ}$, the approach phase is activated otherwise. The pseudo code for the Interest packet processing in each phase is shown in Algorithm 1 and 2. In all phases, if a matched data can be found from the Content Store (CS) of an intermediate node, these data will be propagated directly back to the consumers.

During the approach phase, before rebroadcasting an Interest packet, each intermediate node determines its current distance to the requested AoI and the current number of its total broadcast Interest packets. These attributes are included in the new Interest packet before broadcasting (cf. Figure 3).

Upon receiving an Interest packet, a node extracts the distance of the previous node to the AoI from the Interest packet, then compares with its current distance to the AoI. Each node only rebroadcasts an Interest packet if all of the following conditions are met: 1) The current distance to the AoI is less than the distance of the previous node extracted from the Interest packet; 2) The node is either static or is moving towards the AoI. Choosing a static node binds an Interest packet at a fixed location, which can be later forwarded to mobile nodes with higher chance to reach the AoI, i.e., nodes that move nearer to the AoI. Moving direction of mobile devices can be determined using built-in sensors such as accelerometer. In addition to the current distance to the AoI and moving direction, the residual energy and moving speed of the nodes are also accounted for. An energy threshold is set to make sure that the forwarders still have enough energy to reach the destination, and a speed threshold is set to ascertain that the Interest is forwarded in the fastest way. Altogether, these conditions ensure that the best forwarders are chosen.

To schedule the propagation of Interest packet at the forwarder node N_i, we use the timer DT_{Int} adapted from [7] as follows:

$$DT_{Int} = Time_{DeferSlot} * (T_d + T_e + T_s + T_{md} + T_{Random}) \quad (1)$$

$$\begin{aligned} T_d &= w_d * \frac{d_{max} - d_{N_i\rightarrow}}{d_{max}}, T_e = w_e * \frac{e_{max} - e_{N_i}}{e_{max}} \\ T_s &= w_s * \frac{s_{max} - s_{N_i}}{s_{max}}, T_{md} = w_{md} * \frac{\Psi - \Theta}{\Psi} \end{aligned} \quad (2)$$

In the above equations, $Time_{DeferSlot}$ is a fixed maximum possible value for the defer time. w_d, w_e, w_{md}, w_s are weighting values, indicating the importance of each factor. T_d is the distance-related factor, calculated using the current distance from the forwarder node to the AoI ($d_{N_i\rightarrow}$) and the maximum distance towards the AoI among the consumers, i.e., $d_{max} = max(d_{c\rightarrow})$, which can be extracted from the Interest packet. T_e, T_s, T_{md} are the corresponding time factors, indicating the dependence of defer time on current energy, speed, and moving direction of a forwarder with threshold values. In T_{md}, Θ is the angle between moving direction vector $\overrightarrow{md}$ of forwarder N_i and the vector $\overrightarrow{N_i}$, that represents the straight direction from the current location of N_i towards the center of the AoI. Determining the defer time in such way ensures that, a forwarder will schedule to propagate an Interest packet faster, if it is located nearer to the AoI, moves faster and more straight to the AoI, while possesses sufficient energy. This is designed to reduce the time needed to forward Interest packets to the AoI. Furthermore, we also add T_{Random} to defer time DT to avoid congestion.

When a forwarder N_i enters the wait phase, i.e., $d_{N_i \rightarrow} < R_{BZ}$, the component d_{max} in the T_d factor is replaced with the buffer zone radius R_{BZ}. As such, the defer time will be shorter if a forwarder reaches the buffer zone, getting closer to the AoI. A shorter defer times near the AoI increases the chance of an Interest packet to be disseminated to a mobile producer upon its appearance. Within the AoI, d_{max} is replaced with R_{AoI}. Additionally, an Interest packet is rebroadcast n_{REP} times, when a forwarder is within the buffer zone (i.e., a forwarder in the wait phase), where $n_{REP} = n_{max} * \frac{R_{BZ} - d_{N_i \rightarrow}}{R_{BZ}}$ and n_{max} is a maximum replication number. This replication number can be chosen, e.g., based on the available resources at each node. Obviously, rebroadcasting multiple times further increases the chance to find a capable mobile producer within the AoI

In decentralized ad-hoc networks that consist of mobile devices, especially in scenarios like disaster relief, the conservation of energy and reduction of overhead are important. The life time of such a network relies on the life time of individual nodes, which is influenced by the energy consumed for processing tasks such as forwarding, sensing, etc. Thus, another target of our forwarding approach is to increase fairness in forwarding contribution among nodes. This makes sure that no single node is overused in forwarding. To achieve this, we include the total number of broadcast packets b_i by the current node N_i into the Interest packet before broadcasting. Each node overhearing Interest packets from its neighbours can extract the value of the field *total number of broadcasts packets* from these Interests. Accordingly, the current node determines the median of these values (m_b). The median value is used to minimize the affect of possible outliers, e.g., when having a fresh node joined the network that has not participated in forwarding and a node that has contributed considerably in forwarding as neighbours at the same time. Having its own number of broadcast packets less than other observed numbers ($b_i < m_b$) indicates, that a node is still underutilized for forwarding. In this case, such node adapts its current defer time (DT) to a shorter defer time $DT * \frac{b_i}{m_b}$.

C. Overall Workflow to retrieve Information

Having elaborated on the system model (cf. Section III), the naming scheme (cf. Subsection IV-A) and the Interest forwarding approach (cf. Subsection IV-B), in this subsection, we consolidate the components of our framework and describe the overall workflow to retrieve information. Upon receiving a sensing task from the tasking server, the gateway devices as information consumers initialize an Interest request with its corresponding requirements according to our proposed naming scheme. When an intermediate node receives an Interest packet, the geographical requirement for this Interest is checked against the cached Data in the Content Store of this node. If the Interest and the cached Data are matched, the intermediate node will schedule to propagate the Data directly back to the consumers, without further forwarding towards the AoI. The propagation of Data packet at node N_i is scheduled with the timer DT_{data} as follows:

$$DT_{data} = Time_{DeferSlot} * (\frac{d_{N_i \rightarrow}}{d_{max}} + T_{Random}) \tag{3}$$

where $d_{N_i \rightarrow}$ denotes the current distance between node N_i and the AoI, and d_{max} is the maximum distance between the consumers and the AoI. As a result, the proposed defer time for propagating Data packet is shorter for nodes that are nearer to the AoI. This design is due to the fact that, Data cached nearer to the AoI tend to be more relevant w.r.t. time requirement. Also, for intermediate nodes located farther from the AoI, but nearer to the consumers, a longer defer time increases the chance for aged Data in the Content Store to be replaced by more current Data, e.g., when such Data come back from the information producers. Additionally, we set the priority for propagating Data packets higher than propagating Interest packets on every nodes. This ensures that, Data will always be broadcast before Interest. Thus, the requested Data can reach the consumers as fast as possible. If no matched Data can be found in the Content Store, a node will proceed to rebroadcast the Interest according to our context aware two-phases forwarding approach as discussed in Subsection IV-B. Finally, when the Interest reaches the capable information producers within the AoI, the information producers will execute the sensing task according to the specified requirements (type of sensors, sensing frequency) and propagate the Data packets containing the requested information backwards to the consumers.

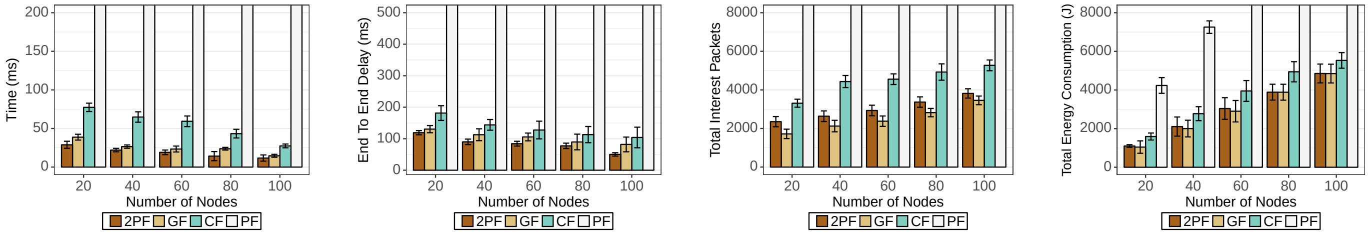
V. EVALUATION

In this section, we present the results obtained from the simulative evaluations. First, we elaborate on the methodology, the evaluation setup, and the performance metrics. Next, we discuss the results with regard to performance of the forwarding approach. We study the effect of buffer zone radius introduced in our two phases forwarding approach. Last, we investigate on fairness among forwarders.

A. Methodology and Evaluation Metrics

We implemented our proposed two phases forwarding approach using NDNSim [25]. Furthermore, we included three other forwarding mechanisms to compare against our approach. These are pure *geo forwarding, controlled forwarding, and pure flooding*. We implemented the pure geo forwarding mechanism by deactivating the buffer zone in the two phase forwarding approach. By doing this, the Interest packets are always forwarded by the nodes with shorter distance to the AoI. Controlled flooding is based on the random defer time approach proposed by Amadeo et al. [7]. Pure flooding uses the default flooding in the forwarding strategy implemented in NDNSim.

In our scenario, the nodes can communicate with each other using the 802.11 WiFi model. To simulate a NDN scenario for wireless network environments, in which multi-hop transmission of Interest packet is possible, we used the patch for NDNSim developed by Amadeo et al. [7]. The simulation area size is $800 \times 800\,m^2$, where the AoI radius



(a) Time to Find Producer

(b) Time to Deliver Data

(c) Message Overhead

(d) Energy Consumption

Fig. 4: Performance and Overhead Comparison among Interest Forwarding Approaches.
2PF denotes our context-aware two-phases forwarding; GF denotes geo forwarding, CD denotes controlled flooding, and PF denotes pure flooding.

is $50\,m$. We use the `RandomWalk2dMobilityModel` as the mobility model for the nodes. Each node can move between 2 and $5\,m/s$; this value corresponds to the movement speed of a pedestrian. With regard to the energy consumption of the nodes, we use the `WifiRadioEnergyModel` and `BasicEnergySource`. With these two models, the energy on each node is drained both over the time and when the node broadcasts. All the nodes are set up with the NDN stack and thus can perform Interest forwarding according to our approach. We created 12 information consumers and 25 information producers. The consumers are static nodes, while the producers are mobile; due to this setup, the producers in our simulation are available inside the AoI arbitrarily. The most important simulation parameters are summarized in Table I.

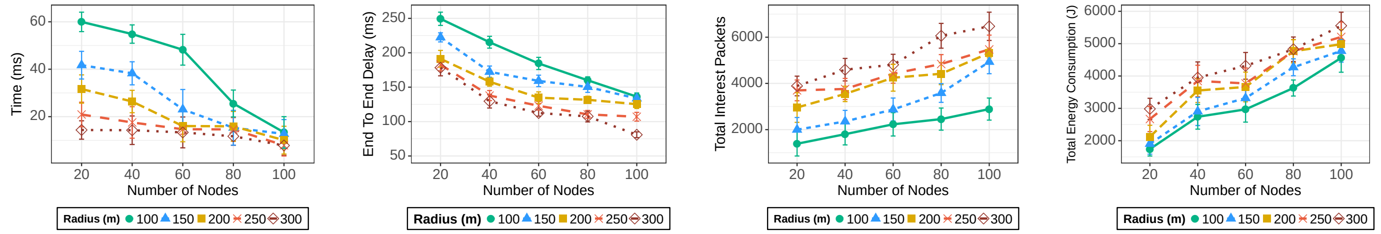
To obtain reliable results for the evaluation, we repeated the simulation 100 times for each configuration setup, which varies in the forwarding approaches, in the number of nodes, and in the radius of the buffer zone. We use four main metrics to measure the performance of our forwarding approach, i.e., time to find a producer, time to deliver data, message overhead, and total energy consumption. Time to find a producer is the metric to measure the time after the first Interest packet is sent out until the first producer is found in the AoI. This metric is the most important performance metric in our scenario, since it determines the efficiency of the forwarding approach. Time to deliver data is defined as the time elapsed between transmission of the first Interest packet and the delivery of the requested information to one of the consumers. Message overhead is defined as the total number of Interest packets that have to be propagated/rebroadcast through the network after sending the first Interest packet until the first producer is found within the AoI. Another overhead metric is the total energy consumption of all nodes. Altogether, observing these four metrics identifies trade-offs between the performance and the overhead caused by different forwarding approaches.

B. Performance Evaluation

Figure 4 shows the performance and overhead comparison of our proposed approach (context-aware two-phases forwarding—2PF in the figures) against other forwarding approaches. The results for time-related metrics, i.e., time to find producer and time to deliver data can be found in Figure 4a and 4b. It can be observed that our proposed approach can forward an Interest to a capable mobile producer faster compared to geo forwading, controlled flooding, and pure flooding. The very long time when using pure flooding concept is expected, since pure flooding does not use any congestion control mechanisms. Accordingly, the congestion rate and Interest drop rate are high. Controlled flooding uses a random defer time to reduce congestion rate, thus, it performs better compared to pure flooding. Nevertheless, due to mobility of producers, an Interest forwarded by controlled flooding can still miss the producers when they are within the requested AoI. Instead, an Interest packet may reach the producers outside the AoI, thus making it unusable. For performance evaluation, we use a buffer zone radius of $150\,m$ in our context-aware forwarding approach. Due to the use of a buffer zone to bind an Interest packet geographically near the AoI, the chance for this Interest packet to reach a capable producer is higher, leading to less time to find a producer. Consequently, buffer zone concept performs better than pure geo forwarding. Furthermore, the results show that the number of nodes can influence the performance. The time to find a producer is shorter with high density of nodes. This trend is common for all forwarding concepts. Nevertheless, our forwarding concept outperforms others despite the variation in node density. The results for overhead metrics, i.e., message overhead and total energy consumption are shown in Figure 4c and 4d. Obviously, the total number of sent Interest packets and energy consumption units increase with the number of nodes. Despite the better performance with regard to time-related metrics, our approach and geo forwarding do not generate high overhead compared to controlled flooding and pure flooding. Having said that, compared to pure geo forwarding, our approach

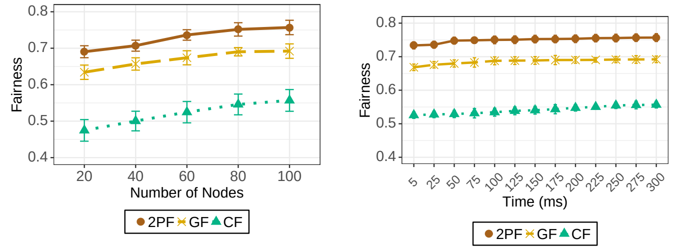
TABLE I: Simulation Parameters

Parameter	**Value**
Simulated Area Size	$800 \times 800\,m^2$
Number of Nodes	$20,\ 40,\ 60,\ 80,\ 100$
AoI Radius	$50\ m$
Transmission Range	$100\ m$
Node Speed	2–5 m/s
Mobility Model	`RandomWalk2dMobilityModel`
Energy Model	`WifiRadioEnergyModel`, `BasicEnergySource`
Buffer Zone Radius	$100\,m$, $\mathbf{150\,m}$, $200\,m$, $250\,m$, $300\,m$



(a) Time to Find Producer (b) Time to Deliver Data (c) Message Overhead (d) Energy Consumption

Fig. 5: Influence of Buffer Zone Radius



(a) Fairness Index w.r.t. Density of Nodes (b) Fairness Index w.r.t. Time

Fig. 6: Fairness Index Comparison among Forwarding Approaches. 2PF denotes our context-aware two-phases forwarding, GF denotes geo forwarding, CF denotes controlled flooding.

generates more overhead. The reason is, in our approach, the Interest propagation rate increases more within the buffer zone. In summary, our approach and geo forwarding perform better than flooding while generate less overhead. In comparison to geo forwarding, our concept has to trade-off overhead, but results to shorter response time, which is the most important metric in our targeted scenario.

C. *Effect of Buffer Zone*

Our forwarding approach relies on two phases, i.e., approach phase and wait phase for disseminating sensing tasks (Interests) to capable producers. The wait phase exploits a buffer zone to increase the propagation rate, aiming to search and wait for mobile producers. The buffer zone size can be set by either the tasking server or the gateway devices. Finding an optimal size of the buffer zone is challenging due to the lack of information. Obviously, the buffer zone radius will affect the performance as well as the overhead of our two-phases forwarding approach. Due to this reason, we study the effect of different buffer zone's radii. The results are shown in Figure 5. Again, we use the 4 evaluation metrics introduced above. Similar to the results of the forwarding performance evaluation, we can also observe a trade-off between performance and overhead. The performance time is faster with larger buffer zone radius, but a larger buffer zone radius also generates more overhead in both number of generated messages and total energy consumption, as shown in Figure 5c and 5d. Interestingly, there is a correlation between the number of nodes and the buffer zone radius. Figure 5a shows that the time to find producer with lower density can benefit from larger buffer zone radius. With 20 nodes, the buffer zone radius of 300 m clearly outperforms the buffer zone radius of 100 m. The performance gap is smaller with the increasing number of nodes. For instance, with 100 nodes, the performance time gap among different buffer zone's radii is marginal in comparison to with 20 nodes. With this information, the size for buffer zone's radius can be adapted with regard to node density.

D. *Fairness*

We conducted the fairness evaluation to investigate whether the forwarding contributions among nodes are fair in our approach. We use Jain's fairness index introduced by Bukh and Jain [26], which is calculated as follows:

$$J(x_1, x_2, \ldots, x_n) = \frac{(\sum_{i=1}^{n} x_i)^2}{n * \sum_{i=1}^{n} x_i^2} \tag{4}$$

Jain's fairness index generates a value between 0 and 1 to quantify the fairness with regard to the resource consumption among n instances $x_1 \ldots x_n$. For our evaluation, x_i is the total number of broadcast Interest packets observed at node N_i. The closer to 1 the index is, the fairer the resources are consumed among participating nodes. We compare the fairness index of our forwarding approach against geo forwarding and controlled flooding. The results are summarized in Figure 6. Figure 6a shows the correlation between the fairness index and the number of nodes, while Figure 6b shows the correlation between fairness index and simulation time. Regardless of node density and simulation time, our context-aware forwarding approach achieves better fairness index than geo forwarding and controlled flooding. Furthermore, the fairness index increases with higher node density; since, higher density makes sharing counters for average number of broadcast packets through the network faster, leading to faster convergence of a fair Interest propagation rate. The fairness index also improves with simulation time. The longer the simulation runs, the more packets broadcast counters can be shared through the network, leading to fairer forwarding contribution. All in all, fairness can be achieved in a distributed manner, using only attributes extracted from an Interest packet.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a distributed framework for information retrieval through crowd sensing based on the Named Data Networking paradigm. Our proposed framework consists of a system model, a naming scheme for crowd sensing, as well as a two-phase context-aware Interest forwarding approach. We implemented and evaluated the proposed framework using the NDNSim simulator. The results obtained through in-depth evaluation showed that our approach copes with the mobility of information producers, successfully disseminates sensing Interests to capable producers in an area of interest with less time and minimal overhead compared to pure geo forwarding and flooding approaches. Furthermore, we demonstrated the successful use of context attribute tags piggybacking on propagated Interest packets, not only for forwarding decisions, but also for achieving fairer resources consumption among nodes.

The sensing tasks as defined in this work contain a query to request for information, which can be obtained by a single type of sensor. We plan to extend our concept so that complex queries, that require multiple type of sensors, can be served. The first step is to combine the concept developed in this paper with complex multi-staged in-network processing technique [27]. This combination allows for distributed processing of sensing data directly in the network. Thus, it is possible to decompose a complex query into several simple sensing tasks and several in-network operations. Moreover, to study our proposed framework in more realistic conditions, we plan to implement our approach on real hardwares, e.g., using NDN forwarding on Android devices[1].

ACKNOWLEDGMENT

This work has been co-funded by the LOEWE initiative (Hessen, Germany) within the NICER project, by the German Federal Ministry of Education and Research (BMBF) Software Campus project "OppEPM" [01IS12054], and by NSF-CAREER.

REFERENCES

[1] H. Ma, D. Zhao, and P. Yuan, "Opportunities in Mobile Crowd Sensing," *IEEE Communications Magazine*, vol. 52, pp. 29–35, 2014.

[2] D. Zhang, L. Wang, H. Xiong, and B. Guo, "4W1H in Mobile Crowd Sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 42–48, 2014.

[3] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick, "A Survey on Privacy in Mobile Participatory Sensing Applications," *Journal of Systems and Software*, vol. 84, no. 11, pp. 1928–1946, 2011.

[4] D. Zhao, H. Ma, and L. Liu, "Energy-efficient opportunistic Coverage for People-centric Urban Sensing," *Wireless networks*, vol. 20, no. 6, pp. 1461–1476, 2014.

[5] L. Zhang, A. Afanasyev, J. Burke, V. Jacobson, P. Crowley, C. Papadopoulos, L. Wang, B. Zhang *et al.*, "Named Data Networking," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66–73, 2014.

[6] Y. Lu, X. Li, Y.-T. Yu, and M. Gerla, "Information-centric Delay-tolerant Mobile Ad-hoc Networks," in *Computer Communications Workshops (INFOCOM Workshop)*. IEEE, 2014.

[7] M. Amadeo, C. Campolo, and A. Molinaro, "Forwarding Strategies in Named Data Wireless Ad Hoc Networks: Design and Evaluation," *Journal of Network and Computer Applications*, vol. 50, pp. 148–158, 2015.

[8] W. Moreira and P. Mendes, "Pervasive Data Sharing as an Enabler for Mobile Citizen Sensing Systems," *IEEE Communications Magazine*, vol. 53, no. 10, pp. 164–170, 2015.

[9] G. Tyson, N. Sastry, I. Rimac, R. Cuevas, and A. Mauthe, "A Survey of Mobility in Information-Centric Networks: Challenges and Research Directions," in *Proceedings of the 1st ACM workshop on Emerging Name-Oriented Mobile Networking Design-Architecture, Algorithms, and Applications*. ACM, 2012, pp. 1–6.

[10] A. Rajaei, D. Chalmers, I. Wakeman, and G. Parisis, "GSAF: Efficient and Flexible Geocasting for Opportunistic Networks," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 2016, pp. 1–9.

[11] G. S. Tuncay, G. Benincasa, and A. Helmy, "Autonomous and distributed Recruitment and Data Collection Framework for Opportunistic Sensing," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 16, no. 4, pp. 50–53, 2013.

[12] S. Lu and Y. Liu, "Geoopp: Geocasting for opportunistic Networks," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*. IEEE, 2014, pp. 2582–2587.

[13] V. N. Soares, J. J. Rodrigues, and F. Farahmand, "GeoSpray: A geographic Routing Protocol for Vehicular Delay-tolerant Networks," *Information Fusion*, vol. 15, pp. 102–113, 2014.

[14] Y. Cao, Z. Sun, H. Cruickshank, and F. Yao, "Approach-and-Roam (AaR): A geographic Routing Scheme for Delay/Disruption Tolerant Networks," *IEEE transactions on Vehicular Technology*, vol. 63, no. 1, pp. 266–281, 2014.

[15] M. Amadeo, A. Molinaro, and G. Ruggeri, "E-CHANET: Routing, forwarding and transport in Information-Centric Multihop Wireless Networks," *Computer communications*, vol. 36, no. 7, pp. 792–803, 2013.

[16] G. Deng, X. Xie, L. Shi, and R. Li, "Hybrid Information Forwarding in VANETs through Named Data Networking," in *Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2015.

[17] W. Drira and F. Filali, "NDN-Q: An NDN Query Mechanism for efficient V2X Data Collection," in *Sensing, Communication, and Networking Workshops (SECON Workshops), 2014 Eleventh Annual IEEE International Conference on*. IEEE, 2014, pp. 13–18.

[18] M. Kuai, X. Hong, and Q. Yu, "Density-Aware Delay-Tolerant Interest Forwarding in Vehicular Named Data Networking," in *Vehicular Technology Conference (VTC)*. IEEE, 2016.

[19] S. H. Ahmed, S. H. Bouk, M. A. Yaqub, D. Kim, H. Song, and J. Lloret, "CODIE: Controlled Data and Interest Evaluation in Vehicular Named Data Networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3954–3963, 2016.

[20] R. A. Rehman and K. Byung-Seo, "Location-Aware Forwarding and Caching in CCN-Based Mobile Ad Hoc Networks," *IEICE TRANSACTIONS on Information and Systems*, vol. 99, no. 5, pp. 1388–1391, 2016.

[21] T. A. B. Nguyen, F. Englert, S. Farr, C. Gottron, D. Böhnstedt, and R. Steinmetz, "Hybrid Communication Architecture for Emergency Response An Implementation in Firefighter's Use Case," in *Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2015.

[22] D. Pesavento, G. Grassi, C. E. Palazzi, and G. Pau, "A Naming Scheme to represent Geographic Areas in NDN," in *IFIP Workshop Wireless Days (WD)*. IEEE, 2013.

[23] P. Cégielski and D. Richard, "Decidability of the Theory of the natural Integers with the Cantor Pairing function and the Successor," *Theoretical Computer Science*, vol. 257, no. 1-2, pp. 51–77, 2001.

[24] T. De Wolf and T. Holvoet, "Design Patterns for Decentralised Coordination in Self-organising Emergent Systems," in *International Workshop on Engineering Self-Organising Applications*. Springer, 2006.

[25] A. Afanasyev, I. Moiseenko, L. Zhang *et al.*, "ndnSIM: NDN Simulator for NS-3," *University of California, Los Angeles, Tech. Rep*, 2012.

[26] P. N. D. Bukh and R. Jain, "The Art of Computer Systems Performance Analysis, Techniques for experimental Design, Measurement, Simulation and Modeling," 1992.

[27] T. A. B. Nguyen, C. Meurisch, S. Niemczyk, D. Böhnstedt, K. Geihs, M. Mühlhäuser, and R. Steinmetz, "Adaptive Task-Oriented Message Template for In-Network Processing," in *International Conference on Networked Systems (NetSys)*. IEEE, 2017.

---

[1] https://github.com/named-data-mobile/NFD-android