Node Cardinality Estimation Using a Mobile Base Station in a Heterogeneous Wireless Network Deployed Over a Large Region

Sachin Kadam
Signal Processing and Communication Research Lab
Lehigh University
Bethlehem, PA 18015 USA
Email: sak420@lehigh.edu

Gaurav S. Kasbekar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai 400076, India
Email: gskasbekar@ee.iitb.ac.in

***Abstract*—We consider the problem of estimation of the node cardinality of each node type in a heterogeneous wireless network with T types of nodes deployed over a large region, where $T \geq 2$ is an integer. A mobile base station (MBS), such as that mounted on an unmanned aerial vehicle (UAV), is used in such cases since a single static base station is not sufficient to cover such a large region. The MBS moves around in the region, and makes multiple stops, and at the last stop, it is able to estimate the node cardinalities for the entire region. In this paper, we propose two schemes, viz., HSRC-M1 and HSRC-M2, to rapidly estimate the number of nodes of each type. Both the schemes have two phases and they are performed at each stop. We prove that the node cardinality estimates computed using our proposed schemes equal and hence are as accurate as the estimates that would have been obtained if a well known estimation protocol designed for homogeneous networks in prior work is separately executed T times. Using simulations, we show that the numbers of slots required by the proposed schemes, viz., HSRC-M1 and HSRC-M2, for computing the node cardinality estimates are significantly less than the number of slots required for T separate executions of the above estimation protocol for homogeneous networks.**

I. Introduction

Mobile base stations (MBSs), such as those mounted on unmanned aerial vehicles (UAVs), are robust, highly mobile and agile, have wide coverage capability and high battery backup capacity, and are being extensively deployed in various military and civilian applications [1]. MBSs can be used for estimating the number of active nodes such as moving vehicles in traffic control systems [2], [3], agricultural field monitoring sensors [4], people affected in disasters such as floods, wild fires [5] etc. In [2], MBSs are used to estimate the number of vehicles moving on some congested roads per hour, so that an efficient traffic controller can be designed. An MBS mounted on a UAV traverses over the given set of busy roads and stops at specific spots for this purpose. The estimated data helps in dynamically fixing the traffic signal ON/OFF durations based on the vehicle density. Consider an agricultural field in which sensors that measure various parameters (e.g., soil moisture, temperature etc.) are deployed. Before collecting the actual data from the active sensors,[1] an MBS moves around the agricultural field and stops at predetermined spots to estimate the number of active sensors [4]. This enhances the efficiency of the data collection process, since the MBS can decide the amount of time it needs to spend at each stop, when it again visits the same spots to collect the actual data and it can inform the active sensors when to be available for sending the data, based on the estimates. During natural disasters such as floods, wild fires etc., MBSs hover over the affected region to find estimates of the number of affected people. These estimates are used to achieve efficient management of disaster relief teams and distribution of relief materials [5]. Node cardinality estimation schemes are also useful in the design of medium access control protocols for wireless networks; in particular, the optimal contention probabilities, contention period and data transmission period durations etc., can be computed as functions of the computed estimates [6]–[8].

Node cardinality estimation schemes for homogeneous Machine-to-Machine (M2M) networks and Radio Frequency Identification (RFID) systems have been proposed in [9]–[11]. A similar estimation problem has been addressed for homogeneous RFID systems deployed over a large region in which, when the entire region is not in the coverage range of a reader, it sequentially moves to different locations in order to estimate the total number of active tags present in the entire region [12]–[14]. However, the above schemes cannot be used to efficiently estimate the node cardinalities of different types of nodes in a *heterogeneous* network with T types of nodes, where $T \geq 2$ is an integer.

Node cardinality estimation schemes for heterogenous M2M networks have been proposed in [7], [8], [15], [16]. In these works, the authors have considered the case where a single static base station (BS) can cover the entire region. In this paper, we address the problem of rapidly estimating the node cardinality of each node type in a heterogeneous network with T types of nodes, which are distributed over a large region; in this case, a single static BS is not sufficient. Hence, we consider an MBS moving around and making multiple stops to cover the region, so that the union of the coverage ranges of the MBS at the set of all stops is the entire region (see Section II-A). In Section II-B, we provide a brief review of the multiple-set Simple RFID Counting (SRC_M) protocol proposed in [12] for tag cardinality estimation in a homogeneous RFID system, which we extend to design our node cardinality estimation schemes for a heterogeneous wireless network. We propose two node cardinality estimation schemes, viz., HSRC-M1 and HSRC-M2; using any one of these, the MBS, at the end of covering the entire region, can find the active node cardinality estimates of all T types. These two schemes are described in Section III. Note that a challenge that needs to be overcome is that the coverage regions at different stops may overlap, hence the proposed schemes are designed to prevent multiple counting of the same active nodes. The schemes, HSRC-M1 and HSRC-M2, proposed in this paper, are generalizations of the HSRC-1 and

S. Kadam worked on this research while he was with IIT Bombay. The contributions of S. Kadam and G. Kasbekar have been supported by SERB grant SB/S3/EECE/157/2016.

[1]Sensors become active whenever they have some new data to send to the MBS.

978-1-7281-8895-9/20/$31.00 © 2020 IEEE

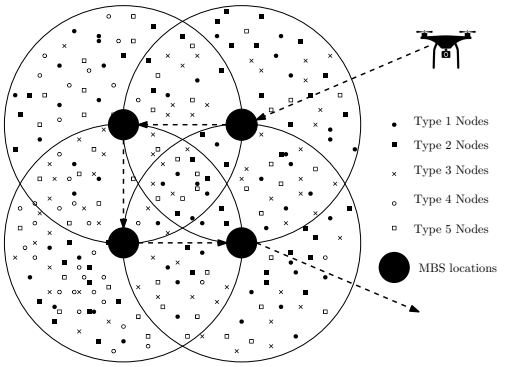


Figure 1: The figure shows $M = 4$ locations (stops) of a mobile base station (MBS) and $T = 5$ types of nodes in a region. The coverage range of the MBS at a stop is the area inside the circle with that stop as the center.

HSRC-2 schemes proposed in [16]. We prove that the node cardinality estimates computed using any one of our proposed schemes, HSRC-M1 and HSRC-M2, *equal and hence are as accurate as* the estimates that would have been obtained if the SRC$_M$ protocol were separately executed T times to estimate the cardinalities of the T node types. In Section IV, using simulations, we show that the number of time slots required by each of the proposed schemes, viz., HSRC-M1 and HSRC-M2, for computing the node cardinality estimates is significantly less compared to the number of slots required for T separate executions of the SRC$_M$ protocol.

II. Network Model, Problem Formulation and Background

A. Network Model and Problem Formulation

Consider a heterogeneous wireless network consisting of a mobile base station (MBS), which moves around in a given region and stops at M different locations, and T different types, say Type 1 ($\mathscr{T}_1$), ..., Type T ($\mathscr{T}_T$), of wireless devices (nodes) in the region, where $T \geq 2$ is an arbitrary integer. We assume that all the nodes lie within the union of the coverage ranges of the MBS at the M stops. Also, the different types of nodes may, e.g., be nodes that send emergency, periodic, normal type data etc. Fig. 1 illustrates such a network for the case $T = 5$ and $M = 4$. We denote the sets of nodes of $\mathscr{T}_1, \ldots, \mathscr{T}_T$ by $\mathscr{N}_1, \ldots, \mathscr{N}_T$, respectively; let $|\mathscr{N}_b| = N_b$, $b \in \{1,\ldots,T\}$.[2] Only a subset of the set of all nodes, i.e., $\bigcup_{b=1}^{T} \mathscr{N}_b$, are *active*, i.e., have data to send. Let $\mathscr{A}_b^{(m)}$ be the set of active nodes of $\mathscr{T}_b$, $b \in \{1,\ldots,T\}$, within the coverage range of the MBS when it is at stop $m \in \{1,\ldots,M\}$ and $n_b^{(m)} = |\bigcup_{i=1}^{m} \mathscr{A}_b^{(i)}|$. So, the total number of active nodes of $\mathscr{T}_b$ in the entire region is $n_b = n_b^{(M)} = |\bigcup_{m=1}^{M} \mathscr{A}_b^{(m)}|$. Our objective is to rapidly estimate the values of n_b, $b \in \{1,\ldots,T\}$. In particular, let $\hat{n}_b^{(m)}$ (respectively, $\hat{n}_b$) be the estimated value of $n_b^{(m)}$ (respectively, n_b). Let ε, the desired relative error bound, and δ, the desired error probability, be the user specified accuracy requirements, i.e., the parameters with which the estimate $\hat{n}_b$ needs to be obtained. Our objective is to rapidly find estimates $\hat{n}_b$ for n_b, $b \in \{1,\ldots,T\}$, such that $P(|\hat{n}_b - n_b| \leq \varepsilon n_b) \geq 1 - \delta$, $\forall b \in \{1,\ldots,T\}$.

Note that a node may lie in the coverage region(s) of the MBS at one or more stopping locations. We cannot simply use an estimation protocol designed for a network with a static base station [7], [8], [15], [16], to separately estimate the node cardinality of a given type at each stop m and add up the estimates, since a node that is in the range of the MBS at multiple locations would appear multiple times in the estimate.

B. Review of the Multiple-set Simple RFID Counting (SRC$_M$) Protocol

The SRC$_M$ protocol was proposed in [12] for node cardinality estimation in a homogeneous multiple-set network; our proposed schemes extend it for node cardinality estimation in a heterogeneous wireless network with T types of nodes using an MBS.[3] The SRC$_M$ protocol estimates the number of active nodes in a homogeneous network, within given accuracy requirements, ε and δ. Note that the network model and objective are those described in Section II-A with $T = 1$. In the SRC$_M$ protocol, the MBS moves around and makes $m \in \{1,\ldots,M\}$ stops and at each stop, it estimates the number of active nodes, say $n_1^{(m)}$, present in the union of its coverage regions till stop m. The SRC$_M$ protocol consists of two phases and it executes both the phases at each stop $m \in \{1,\ldots,M\}$ (see Fig. 2). At stop m, at the end of phase 1 (respectively, phase 2), it finds a rough estimate $\tilde{n}_1^{(m)}$ (respectively, the final estimate $\hat{n}_1^{(m)}$) of $n_1^{(m)}$ [12]. Phase 1 (respectively, phase 2) of the protocol consists of a sequence of trials (respectively, a single trial), and each trial consists of multiple slots (see Fig. 2). The number of slots in a trial is called the length of the trial. After a trial, a slot can be in one of the following three states: (i) *Empty*: No node transmitted in that slot, (ii) *Success*: Exactly one node transmitted in that slot, (iii) *Collision*: More than one node transmitted in that slot. We provide a brief review of phase 1 (respectively, phase 2) of the SRC$_M$ protocol in Section II-B1 (respectively, Section II-B2).

1) Review of Phase 1 of the SRC$_M$ Protocol: Let $T = 1$, n_{all} be the total number of nodes manufactured and $t = \lceil \log_2 n_{all} \rceil$.[4] At each stop $m \in \{1,\ldots,M\}$, phase 1 of the SRC$_M$ protocol executes W independent trials, each consisting of t time slots and W is determined based on the desired error probability δ. For example, for $\delta = 0.2$, $W = 30$ is used [12]. Let

$$\bar{p}^{(m)}(i) = \begin{cases} 2^{-i}, & \text{for } i \in \{1,\ldots,t-1\}, \\ 2^{-(t-1)}, & \text{for } i = t. \end{cases} \quad (1)$$

Suppose in trial $w \in \{1,\ldots,W\}$, each active node in the coverage range of the MBS at stop m independently transmits in a slot i, $i \in \{1,\ldots,t\}$, with probability $\bar{p}^{(m)}(i)$ (see Fig. 2). Let $S_w^{(m)}(i)$, $i \in \{1,\ldots,t\}$, be a bit vector of length t at stop $m \in \{1,\ldots,M\}$ and trial $w \in \{1,\ldots,W\}$, whose i^{th} bit is 0 if no active node transmits in the i^{th} slot, else it is 1. We find another bit vector using the following equation: $Y_w^{(m)}(i) = Y_w^{(m-1)}(i) \vee S_w^{(m)}(i)$,[5] where $Y_w^{(0)}(i)$ is a bit vector, all of whose elements are zeros. Now, phase 1 searches $Y_w^{(m)}(i)$, at bit positions $i = 2^{j-1}$, $j \in \{1,\ldots,t\}$, until it encounters a 0 bit at $j = j'$ (say) for the first time.[6] Then it uses the binary search algorithm over the

[2] $|A|$ denotes the cardinality of set A.

[3] Note that in [12], the SRC$_M$ protocol is designed for a Radio Frequency Identification (RFID) system, which consists of a reader and several tags in a region. In this paper, we use the SRC$_M$ protocol in the wireless network context, so while reviewing the SRC$_M$ protocol we use the terms "MBS" and "node" in place of "reader" and "tag", respectively.

[4] $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

[5] $\vee$ denotes the bitwise OR operator.

[6] If no bit is 0, then we take $v_w^m = t$.

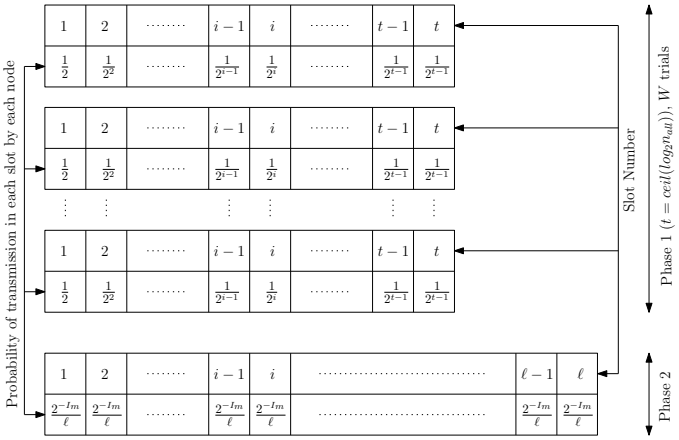


Figure 2: The figure shows the frame structure used in the SRC$_M$ protocol when the MBS is at stop m.

set $\{2^{j'-2},\ldots,2^{j'-1}-1\}$, to find the maximum integer (slot number) v_w^m, $w \in \{1,\ldots,W\}$, such that the bit $Y_w^{(m)}(v_w^m)$ is 1. At stop m, at the end of all W trials, the estimate of $n_1^{(m)}$ is computed as [12]:

$$\tilde{n}_1^{(m)} = 0.794 \times 2^{\left(\Sigma_{w=1}^{W} v_w^m\right)/W}. \tag{2}$$

2) Review of Phase 2 of the SRC$_M$ Protocol: At each stop $m \in \{1,\ldots,M\}$, phase 2 of the SRC$_M$ protocol uses the "balls-and-bins" (*BB*) method [12]. In this method, each active node independently chooses a slot out of a fixed number, say ℓ, of slots uniformly at random, transmits in that slot with a fixed probability assigned to it and otherwise does not transmit (see Fig. 2). It determines the probability of participation 2^{-I_m}, by rounding off p_m (see (3)) to the nearest integer I_m, as explained below. The parameter ℓ is a function of the desired relative error ε and it is found from a numerical lookup table, which is constructed by executing the SRC$_M$ protocol for different values of $n_1^{(m)}$, and finding the value of ℓ required to achieve a given value of ε [12]. For each $m \in \{1,\ldots,M\}$, the values of p_m and I_m are computed using the following equations:

$$p_m = \min\left(1, 1.6\ell/\tilde{n}_1^{(m)}\right), \tag{3}$$

$$I_m = \underset{j\in\{1,2,3,\ldots\}}{\operatorname{argmin}} |2^{-j} - p_m|. \tag{4}$$

Let z be the number of slots, out of the ℓ slots, which are empty in the phase 2 executions at all stops $m \in \{1,\ldots,M\}$. The protocol counts the number of empty slots, z, and computes the final estimate as follows [12]:[7]

$$\hat{n}_1 = \hat{n}_1^{(M)} = \frac{\ln(z/\ell)}{\ln\left(1 - p_M/\ell\right)}. \tag{5}$$

III. Proposed Node Cardinality Estimation Schemes

We now describe the proposed schemes, which are extensions of the SRC$_M$ protocol for estimating the number of active nodes of each type in the model with an MBS with M stops and T different types of nodes in the union of its coverage ranges described in Section II-A. The proposed schemes are Heterogeneous SRC$_M$-1 (HSRC-M1) and Heterogeneous SRC$_M$-2 (HSRC-M2) and both consist of two phases– they correspond to the two phases of the SRC$_M$ protocol (see Section II-B). Each of HSRC-M1 and HSRC-M2 executes both its phases at each stop $m \in \{1,\ldots,M\}$. Phase 2 of HSRC-M1 consists of 3 steps and that of HSRC-M2 consists of 2 steps (except for $T=2$ and $T=3$). So henceforth, we refer to them as "The 3-Step Protocol" (3-SP) (see Section III-A) and "The 2-Step Protocol" (2-SP) (see Section III-B), respectively.

[7]If $z=0$, then $\hat{n}_1$ is set to an arbitrary integer.

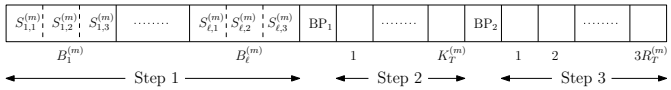


Figure 3: The figure shows the frame structure used in the 3-Step Protocol (3-SP) for the case $T=4$.

At each stop $m \in \{1,\ldots,M\}$, let the MBS have an array of length ℓ (number of slots used in phase 2 of the SRC$_M$ protocol) assigned to nodes of each $\mathscr{T}_b$, $b \in \{1,\ldots,T\}$; also, each element of the array takes value 0 or 1. We refer to this array as the bit pattern of $\mathscr{T}_b$, $b \in \{1,\ldots,T\}$, and denote it by $X^{(m)}(b,i)$. At each stop $m \in \{1,\ldots,M\}$, during phase 1 of HSRC-M1 or HSRC-M2, phase 1 of SRC$_M$ is separately executed T times to compute rough estimates, $\tilde{n}_1^{(m)},\ldots,\tilde{n}_T^{(m)}$, of the active node cardinalities of the T node types. Note that this requires execution of W independent trials for each node type, i.e., a total of WT independent trials at each stop m, in phase 1.[8]

A. Phase 2 of the Heterogeneous SRC$_M$-1 (HSRC-M1) Scheme

Recall that in phase 2 of the HSRC-M1 scheme, 3-SP is used, which we describe now. At each stop $m \in \{1,\ldots,M\}$, step 1 of 3-SP consists of ℓ blocks, say $B_h^{(m)}$, $h \in \{1,\ldots,\ell\}$ (see Fig. 3). Each block, $B_h^{(m)}$, is divided into $(T-1)$ slots $S_{h,1}^{(m)}$, …, $S_{h,T-1}^{(m)}$. Each active node of each of the T types independently chooses a block $B_h^{(m)}$ out of the ℓ blocks uniformly at random and transmits in that block with probability $2^{-I_{b,m}}/\ell$, where $I_{b,m}$ is obtained from the following equations:[9]

$$p_{b,m} = \min\left(1, 1.6\ell/\tilde{n}_b^{(m)}\right), \tag{6}$$

$$I_{b,m} = \underset{j\in\{1,2,3,\ldots\}}{\operatorname{argmin}} |2^{-j} - p_{b,m}|. \tag{7}$$

$\mathscr{T}_1$ active nodes whose chosen block is $B_h^{(m)}$ transmit symbol α in all $(T-1)$ slots, i.e., $S_{h,1}^{(m)}$, …, $S_{h,T-1}^{(m)}$, of block $B_h^{(m)}$. $\mathscr{T}_2$ (respectively, $\mathscr{T}_3,\ldots,\mathscr{T}_T$) active nodes whose chosen block is $B_h^{(m)}$ transmit symbol β in slot $S_{h,1}^{(m)}$ (respectively, $S_{h,2}^{(m)}$, …, $S_{h,T-1}$) and do not transmit in the other slots of block $B_h^{(m)}$. For example, for $T=4$, $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{T}_3$, and $\mathscr{T}_4$ active nodes whose chosen block is $B_h^{(m)}$ transmit symbols (α,α,α), $(\beta,0,0)$, $(0,\beta,0)$, and $(0,0,\beta)$, respectively, in the $(T-1)=3$ slots of $B_h^{(m)}$. The outcome in each slot can be any of the following: (i) Empty (E), (ii) Success (α or β), (iii) Collision (C). Step 1 concludes with this. Now, it can be shown that if the outcome at a given block $B_h^{(m)}$ is C in at most $(T-2)$ slots, then the

[8]We use this simple scheme of T separate executions of phase 1 of SRC$_M$ in phase 1 of HSRC-M1 and HSRC-M2 since the number of slots used in phase 1 of SRC$_M$ is negligible compared to the number of slots used in phase 2. The performances of the proposed schemes can be slightly improved by instead using schemes similar to those described in Sections III-A and III-B in phase 1.

[9]Note that the values of $\tilde{n}_b^{(m)}$, $m \in \{1,\ldots,M\}$, $b \in \{1,\ldots,T\}$, are available from phase 1 of the HSRC-M1 protocol.

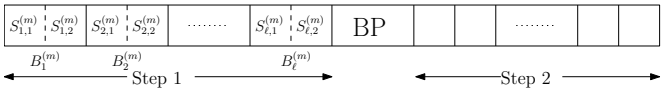


Figure 4: The figure shows the frame structure used in the 2-Step Protocol (2-SP) for the case $T = 4$.

set of types of nodes that transmitted in block $B_h^{(m)}$ can be unambiguously inferred by the MBS.[10] However, for some blocks of step 1, collisions in all $(T-1)$ slots of the block $B_h^{(m)}$ may occur; in this case, we can see that the MBS has ambiguity about the types of nodes that transmitted in those particular blocks. To resolve the ambiguity, after the end of step 1, the MBS transmits a broadcast packet (BP), say BP_1 (see Fig. 3), in which the list of the numbers of all blocks in which collisions in all $(T-1)$ slots occurred is encoded.

In step 2, there are $K_T^{(m)}$ slots (see Fig. 3), where $K_T^{(m)}$ is the number of blocks in step 1 in which collisions occurred in all $(T-1)$ slots. For $i \in \{1,\ldots,K_T^{(m)}\}$, in the i^{th} slot of step 2, $\mathscr{T}_1$ nodes that transmitted in the i^{th} block of step 1 in which collisions occurred in all $(T-1)$ slots, transmit symbol α. $\mathscr{T}_2,\ldots,\mathscr{T}_T$ nodes do not transmit in step 2. Now, it is easy to see that at the end of step 2, the MBS unambiguously knows the set of block numbers of step 1 in which $\mathscr{T}_1$ nodes transmitted. However, if in step 2, there are collisions in some of the slots, ambiguity remains with the MBS on whether $\mathscr{T}_2,\ldots,\mathscr{T}_T$ nodes transmitted in the corresponding blocks of step 1. To resolve this ambiguity, after the end of step 2, the MBS transmits a BP, say BP_2 (see Fig. 3), in which it encodes, the list of block numbers of step 1 for which collisions occurred in the corresponding slots of step 2. Suppose there are $R_T^{(m)}$ blocks in this list.

In step 3, $(T-1)R_T^{(m)}$ slots are used (see Fig. 3). For $i \in \{1,\ldots,R_T^{(m)}\}$, $\mathscr{T}_2$ (respectively, $\mathscr{T}_3,\ldots,\mathscr{T}_T$) active nodes corresponding to the i^{th} block in the above list transmit symbol β in the $((i-1)(T-1)+1)^{th}$ (respectively, $((i-1)(T-1)+2)^{th}$, $\ldots$, $(i(T-1))^{th}$) slot of step 3. It is easy to see that for each $b \in \{1,\ldots,T\}$, at the end of step 3, the MBS unambiguously knows the set of block numbers of step 1 in which $\mathscr{T}_b$ nodes transmitted. For each $b \in \{1,\ldots,T\}$ and $i \in \{1,\ldots,\ell\}$, $X^{(m)}(b,i)$ is set to 1 if at least one node of type b transmitted in block i of step 1, and to 0 otherwise.

B. Phase 2 of the Heterogeneous SRC_M-2 (HSRC-M2) Scheme

Recall that in phase 2 of HSRC-M2, 2-SP is used, which we describe now. For $T = 2$ and $T = 3$, 2-SP is identical to 3-SP. For $T \geq 4$, 2-SP is a more sophisticated scheme than 3-SP and has only two steps. For ease of exposition and due to space constraints, we explain the operation of 2-SP only for $T = 4$ in this paper. At each stop $m \in \{1,\ldots,M\}$, step 1 of 2-SP consists of ℓ blocks (see Fig. 4). Each block, $B_h^{(m)}$, $h \in \{1,\ldots,\ell\}$, is divided into $4/2 = 2$ slots, i.e., $S_{h,1}^{(m)}, S_{h,2}^{(m)}$. Each active node of each of the T types independently chooses a block $B_h^{(m)}$ out of the ℓ blocks uniformly at random and transmits in that block with probability $2^{-I_{b,m}}/\ell$, where $I_{b,m}$ is obtained using (6) and (7). $\mathscr{T}_1, \mathscr{T}_2, \mathscr{T}_3$, and $\mathscr{T}_4$ active nodes whose chosen block is $B_h^{(m)}$ transmit symbols $(\alpha,0)$, (α,α), $(0,\beta)$, and (β,β), respectively, in the two slots of $B_h^{(m)}$. Now, it is easy to see that if collisions do not occur in any of the slots of a block $B_h^{(m)}$, then the set of types of nodes that transmitted in block $B_h^{(m)}$ can be unambiguously inferred by the MBS.[11] In case of collisions in at least one slot, but not all slots, of a block $B_h^{(m)}$, ambiguity may remain and it is resolved in step 2.[12] In case of collisions in all the slots of a block $B_h^{(m)}$, ambiguity remains for nodes of each of the node types about whether or not they transmitted. In this case, the set of all node types $\{1,\ldots,4\}$ is divided into two groups: $\{1,2\}$ and $\{3,4\}$ and each of these groups recursively uses the step 1 protocol in step 2 to resolve the ambiguity.

A broadcast packet (BP) is sent by the MBS after step 1 (see Fig. 4), which contains instructions that the active nodes should follow to resolve the remaining ambiguity, if any, in step 2. It can be shown that for each $b \in \{1,\ldots,T\}$, at the end of step 2, the MBS unambiguously knows the set of block numbers of step 1 in which $\mathscr{T}_b$ nodes transmitted. For each $b \in \{1,\ldots,T\}$ and $i \in \{1,\ldots,\ell\}$, $X^{(m)}(b,i)$ is set to 1 if at least one node of type b transmitted in block i of step 1, and to 0 otherwise.

C. Computation of Node Cardinality Estimates

After the end of 3-SP or 2-SP at stop $m \in \{1,\ldots,M\}$, the MBS knows the bit patterns $X^{(m)}(b,i)$, $\forall b \in \{1,\ldots,T\}$, $\forall i \in \{1,\ldots,\ell\}$, of all the types of nodes. Let z_b be the number of zeros in $X(b,i) = \vee_{m=1}^{M} X^{(m)}(b,i)$, $i \in \{1,\ldots,\ell\}$, at the last stop M of the MBS.[13] Then for each $b \in \{1,\ldots,T\}$, the final estimate generated by the protocol is [12]:

$$\hat{n}_b = \frac{\ln(z_b/\ell)}{\ln(1-p_{b,M}/\ell)}. \tag{8}$$

Theorem 1: The final node cardinality estimate, $\hat{n}_b$, of each type $b \in \{1,\ldots,T\}$, obtained using any one of the proposed schemes, viz., HSRC-M1 and HSRC-M2, *equals and hence is as accurate as* the estimate that would have been obtained if phases 1 and 2 of the SRC_M protocol were separately executed T times at each stop $m \in \{1,\ldots,M\}$ to estimate the number of active nodes of each type $b \in \{1,\ldots,T\}$.

The proof of Theorem 1 is omitted due to space constraints.

IV. Simulation Results

Consider the network model described in Section II-A with $M = 4$. Let q_b, $b \in \{1,\ldots,T\}$, be the probability with which a given node of $\mathscr{T}_b$ is active. In this section, for simplicity, we assume that $|\mathscr{N}_1| = \ldots = |\mathscr{N}_T| = D$ (say) and

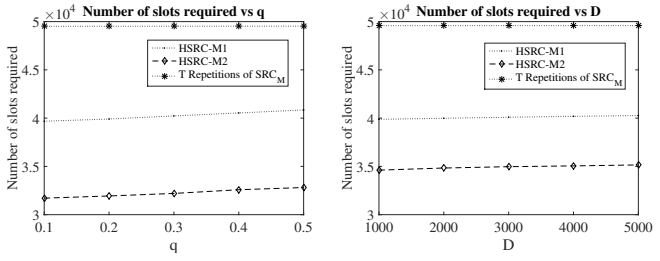
[10]The proof of this claim is omitted due to space constraints, but we provide two simple examples to support our claim. (i) If the outcome is $(\beta,\beta,\ldots,\beta)$, then it implies that exactly one node from each of $\mathscr{T}_2,\ldots,\mathscr{T}_T$ has transmitted and no node from $\mathscr{T}_1$ has transmitted. (ii) If the outcome is $(\alpha,C,C,\ldots,C,C,\alpha)$, then it implies that exactly one node from $\mathscr{T}_1$, at least one node from $\mathscr{T}_3,\mathscr{T}_4,\ldots,\mathscr{T}_{T-2},\mathscr{T}_{T-1}$ have transmitted and no node from $\mathscr{T}_2$ and $\mathscr{T}_T$ has transmitted.

[11]The proof of this claim is omitted due to space constraints, but we provide the following example to support our claim. If the slot results are (α,β), it implies that exactly one node each from $\mathscr{T}_1$, $\mathscr{T}_3$ and no nodes from $\mathscr{T}_2$ and $\mathscr{T}_4$ have transmitted.

[12]For example: In step 1, if slot 1 results in β and slot 2 results in C, then the MBS unambiguously infers that at least one node of $\mathscr{T}_3$ and exactly one node of $\mathscr{T}_4$ transmitted, and no $\mathscr{T}_1$ and $\mathscr{T}_2$ nodes transmitted. Step 2 is not required in this case. Similarly, if slot 1 results in α and slot 2 results in C, then the MBS infers that at least one node of $\mathscr{T}_3$ transmitted, no node of $\mathscr{T}_4$ transmitted, and exactly one node of either $\mathscr{T}_1$ or $\mathscr{T}_2$ transmitted. Step 2 is required in this case to resolve the ambiguity about whether a $\mathscr{T}_1$ or $\mathscr{T}_2$ node transmitted.

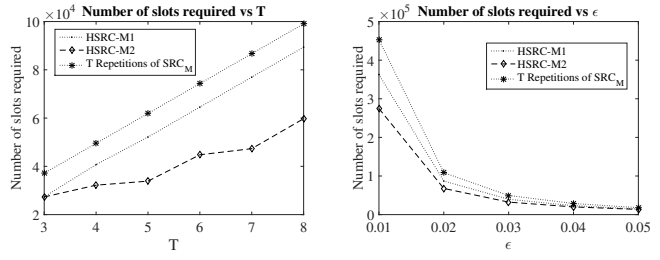
[13]If $z_b = 0$, then $\hat{n}_b$ is set to be an arbitrary integer.

$q_b = q$, $\forall b \in \{1,\ldots,T\}$. For the simulations, a network is generated in a square of dimensions $((0,1)\times(0,1))$ with the MBS making stops at the following locations: $(0.75,0.75)$, $(0.25,0.75)$, $(0.25,0.25)$, and $(0.75,0.25)$ (see Fig. 1). Nodes of all T types are placed inside this square at locations that are chosen uniformly at random from $((0,1)\times(0,1))$. The coverage range of the MBS is $R = \pi/4$ units and the locations at which the MBS makes stops are chosen in such a way that all the nodes of all T types inside the square of size 1×1 are covered by the MBS.



(a) $D = 300$. (b) $q = 0.6$.

Figure 5: These plots show the average numbers of slots required by various estimation schemes versus q and D, respectively. The following common parameter values are used in these plots: $T = 4$ and $\varepsilon = 0.03$.



(a) $\varepsilon = 0.03$. (b) $T = 4$.

Figure 6: These plots show the average numbers of slots required by various estimation schemes versus T and ε, respectively. The following common parameter values are used in these plots: $D = 300$ and $q = 0.3$.

Using simulations, we compare the performances of the proposed schemes, viz., HSRC-M1 and HSRC-M2, with that of the scheme in which the SRC$_M$ protocol proposed in [12] is separately executed T times to estimate the active node cardinality of each node type. For a fair comparison, all the schemes are executed as many times as required to achieve the same accuracy level $\delta = 0.2$.

Fig. 5a (respectively, Fig. 5b) shows a plot of the number of slots required by various estimation schemes versus q (respectively, D). Fig. 5a and Fig. 5b show that the proposed schemes significantly outperform the scheme in which the SRC$_M$ protocol is executed T times. In Fig. 5a, HSRC-M2 (respectively, HSRC-M1) outperforms the T repetitions of SRC$_M$ protocol by 34.88% (respectively, 18.74%) on average. Also, in Fig. 5b, HSRC-M2 (respectively, HSRC-M1) outperforms the T repetitions of SRC$_M$ protocol by 29.56% (respectively, 19.16%) on average. Among the proposed schemes, HSRC-M2 performs better than HSRC-M1.

Fig. 6a (respectively, Fig. 6b) shows a plot of the number of slots required by various estimation schemes versus T (respectively, ε). Both Figs. 6a and 6b show trends that are similar to those in Figs. 5a and 5b. In Fig. 6a, HSRC-M2 (respectively, HSRC-M1) outperforms the T repetitions of SRC$_M$ protocol by 38.59% (respectively, 15.68%) on average. Also, in Fig. 6b, HSRC-M2 (respectively, HSRC-M1) outperforms the T repetitions of SRC$_M$ protocol by 33.47% (respectively, 19.74%) on average.

V. Conclusions and Future Work

We proposed two schemes, viz., HSRC-M1 and HSRC-M2, to rapidly estimate, using an MBS, the number of active nodes of each type in a heterogeneous network deployed over a large region. We proved that the node cardinality estimates computed using our proposed schemes *equal and hence are as accurate as* the estimates that would have been obtained if the SRC$_M$ protocol were separately executed T times. Using simulations, we showed that the numbers of slots required by both HSRC-M1 and HSRC-M2 for computing the node cardinality estimates are significantly less compared to the number of slots required by the T separate executions of the SRC$_M$ protocol. A direction for future research is to address the information theoretic question of finding a lower bound on the number of time slots that a protocol requires for finding separate estimates of the numbers of active nodes of each type in a heterogeneous network deployed over a large region.

References

[1] M. Mozaffari, W. Saad, M. Bennis, Y. Nam and M. Debbah, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," *IEEE Commun. Survey Tuts.,* vol. 21, no. 3, pp. 2334–2360, Third Quart. 2019.

[2] K. Kanistras, G. Martins, M. J. Rutherford and K. P. Valavanis, "A Survey of Unmanned Aerial Vehicles (UAVs) for Traffic Monitoring," in *Proc. ICUAS*, pp. 221–234, 2013.

[3] R. Ke, Z. Li, S. Kim, J. Ash, Z. Cui and Y. Wang, "Real-Time Bidirectional Traffic Flow Parameter Estimation From Aerial Videos," *IEEE Trans. on Intel. Transport. Syst.,* vol. 18, no. 4, pp. 890–901, 2017.

[4] G. Giambene, E. O. Addo and S. Kota, "5G Aerial Component for IoT Support in Remote Rural Areas," in *Proc. IEEE 2nd 5GWF,* pp. 572–577, 2019.

[5] T. D. Dinh, D. T. Le, T. T. T. Tran, and R. Kirichek, "Flying Ad-Hoc Network for Emergency Based on IEEE 802.11p Multichannel MAC Protocol," in *Proc. Springer ICDCCN,* pp. 479–494, 2019.

[6] J. T. Liew, F. Hashim, A. Sali, A. Rasid, A. Jamalipour, "Probability-based Opportunity Dynamic Adaptation (PODA) of Contention Window for Home M2M Networks," *Journal of Network and Computer Applications,* vol. 144, pp. 1–12, 2019.

[7] S. Kadam, C. S. Raut, and G. S. Kasbekar, "Fast Node Cardinality Estimation and Cognitive MAC Protocol Design for Heterogeneous M2M Networks," in *Proc. IEEE GLOBECOM*, pp. 1–7, 2017.

[8] S. Kadam, C. S. Raut, A. Meena, and G. S. Kasbekar, "Fast Node Cardinality Estimation and Cognitive MAC Protocol Design for Heterogeneous Machine-to-Machine Networks," *Springer Wireless Networks*, vol. 26, no. 6, pp. 3929–3952, Aug. 2020.

[9] C. Qian, H. Ngan, Y. Liu, and L. M. Ni, "Cardinality Estimation for Large-Scale RFID Systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 9, pp. 1441–1454, 2011.

[10] A. H. Bui, C. T. Nguyen, T. C. Thang and A. T. Pham, "A Novel Effective DQ-Based Access Protocol with Load Estimation for Massive M2M Communications," in *Proc. IEEE GLOBECOM Workshops,* pp. 1–7, 2017.

[11] L. Arjona, H. Landaluce, A. Perallos and E. Onieva, "Scalable RFID Tag Estimator With Enhanced Accuracy and Low Estimation Time," *IEEE Signal Processing Letters,* vol. 24, no. 7, pp. 982–986, 2017.

[12] Z. Zhou, B. Chen, and H. Yu, "Understanding RFID Counting Protocols," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 312–327, 2016.

[13] H. Han, B. Sheng, C.C. Tan, Q. Li, W. Mao, and S. Lu, "Counting RFID Tags Efficiently and Anonymously," in *Proc. IEEE INFOCOM*, pp. 1028–1036, 2010.

[14] Q. Xiao, Y. Zhang, S. Chen, M. Chen, J. Liu, G. Cheng, and J. Luo, "Estimating Cardinality of Arbitrary Expression of Multiple Tag Sets in a Distributed RFID System," *IEEE/ACM Trans. Netw.*, vol. 27, no. 2, pp. 748–762, 2019.

[15] S. Vivek Y., P. H. Prasad, R. Kumar, S. Kadam, and G. S. Kasbekar, "Rapid Node Cardinality Estimation in Heterogeneous Machine-to-Machine Networks," in *Proc. IEEE VTC D'IoT*, pp. 1–7, Spring 2019.

[16] S. Kadam, S. Vivek Y., P. H. Prasad, R. Kumar, and G. S. Kasbekar, "Rapid Node Cardinality Estimation in Heterogeneous Machine-to-Machine Networks," Tech. Rep., 2019. [Online]. Available: https://arxiv.org/pdf/1907.04133.pdf