Source Selection Algorithms for Routing in Integrated Cellular Networks

Yumin Wu, Kun Yang
ywud@essex.ac.uk, kunyang@essex.ac.uk
University of Essex, Department of Electronic Systems
Engineering, Colchester CO4 3SQ, UK

Abstract: Integrated Cellular Networks (ICNs), as part of heterogeneous networks, are generally constructed by adding ad hoc overlay on cellular networks to solve present issues and improve network performance. Routing plays an important role in such systems. In some cases, source nodes are unknown by the system, and need to be pre-decided prior to a route discovery process. Three Source Selection Procedures (SSP’s) are proposed for routing process in ICNs. By evaluating the performance of these three SSP’s, a routing protocol in ICNs can choose one according to its own design purposes.

I. INTRODUCTION

Current cellular systems are widely deployed, but with limited flexibilities in that cellular architectures are fixed and bandwidth allocation cannot dynamically adapt to instant network situations [1]. As a result, some cells could be congested during busy time. The demerits existing in current cellular networks (CNs) encourage people to integrate CNs with other networks (such as ad hoc networks), such as iCAR [2] (Integrated Cellular and Ad Hoc Relaying System) and MADF [3] (Mobile-Assisted Data Forwarding). Also, the increasing popularity of multiple communication interfaces (e.g., cellular and IEEE 802.11) implemented in mobile devices eases the way of constructing integrated network architectures. iCAR-FA (Integrated Cellular and Ad Hoc Network with Flexible Access) [4] as an ICN is proposed to solve the congestion problem by diverting overloaded traffic from a congested cell to a non-congested cell through relaying routes. To setup relaying routes, routing protocols has to accommodate the heterogeneous network architecture (such as DARP in [5]).

Although several routing protocols for heterogeneous networks have been proposed [2], [3], the source selection in heterogeneous network routing needs to take further steps for improving the performance of ICNs. For instance, iCAR points out the need for secondary source selection but leaves the implementation open [2]. Whereas in MADF [3], the reallocation of the bandwidth occupied by secondary source is not involved in traffic diversion, and routing protocols can only be applied to original sources, i.e., the source whose initial call is blocked due to insufficiency of bandwidth at the base station. In UCAN [6], a route discovery process can start from either a source node or a proxy node. However, before the start of a route discovery process, a source needs to be identified. And a proper choice of such a source node plays a critical role. This paper proposes three algorithms for Source Selection Procedure (SSP) and investigates their performance under different network circumstances. The SSP’s proposed in this paper could also be utilized by diverse routing protocols such as these in iCAR, iCAR-FA and MADF. Refer to [4] for the routing protocols for iCAR-FA.

The rest of this paper is organized as follows. Section 2 introduces the background of the iCAR-FA system and routing issues in iCAR-FA. Section 3 details three SSP’s. Section 4 presents numerical analysis of the three SSP’s, mainly on the request rejection rate (RRR) of the system. According to the evaluation results given in Section 5, the paper finally concludes in Section 6.

II. PRELIMINARY

iCAR-FA is proposed to solve congestion problems in hot cells. Similar to iCAR, iCAR-FA deploys TDS’s in managed locations [7] so that MH’s in a hot cell can utilize the bandwidth from a cell by accessing to TDS’s and through relaying routes (which are constructed by TDS’s and MH’s). As shown in Figure 1, two air interfaces are utilized for the communication between nodes: C (Cellular) interface that operates at a cellular network frequency (in-band, e.g., 2G), and A (Ad-hoc) interface that operates at an ad-hoc network frequency (out-of-band, e.g., IEEE 802.11). Different from iCAR, TDS’s in iCAR-FA are designed with more flexibility in that MH’s with and without A-interface are both allowed to access to TDS’s. However, the communication between TDS’s and MH’s through C-interface is at the cost of in-band frequencies. For a more efficient use of out-of-band frequencies, TDS’s should prefer communicating through A-interface rather than through C-interface. Besides, relaying routes in iCAR-FA are composed of both TDS’s and MH’s with A-interface, not just TDS’s. By this means, the number of TDS’s added in iCAR-FA can be reduced, and relaying routes could be constructed more easily and flexibly.
III. BANDWIDTH REALLOCATION VIA SOURCE SELECTION PROCEDURE

This section discusses the process of reallocating the bandwidth released by the pseudo source and the operation of SSP’s.

A. Bandwidth Reallocation of Pseudo Sources

In a congested cell, in order to allocate a bandwidth for the use of a MH uncovered by TDS’s, a pseudo source is chosen by the home BS to release its occupied bandwidth without interrupting the present communication of the pseudo source. Figure 2 shows the main steps of bandwidth reallocation in a certain cell.

![Figure 2: Operation of Bandwidth Reallocation](image)

If a MH in congested cell is uncovered by any TDS’s, it sends a Pseudo Source Request Packet (PSREQ) to the home BS as trying to make a call (see step 1 in Figure 2). After receiving a PSREQ, the home BS broadcasts a MH List Request Packet (MLREQ) to all TDS’s within the home cell (see step 2 in Figure 2). Once a TDS receives a MLREQ, it broadcasts a Neighbor Discovery Request Packet (NDREQ) to all MH’s within its coverage (see step 3 in Figure 2), which respond Neighbor Reply Packets (NREP) to the TDS (see step 4 in Figure 2). Then, the TDS return a list of MH’s to the home BS by sending a MH List Reply Packet (MLREP) (see step 5 in Figure 2), which also contains the bandwidth status of the TDS. After receiving the MLREP’s from all the TDS’s upper bounded to a timeout, the home BS applies a rational SSP to analyze the information included in MLREP’s so as to choose a proper pseudo source (see step 6 in Figure 2). Following this, the home BS sends a Pseudo Source Reply Packet (PSREP) to the decided pseudo source (see step 7 in Figure 2) so that the pseudo source starts a route discovery process by broadcasting Route Request Packets (RREQ) (see step 8 in Figure 2). After receiving Route Reply Packets (RREP) (see step 9 in Figure 2), the pseudo source releases its occupied bandwidth and starts diverting its calling traffic through a relaying route (see step 10 in Figure 2). Finally, the home BS reallocates the released bandwidth to the original source (see step 11 in Figure 2).

B. Three Source Selection Procedures

As mentioned above, SSP’s are designed to run in BS’s. The objective of SSP is to choose pseudo sources to improve the overall performance of the system. In other words, pseudo sources chosen by the home BS should have the most possibility of discovering a relaying route, and a large number of pseudo sources will not partially block TDS’s. Additionally, the operation of SSP depends on the information contained in MLREP packets. An example of MLREP packet contains the following fields:

- `<TDS ID, MH ID List, Bandwidth Info>`

In the MLREP packet, “TDS ID” refers to the address of a TDS in the home BS. “MH ID List” includes a list of the addresses of MH’s within the transmission range of the TDS. “Bandwidth Info” is the available free bandwidth (or channel) in the TDS.

After receiving MLREP packets, the home BS gets a big list of MH’s. In fact, a pseudo source should fulfill the requirement that a pseudo source is an in-service MH within the transmission range of the home cell (relevant information can be obtained by checking the register information of MH’s). Hence, the home BS chooses all the MH’s according to the basic requirements of pseudo sources to build up a table of available pseudo sources. As shown in Table 1, Bij is the address of an available pseudo source. TDSij is the address of a reachable TDS of MHij. Bij contains the available bandwidth information of TDSij. Also, Table 1 gives some examples according to Figure 3. In order to choose a final pseudo source, three SSP’s are proposed according to different design purposes.

<table>
<thead>
<tr>
<th>MH ID-MHij</th>
<th>Home TDS ID: TDSij</th>
<th>Bandwidth Info: Bij</th>
</tr>
</thead>
<tbody>
<tr>
<td>MH1</td>
<td>TDS1</td>
<td>B_{TDS1}</td>
</tr>
<tr>
<td>MH2</td>
<td>TDS2</td>
<td>B_{TDS2}</td>
</tr>
<tr>
<td>MH3</td>
<td>TDS3</td>
<td>B_{TDS3}</td>
</tr>
<tr>
<td>MH4</td>
<td>TDS4</td>
<td>B_{TDS4}</td>
</tr>
</tbody>
</table>
SSP1 chooses pseudo sources mainly according to the number of reachable TDS’s of a MH, which is counted depending on the information in Table 1 (such as MH1 with one reachable TDS, and MH2 with 3 reachable TDS’s). The main idea is to choose the MH’s in Table 1 with the maximum number of reachable TDS’s. As shown in Figure 3, MH1 can only broadcast RREQ to TDS1, but MH2 can broadcast RREQ’s to TDS2, TDS3 and TDS4. MH2 stands a better chance to find a more effective relaying route so as to release its bandwidth to the original source.

![Figure 3: RRR Analysis of Source Nodes](image)

Considering the Request Rejection Rate (RRR), the RRR is the probability of a MH or a TDS failing to find a relaying route after broadcasting RREQ’s. Assuming the RRR of TDSij is $P_{ij}$ ($P_{ij} \leq 1$), then the RRR of MH1 is $P_{MHi} = P_{11}$ ($P_{11}$ denotes the RRR of TDS1 in Figure 3). Also, the RRR of MH2 is $P_{MHi} = P_{21}P_{22}P_{23}$ ($P_{21}$, $P_{22}$ and $P_{23}$ denote the RRR’s of TDS2, TDS3 and TDS4 in Figure 3 respectively). If $P_{11} = P_{21} = P_{22} = P_{23}$, then $P_{MHi} \geq P_{MHi}$. In general, if the number of reachable TDS’s of MHi is $Ni$, the RRR of MHi is $P_{MHi} = \prod_{j=1}^{Ni} P_{ij}$. Thus, the pseudo sources chosen in SSP1 is the MH with minimum RRR, namely $\min(P_{MHi},...,P_{MHi})$ (assuming the number of available MH’s in Table 1 is $n$). To simplify the notation, we assume that the RRR of each TDS is identical at $P$. Then, the RRR of MHi is:

$$P_{MHi} = \prod_{i=1}^{Ni} P_{ij} = (P)^{Ni}, \quad (3.1)$$

Thus, SSP1 just chooses MH’s with the maximum number of reachable TDS’s (namely $\max(N_1,...,N_n)$), because they have the minimum RRR. One problem of SSP1 is that a temporary event happening in the field uncovered by TDS’s could cause that SSP1 only chooses MH’s within a specific cross area with most reachable TDS’s. The result is that the bandwidth of TDS’s covering the specific cross area is firstly consumed. Then, these TDS’s could become congested but other TDS’s still have lots of free bandwidth unused. For example, in Figure 3, the bandwidths of TDS2, TDS3 and TDS4 are first cost because MH2 is chosen as a pseudo source prior to MH1. Therefore, the bandwidth of TDS1 is left to be unused.

SSP2 tries to balance the bandwidth consumption amongst TDS’s when choosing pseudo sources. Using the information in Table 1, the home BS can get the bandwidth status of all reachable TDS’s of available pseudo sources. $B_{ij}$ refers to the free bandwidth of the reachable TDSij of MHi. Then, the average bandwidth of all reachable TDS’s of an available pseudo source is:

$$B_{MHi} = \frac{\sum_{j=1}^{Ni} B_{ij}}{Ni}, \quad (3.2)$$

To achieve the balance of TDS bandwidth consumption, SSP2 chooses pseudo sources with maximum $B_{MHi}$, namely $\max(B_{MHi},...,B_{MHi})$. As a result, pseudo sources consume the bandwidth of TDS’s in average, and do not partially congest TDS’s. However, such a selection of pseudo sources may lead to a relatively higher RRR, because SSP2 is designed not to find a pseudo source with a minimum RRR but to choose a pseudo source with a maximum average bandwidth of reachable TDS’s. Then, the instantaneous RRR in SSP2 is determined by the RRR of current pseudo sources with a random number of reachable TDS’s. Thus, compared with SSP1, SSP2 can avoid partially blocking TDS’s by balancing the consumed bandwidth of TDS’s, but the RRR in SSP2 could be higher due to the unpredictable number of reachable TDS’s of a pseudo source.

SSP3 takes both the number and the average free bandwidth of reachable TDS’s into consideration. Therefore, SSP3 can achieve a balance of the bandwidth consumption of TDS’s without highly increasing the RRR of the system during the selection of pseudo sources. In SSP3, node $MHi$ has two weights for selecting pseudo sources, namely $(P_{MHi}, B_{MHi})$. The combined weight considered by SSP3 is $Wi = B_{MHi}(1 - P_{MHi})$ ($1 - P_{MHi}$ refers to the possibility of successfully discovering a relaying route). Also, the weight ($Wi$) can be named as the average free bandwidth of reachable TDS’s of $MHi$ in probability. According to formula (3.1) and (3.2) shown above, the combined weight is:

$$Wi = \frac{\sum_{j=1}^{Ni} B_{ij}}{Ni} \left[1 - (P)^{Ni}\right], \quad (3.3)$$

SSP3 chooses pseudo sources with the maximum $Wi$, namely $\max(Wi,...,Wi)$. Because $P$ in the above formula is the average RRR of TDS’s, the calculation of $P$ depends on the statistical results calculated by BS’s and TDS’s. To reduce the amount of calculation in BS’s and TDS’s, the combined weight of a $MHi$ can be simplified as following. For the MH’s
given in Table 1, both the number of reachable TDS’s ($N_i$) and the average free bandwidth of reachable TDS’s ($B_{MHi}$) are sorted in order. Then, instead of ($P_{MHi}$, $B_{MHi}$), the weight of $MHi$ in SSP3 is ($N_r$, $N_a$) ($N_r$ refers to the order of $MHi$ in terms of the number of reachable TDS’s, and $N_a$ refers to the order of $MHi$ in terms of the average free bandwidth of reachable TDS’s). Then, the combined weight of $MHi$ in SSP3 is simplified as:

$$W_i = N_r N_a.$$ (3.4)

SSP3 could just choose pseudo sources with the simplified minimum $W_i$.

After applying a SSP, more than one pseudo sources may still exist. Then, among these pseudo sources, the MH with the furthest distance from the home BS are chosen as pseudo sources. The reason is that the MH which is furthest from the home BS has the most possibility of moving outside the home cell, and make use of the bandwidth in a neighbor cell.

IV. ALGORITHMS ANALYSIS

This section evaluates the performance of proposed SSP’s in terms of the average request rejection rate (RRR) of the whole system. To simplify the analysis, we assume that MH’s randomly choose one of their reachable TDS’s as its first hop to divert traffic, and each TDS has the same request rejection rate under the same networks environment. Also, whatever temporary events (or overloaded traffic) happen inside or outside the area covered by TDS’s, traffic diversion services are offered to MH’s only after applying SSP’s. Table 2 shows the parameters used for RRR analysis

<table>
<thead>
<tr>
<th>$P_{R1}$</th>
<th>RRR in a certain cell by applying SSP1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{R2}$</td>
<td>RRR in a certain cell by applying SSP2</td>
</tr>
<tr>
<td>$P_{R3}$</td>
<td>RRR in a certain cell by applying SSP3</td>
</tr>
<tr>
<td>$P$</td>
<td>Average RRR of TDS’s</td>
</tr>
<tr>
<td>$N_r$</td>
<td>Overall number of TDS’s in a certain cell</td>
</tr>
<tr>
<td>$N_C$</td>
<td>Number of current TDS’s with enough free bandwidth in a certain cell</td>
</tr>
<tr>
<td>$T_T$</td>
<td>Maximum traffic that a TDS can burden</td>
</tr>
<tr>
<td>$T_N$</td>
<td>Overall available bandwidth of current TDS’s which cover a MH with $i$ reachable TDS’s</td>
</tr>
<tr>
<td>$T_{MAX}$</td>
<td>Maximum traffic that a BS can burden without traffic diversion</td>
</tr>
<tr>
<td>$T_C$</td>
<td>Current traffic in a certain cell</td>
</tr>
<tr>
<td>$N$</td>
<td>Maximum number of reachable and available TDS’s of a MH in a certain cell without traffic diversion</td>
</tr>
<tr>
<td>$N_a$</td>
<td>Current maximum number of reachable and available TDS’s of a MH in a certain cell after traffic diversion</td>
</tr>
<tr>
<td>$N_i$</td>
<td>The number of TDS’s covering MH’s with $i$ reachable TDS’s</td>
</tr>
</tbody>
</table>

According to the parameters given in Table 2, we have:

$$T_{Ti} = N_r T_T.$$ (4.1)

SSP1 focuses on finding pseudo sources with minimum RRR, but may cause that diversion traffic floods some TDS’s and lets the bandwidth of other TDS’s unused. As shown in Figure 4, TDS1, TDS2 and TDS3 have the highest priority of being chosen to divert traffic, because they cover MH’s with 3 reachable TDS’s (such as MH1 in Figure 4). Following them, TDS4, TDS5, TDS6 and TDS7 have the same priority. At last, TDS8 has the lowest priority, because it does not have a cross-covered field. Hence, TDS1, TDS2 and TDS3 consume their traffic firstly, and TDS8 consumes its traffic lastly.

Because the current diversion traffic in the home cell is $T_C - T_{MAX}$ ($T_C > T_{MAX}$), then:

If $T_C - T_{MAX} \leq \sum_{i=\alpha}^N (N_i T_T)$, then $N_a = \alpha$.

If $T_C - T_{MAX} \geq \sum_{i=\alpha}^N (N_i T_T)$, then $N_a = 0$.

Considering the average RRR of TDS’s ($P$), $P$ is actually a function of the number of current available TDS’s in a cell ($N_C$), namely $P = f(N_C)$. Then,

Assuming $P = P_{\text{default}} \frac{\beta N_L}{N_C}$, (4.2)

In SSP1, due to the partial use of TDS’s, $N_C$ decreases with the rise of diversion traffic. Then,

Assuming $N_C = \frac{N_r T_T - T_C - T_{MAX}}{T_T}$, (4.3)

In formula (4.2) and (4.3), $\beta$ and $\chi$ are constants related to the distribution density and bandwidth of TDS’s, and $P_{\text{default}}$ is the default value of $P$ without the effect of diversion traffic. Thus, according to formula (3.1), the RRR in a certain cell by applying SSP1 ($P_{R1}$) is:
\[ P_{R1} = \left( \frac{B_{N_{ii}}}{N_{C}} \right)^{\alpha}, \quad (4.4) \]

**SSP2** aims at balancing diversion traffic among TDS’s. According to formula (3.2), SSP2 chooses pseudo sources which have the maximum average bandwidth of reachable TDS’s. In other words, SSP2 selects pseudo sources by considering only the bandwidth effects of TDS’s. Hence, every TDS has the same opportunity of being chosen to divert a call. In fact, the instant RRR of SSP2 fluctuates between the maximum value of RRR (\( P_{RR1}^{N} \)) and the minimum value of RRR (\( P^{i} \)). According to the parameters shown in Table 2, the average RRR in a certain cell by applying SSP2 (\( P_{R2} \)) is:

\[ P_{R2} = \sum_{i=1}^{N} \left( \frac{N_{i}}{N_{T}} P^{i} \right), \quad (4.5) \]

In formula (4.5), the number of current available TDS’s (\( N_{C} \)) does not change like the way in SSP1. Actually, \( N_{C} \) starts to decrease only as every TDS’s in a certain cell nearly exhausts its own bandwidth, because the bandwidth consumption of each TDS during traffic diversion is balanced by SSP2. Additionally, the change of \( P \) in SSP2 also depends on formula (4.2).

**SSP3** takes both the RRR and bandwidth of TDS’s into consideration. According to formula (3.2), (3.3) and (3.4), SSP3 uses a combined weight (\( W \)) (related to the average free bandwidth of reachable TDS’s (PMhi)) and the RRR of each available pseudo source (\( P_{DMi} \)) to choose pseudo sources. Then, SSP3 chooses pseudo sources with maximum \( W \). By this means, pseudo sources will not partially congest TDS’s, but still keep a relatively low RRR. In SSP3, the number of current available TDS’s (\( N_{C} \)) changes like the way in SSP2. Hence, the RRR of SSP3 (\( P_{R3} \)) is mainly related to the number of reachable TDS’s. The probability of consuming the bandwidth in the area covered by \( i \) TDS’s is:

\[ P_{Ti} = \frac{i}{N}, \quad (4.6) \]

According to the assumptions given above, TDS’s, which covers the same field, have the same probability of being chosen to divert traffic. Hence, we assume that TDS’s covering the same field have the same free bandwidth (\( T_{fi} \)). According to formula (3.1), the RRR of a MH drops with the increase of the number of reachable TDS’s of the MH. Thus, the MH with maximum \( W_{i} \) is also a MH with maximum value of \( i T_{fi} \). Moreover, assuming the current pseudo source has \( d \) reachable TDS’s, and it has not start diverting traffic, then:

\[ \text{if} \ nT_{Ti} = \max(T_{T1}, \ldots, T_{TN}), \text{then} \delta = n. \]

Thus, according to formula (3.1), the current RRR in SSP3 is:

\[ P_{R3} = (P)^{\delta}, \quad (4.7) \]

Considering formula (4.5), we have:

In terms of formula (4.4), \( \alpha \) changes from \( N \) to 1 as the traffic of the system rises, and \( P \) drops with the decrease of \( N_{C} \) (formula (4.2)). Therefore, \( P_{R1} \) could be smaller than \( P_{R2} \) at a low level of system traffic, and higher than \( P_{R2} \) at a high level of system traffic because of the higher TDS’s block rate caused by a partial selection of pseudo sources in SSP1. In SSP3, because \( \delta \) (formula (4.7)) fluctuates between \( N \) and 1, \( P_{R1} \) is an intermediate value between \( P_{R2} \) and \( P_{R3} \) in general. Hence, if a network has a low overloaded traffic, or the number and the bandwidth of TDS’s is enough, we can use SSP1 to choose pseudo sources. By contrast, SSP2 is suitable for networks with limited number of TDSs and high overloaded traffic. SSP3 could be applied to networks with intermediate overloaded traffic.

![V. SIMULATION RESULTS](image)

In this section, we evaluate the three SSP’s in terms of the average RRR of the overall network, by plugging in reasonable values of parameters in formula (4.1) through (4.7). The results are based on the numerical analysis in Section 4. The general cell topology is similar to that in Figure 4. The maximum traffic that a BS can carry without traffic diversion is \( T_{MAX}=100 \) Erlangs. The maximum traffic that a TDS can burden is \( T_{f}=10 \) Erlangs. The average traffic of each call is 1 Erlangs. The average RRR of TDS’s is \( P=0.5 \). Because the RRR in SSP3 (\( P_{R3} \)) fluctuates between a maximum value and a minimum value, we only compute the average value of \( P_{R3} \) during the evaluation.

According to Figure 5, the average RRR in SSP3 keeps stable during the evaluation, because the pseudo sources are chosen only according to the bandwidth status of TDS’s. When the diversion traffic is low, The RRR in SSP1 and SSP3 stays at a very low level, compared with that in SSP2. However, as the diversion traffic in the home cell goes up, the average RRR in SSP3 jumps to an intermediate stage due to the bandwidth balancing amongst TDS’s. In SSP1, with the
increase of the diversion traffic, more and more TDS’s are blocked due to the partial selection of TDS’s. The congestion of TDS’s causes a high average RRR of TDS’s. Thus, the RRR in SSP1 increases rapidly with the diversion traffic. In addition, the RRR in both SSP2 and SSP3 only rises sharply as the diversion traffic increases to a very high level, because the balance of the bandwidth consumption of TDS’s result in that only very high diversion traffic can congest TDS’s. Hence, if TDS’s have limited available bandwidth, SSP2 or SSP3 could be used to choose pseudo sources to avoid congesting TDS’s. If the diversion traffic in a network is not high, SSP1 can be adapted to achieve a low RRR of the system.

VI. CONCLUSION AND FUTURE WORKS

This paper focuses on the SSP part of routing protocols in ICNs (such as iCAR and MADF). In order to solve the problem of unknown sources during route discovery, three SSP algorithms are designed to run in BS’s to choose pseudo sources. SSP’s proposed in this paper aim to choose pseudo sources that have the most possibility of successfully discovering relaying routes and releasing its occupied bandwidth. Moreover, the selection of pseudo sources in SSP’s also tries to balance the bandwidth of TDS’s, which is used for traffic diversion. The evaluation of these algorithms has been presented and some guidelines as to when to use which algorithm have also been discussed.

In this paper the design of SSP’s is mainly driven by the need of decreasing the request rejection rate (RRR) of a cellular network system. With more network information collected by the home BS, a future SSP could further reduce the RRR and help to build more stable relaying routes.

REFERENCES


