Leveraging P2P Interactions for Efficient Location Privacy in Database-driven Dynamic Spectrum Access

Erald Troja
The Graduate Center
City University of New York
etroja@gc.cuny.edu

Spiridon Bakiras
John Jay College
City University of New York
sbakiras@jjay.cuny.edu

ABSTRACT

In the database-driven DSA model, clients learn their geographic location through a GPS device and use this location to retrieve a list of available channels from a centralized white-space database. To mitigate the potential privacy threats associated with location-based queries, existing work has proposed the use of private information retrieval (PIR) protocols when querying the database. Nevertheless, PIR protocols are very expensive and may lead to significant costs for highly mobile clients. In this paper, we propose a novel method that allows wireless users to collaborate in a peer-to-peer (P2P) manner, in order to share their *cached* channel availability information that is obtained from previous queries. Our experimental results with a real-life dataset show that our methods reduce the number of PIR queries by 50% to 60%, while incurring low computational and communication costs.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Application—*spatial databases and GIS*; C.2.4 [**Computer-communication Networks**]: Distributed Systems—*distributed applications*

General Terms

Algorithms, Security, Experimentation

Keywords

Location Privacy, White-Space Database, Dynamic Spectrum Access, Private Information Retrieval, Anonymity

1. INTRODUCTION

Dynamic spectrum access (DSA) is a novel communication paradigm that enables wireless clients to utilize statically allocated radio channels when not in use by their licensed owners. DSA is accomplished through cognitive radio (CR), an intelligent wireless communication system that is aware of its operating spectral environment [4].

ACM GIS'14 November 04 - 07 2014, Dallas/Fort Worth, TX, USA

http://dx.doi.org/10.1145/2666310.2666477.

Our work revolves around the database-driven model in which mobile clients learn their spectral surroundings by querying a centralized white-space database (WSDB). In this model, mobile clients simply send their GPS coordinates to the database and receive the centrally fused repository report for that area.

To mitigate the potential privacy threats associated with location-based queries, existing work has proposed the use of private information retrieval (PIR) protocols when querying the database [2]. A PIR protocol enables a user to retrieve a record from a database server, while maintaining the identity of the record secret from the server. However, PIR protocols are very expensive and may lead to significant costs in the case of highly mobile clients that issue numerous queries throughout their trajectories.

Since typical PIR protocols offer a trade-off between computational and communication complexity, we argue that any location privacy method for the database-driven DSA model is bounded by the limitations of the underlying PIR protocol. As such, it is desirable to identify new mechanisms for users to acquire the necessary spectral knowledge. Our intuition is that, in a white-space TV band network, mobile TV band device (TVBD) users will gradually develop a trajectory-specific spectrum knowledge *cache*, through a series of PIR requests. Therefore, we propose that mobile users that are within communication range interact in a peer-to-peer (P2P) manner, in order to privately exchange their cached spectrum knowledge for the surrounding area. We leverage the *anonymous veto network* (AV-net) protocol of Hao and Zieliński [3] that anonymizes the exchange of information among a group of users. Our experimental results with Microsoft's GeoLife trajectory dataset [5] show that our methods reduce the number of PIR queries by 50% to 60%, while incurring low computational and communication costs for the mobile clients.

The rest of this paper is organized as follows. Section 2 describes the details of our P2P protocol and Section 3 presents the results of the experimental evaluation. Finally, Section 4 concludes our work.

2. P2P PROTOCOL

System Architecture. Similar to previous work [2], we assume a fixed grid of $n \times n$ cells each of 100m x 100m, where mobile users can communicate through white-space TV bands, while maintaining their location privacy. Mobile TVBDs are allowed to communicate only in the frequency ranges 512-608 MHz (TV channels 21-36) and 614-698 MHz (TV channels 38-51), i.e., there are a total of 31 possible

white-space TV band channels that can be accessed in a DSA manner. Therefore, we represent the daily channel availability as 32 bits (per cell), where bit 0 represents a busy channel and bit 1 represents an idle channel.

In our model, we assume an out-of-band common control channel (CCC) through a dedicated transceiver which enables mobile users to exchange concurrently both control and data messages. Out-of-band CCC coordination can be realized over the 802.11 protocol in *ad-hoc* mode or through any of the methods proposed in [1].

We assume a PIR protocol that retrieves channel information for multiple cells with a single query[1]. As a proof of concept, we consider a fixed grouping of the available cells into 4×4 blocks. Therefore, we assume that each PIR query retrieves the 16-cell block that contains the user's current cell.

Fig. 1 shows an example of this approach. The black colored cells signify the locations where a new PIR query is issued, due to lack of spectrum availability knowledge. The alternating white and grey colored cells identify the different blocks, with the block *id* shown in the lower-left corner of the block.

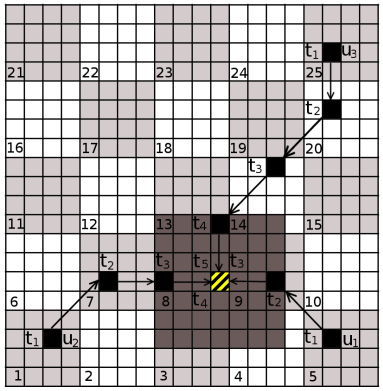


Figure 1: Three mobile users invoking the AV-net protocol for the region identified by the darker shaded cells

As illustrated in Fig. 1, each of the three mobile TVBDs gradually builds a spectrum knowledge *cache* containing channel availability information from their respective trajectories. When the users eventually meet at the diagonally striped cell, it may be beneficial to all of them to exchange their cached information. To maximize utility for all participating users, the sharing of spectrum information involves the area surrounding the current location (as users may continue their trajectories towards any direction). In particular, the TVBD nodes agree on the number of surrounding rings (AR) that they wish to explore during the protocol invocation. (Table 1 summarizes the symbols used in the remainder of this paper, along with the values tested in the experimental evaluation.) In the example of Fig. 1, $AR = 3$, and the explored region is shown in a darker shade.

To illustrate the location privacy leakage from a *plaintext* exchange of spectrum availability information (i.e., without the invocation of the AV-net protocol), consider the example of Fig. 1. We can infer that u_1 arrived at the current cell through block 9, while u_3 visited blocks 13 and 14. On the other hand, u_2's trajectory contains some uncertainty, as u_2 may have arrived at the current cell through blocks 2, 7, or 12. Furthermore, if two users participate in the same group multiple times (at different locations), they can derive more information about each other's movement patterns.

[1]All existing PIR schemes can retrieve multiple records from a database.

Table 1: Summary of symbols

Symbol	Description	Range
GS	Group size	3-10
BS	Number of cells in a PIR block	16
AR	Ring(s) explored through AV-net invocations	1-3
AP	AV-net participation probability (fixed)	0-1
PI	AV-net participation probability increment (TCP)	0.05-0.2
K	AV-net initiation threshold	0.2-0.8
AK	Actual knowledge of the AR area	0-1

We assume that intersecting users remain within communication range for ample periods of time (e.g., 1-2 minutes). However, they do not need to reside in the same cell continuously. The three conditions that control a successful invocation of our protocol are (i) protocol *initiation*, (ii) protocol *participation*, and (iii) successful *group formation*. Group formation is dependent on at least three users willing to engage in the P2P protocol, such that at least one of the engaging users is an *initiator*. We examine each condition separately in the following sections.

Protocol Initiation. Ideally, a mobile TVBD would like to maintain DSA connectivity throughout its trajectory, without any disruptions. As such, whenever the TVBD moves into a new cell, it measures the ratio of knowledge (AK) in the surrounding area. If that ratio falls under a system-defined threshold K, it initiates the protocol that triggers the group formation algorithm (described later). Algorithm 1 shows the detailed protocol initiation procedure.

Algorithm 1 Protocol initiation Algorithm

```
1: procedure INITIATE-PROTOCOL
2:
3:     bool initiate = false;
4:     double AK = 0.0;
5:
6:     if no spectrum information for current cell then
7:         initiate = true;
8:     else
9:         for i = 1 to AR do
10:            for all cells c_j in ring i do
11:                if no spectrum information for c_j then
12:                    AK = AK + 1.0/(AR · 8 · i);
13:                end if
14:            end for
15:        end for
16:        if AK ≤ K then
17:            initiate = true;
18:        end if
19:    end if
20:    return initiate;
21: end procedure
```

Protocol Participation. Participation is defined as the selfless event, where one or more users in the group decide to participate in the AV-net protocol for the purpose of disseminating (and also collecting) channel information about the surrounding area. We propose the following three AV-net participation methods.

(i)Fixed probability. This is the simplest approach where, whenever a protocol is initiated, a nearby TVBD always chooses to participate with probability AP. Larger AP values produce a greedy behavior that is optimal in terms of PIR query savings. On the other hand, this may also lead to numerous AV-net invocations in close (spatial) proximity, which are redundant in terms of gained knowledge.

(ii)TCP-like approach. In the second method, we borrow from TCP Reno's congestion control mechanism. In particular, a mobile user starts with a participation probability $AP = 1.0$. At each successful AV-net participation, AP is cut by half. Otherwise, if there is a protocol initiation but the TVBD does not participate, AP is incremented by PI units. This technique is expected to be the most conservative one, due to its aggressive back-off behavior.

(iii)Weighted sliding window. The final method is based on the weighted sliding window (SW) projection. We experimented with different window sizes, and decided to utilize a model with five entries, such that $W_1 = 0.5$, $W_2 = 0.25$, $W_3 = 0.15$, $W_4 = 0.07$, $W_5 = 0.03$, and $\sum_i W_i = 1$. (W_1 corresponds to the most recent entry.) The current window snapshot is stored as a 5-bit array, where '0' represents participation and '1' represents non-participation.

Group Formation. When Algorithm 1 (protocol initiation) returns *true*, the underlying TVBD initiates an invocation of the AV-net protocol. This is done by broadcasting its interest in the lowest out-of-band CCC channel. Assuming 801.11 as our out-of-band CCC implementation, any potential initiators will broadcast their unique MAC addresses, their current cell id, their *rendezvous* channel id[2], and an initiation flag over 802.11 channel 1. Users that are already engaged in an AV-net invocation/transmission will not hear such broadcast. We assume standard 802.11 MAC contention mechanisms are in place. We coin as "root" the first mobile user that successfully broadcasts the AV-net initiation control packet, regarding a specific cell id. Any other users (including other potential initiators situated in the same cell) that receive the first successful broadcast from a root node, and whose cell id *matches* the broadcast cell id, will use a simple three-way handshake group formation protocol.

Mobile users that decide to participate (based on the methods described earlier) or had attempted to initiate an AV-net invocation themselves, will first switch to the rendezvous channel. They will announce to the root user, through broadcast communication, their willingness to engage. We coin as "children" any of the users that have successfully rendezvoused in the channel id specified by the root user. The three-way handshake broadcast MAC protocol is summarized in Algorithm 2.

Algorithm 2 Group formation Algorithm

```
1: procedure THREE-WAY-HANDSHAKE
2:
3:     [all children broadcast] Send Request To Join Group
4:
5:     while group size < GS do
6:         [root] randomly pick a child
7:         [root] send Clear to Join Group
8:         [root] increment group size counter
9:     end while
10:
11:    for all other children who sent a Request to Join do
12:        [root] send Reject to Join Group
13:    end for
14:    [all who received Clear to Join Group] Send Confirm
       To Join Group
15:
16:    [root] send ABORT if group size counter ≤ 3
17: end procedure
```

[2]Assuming there is a free channel in the out-of-band CCC range.

AV-net Protocol. When a group is formed, the nodes therein execute the AV-net protocol [3] for each bit of information that they want to share. However, to avoid excessive network delays due to the 2-round nature of the AV-net protocol, we group all individual invocations into two aggregate rounds, as shown in Algorithm 3. Specifically, the users first agree on the the specific order in which the cell information is transmitted, and then each user broadcasts its aggregate data to the rest of the group. The broadcast order can be arranged based on the unique MAC addresses of the TVBDs.

Algorithm 3 AV-net Protocol

```
1: procedure AV-NET(G, g)
2:
3:     for all users i in the group do
4:         for all bits b in the explored area do
5:             compute g^{y_{i_b}};
6:         end for
7:     end for
8:
9:     for all users i in the group do
10:        for all bits b in the explored area do
11:            compute g^{c_{i_b} y_{i_b}};
12:        end for
13:        broadcast all exponentiations for user i;
14:    end for
15:
16:    for all users i in the group do
17:        for all bits b in the explored area do
18:            compute r_b = Π_i g^{c_{i_b} y_{i_b}};
19:            if r_b ≠ 1 then
20:                mark the corresponding channel as free;
21:            end if
22:        end for
23:    end for
24: end procedure
```

3. EXPERIMENTAL EVALUATION

In this section we evaluate experimentally the performance of our methods.

Experimental Setup. We developed our experiments in Java SDK, running on a Ubuntu 10.4 LTS machine. To simulate the mobile TVBD users, we utilized Microsoft's GeoLife GPS Trajectories[3], which is an excellent dataset containing real-life trajectories from users traveling around Beijing, China. The GeoLife dataset [5] was collected as part of the Microsoft Research Asia GeoLife project, by monitoring numerous users for a period of over five years (from Apr. 2007 to Aug. 2012). A GPS trajectory from this dataset is represented as a sequence of time-stamped points, each containing information regarding the user's latitude, longitude, and altitude.

In addition to the simulation results, we also implemented the basic cryptographic operations of the AV-net protocol on an iPhone 5, running iOS 7.1. Specifically, we cross compiled the GMP[4] multiple precision arithmetic library for the ARM architecture, and built a benchmark app to measure the cost of these operations on a handheld device. We generated a cyclic group G of prime order q, where q is a 160-bit number. The group modulus was chosen as a 64-byte prime. Table 2 shows the cost of these operations.

Results. Fig. 2a illustrates the projected CPU time needed to run the AV-net protocol (Algorithm 3) on a handheld device. This cost is dominated by the expensive modular

[3]http://research.microsoft.com/en-us/projects/GeoLife/

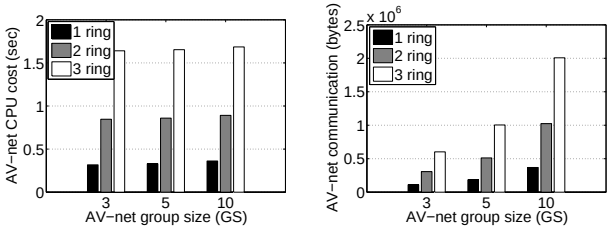
[4]http://gmplib.org

Table 2: Cost of cryptographic primitives

Operation	Cost
Modular multiplication	0.004 ms
Modular exponentiation	0.518 ms

exponentiation operations and is, thus, unaffected by the group size GS. The major factor that determines this cost is the number of surrounding rings (AR) that are explored during a protocol invocation, since each cell contributes 32 modular exponentiations. Nevertheless, even for a value of $AR = 3$, the total CPU time is around 1.65 sec, which is an acceptable cost.

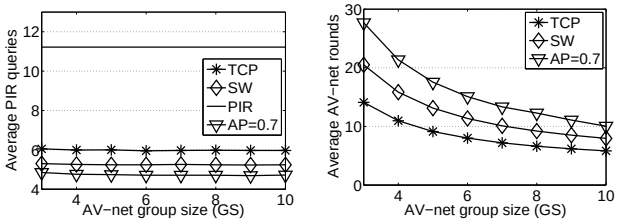
Fig. 2b shows the total number of bytes that are broadcast during an AV-net protocol invocation. Clearly, the communication cost is linear in GS, as each group member needs to broadcast its own input to the protocol. We believe that $GS = 5$ is a very reasonable value for anonymity purposes, in which case the communication cost remains below 1 MB. While this cost might appear significant, we stress that, AV-net broadcasts occur over the 802.11 CCC band and do not involve the cellular network infrastructure.



(a) (b)

Figure 2: Cost of AV-net protocol on handheld devices (a) CPU cost (b) Communication cost

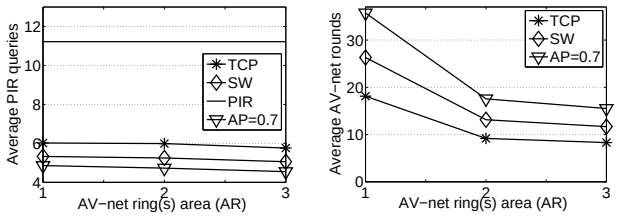
Fig. 3 demonstrates the effect of the group size (GS) on the different methods ($AR = 2$, $K = 0.5$, $PI = 0.1$). As Fig. 3a implies, larger groups do not contribute more information during the P2P data exchange. Therefore, the average number of PIR queries remains fairly constant. Nevertheless, users may still opt for larger groups, in order to gain more privacy. On the other hand, a larger group size reduces the number of AV-net invocations (Fig. 3b), because some groups may fail to form due to insufficient number of members. Among the three participation algorithms, the sliding window (SW) approach strikes a good balance between PIR savings (53%) and AV-net overhead (13 rounds, for $GS = 5$).



(a) (b)

Figure 3: Effect of varying the AV-net group size (a) Average number of PIR queries (b) Average number of AV-net invocations

Finally, Fig. 4 illustrates the effect of the number of surrounding rings (AR) that are explored during an AV-net protocol invocation ($K = 0.5$, $GS = 5$, $PI = 0.1$). The first observation, is that the number of PIR queries remains almost constant (Fig. 4a). The reason is that, as shown in Fig. 4b, exploring one ring at a time merely results in more AV-net rounds, since users invoke a new AV-net protocol once they move further away from their current position. However, the overall PIR reduction is not affected, because users still get most of their spectrum knowledge from the P2P protocol. A value of $AR = 2$ seems like the best choice, given that the number of AV-net rounds does not decrease significantly from 2 to 3 rings.



(a) (b)

Figure 4: Effect of varying the AV-net exploration area (a) Average number of PIR queries (b) Average number of AV-net invocations

4. CONCLUSIONS

We proposed a novel approach that allows database-driven DSA mobile users to share anonymously their cached channel availability information that is obtained from previous PIR queries. Our experiments with a real-life dataset, indicate that our methods reduce the number of PIR queries by 50% to 60%. Furthermore, they are efficient in terms of both computational and communication cost.

Acknowledgments

This research has been funded by the NSF CAREER Award IIS-0845262.

5. REFERENCES

[1] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury. CRAHNs: Cognitive radio ad hoc networks. *Ad Hoc Networks*, 7(5):810–836, 2009.

[2] Z. Gao, H. Zhu, Y. Liu, M. Li, and Z. Cao. Location privacy in database-driven cognitive radio networks: Attacks and countermeasures. In *IEEE INFOCOM*, pages 2751–2759, 2013.

[3] F. Hao and P. Zieliński. A 2-round anonymous veto protocol. In *Security Protocols*, pages 202–211, 2009.

[4] J. Mitola III. Cognitive radio: An integrated agent architecture for software defined radio. *Doctoral Dissertation, KTH, Stockholm, Sweden*, May 2000.

[5] Y. Zheng, L. Wang, R. Zhang, X. Xie, and W.-Y. Ma. GeoLife: Managing and understanding your past life over maps. In *IEEE MDM*, pages 211–212, 2008.