ISSN: 2278-4632
Vol-10 Issue-7 No. 11 July 2020

LOCATION PRIVACY PRESERVATION IN WIRELESS COGNITIVE NETWORKS THROUGH ENCRYPTED PROBABILISTIC DATA STRUCTURES

Mr. G. DILIP KUMAR M.Tech

Assistant Professor, Department of CSE, Sri Mittapalli College of Engineering, Tummalapalem, NH-16, Guntur, Andhra Pradesh, India.

1) S.RAVITEJA , 2) G.UMESH CHANDRA , 3) V. M.GOPI REDDY , 4) K.B. SUBBARAO

B. Tech Students, Department of CSE, Sri Mittapalli College of Engineering, Tummalapalem, NH-16, Guntur, Andhra Pradesh, India.

Abstract—In this paper, we propose new location privacy preserving schemes for database-driven cognitive radio networks (CRN s) that protect secondary users' (SU s) location privacy while allowing them to learn spectrum availability in their vicinity. Our schemes harness probabilistic set membership data structures to exploit the structured nature of spectrum databases (DBs) and SU s' queries. This enables us to create a compact representation of DB that could be queried by SU s without having to share their location with DB, thus guaranteeing their location privacy. Our proposed schemes offer different cost performance characteristics. Our first scheme relies on a simple yet powerful two-party protocol that achieves unconditional security with a plausible communication overhead by making DB send a compacted version of its content to SU which needs only to query this data structure to learn spectrum availability. Our second scheme achieves significantly lower communication and computation overhead for SU s, but requires an additional architectural entity which receives the compacted version of the database and fetches the spectrum availability information in lieu of SU s to alleviate the overhead on the latter. We show that our schemes are secure, and also demonstrate that they offer significant advantages over existing alternatives for various performance and/or security metrics.

I. INTRODUCTION

Cognitive radio networks (CRN s) have emerged as a key technology for addressing the problem of spectrum utilization inefficiency. CRN's allow unlicensed users, also referred to as secondary users (SU s), to access licensed frequency bands opportunistically, so long as doing so does not harm licensed users, also referred to as primary users (PU s). In order to enable SU s to identify vacant frequency bands, also called white spaces, the federal communications commission (FCC) has adopted two main approaches: spectrum sensing based approach and geo-location database-driven approach. In the sensing-based approach, SU s themselves sense the licensed channels to decide whether a channel is available prior to using it so as to avoid harming PU s. In the database driven approach, SU s rely on a geo-location database (DB) to obtain channel availability information. For this, SU s are required to be equipped with GPS devices so as to be able to query DB on a regular basis using their exact locations. Upon receipt of a query, DB returns to SU the list of available This manuscript was published in: Computer Networks and Information Security (WSCNIS), 2015 World Symposium on. channels in its vicinity, as well as the transmission parameters that are to be used by SU . This database-driven approach has advantages over the sensing-based approach. First, it pushes the responsibility and complexity of complying with spectrum policies to DB. Second, it eases the adoption of policy changes by limiting updates to just a handful number of databases, as opposed to updating large numbers of devices.

ISSN: 2278-4632
Vol-10 Issue-7 No. 11 July 2020

A. Our Contribution

In this paper, we propose two location privacy-preserving schemes for database-driven CRN s with different performance and architectural benefits. The first scheme, location privacy in database-driven CRNs (LPDB), provides optimal location privacy to SU s within DB's coverage area by leveraging set membership data structures (used to test whether an element is a member of a set) to construct a compact version of DB. The second scheme, LPDB with two servers (LPDBQS), minimizes the overhead at SU 's side at the cost of deploying an additional entity in the network. The cost performance tradeoff gives more options to system designers to decide which topology and which approach is more suitable to their specific requirements.

II. METHODOLOGY

Despite its importance, the location privacy issue in CRN s only recently gained interest from the research community. Some works focused on addressing this issue in the context of collaborative spectrum sensing while others focused on addressing it in the context of dynamic spectrum auction. However, these works are not within the scope of this paper as we focus on the location privacy issue in database-driven CRN s.

III. SYSTEM MODEL AND SECURITY ASSUMPTIONS

A. Database-driven CRN Model

We first consider a CRN that consists of a set of SU s and a geo-location database (DB). SU s are assumed to be enabled with GPS and spectrum sensing capabilities, and to have access to DB to obtain spectrum availability information within its operation area. To learn about spectrum availability, a SU queries DB by including its location and its device characteristics. DB responds with a list of available channels at the specified location and a set of parameters for transmission over those channels. SU then selects and uses one of the returned channels. While using the channel, SU needs to recheck its availability on a daily basis or whenever it changes its location by 100 meters as mandated by PAWS .

B. Security Model and Assumptions DB and QS are assumed to be honest but curious. That is, DB and QS follow the protocol honestly but may try to infer information on the input of other parties beyond what the output of the protocol reveals. Specifically, our objective is to prevent these two entities from learning SU s' location. Therefore, our security assumptions are as follows: Security Assumption 1. DB and QS do not modify the integrity of their input. That is, (i) DB does not maliciously change SU 's query's content; (ii) QS does not modify the input that it receives from DB or SU . Security Assumption 2. DB and QS do not collude with each other to infer the location of SU s from their queries.

IV. SET MEMBERSHIP DATA STRUCTURES

Our proposed privacy-preserving schemes utilize set membership data structures to exploit the highly structured property of DB. There are several data structures that are designed for set membership tests, e.g. bloom filter, cuckoo filter, etc. However, in this paper, we opt for cuckoo filter as the building block of our schemes. We use cuckoo filter to construct a compact representation of the spectrum geolocation database as explained in Sections V-A & V-B. What motivates our choice is that cuckoo filter offers the highest space efficiency among its current well known alternatives, such as bloom filters. Besides, it has been proven to be more efficient than these alternatives especially for large sets

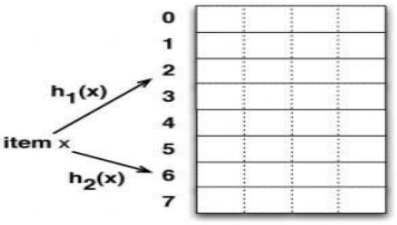


Fig: Cuckoo Filter: 2 hashes per item, 8 buckets each containing 4 entries

V. PROPOSED SCHEMES

In this section, we describe our proposed schemes. The first scheme, LPDB, is simple as it involves only two parties, SU s and DB, and provides unconditional location privacy to SU s within the coverage area of DB. The second scheme, LPDBQS, offers computational privacy with a significantly reduced overhead on SU s' side compared to LPDB, but at the

ISSN: 2278-4632
Vol-10 Issue-7 No. 11 July 2020

cost of introducing an extra architectural entity.

A. LPDB

In this section, we describe our basic scheme, which is referred to as location privacy in database-driven CRN s (LPDB). The novelty of LPDB lies in the use of set membership data structures to construct a compact (space efficient) representation of DB that can be sent to querying SU s to inform them about spectrum availability. In our scheme, instead of sending its location, a SU sends only its characteristics (e.g., its device type, its antenna type, etc.), as specified by PAWS to DB, which then uses them to retrieve the corresponding entries in all possible locations. DB then puts these entries in a cuckoo filter and sends it to SU . Upon receiving this filter, SU constructs a query that includes its characteristic information, its location, and one of the possible channels with its associated parameters. SU then looks up this query in the received cuckoo filter to see whether that channel is available in its current location.

B. LPDBQS

In this section, we propose a new scheme, LPDBQS, which offers better performance at SU s' side than that of LPDB. This comes at the cost of deploying an additional entity, referred to as query server (QS), and having a computational security as opposed to unconditional. QS is introduced to handle SU s' queries instead of DB itself, which prevents DB from learning information related to SU s' location information. QS learns nothing but secure messages sent by SU s to check the availability of a specific channel. • Intuition: We introduce QS to avoid sending CF, which might be large, to SU . Instead, CF, that contains HMAC secure entries inserted by DB using a secret key provided by SU , is sent to QS through a high throughput link pre-established with DB. SU just needs to query, using HMAC messages, QS which looks for its queries in CF. Using HMAC , SU can hide the content of the query string, which includes its location information, among others, from QS which ignores the key used to construct the hashed query and the CF. This not only prevents QS from learning the query's content but also the entry that matches it in the filter. As most of the computation and communication overhead are incurred by both DB and QS, this scheme is the most efficient in terms of overhead incurred by SU's.

VI. SECURITY ANALYSIS

In this section, we analyze the security of our proposed schemes LPDB and LPDBQS. Theorem 1. Under Security Assumptions 1 and 2, LPDB does not leak any information on SU s' location. Proof: We construct a history list H of each entity's knowledge about SU s' information during the execution of LPDB. SU . A SU cannot learn anything about other SU s information nor the filters {CFi,t} n−1,tf i=1,t=t0 that they receive from DB as the communication between each SU and DB is secured, i.e. HSU = ∅. Note that, even if, a SU would learn the filters of other SU s, i.e. HSU = {CFi,t} n−1,tf i=1,t=t0 , HSU includes no information about SU s' location. DB. In Step 1 of Algorithm 1, DB learns HDB = {char} n i=1 which contains the characteristics of the querying SU s. HDB may include information like frequency ranges in which SU can operate, antenna characteristics, etc. This information is not related to the querying SU s' location. This shows that the knowledge that DB gains during the execution of LPDB does not allow it to infer SU s' location when they try to learn about spectrum opportunities. LPDB offers an unconditional privacy in the sense that DB's knowledge about SU s' location, during the execution of LPDB, does not increase compared to its initial knowledge, which is necessarily the coverage area of DB. Theorem 2. Under Security Assumptions 1 and 2 LPDBQS does not leak any information about SU s' location beyond κ − HMAC secure values. Proof: We construct a history list of each entity's knowledge during the execution of LPDBQS. SU . As the communication between different entities is secured, SU s cannot learn any information about the communicated information of other entities, i.e. HSU = ∅. DB. In Line 1 of Algorithm 2, DB learns HDB = {ki,t, char i , ts t} n,tf i=0,t=t0 . Obviously, SU s' secret keys {ki,t} n,tf i=0,t=t0 and timestamp values {ts t} tf t=t0 cannot leak any information about SU s' location since these values are not correlated to their physical location. Similarly, their characteristics {char i} n i=1 contain information about SU s' devices capabilities, like their possible transmit powers, antennas

ISSN: 2278-4632
Vol-10 Issue-7 No. 11 July 2020

height, etc, which cannot be used to localize them. This proves that DB's knowledge about SU s' location during the execution of LPDBQS does not differ from its initial knowledge; i.e. that SU s are within DB's covered area. QS. As indicated in Lines 8 & 12 of Algorithm 2, the only information that QS can learn during the execution of LPDBQS, is HQS = {yki,t , CFki,t } n,tf i=1,t=t0 . {yki,t } n,tf i=1,t=t0 are as secure as HMAC . The elements of {CFki,t } n,tf i=1,t=t0 are computed using a pseudo random function (as an HMAC is also a pseudo random function) with SU s' secret keys {ki,t} n,tf i=1,t=t0 , where {ki,t} n,tf i=1,t=t0 $\leftarrow$ {0, 1} κ and κ is the security level. {yki,t } n,tf i=1,t=t0 are independent from each other. The same applies to {CFki,t } n,tf i=1,t=t0 . Each query from {yki,t } n,tf i=1,t=t0 has a corresponding HMAC key, which means that even for the same SU querying the same information, there will be randomly independent and uniformly distributed outputs generated by DB and SU s. Since only SU s and DB know the keys {ki,t} n,tf i=1,t=t0 and that these keys are updated for every query made by SU s, QS cannot learn any information about SU s' location as long as it does not collude with DB as stated in Security Assumption 2. Correlating queries {yki,t } n,tf i=1,t=t0 to SU s' physical location is equivalent to breaking the underlying HMAC or P RF, which is of probability 1/2 κ . We can conclude that LPDBQS is as secure as the underlying HMAC .

VII. EVALUATION AND ANALYSIS

In this section, we evaluate the performance of our proposed schemes. We consider that DB's covered area is modeled as a √ m × √ m grid that contains m cells each represented by one location pair (locX ,locY) in DB. We use the efficient cuckoo filter implementation provided in for our performance analysis with a very small false positive rate = 10−8 and a load factor α = 0.95. In addition, since personal/portable TVBD devices of SU s can only transmit on available channels in the frequency bands 512 − 608 MHz (TV channels 21−36) and 614-698 MHz (TV channels 38−51), this means that users can only access 31 white-space TV band channels in a dynamic spectrum access manner. Therefore, in our evaluation we set the number of TV channels s = 31.

A. Communication and Computation Overhead

1) Communication Overhead: We provide analytical expressions of the communication overhead of these schemes in Table II. For LPDB, we provide two expressions of the overhead with respect to two scenarios: (i) when SU s do not reveal one of their coordinates, (ii) when one of the coordinates is revealed by SU s. In both scenarios the data transmitted consist basically of query, sent by SU , and the response of DB to it.

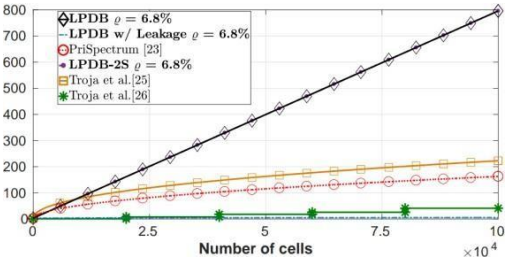


Fig: Communication Overhead

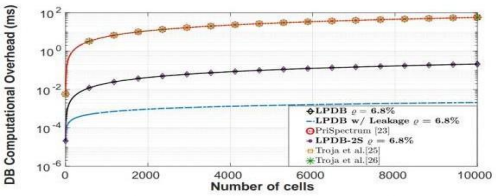
2) Computational Overhead:

We also investigate the efficiency of our proposed schemes in terms of their computational overhead. We evaluate the computation required at each entity separately, and we provide the corresponding analytical expression of the overhead.

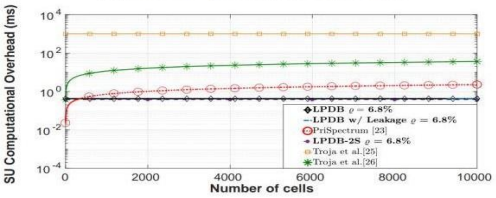
Again we provide two estimated costs for both scenarios of LPDB. The computation of DB is given in terms of the number of insertions it has to perform into CF. This depends on the number of DB entries that comply with query considering only the available channels.

This number is equal to %·s ·m in LPDB and reduces to % · s · √ m in LPDB with leakage. For the computational cost at the SU 's side, LPDB's overhead depends solely on the number of possible channels, s, and the cost of one Hash and one Lookup operations.

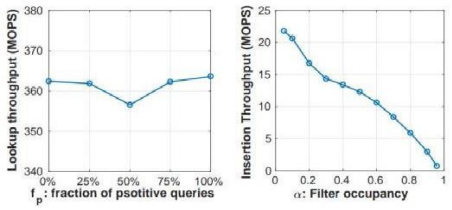
ISSN: 2278-4632
Vol-10 Issue-7 No. 11 July 2020



(a) DB Computational Overhead.



(b) SU Computational Overhead.



(a) (b)

Fig: Lookup Performance

We also assess the insertion throughput that DB experiences to construct the CF as a function of the load factor α as shown in Figure. As opposed to the lookup throughput shown in Figure a, CF has a decreasing insert throughput when it is more filled (though their overall construction speed is still high). This is mainly due to the fact that CF may have to move a sequence of existing fingerprints recursively before successfully inserting a new item, and this process becomes more expensive when the load factor grows higher.

3) Impact of varying the percentage % of entries with available channels: We also study the impact of % on the overhead incurred by our schemes. For this, we plot in Figure 6 the communication and the system computational overheads for different values of %. We plot only LPDB and LPDB with leakage as LPDBQS has almost the same overhead as LPDB.

B. Location privacy

We compare our schemes to existing approaches in terms of location privacy level by presenting the security problems on which they rely as illustrated in Table III. We also precise the localization probability of SU s under these schemes. The best probability that could be achieved is 1/m, i.e. SU s are within DB coverage area. If one of the schemes is broken then this probability increases considerably. LPDB offers unconditional security, as SU s do not share any information that could reveal their location. LPDB could be seen as a variant of PIR in which the server sends a whole copy of the database to the user and this is the only way to achieve information theoretic privacy (i.e. cannot be broken even with computationally unbounded adversary) in a single server setting. Even if one of the coordinates is intentionally revealed by a SU , its location is still indistinguishable from $\sqrt{m} - 1$ remaining possible locations.

TABLE III: Location privacy

Scheme	Security level	Localization probability
LPDB	Unconditionally secure	$1/m$
LPDB w/ leakage	Uncond. within 1 coordinate	$\sqrt{1/m}$
PriSpectrum [23]	Computational *PIR*	$1/m$
Troja et al [26]	Computational *PIR*	$1/m$
Troja et al [25]	Computational *PIR*	$1/m$
LPDBQS	$\kappa - HMAC$	$1/m$
Zhang et al. [22]	*k-anonymity*	$1/k$
Zhang et al. [27]	Geo-Indistinguishability	$1/r$

Variables: r is the radius of the ϵ-geo-indistinguishability mechanism in [27].

VIII. CONCLUSION

In this paper, we have proposed two location privacy preserving schemes, called LPDB and LPDBQS, that aim to preserve the location privacy of SU s in database-driven CRN s. They both use set membership data structures to transmit a compact representation of the geo-location database to either SU or QS, so that SU can query it to check whether a specific channel is available in its vicinity. These schemes require different architectural and performance tradeoffs.

IX. REFERENCES

[1] M. Grissa, A. A. Yavuz, and B. Hamdaoui, “Cuckoo filter-based location-privacy preservation in database-driven cognitive radio networks,” in WSCNIS. IEEE, 2015, pp. 1–7.

[2] “Spectrum policy task force report,” Federal Communications Commission, Tech. Rep. ET Docket No.02-135, 2002.

[3] B. Khalfi, M. B. Ghorbel, B. Hamdaoui, and M. Guizani, “Optimal power allocation for smart-grid powered point-to-point cognitive radio system,” in ComComAp, 2014 IEEE, pp. 316–320.

[4] N. Adem and B. Hamdaoui, “The impact of stochastic resource availability on cognitive network performance: modeling and analysis,” Wireless Communications and Mobile Computing,2015.