Agent-based Channel Selection Scheme against Location Inference Attack in Cognitive Radio Networks

Hongning Li¹, Qingqi Pei^{1,2} and Yao Liu²

¹The State Key Laboratory of Integrated Services Networks, Xidian University, Shaanxi, Xi'an, China ²Department of Computer Science and Engineering, University of South Florida, Tampa, FL, USA hnli@xidian.edu.cn, qqpei@mail.xidian.edu.cn, yliu@cse.usf.edu

Abstract-In database-driven cognitive radio networks, location inference attack may cause location privacy leakage of secondary users. Malicious entities could geo-locate secondary users with spectrum utilization information. The efficiency of existing solution largely depends on the stability of available spectrum. Most solutions are not effective because of the unpredictable return of primary users. To prevent location inference attack, we propose an Agent-based Channel Selection(ACS) scheme in this paper. In this scheme, the database allocates spectrum with self-coexistence mechanism to avoid interference, and each base station registers spectrum information as an agent in the database instead of secondary users (SUs). Thus, the proposed scheme can decrease the probability of being geo-located for SUs effectively. According to the simulation, we analyze the relationship among the factors that impact the probability and the efficiency of the proposed scheme.

Index Terms—Cognitive Radio Networks, spectrum allocation, location privacy, database.

I. INTRODUCTION

Cognitive radio networks (CRNs) has been widely accepted as an effective way to improve spectrum efficiency of wireless communication systems. It allows SUs to communicate with each other within the tolerance of primary users (PUs)[1], such that limited spectrum can be utilized more efficiently. As a precondition of CRNs, spectrum sensing can collect the states of licensed spectrum to get the availability spectrum information. Then the system can analyze and allocate the idle spectrum to SUs. Different kinds of sensing technologies have been proposed and lay a good foundation for the usability of CRNs [2], [3], [4], [5], [6], [7]. Most assume that there is an entity whose behaves like a fusion center to collect the sensing data from all SUs and make final decision. There are two ways to determine which licensed channels are locally available for SUs, namely, one is spectrum sensing, the other is using white space database [8]. In spectrum sensing, SUs should sense channels and decide whether the sensed channels are used by PUs. In white space database, SUs can query the database to get the channel availability information at their location. We call the latter a database-driven CRN. In database-driven CRNs, to access an idle channel, a SU has to first query database (DB) to obtain the Spectrum Availability Information (SAI), and then register the information of the channel which has been allocated so that DB can allocate other channels without interference. Since available channels is different in different areas or cells, DB should allocate channels according to SUs' location. Thus the channel utilization information has close relation to SUs' location privacy. For example, the regional distribution feature of available channels would reveal the fact that there exists an inalienable relationship between the location of secondary users and the available channels that they can access. Adversaries can infer secondary users' location through their used channels. With this feature, attackers can find the locations of SUs and launch attacks. We call this kind of attack location inference attack. Even the final goal of SUs in CRNs is to access idle channels, they are unwilling to leak private information. In dynamic spectrum access environment, protecting the location privacy is a big challenge because in the current database-driven CRNs, the location information has direct relationship to the channel usage. Adversaries can launch location inference attack easily. Therefore, users' property or even personal safety can not be guaranteed if location information has been leaked. For instance, a service provider or an adversary can track the whereabouts of a user and discover his/her personal habits. These sensitive information can be sold to third parties without the user's consent. With users' location information, an adversary can monitor the behavior of individuals, or know the places they have visited, even burglary when users are absent. Therefore, location privacy protection should be investigated in CRNs.

Recently, in [8], Gao et al. proposed a scheme to protect SUs' location in database-driven CRNs. Experiment results showed that the proposed protocols can significantly improve location privacy, but it mainly relies on the stability of channels selected by SUs. Besides, each SU must store a used channel list and a prediction list in order to select stable channels. In this paper, we propose an Agent-based Channel Selection (ACS) scheme for database-driven CRNs, which can decrease the probability of location leakage effectively by using self-coexistence (we define the coexistence among secondary users as self-coexistence in order to distinguish the coexistence between primary users and secondary users) mechanism and agent-based registration. It is comprised of two modules, namely, secure query and unified channel allocation. Secure query can guarantee the legitimate SUs to pass the authentication without leaking their identities. Unified channel allocation can hide the relationship between SUs and the information of their used channels.

The rest of this paper is organized as follows. Section II introduces related work and explains the privacy leakage problem in database-driven CRNs. Section III gives the system architecture and the model of location inference attack of CRNs. Section IV presents the proposed ACS scheme. Section V discusses and analyzes the proposed scheme. Section VI concludes this paper.

II. RELATED WORK

More and more users are paying close attention to their privacy recently [9], [10], [11], [12]. There are many services such as location-based services, friend finder services (e.g. Loopt, which determines all friends in the vicinity of a user)[13], or social networks as Facebook, where users can show their pictures to other users online. Although these services are very popular, their usage raises severe privacy concerns. For example, by collecting location information of users, adversaries can infer sensitive privacy information about users, such as their home location, life styles, work place and health conditions [14]. Therefore, it is important to design some strategies to protect users' privacy in communication networks.

In CRNs, data leakage of SUs is a serious problem, from which attackers can locate SUs and get sensitive information. Spectrum sensing and spectrum allocation are two main stages that may leak SUs' privacy. During spectrum sensing, SUs' sensing data is related to their location. Inside adversaries can infer the location of a SU from its sensing report. To address the location leakage problem in cooperative spectrum sensing, Li et al. proposed a Privacy Preserving collaborative Spectrum Sensing (PPSS) scheme [15], which includes two primitive protocols: Privacy Preserving Sensing Report Aggregation protocol (PPSRA) and Distributed Dummy Report Injection Protocol (DDRI). It can significantly improve SUs' location privacy with a reasonable security overhead in collaborative sensing. However, once there exists one or more SUs whose sensing reports are failed during transmission due to some reason, such as channel fading, the fusion center cannot aggregate all the reports, and thus the protocol would become invalid.

During the spectrum allocation, if the location of SUs is relatively fixed, given PUs' states, we can get the list of available channels in different areas. Based on the given information, attackers can infer SUs' location by its used channel information. Gao et al. proposed a novel prediction based Private Channel Utilization protocol to mitigate the possibilities of location privacy leakage by choosing the most stable channel [8]. However, in this scheme, each SU should store and update the stability level of all the channels and select higher one to access, and the efficiency mostly depends on the stability of channels.

In some scenarios, PUs will share licensed spectrum with SUs and SUs need to pay to PU according to special rules or agreement. Therefore, SUs' detailed usage information, such as when and how long the licensed spectrum is used is needed for PUs to calculate the payment. However, the information



Fig. 1: Process of channel application in database-driven CRNs

may compromise SUs' privacy. To solve this dilemma, in [16], Qin et al. proposed a novel privacy preserving mechanism for cognitive radio transactions through commitment scheme and zero-knowledge proof to preserve SUs' privacy while using a random-checking monitor to protect the primary user's interests. The result shows that the proposed scheme can guarantee that the payment is correctly calculated with minimal SU's usage information, but the scenario is limited.

In this paper, to prevent location leakage from channel usage, we propose an agent-based channel selection scheme. Firstly, considering the increasing number of SUs and the limited available spectrum, the proposed scheme achieves a non-interference allocation with self-coexistence mechanism and power control method. Secondly, to increase adversaries' difficulty to infer SUs' location from channel usage information, the proposed scheme uses agent to register instead of SUs to cut off the relationship between SUs' location and register data. With the same times of channel switches, our scheme can protect SUs' location privacy more effectively.

III. LOCATION INFERENCE ATTACK IN DATABASE-DRIVEN CRNs

A. Basic system architecture

A typical database-driven CRN consists of a set of PUs, a set of SUs and a DB. Each single SU can request channels through DB, as shown in Figure 1. FCC allows SUs to periodically access DB to obtain a list of available channels for their services. If there is a SU, say, Alice, who wants to communicate with others, first, Alice should connect to the DB and send a query to get spectrum availability information which can be used in the current area where Alice is located. The query includes the location of Alice. After receiving the list of available channels, Alice selects one channel and registers it in the DB so that the DB can reallocate the channel to other SUs without interference among SUs.

B. Attack model in database-driven CRNs

In Figure 1, there are two kinds of communication entities: PUs and SUs. We assume that:

(i) The DB is semi-honest, i.e., curious but not malicious, which means that it exactly follows the protocol but may learn some information about the registered data, or even leak the registered data to adversaries to get benefits.



Fig. 2: The process of Location Inference Attack

- (ii) The adversary knows the SU's channel usage information but does not know the SU's query and identity.
- (iii) The adversary knows the spectrum availability information (SAI) of the whole network.
- (iv) The location of SUs is relatively fixed during a certain interval.

As mentioned earlier, [8] showed that the adversary can infer a SU's location by its channel usage information after changes of available channels. Gao et.al named this kind of attack as location inference attack. With this attack, adversaries can geolocate a SU. The process can be described in Figure 2. The unavailable area of a channel is the area being used by the PU, while the available area of the channel is the complement. If an attacker obtains the SAI, it can infer in which area the channel can be used. With the SU's channel usage information, the attacker can infer the SU's location. For example, in Figure 2, the available areas of the channels vary over three time slots t_1 , t_2 and t_3 . In time slot t_1 , SUs in the available areas can use ch_1 but SUs in other areas can not. In time slot t_2 , SUs in the available areas can use ch_2 but SUs in other areas can not. In time slot t_3 , SUs in the available areas can use ch_3 but SUs in other areas can not. If a SU is using ch_1 , ch_2 and ch_3 in time slot t_1 , t_2 , t_3 respectively, adversaries can infer its location by the intersection of the available areas in the three time slots. After three time slots, the SU can be positioned. [8] also described two other kinds of attacks, namely, Primary User Coverage Complement Attack (PUCC) and Enforced Channel Switch Attack (ECS), respectively. Attackers can infer the SU based on the change of the channel state or the switch of the SU.

From above observation, we conclude that given channel state (occupied by PU or not), DB or malicious users can locate SUs by their used channels. Therefore, how to effectively hide channel selection information to prevent location privacy leakage is an important issue in database-driven CRNs.

IV. AGENT-BASED CHANNEL SELECTION SCHEME

To defense location inference attack, in this section, we propose an Agent-based Channel Selection (ACS) scheme. The target of ACS scheme is to obtain a non-interference channel allocation with small probability of location privacy leakage of SUs.

The agent-based channel selection (ACS) scheme includes two modules, which are shown in Figure 3. The first part is the secure query. All legitimate SUs can pass and get the rights to



Fig. 4: Main architecture of database-driven CRNs

request for available channels. The second part is the dummy channel allocation with self-coexistence mechanism. During this process, DB can allocate channels to users with selfcoexistence mechanism. This can not only improve channel usage, but also confuse adversaries to obtain channel usage information of SUs. The last part is channel registration, in which each BS works as an agent to help DB to allocate spectrum more effectively. In this paper, we consider the system architecture as shown in Figure 4. The whole area covered by CRNs is defined as C. There are N base stations in the network. C_i represents the area covered by BS_i , and we also call the area covered by a BS a cell. All SUs in C_i can request channels from DB through BS_i . In the architecture shown in Figure 4, SUs can communicate with limited power to guarantee that they do not interfere with other areas. Adjacent areas can be allocated with different channels, and adopt self-coexistence mechanism during spectrum allocation. The object function of ACS scheme is an allocation matrix A_L ,

$$A_{L} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1q} \\ A_{21} & A_{22} & \dots & A_{2q} \\ \dots & \dots & \dots & \dots \\ A_{p1} & A_{p2} & \dots & A_{pq} \end{bmatrix}$$
(1)

where $p \times q = N$ and N is the total number of cells. A_{ij} is the set of channels allocated to the cell. The constraint condition of A_L is that: if $ch \in A_{ij}$, $1 \le i \le p, 1 \le j \le q$, $ch \notin (A_{i\pm 1,j} \cup A_{i,j\pm 1})$.

AS shown in Figure 5, our ACS scheme includes channel request, channel allocation and unified register. In ACS scheme, DB can allocate idle spectrum with self-coexistence mechanism and each BS acts as an agent to register channels in the DB to cut off the relation between registered information and SUs' identities. The self-coexistence and the unified register with dummy injection can increase the spectrum utilization and decrease the location leakage probability simultaneously.

The steps of ACS scheme are described as follows:



Fig. 5: Process of ACS scheme

Channel request: SU_j in the area C_j sends the encrypted message to BS_j located in C_j . We use public key encryption in this paper just for demonstration.

After BS_j receives all channel request messages at time t, it will decrypt the messages and send the total requests of the SUs in C_i with dummy information to the DB.

Channel allocation: The DB allocates idle spectrum to all BSs requested with self-coexistence mechanism. In our paper, we use coloring model to achieve the self-coexistence channel allocation. Besides, to avoid interference, all SUs must obey the power control rules in CRNs. Then, BS_j allocates channels to SUs in C_j according to the available channel list and users' requests.

The available channel list can be formed with coloring model according to spectrum availability information (SAI). It is a subset of SAI. Each available channel list is different from the available channel lists in adjacent areas/cells. Thus, we can achieve the non-interference spectrum allocation in the whole network.

Unified Register: After allocation, BS_i will register the channels allocated to itself in order to help the DB to allocate spectrum without interference. In this step, the channels allocated to BS_i include not only the channels used by SUs in C_j , but also the dummy information. The process is shown in Figure 5. During channel request in ACS scheme, SUs request channels through the BSs that they belong to. After channel allocation, BSs would register the selected channels instead of SUs so that DB can assign channels to other SUs in other cells without interference. In the scheme, DB does not store SUs' used channel information, which can prevent privacy leakage by inferring the location from used channel. Besides, we adopt hash matching, which can not only verify SUs' identities, but also increase difficulty of attackers. Even if knowing someone's location, attackers have no knowledge about the identity of the user.

In our scheme, there are two important points: firstly, the same channel can be used in multi-cells/multi-areas, the selfcoexistence mechanism and power control can guarantee noninterference. It also helps the system to make full use of available spectrum. Secondly, each BS works as an agent to register channel usage information with dummy injection instead SUs, and this can confuse adversaries or cut off the directly relation between SUs' location and the register data.

V. ANALYSIS OF ACS SCHEME AGAINST LOCATION INFERENCE ATTACK

In this section, we first evaluate the efficiency of our ACS scheme against location inference attack and give the comparison with the solution existed. Then we analyze the factors/parameters that can affect the average probability of geo-location and give the variation curves with different sets of the parameters mentioned above.

A. Efficiency of ACS scheme against location inference attack

In traditional channel selection schemes in database-driven CRNs, attackers can infer the locations of SUs from their used channels and the list of available channels of each area. The more channels a SU switches or the more frequently the available channel states change, the more quickly of the SU will be fixed. To infer a SU's location, attackers need the following information: the list of available channels for each area, identity of the targeted SU, used channel information of the SU. By collecting all information above, attackers can infer SU's location. The first step taken by attackers is to obtain distribution of available list in different areas. In the Spectrum Availability Information Retrieval scheme proposed in [8], Gao et.al utilized a blind factor to hide the location of each SU. Each SU can only retrieve the available channel list at its own location. In this case, malicious users can send multi-query to obtain SAI of different areas easily even if DB is trusted. They can obtain SAI of all C_i with the query blinding factor. In our scheme, only BSs query DB to obtain the SAI of their coverage area. Any SU has no knowledge of the SAI in the coverage area of the BS it belongs to. It is impossible for SUs to send query to DB.

In our ACS scheme, if location inference attackers obtain the registered data in DB, they can only get the used channel information of each BS, but have no knowledge of the used channel information of each SU. Therefore, knowing only the registered data, attackers cannot infer SU's location, they can only obtain how many SUs has requested channels under a certain base station, but cannot distinguish the SUs within the coverage of the same base station.

Given ID_A , the identity of SU_A and the registered data in DB, the probability that an attacker can infer the SU's location is 1/N, where N is the number of base stations in the whole network. The positioning precision function as follows:

$$F_{Loc} = \frac{1}{N} f(\mathbf{S}_N / r^2) \tag{2}$$

 $f(S_N/r^2)$ is an increasing function of S_N/r^2 , and it was used to normalize F_{Loc} , which is inversely proportional to N. S_N is the coverage area of the whole network and r is the radius of each base station. The larger N is, the smaller F_{Loc} is for a SU. The variation tendency is shown in Figure 6. For illustration we set $S_N = 1$.

If an attacker captures a base station BS_j or the base station BS_j itself is malicious, it can obtain the used channel information of SU_A , but it cannot distinguish which SU. Thus, the probability of that the secondary user is the targeted user is $1/N_{user}$, where N_{user} is the number of SUs in the whole



Fig. 6: The variation tendency of positioning precision



Fig. 7: Comparison of the probability to be located

network. Besides, the location precision of the SU is also related to the size of area covered by base station BS_{j} .

DB allocates channels based on coloring model and SUs? requests. The same channel can be reused among different base stations without interference. The number of available channels, the coverage area and the number of base stations may affect the probability of positioning precision. Even obtaining the identity of SUA and its used channels, the attacker can only locate the SU_A with probability of $1/N_{ch}$ at one time, where, N_{ch} is the number of base stations that register the same channel ch. Due to the uncertainty of PUs' signal, the state of the same channel is also changing with time. For our study here, we set that there are 20 channels which are open to SUs, and the states of the channels change randomly under the condition that there are 10 channels are available for SUs to access each time. Under this condition, we compare our ACS scheme with the PCU algorithm [8] in this section. From Figure 7 we can see that the average probability of SUs to be positioned is increasing to 1 with the increasing number of channel switches. In our simulation, we assume that the SAI changes each hour which means after one hour, the channel states will be changed. Therefore, after some hours(or handoffs), the average probabilities of both schemes reach 1. However, in the real scenario, it is impossible for SUs to switch so frequently. With fewer times of switches, the average probability with ACS scheme is much lower than the PCU algorithm in the same scenario of CRNs.



Fig. 8: The average probability of geo-location with different No. of channels allocated to each BS, $((i \ j \ k))$ means there are *i* BSs and *j* available channels, each BS has been allocated *k* channels)

B. Impacts of different factors on the average probability of geo-location

The probability of SUs to be located is related to the total number of available channels, the allocation model (coloring model in this paper) and the number of base stations. It is also partly depends on the states of PUs. We assume the states of channels are randomly changed in this paper. For differen number of channels allocated to each area, different number of available channels and different number of base stations, we give the variation curves in Figure 8, Figure 9, and Figure 10, respectively. p is the average probability of a SU positioned to a certain base station. From Figure 8, we can see that if the number of base stations remains unchanged. the times of the same channel will be multiplexed more with the increasing average number of channels allocated to each base station. Given the data registered in DB, the more times the same channel multiplexed, the smaller of the probability of positioning precision for the attacker. From Figure 9, we can obtain the relationship between the number of available channels and the average probability of positioning precision. If the number of base stations and the number of channels allocated to each base station are fixed, the average probability of positioning precision increases as the available channels increase. Figure 10 shows that with the fixed number of available channels of the whole network and the fixed number of available channels allocated to each base station, the average probability of positioning precision decreases as the number of base stations increases. All the simulations in 8, Figure 9, and Figure 10 assume that the number of available channels is less than the number of base stations.

Under the condition that SUs are fixed, attackers can infer SUs' locations based on multi-switch of SUs. However, in real networks, many SUs are mobile. If a SU moves to another place before attackers infer its location, the effect of the location inference attack will be decreased. Therefore, the probability of positioning precision is also related to the move speed of SUs.

Compared to the scheme proposed in [11], the ACS scheme proposed in this paper has the following advantages. First, it



Fig. 9: The average probability of geo-location with different No. of available channels



Fig. 10: The average probability of geo-location with different No. of BSs

adopts self-coexistence mechanism and dummy injection, with the increasing number of available channels, the probability to be geo-located of SUs will be much smaller. Besides, it does not so strongly depend of the stability of channel states. Second, the scheme adopts unified register by base stations as agents during channel allocation. The register information only includes the used channels and dummy information of each agent, and no information about the SUs who are using the channels. Therefore, attackers cannot infer SUs' location only with the registered data in DB. Moreover, selfcoexistence mechanism is used in the scheme to allocate spectrum resource among different agents (base stations) with dummy information, and it can guarantee SUs to communicate without interference simultaneously and maximize the resource utilization with high QoS.

VI. CONCLUSION

We proposed an ACS scheme to defense location inference attack. Our scheme can protect SUs' location privacy against untrusted DB by decreasing location inferring probability of attackers. We also identified the main factors that can affect the average probability of positioning precision through simulation results. By self-coexistence mechanism and agent-based register, the proposed ACS scheme can decrease the probability to be geo-located. With the same stability of multiple channel states, our scheme can protect SUs' location more efficiently.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grant No.61373170) and the Key Program of NSFC-Guangdong Union Foundation (Grant U1401251).

REFERENCES

- I. Mitola, J. and J. Maguire, G. Q., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.
- [2] X. Sun, L. Chen, and D. H. K. Tsang, "Energy-efficient cooperative sensing scheduling for heterogeneous channel access in cognitive radio," in *Proc. IEEE Computer Communications Workshops (INFOCOM WKSHPS)*, 2012, pp. 145–150.
- [3] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE New Frontiers* in *Dynamic Spectrum Access Networks*, 2005, pp. 131–136.
- [4] M. Abdelhakim, L. Zhang, J. Ren, and T. Li, "Cooperative sensing in cognitive networks under malicious attack," in *Proc. IEEE Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3004–3007.
- [5] G. Ganesan and L. Ye, "Cooperative spectrum sensing in cognitive radio, part ii: Multiuser networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2214–2222, 2007.
 [6] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing
- [6] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. IEEE INFOCOM*, ser. 2008 Proceedings IEEE INFOCOM. IEEE, April 2008, pp. 1–5.
- [7] P. Kaligineedi, M. Khabbazian, and V. K. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *Proc. IEEE ICC*, 2008, pp. 3406–3410.
- [8] Z. Gao, H. Zhu, Y. Liu, M. Li, and Z. Cao, "Location privacy in database-driven cognitive radio networks: Attacks and countermeasures," in *Proc. IEEE INFOCOM*, 2013, pp. 2751–2759.
- [9] B. Gedik and L. Liu, "Protecting location privacy with personalized kanonymity: Architecture and algorithms," *IEEE Transactions on Mobile Computing*, vol. 7, no. 1, pp. 1–18, 2008.
- [10] W. Zhang, X. Cui, D. Li, D. Yuan, and M. Wang, "The location privacy protection research in location-based service," in *Proc. 18th International Conference on Geoinformatics*, 2010, pp. 1–4.
- [11] K. Mehta, D. Liu, and M. Wright, "Protecting location privacy in sensor networks against a global eavesdropper," *IEEE Transactions on Mobile Computing*, vol. 11, no. 2, pp. 320–336, 2012.
- [12] Y. Bidi and D. Makrakis, "Protecting location privacy with clustering anonymization in vehicular networks," in *Proc. IEEE Computer Communications Workshops (INFOCOM WKSHPS)*, 2014, pp. 305–310.
- [13] L. Barkhuus, B. Brown, M. Bell, S. Sherwood, M. Hall, and M. Chalmers, "From awareness to repartee: sharing location within social groups," in *proceedings of the SIGCHI conference on human factors in computing* systems. ACM, 2008, pp. 497–506.
- [14] G. Gabriel, "Private queries and trajectory anonymization: a dual perspective on location privacy," *Transaction on data privacy*, pp. 3–19, 2009.
- [15] S. Li, H. Zhu, Z. Gao, X. Guan, K. Xing, and X. Shen, "Location privacy preservation in collaborative spectrum sensing," in *Proc. IEEE INFOCOM*, 2012, pp. 729–737.
- [16] Z. Qin, S. Yi, Q. Li, and D. Zamkov, "Preserving secondary users" privacy in cognitive radio networks," in *Proc. IEEE INFOCOM*, 2014, pp. 772–780.