An Anti-sniffing Protocol for Location-based Services in Wireless Networks

Greg Yera
Department of Computer Science
Lamar University
Beaumont, USA
gyera@lamar.edu

Dr. Xingya Liu
Department of Computer Science
Lamar University
Beaumont, USA
xliu@lamar.edu

***Abstract* — In Wi-Fi networks, a user's location can be determined through the access points (APs) in the network. These kinds of location-based services, LBS, use a calculation of trilateration from the distance away from three routers. This distance is most often determined with received signal strength indication, RSSI. Wi-Fi communications are broadcasted, and RSSI is contained in the header of a packet. As a consequence, simple packet sniffing attacks make it relatively easy for malicious users to track a user's location. In order to increase the security of using LBS, this paper proposes a new algorithm for channel hopping of the routers. In this work, the algorithm is tested for its effectiveness at increasing the difficulty for the malicious user to determine the location of the LBS user. The simulation results support the hypothesis that channel hopping significantly increases the effort needed for the malicious user to track the user, at an acceptable cost of LBS efficiency.**

Keywords—RSSI, Location Based Service, Channel Hopping, Security

I. INTRODUCTION

A network is a collection of devices that are connected in a way that allows them to communicate with each other. Wireless devices, which are physically separated from each other, communicate by means of radio signals. One of the most common occurrences of this is the Wi-Fi networks many people use every day. With the rise in prominence of these Wi-Fi networks, there is also a rise in uses for them. One use is for Location Based Services (LBS). LBS use methods like trilateration to help a mobile device determine its location with respect to stationary routers, similar to a Global Positioning System (GPS). However, in these networks, since the communications are broadcast, this makes them naturally more vulnerable than traditional wired networks[1]. This vulnerability needs to be taken into account when providing LBS, in order to protect user privacy as much as possible.

Currently, one of the more common method for calculating location in wireless networks is a combination of trilateration with Received Signal Strength Indication (RSSI)[2]. Using the difference between the transmission power and the strength of the signal received, a device can estimate its distance from a router, which has a fixed location that it knows and can share with the device. If the device can make this estimate with three different routers, then it can determine its position from the three known points. This method of location approximation has drawbacks due to the unknown amount of interference, and loss of signal strength caused by possible walls, obstacles, and other devices between the device and a given router. Despite all of these drawbacks, it is still considered one of the more reliable methods of calculating relative location. This also presents a security problem though, because received signal strength is recorded and stored in the link layer header of a packet, which is not generally encrypted. This ultimately means that basic packet sniffing is a sufficient attack to obtain a devices location.

The interference between the routers of the network is attempted to be handled in IEEE 802.11 standard by assigning specific ranges of bandwidths, known as channels, to be used in Wi-Fi with a specific pattern. Thus, routers are most commonly configured to use three channels, (1, 6, and 11) in order to allow for a topology like the one seen in Fig. 1, where no routers have an overlapping area that uses the same channel. However, in a situation where the routers are not coordinated, you may see a topology like that in Fig 2, where all routers are using the same channel. This can happen in places like apartment buildings, or malls, where routers are set up with manufacture defaults, which often means that the same channel is always used unless it is manually configured. In both figures though, none of the routers can directly detect the others, as they are outside of each other's range.

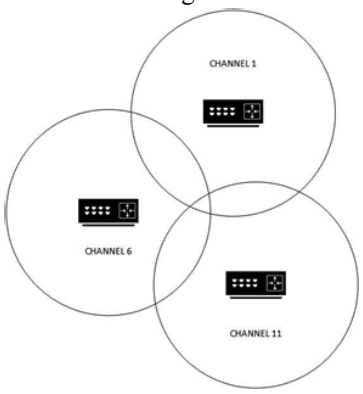

Fig. 1 Optimized Channel Distribution

978-1-7281-4905-9/20/$31.00 ©2020 IEEE

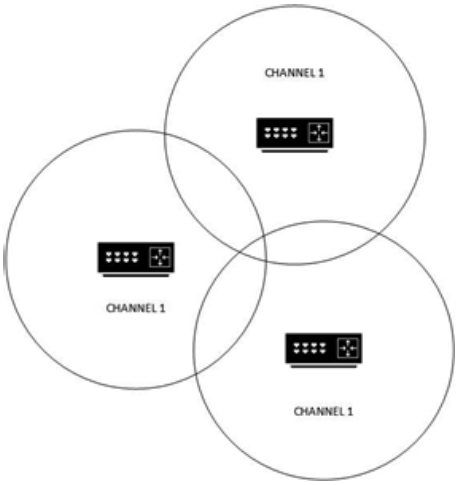

Fig 2 Worst Case Channel Distribution

The ability to communicate with multiple routers and devices, combined with the insecure nature of broadcasting, raises many challenges in providing privacy to the users of the network. In the related work chapter, it is discussed how other protocols attempt to maintain the privacy of the data sent through different methods. However, in this paper, a method is proposed for increasing the location privacy in these networks. With a combination of packet sniffing and a similar process to how the user's device determines its location, a malicious user can obtain and track the location of said user. This paper proposes a method of randomized channel hopping, where routers switch between a randomly chosen topology from a set at pseudo-randomly chosen intervals to increase location privacy for users. Both the attack and the proposed method for deterring it are further discussed in the System Model chapter.

II. Related Work

Previous research on privacy over Wi-Fi networks focuses mostly on protecting the privacy of the data, and the anonymity of the sender. The most common solution to protecting data is through encryption, which relies on the secure key distribution [3]. Most of these protocols can be considered secure, and are generally used in 802.11 standards, but since location is commonly determined with RSSI, these protocols are ineffective. This is because signal strength received is kept in the link layer header, which would be sniffed before any link layer encryption could take place. However, if the request is just a location request, then link layer encryption may not even take place, since the packet will not be forwarded.

The other approach to securing a user's information is anonymity. This is the practice of concealing the identity of the device making a query. One proposed method for providing anonymity is a system of intermediate packet hops before being sent to what is referred to as the sink node, which then forwards the packet normally [4]. This method works well enough in a setting where the routers have vastly more resources than the devices using them, because the extra forwards add artificial traffic to the network. In addition, while this method can provide a degree of anonymity once forwards begin, it also relies on the packet making it through the first router before it is read by a malicious user. This renders it as effective as any method of link layer encryption.

A second proposed method of providing anonymity iscalled Adaptive Nearest Neighborhood Cloaking (ANNC). This method uses the idea of cloaking regions, where multiple devices within a given proximity are provided an nth level of anonymity by the router through a grouping of the queries of other local devices [5]. This method can be effective at providing location privacy, but it relies on a minimum level of users to be querying for location in order to be effective. There is also a slight drawback in that it assumes all routers and devices are using the proper channel at a given time, which means that on a large scale it would require coordination between all devices.

III. System Model

Taking into account the drawbacks of current security measures, this paper proposes a method of organized channel hopping by the routers in a network. In theory, this should cause lapses in information to the malicious user, referred to as Eve in this experiment, as a typical device can only monitor a single channel at a time. In this experiment there were five stationary wireless components used in testing. There are three objects representing the routers of a wireless network, one object sending generic location requests, and one Eve monitoring the location of the requesting user. Although in a real world scenario, the user would be moving, this experiment keeps the location of all objects constant so as to isolate the effectiveness of the channel hopping as much as possible.

Three different scenarios were simulated and compared. The first, worst case scenario, was done in the network simulator ns-2, because by default, there is only one network interface used for all nodes in the simulation. The second and third scenarios were simulated using custom code written in Java. The Java nodes simulate only the behavior of a router at the link layer level in an effort to isolate any variables not used for calculations, as the objective is to test the concept of the theory only. In the second scenario, the routers use the same channel they start with throughout the test. Finally, in the third scenario, the routers were coordinated by a controller class, simulating a common server the nodes may share. This server has a randomized timer that sends a message to each router to change its channel. The message contains either a 1, 6, or 11, to signal each router which channel to use for that cycle. The signal of 1, 6, or 11 is pseudo- randomly chosen, in that a router cannot keep the same channel for more than two consecutive cycles, and all three routers must have a different channel on every cycle. This format was designed to maximize unpredictability of the hops while maintaining the integrity of the network, and minimizing interference.

The 802.11 standard of channels 1, 6, 11 with MIMO (multiple input/multiple output) antennae are simulated for both the routers and devices. The routers are situated in a standard triangular pattern that would be typical of a wireless network using any sort of multilateration distance bounding [6].The requesting node is located near the center of the triangle sending

out location requests at random intervals, and Eve is located slightly outside of the triangle. Fig. 3 shows the topology used.

Considering a potential hacking attempt, Eve monitors a single channel at a time, attempting to discern the location of the requesting node. Using a similar method as the trilateration, Eve can obtain a similar estimate of the location of the requesting node. Since received signal strength is contained in the link layer header, Eve can obtain it through any standard form of packet sniffing. Hence, in all of the simulations, it is assumed that Eve has access to the signal strength received by a router, as long as they are on the same channel.

This loss of location privacy is not ideal, so the algorithm proposed is intended to either conceal or interrupt the information available to Eve. The three routers are set to change their channels used at random intervals. In order to preserve the standard, all routers are changed at the same time, and all will change by the same randomized value. In theory, this will either interrupt information to Eve, or provide an incorrect location estimation.

The distance calculated by RSSI is derived from the equation for calculating expected RSSI, and given by a formula [7],

$$d = 10^{(P-R)/(10*n)}$$

where P is the initial transmission power, R is the RSSI, and n is the signal propagation constant. For the purposes of the experiment, and in the interest of isolating the parameters of the study, P was set to -13dBm, and n is set to 2.75. Both of these values are the de-facto industry standard, -18 to -25 dBm for broadcasting [8], and 2 for propagation through open space [9]. Once the distance is estimated from a node, there is a circle of estimated location, as shown in Fig. 4. When the distance is estimated from the second router, the estimated location is narrowed to the overlapping region, with highest probability being at the point of intersection of the two circles, as shown in Fig. 5. Finally, by adding the third distance, as shown in Fig. 6, the estimated location is narrowed to a region between the three closest intersections. This point is the final estimation for the location of the device, and is calculated by the following equations [10].

$$v_a = \frac{(d_a^2 - d_b^2) - (x_a^2 - x_b^2) - (y_a^2 - y_b^2)}{2}$$

$$v_b = \frac{(d_b^2 - d_a^2) - (x_b^2 - x_a^2) - (y_b^2 - y_a^2)}{2}$$

$$y = \frac{vb(xc - xb) - va(xa - xb)}{(ya - yb)(xc - xb) - (yc - yb)(xa - xb)}$$

$$x = \frac{va - y(yc - yb)}{(xc - xb)}$$

However, in this particular experiment, the calculations were not necessary to test the plausibility of the theory. The theoretical simulations were run with custom written Java code in order to isolate the effects of channel hopping exclusively. Simulations were then run with three scenarios:

1) All routers using the same channel the entire time.
2) The three routers all using a different channel the entire time.
3) All routers using a different channel, and changing that channel to a randomly chosen topology at given intervals

For Eve, the experiment assumes that they are within range of two routers and the user, but outside of the triangle formed by the three routers, and thus not within range of the third router. This gives Eve access to three points of reference as well, and forming a trilateration like that shown in Fig. 7.

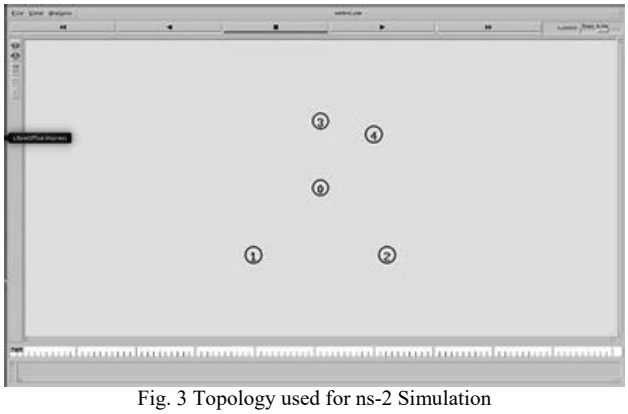

Fig. 3 Topology used for ns-2 Simulation

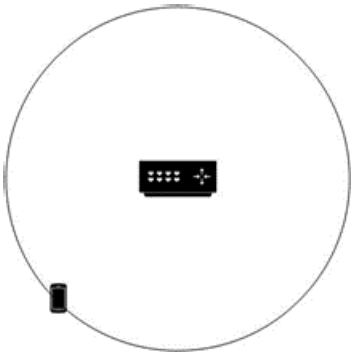
Fig. 4 Distance from Single Router

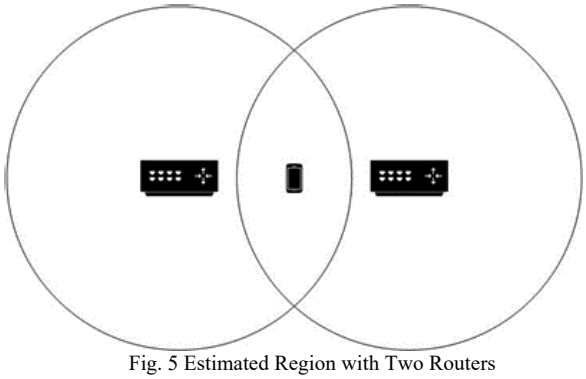
Fig. 5 Estimated Region with Two Routers

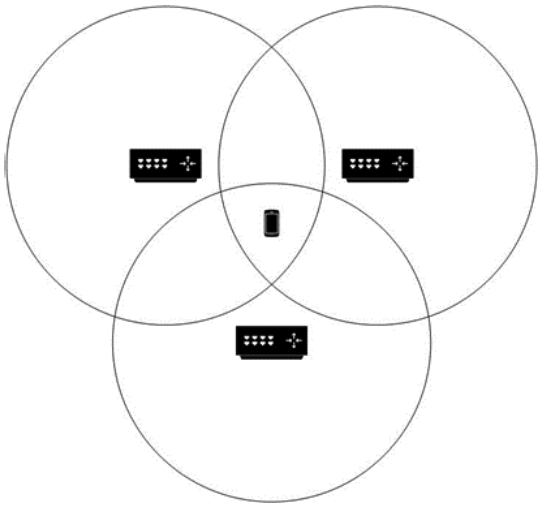

Fig. 6 Estimated Region with Three Routers

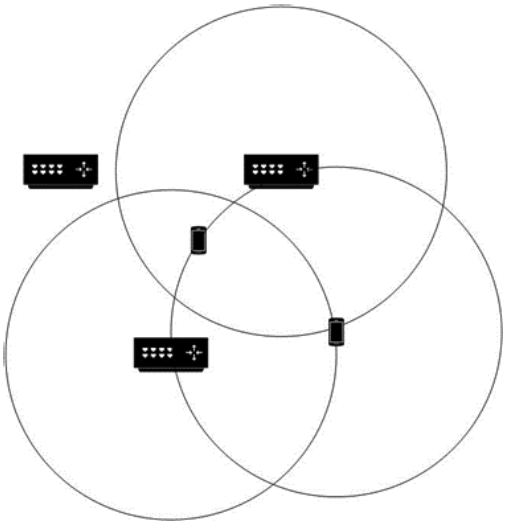

Fig. 7 Location Estimation by Malicious User

III. Performance Evaluation

In the simulations, due to the isolation of the channel hopping, the results of the control simulations were constant. As predicted, when all routers use the same channel, it takes the malicious user two "sniffs" to locate the user, one sniff for each router. In this amount of time, the user is able to retrieve their location twice, because all of the information needed is retrieved with each broadcast. It should be noted though, that in the NS2 simulation, there were recorded packets lost which can be attributed to signal interference on the single channel.

Similarly, when each router has a different channel, but does not change their channel, it took the malicious user four "sniffs." The expected value would be 3 5/6 from a purely statistical approach as follows.

Expected tries to find the first router's channel

$$1/3*1+2/3*(1/2*1+1/2*2)=2\ 1/3$$

Expected tries to find the second router's channel

$$1/2*1+1/2*2=1\ ½$$

The total number of expected tries to find both routers

$$2\ 1/3+\ 1\ 1/2=3\ 5/6$$

Since partial sniffs are not possible, the outcome of four "sniffs" is within theoretical parameters. It should be noted though, that after Eve has determined the channels of the routers, the number of sniffs needed per determining the user's location falls back down to 2 because, Eve knows what channels to use for each router. However, the user still requires 3 broadcasts each time they want to determine; one for each router.

In the simulation where the routers are changing channels to different topologies after each iteration, the results were drastically different. As seen in Figure 8, the number of "sniffs needed by Eve to find user" ranged from 4 to 32, while the "times the user found its location" ranged from 0 to 6.

Taking the average over 30 trials, the user is able to identify their location an average of 2.6 times before Eve does, as seen in Table 1. Also, as shown in Table 2, Eve requires an average of 13.9 "sniffs" to determine the user's location. Comparing this to the other two scenarios, the channel hopping increased the work needed by Eve by an average of 3.475 times from the current coordinated scenario, and 6.95 times the worst case scenario.

From the numbers given in the above tables, there is a ratio of efficiency that can be obtained to evaluate the user's ability to find their location. Dividing the average number of times the user found its location by the number of sniffs needed by Eve, gives what will be considered relative user efficiency. As seen in Table 3, this efficiency starts at 1 while everything is on the same channel, and drops to 0.25 with routers on different channels. The efficiency drops further to 0.19. This means that while the work needed by Eve increases, so does that of the user. This would be expected since neither of them have any more information about the routers than the other.

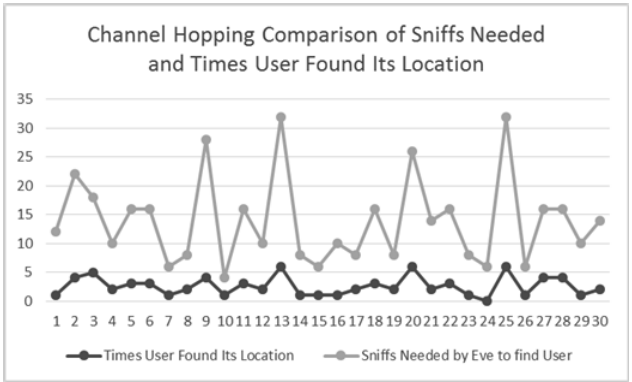

Figure 8 Channel Hopping Comparison of Sniffs Needed and Times User Found Its Location

Table 1 Average Number of Times User Obtained Its Location before Eve Determined the User's Location

Everything on the Same Channel	Routers All on Different Channels	Routers All on Different Channels With Hopping
2	1	2.6

Table 2 Average Number of Sniffs Needed by Eve to Determine User's Location

Everything on the Same Channel	Routers All on Different Channels	Routers All on Different Channels With Hopping
2	4	13.9

Table 3 User Efficiency

Everything on the Same Channel	Routers All on Different Channels	Routers All on Different Channels With Hopping
1	0.25	0.19

IV. Discussion

There are three significant figures to be compared for the performance. In the two measured categories, times user determined its location and sniffs needed by Eve, the proposed method had an impact. With an increase of more than 3 times the effort needed by Eve, and an increase of times the user determined their location before Eve, the data supports that this impact was significant.

The data then supports the theory that channel hopping routers do in fact increase the security of the user's location. However, when looking at the third figure of user efficiency, it can be inferred that the increase in security comes with a loss of LBS performance. Thus, the practicality of the algorithm may need to be assessed.

In conclusion, the data supports the idea that, at least in theory, channel hopping increases location security in a Wi-Fi network. This security does come at a cost to efficiency. So, it would be recommended that further studies be conducted to test the theory in more complex simulations, or even real device experiments.

V. Future Work and Conclusion

There are a few factors that were not accounted for during the simulation. As noted previously, when all routers use the same channel, there is loss of data that is caused by the interference on the channel. Taking this into account, on a simulation that measures the interference, the loss of efficiency may be less significant.

Another variable that was not measured in this simulation was the relevance of the packet. In a real-world scenario, the user is moving constantly, and packets would only be in memory for a short period, so time is a major factor for sniffing. Thus, the proposed method may perform better in real time as successfully finding the current channel of the router may still not yield useful information.

Finally, probably the potential disadvantage factor that this scheme does not address, is that more often, there are multiple users, and not all of them are using LBS. This means that the algorithm may affect normal internet access, unless there are provisions taken by the routers' coordinator. In this scenario, normal users, who are not sending location requests, may find that the channel they were using for their nearest router, is now the channel of the furthest router. While this may not stop access, it could slow it down before the user's device adjusts.

Nevertheless, considering the fact that non-LBS applications can also release their location information, not sniffed in the header, but derived by the responding time from different APs. In fact, it is easier for hackers to crack the location information in this way (bypass the encryption) since data in the header may be highly encrypted in nowadays protocols. There is also a need for nowadays WiFi users, instead of only LBS services, to hide from location sniffing. Therefore, the proposed scheme is a general and promising candidate for location-anonymous communications based on its anti-sniffing features other than encryption.

References

[1] Di Pietro, Roberto, and Alexandre Viejo. "Location privacy and resilience in wireless sensor network querying." Universita di Roma Tre, Dipartimento di Matematica, L.go S. Leonardo Murialdo n.1, 00146 Roma, Italy. 2010

[2] Kuntal, Ashutosh, Madan Lal Tetarwal, Purnendu Karmakar. "A Review of Location Detection Techniques in Wi-Fi." National Seminar on Recent Advances in Wireless Networks and Communications, NWNC. 2014.

[3] Mi, Qi, John A. Stankovic, and Radu Stoleru. "Secure Walking GPS: A Secure Localization and Key Distribution Scheme for Wireless Sensor Networks." Department of Computer Science and Engineering, Texas A&M University, USA. ACM, 2010

[4] Li, Yun, Jian Ren, Snior Member, IEEE, and Jie Wu, Fellow, IEEE. "Quantitative Measurement and Design of Source-Location Privacy Schemes for Wireless Sensor Networks." IEEE Transactions on Parallel and Distributed Systems, Vol. 23, No. 7, July 2012.

[5] Talukder, Nilothpal, Sheikh Iqbal Ahamed. "Preventing Multi-query Attack in Location-based Services." Department of Computer Science, Purdue University, Department of MSCS, Marquette University. ACM, 2010.

[6] Chiang, Jerry T, Jason J. Haas, Yih-Chun Hu. "Secure and Precise Location Verification Using Distance Bounding and Simultaneous Multilateration." WiSec '09, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 2009.

[7] Dong, Qian, and Waltenegus Dargie. "Evaluation of the Reliability of RSSI for Indoor Localization." 2012 International Conference on Wireless Communications in Underground and Confined Areas, 2012, pp. 2-3. Accessed June 25, 2018. doi:10.1109/icwcuca.2012.6402492.

[8] "Understanding Transmit and Receive Levels on Modems." Cisco. August 23, 2015, para. 1. Accessed October 02, 2018. https://www.cisco.com/c/en/us/support/docs/dial-access/asynchronous-connections/15380-trans-rec-15380.html.

[9] Hidayab, Muzaiyanah, Abdul Halim Ali, and Khairul Bariah Abas Azmi. "Wifi Signal Propagation at 2.4 GHz." 2009 Asia Pacific Microwave Conference, 2009. doi:10.1109/apmc.2009.5384182.

[10] Kamila, Narendra Kumar. Handbook of Research on Wireless Sensor Network Trends, Technologies, and Applications. Hershey: IGI Global, 2016.