A Method for Cracking the Password of WPA2-PSK Based on SA and HMM

Liang Ge[1,a], Lianhai Wang[1,b], Lijuan Xu[1,c]

[1]Shandong Provincial Key Laboratory of Computer Networks,Shandong Computer Science Center(National Supercomputer Center in Jinan), Jinan, China 250014

e-mail: [a]gel@sdas.org, [b]wanglh@sdas.org, [c]xulj@sdas.org

***Abstract*—Password recovery of WPA2-PSK is an important problem in digital forensics. Since the encryption mechanism of WPA-PSK is gradually enhanced, it is difficult to deal with this problem by the traditional methods such as brute force, rainbow table, Markov model, and so on. In this paper, we give a new method based on simulated annealing (SA) and hidden markov model (HMM). The main principle of this method is to create the hidden markov model of the known password based on the SA which could be used to generate the password candidates in the wireless network password recovery. It means that the passwords are given by a probability learning of the known password. The tests have shown that this approach could improve the effectiveness of password recovery for the wireless network, comparing with the Markov model which has been shown much more efficiently than the traditional methods such as brute force and dictionary attack.**

Keywords-simulated annealing; hidden markov model; password cracking; WPA2-PSK

I. INTRODUCTION

Nowadays, with the increasing crime of wireless network, wireless network forensics is becoming increasingly important. In order to examine the crime on the wireless network, we must break through the secure scheme of wireless network before using the computer forensics to investigate [1]. There are three security protocols named Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA) and Wi-Fi Protected Access (WPA2) that have been developed by the wi-Fi Alliance to secure wireless network. Since the WEP and WPA protocols have some weakness, the WPA2 protocol is currently the popular way. Usually the WPA2 use the WPA2-PSK encryption schemes [2]. So at present the cracking of WPA2-PSK is an important problem in computer forensics. WPA2-PSK is a subset of IEEE 802.1X WPA/WPA2 that skips the complex task of key distribution and client authentication by assigning every participating party the same pre shared key. It is difficult to crack this protocols, since the AES algorithm is used as the encrypt algorithm and the key is generated by the password through a complex algorithm. Every password which is used in WPA/WP2-PSK network has to contain from 8 to 63 printable ASCII Character. The key is obtained by using the algorithm PBKDF2(password-based key derivation function) [3].

Usually, brute force attack and rainbow table attack are useless for the attack of WPA2-PSK. Dictionary attack is a better approach for the WPA2-PSK key decryption. A typical dictionary attack for decrypting the key of WPA2-PSK is that a list of potential passwords is generated based on the knowledge about natural language, mathematical symbols, logic language or user's habit and then the decryption key is given by the passwords following a password-derived algorithm. It is important to give an efficient password generating method. There have been a few studies for this field. Weir [4-5] proposed a method which uses the Markov model to generate a list of passwords. Probabilistic context-free grammars can be crafted to generate the password guess in [6]. The cracking tool -- John the Ripper [7] generates the password list according to a set of hand-crafted rules that mimic how a human is likely to perturb their password. There are two main hypotheses for the above method. First, all characters in a password are assumed to be identically distributed after the initial character and any additional information about the distribution of characters at each index of the password is discarded. Second, the distribution of passwords and character transition probabilities in the training set is assumed to be representative of passwords in any database. In practice, neither of these assumptions hold true.

In this paper, we give a new dictionary cracking method for WPA2-PSK based on HMM, which is firstly given by Chris Karlof [8]. The HMM can improve the above two hypotheses. But since the Baum-Welch algorithm used as the training tool for HMM is a local search algorithm, the training of HMM which depends on the initial value is easy to fall into local optimum. Then we use the SA algorithm [9] to improve the HMM algorithm. The new method can combine the advantage of HMM and SA to improve the efficiency of WPA2-PSK password cracking.

The rest of this paper is organized as follows. Section 2 presents background information and related work. Section 3 presents our password cracking method based on SA and HMM. Section 4 shows empirical results of our SA and HMM model on some real-world datasets. Finally, Section 5 presents our conclusions.

II. BACKGROUND

In this section we will introduce the password recovery process for WPA2-PSK. WPA2-PSK adopted IEEE 802.1 EAPOL-Key frames, called the four-way handshake to achieve the authentication [10]. In the four-way handshake, the authentication of WPA2-PSK is dependent on the generation of value MIC. MIC is generated by a list of

978-1-5090-2535-0/16 $31.00 © 2016 IEEE
DOI 10.1109/ICISCE.2016.23

values, such as PTK, ANonece, SNonce, SA, AA, PMK, and SSID. If we want to crack the password of WPA2-PSK, we need to sniffer the handshake process and use the given password to compute the MIC. The password recovery for WPA2-PSK is as following:

(1) Obtain the 4 times handshake package and SSID of AP by monitoring the communication between AP and STA;
(2) Generation a password and compute the PMK by PSK=PMK=pbkdf2_SHA1(password, SSID, length of SSID, 4096);
(3) Obtain Anonce, Snonce, AA, SPA by analyzing the handshake package 1 and 2, and use these to compute the MIC_KEY jointed with the PMK as following:
PTK=SHA1-PRF (PMK, 'pairwise key expansion', Min(AA, SPA)||Max(AA, SPA)||Min(ANonce, SNonce)||Max(ANonce, SNonce)),
For(i=0; i<128; i++) MIC_KEY[i]=PTK[i];
(4) Decide the message authentication algorithm following the key descriptor type in EAPOL-key frame. Use the KCK given in 3) to compute the MIC by following:
MIC=HMAC_md5(or MAC_SHA1)(MIC_KEY,16, 802.1x data);
(5) If MIC is equal to the MIC from handshake package 2, then terminate the algorithm and output the password; else go to 2) and continue the calculation.

III. CRACKING THE WPA2-PSK PASSWOED BASED ON SA AND HMM

To crack the WPA2-PSK password, we first use the SA algorithm to obtain the optimum parameters of HMM; then use the HMM to generate the password list by training the known password dictionary; finally obtain the key by taking the password into the key generation, and confirm whether the key is right.

The establishment of password sequence by HMM can be seen as an optimal problem. Given a password sequence $O^{(l)} = O_1^{(l)}, O_2^{(l)}, \cdots, O_{T_l}^{(l)}$, $l = 1,2,\cdots,L$. Let $x = (A, B, \pi)$ which makes the probability $P(O \mid x)$ to take the maximum. Since the SA algorithm is used to find the minimum, we define the energy function $E(x) = -\log P(O \mid x)$. The definitions are following:

Definition 1 Forward probability is a probability that the state S_i happens in t_l time, and $(O_1^{(l)}, O_2^{(l)}, \cdots, O_{t_l}^{(l)})$ happens before t_l time,

$$\alpha_{t_l}^{(l)}(i) = p(q_{t_l} = S_i \mid O_1^{(l)} O_2^{(l)} \cdots O_{t_l}^{(l)}; \lambda), \qquad (1)$$

where $1 \le t_l \le T_l, 1 \le i \le N$.

The steps of computing the forward probability are following:

a) Let $\alpha_1^{(l)}(i) = \hat{\pi}_i \hat{b}_i(O_1^{(l)})$ be the initial value.

b) Calculate $\alpha_{t_l+1}^{(l)}(j) = [\sum_{i=1}^{N} \alpha_{t_l}^{(l)}(i)\hat{a}_{ij}]\hat{b}_j(O_{t_l+1}^{(l)})$ by the induction.

c) Finally give the following value

$$p(O^{(l)} \mid x) = \sum_{i=1}^{N} \alpha_{T_l}^{(l)}(i) .$$

Definition 2 Backward probability is a probability that the state S_i happens in t_l time, and $(O_{t_l+1}^{(l)}, O_{t_l+2}^{(l)}, \cdots, O_{T_l}^{(l)})$ happens after t_l time. The steps of computing the backward probability are following:

a) Let $\beta_{T_l}^{(l)}(i) = 1$ be the initial value.

b) Calculate $\beta_{t_l}^{(l)}(i)$ by the following formula

$$\beta_{t_l}^{(l)}(i) = \sum_{j=1}^{N} \hat{a}_{ij}\hat{b}_j(O_{t_l+1}^{(l)})\beta_{t_l+1}^{(l)}(j) \qquad (2)$$

The steps of obtaining the HMM parameters are following:

Step1. Set the initial value.

Step2. Based on the initial parameters, $P(O \mid x_{old})$ is calculated by formula (1), (2) and (3).

$$P(O \mid x_{old}) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i) = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j) \qquad (2)$$

Then the initial energy function $E(x_{old})$ and x_{old} is obtained. Let $s_{num} = 1$ be the annealing number, and $iter_{num} = 1$ be the iterations.

Step3. Determine whether the termination condition is satisfied. If satisfied, then the algorithm is terminated. Else $s_{num} = s_{num} + 1$, go to next step.

Step4. Generate the new solution $x_{new} = x_{old} + rand$ (where rand is the normal probability distribution of candidate solutions), and the normalization will be used to x_{new} to satisfy restrictions of HMM. $P(O \mid x_{new})$ is given by formula (3),(4),(5),(6).

$$\hat{\pi}_i = \sum_{l=1}^{L} \frac{\alpha_1^{(l)}(i)\beta_1^{(l)}(i)}{p(O^{(l)} \mid x_{new})}, 1 \le i \le N , \qquad (3)$$

$$\hat{a}_{ij}=\frac{\sum_{l=1}^{L}\sum_{t_l=1}^{T_l-1}\alpha_{t_l}^{(l)}(i)a_{ij}b_j(O_{t_l+1}^{(l)})\beta_{t_l+1}^{(l)}(j)\Big/p(O^{(l)}\mid x_{new})}{\sum_{l=1}^{L}\sum_{t_l=1}^{T_l-1}\alpha_{t_l}^{(l)}(i)\beta_{t_l}^{(l)}(j)\Big/p(O^{(l)}\mid x_{new})},1\le i,j\le N \quad (4)$$

$$\bar{b}_{jk}=\frac{\sum_{l=1}^{L}\sum_{t_l=1\cap O_{t_l}=V_k}^{T_l-1}\alpha_{t_l}^{(l)}(i)\beta_{t_l}^{(l)}(j)\Big/p(O^{(l)}\mid x_{new})}{\sum_{l=1}^{L}\sum_{t_l=1}^{T_l}\alpha_{t_l}^{(l)}(i)\beta_{t_l}^{(l)}(j)\Big/p(O^{(l)}\mid x_{new})},$$
$$1\le j\le N,1\le k\le M \quad (5)$$

Then $E(x_{new})$ and x_{new} is obtained.

Step5. If $E(x_{new})<E(x_{old})$, then accept the new solution; else go to step6.

Step6. Obtain the accept probability P from formula (7) and randomly generate a constant c ($0<c<1$). If P>c, then accept the new solution; else reject the new solution.

$$p=\begin{cases}1 & if\ E(x_{new})<E(x_{old})\\ \exp(-\frac{E(x_{new})-E(x_{old})}{T(t)}) & if\ E(x_{new})\ge E(x_{old})\end{cases} \quad (6)$$

Step7. If $|E(x_{new})-E(x_{old})|\le\varepsilon$ (ε is the given convergence accuracy), terminate the algorithm.; else reduce the temperature following the formula:

$T(t)=T_0\cdot e^{-t^3}$, $iter_{num}=iter_{num}+1$, go to step3.

Then the steps of obtaining the password list from the HMM model are following:

a) Let $t=1$, and choose an initial state $q_1=i$ according to the initial probability distribution $\hat{\pi}$.

b) Choose an output value $O_t=V_k$ according to the probability distribution $b_i(k)$ of the state i.

c) Choose a subsequent state $q_{t+1}=j$ according to the transition probability distribution a_{ij}.

d) If $t<T_l$, let $t=t+1$ go to b); else terminate the algorithm.

Finally, the passwords are fed into the WPA2-PSK decryption mechanism given in section 2 to obtain which one is the right password.

IV. EXPERIMENT

In this section we will use the experiment to test the effectiveness of WPA2-PSK password cracking algorithm given in section 3. In order to establish the effective training set of HMM, we use some public leakage password database to study the pattern of password character. In this paper, we use the micro-blog information of sina leaked in 2011 as the data set. First, we will preprocess the data by remove some passwords which contain non keyboard character. So we have 1300000 valid passwords. For simple, we only randomly choose 100000 passwords as the training set, and choose another 10000 passwords as the testing set. We compare the SA HMM model with the Markov model [4-5] which has been shown more effective than brute force to crack the passwords. The result is shown as following:

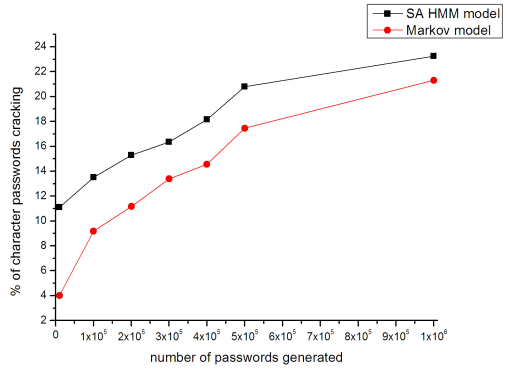

Figure1 comparison with two methods for 10000 passwords testing set

From Figure1, we can see that the SA HMM model can crack more passwords than the Markov model by generating the same number of passwords. It suggests that the SA HMM model given in this paper is more efficient than the Markov model in [4-5].

V. CONCLUSION

At present password cracking of WPA2-PSK is still an important problem in computer forensics. The traditional methods such as brute force attack, rainbow table attack, are difficult to deal with gradually enhanced encryption technology for Wireless network. Since these, we give a new cracking algorithm based on SA and HMM in this paper. The details of this algorithm are shown in Section 3 and experiment comparing with the Markov model is given in Section 4. Experimental results show the password cracking method for WPA2-PSK based on SA and HMM given in this paper has performed very well. This indicates our method has practical significance.

ACKNOWLEDGMENT

The work was supported by the National Natural Science Foundation of China (61572297), by the Shandong Province Outstanding Young Scientists Research Award Fund Project(Grant No. BS2013DX010), by the Natural Science Foundation of Shandong Province, China(Grant No. ZR2014FM003, ZR2013FQ001, ZR2013FM025, ZR2015YL018), and by the Shandong Academy of Sciences Youth Fund Project(Grant No. 2013QN007) .

REFERENCES

[1] L.Garfinkel, “Digital forensics research: The next 10 years,” Digital Investigation, 2010, 7: S64-S73.

[2] IEEE Computer Society, “IEEE Standard for information technology-telecommunications and information exchange between systems-local and metropolitan area networks-specific requirements-part II: wireless LAN medium access control (MAC) and physical layer (PHY) specifications,” IEEE Std 802.11TM-2007, 2007.

[3] I. Marvridis, E. Androulakis, B. Halkias, P. Mylonas, “Real-life paradigms of wireless network security attacks,” in Proceedings 2011 Panhellenic Conference on Informatics, pp. 112-116.

[4] M. Weir, S. Aqqarwal, M. Collins, H. Stern, “Testing metrics for password creation policies by attacking large sets of revealed passwords,” In Proceedings of the 17th ACM conference on Computer and communications security, 2010, pp. 162-175.

[5] M. Weir, “Password cracking using probabilistic context-free grammars,” In Security and Privacy, 2009 30th IEEE Symposium on, pp 391-405.

[6] M. Narayanan, etc,“Fast dictionary attacks on passwords using time-space tradeoff,” In Proceedings of the 12th ACM conference on Computer and communications security, 2005, pp. 364-372.

[7] John the Ripper password cracker, Openwall Project, http://www.openwall.com/john.

[8] C. Karlof, D. Wagner, “Hidden Markov model cryptanalysis,” Lecture Notes in Computer Science, 2003, pp. 17-34.

[9] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, “Optimization by simulated annealing,” Science, 1983, 220(11):650-671.