Secure Protocol Analysis Method with Hush Functions

Yang Jie[1], Ma Xian-heng[2], Luo Wei-liang[3]

1. 2.3. School of software, South China University of technology, Guangzhou, Guangdong Province, China

Dept. name of organization, name of organization, acronyms acceptable, City, Country

1. yjclear@scut.edu.cn, 2,war3shane@gmail.com 3. yjclear@gmail.com

Abstract: Strand space inherits the merit of algebra method, which can prove the security of protocol more precisely for it overcomes the shortcoming of inconvenience. But because of shortage of primitives it can't describe some protocols as IKE, let alone verification. We made improvement to the original Strand Space by adding the description of hash function and some definitions with lemmas. So we can use the intensive model to analyze the protocols using hash function. At last, we analyze one IKE sub-protocol protocol.

Keywords: strand space; IKE; protocol analysis

使用哈希函数的安全协议分析方法

杨捷[1]，马献恒[2]，罗伟良[3]

1. 2. 3. 华南理工大学软件学院，广州，中国，510006

1. yjclear@scut.edu.cn, 2. war3shane@gmail.com, 3. yjclear@gmail.com

【摘要】链空间（**Strand Sp ace**）协议分析模型继承了代数方法的优点，能较为精确地证明协议的安全性同时克服代数方法的不足。但目前的链空间模型原语不足，不能描述和验证所有的协议包括一些著名的协议如**IKE(Internet Key Exchange)**。本文针对这一问题在原有模型的基础上增加了对散列函数的描述，同时给出了相关的定义和引理，并且使用扩充后的模型对包含散列函数的协议进行分析和验证。最后实际分析了 IKE 协议的一个子协议。

【关键词】链空间；IKE；协议分析

1 Introduction

The execution of protocol in the strand space model consists of multiple strands (in the form of such definite sequence <<b_1, a_1>,…, <b_n, a_n>>), and the algebraic structure running in the strand space is called algebra A; The set of strand is strand space, which represent the running of the protocol and is designated by symbol Σ. Among it, $b_i \in \{+,-\}$ (+ for receive, - for send), $a_i \in A$，$1 \leqslant i \leqslant n$. One two-tuple represents one node and is also designated by one natural number. n=<s, i>∈N (natural number). s means the strand, i means the location of the node in s. term (n) represents the node n with signed term, while uns_term (n) for unsigned. If term (n_1) =+a, term (n_2) =-a (a∈A), then designate it by $n_1 \rightarrow n_2$; if n_1=<s, i>, n_2=<s, i+1>, then we mark it by $n_1 \Rightarrow n_2$. If $S \subset \rightarrow \cup \Rightarrow$, then <S represents the transitive closure of S, $\leqslant S$ for reflexive transitive closure.

Definition 1. 1 Assume that $\rightarrow_C \subset \rightarrow$, $\Rightarrow_C \subset \Rightarrow$, C=<$N_C$, ($\rightarrow_C \cup \Rightarrow_C$)> is the sub-graph of <N, ($\rightarrow \cup \Rightarrow$)>. If C satisfies the following four conditions, it is called bundle: (1) C is definite. (2) If $n_2 \in N_C$ and term (n_2) is negative, then there exists the unique n_1 which satisfies that $n_1 \rightarrow_C n_2$. (3) If $n_2 \in N_C$ and $n_1 \Rightarrow n_2$, then $n_1 \Rightarrow_C n_2$. (4) C is not circular. The number of nodes of strand s in bundle C is called the C-height of s. $I \subseteq A$, if node n ∈ N, term (n) =+t (t ∈ I) and for any node n' ahead of n in the same strand, uns_term (n') ∉ I, then we call that node n is the access point of set I.

Proposition 1. 1 If C is a bundle, then $\leqslant_C$ is a partial order, any nonempty set in C nodes have the minimize related to $\leqslant_C$.

This work is supported by S CUT SRP Y1070550, Y1080 150, Y1080160, Y1080110, Y1080170, Y1090150, SCUTKCCP (Y3080020) and NSF 60873078.

Proposition 1.2 If C is a bundle, $S \subseteq C$ which a set of nodes satisfy $\forall m,m'$ uns_term(m) = uns_term(m') $\Rightarrow$ (m $\in$ S iff m' $\in$ S). If n is a minimize related to $\leqslant_C$ is s, then n is positive.

Entities are divided into regular entities and penetrator entities. The action of penetrator (represents the penetrator capability) is described by specific strand in strand space model. The nodes of penetrator strand are penetrator nodes, others are regular nodes. Besides, $K' \subseteq K$, a K'-ideal in the algebra A is a subset I, and it satisfies that for all $h \in I$, $g \in A$, $k \in K'$: hg, gh $\in$ I and $\{h\}_K \in I$. The minimal K'-ideal including h is designated as $I_{k'}[h]$.

M. <+t>, among it $t \in A$

K. <+k>, among it $k \in K_P$ (deadly private key)

F. <-g>, T. <-g, +g, +g>, C. <-g,-h,+gh>

S. <-gh,+g,+h>, E. <-k,-h,$\{h\}_k$>, D. <-k^{-1},-$\{h\}_k$,+h>

Proposition 1.3 If $S \subseteq A$, then $I_k[S] = \cup_{x \in S} I_k[x]$.

Proposition 1.4 If C is a bundle on A, and if m is the minimize of {m $\in$ C: uns_term (m) $\in$ I}, then m is the access point of I.

Definition 1.2 An date term is simple if and only if it is not in the form of ab (a, b $\in$ A).

Proposition 1.5 $k \in K$, $S \subseteq A$, $K' \subseteq K$, for all $s \in S$, s is simple and not in the form of $\{g\}_k$. if $\{h\}_k \in I_{K'}[S]$, then $h \in I_{K'}[S]$.

Proposition 1.6 $S \subseteq A$, $K' \subseteq K$, and for all $s \in S$, s is simple. If gh $\in$ $I_{K'}[S]$, then g $\in$ $I_{K'}[S]$ or h $\in$ $I_{K'}[S]$.

2 Intensive Strand Space Model

(1) Set $T \subseteq A$ represents that the atomic data in A, $T_{name} \subseteq T$ represents that the identity of each entity; $T_{nonce} \subseteq T$ represent the random number in protocol (namely random created data only used for once). T_{name} and T_{nonce} are not intersected.

(2) Set $K \subseteq A$ represents the private key in A, and there is a mapping inv on K: K→K, which represents that from the encryption of private key to the decryption of private key or otherwise. K and T are not intersected.

(3) Three operators: encr: K X A → A; join: A X A → A; hash: A → A. Compared with the original axioms, hash function is conflict-free (Axiom 1), its range and the others of two operations do not intersect with each other, and their values are not atomic data (Axiom 2).

Axiom 1 $\forall$ m, m'$\in$A, $\forall$ k, k'$\in$K, then $\{m\}_k=\{m'\}_{k'}$ $\Rightarrow$ m=m' And k=k'. hash (m)=hash(m') $\Rightarrow$ m=m'

Axiom 2 $\forall$ m_0,m_0',m_1,m_1'$\in$A, $\forall$ k, k'$\in$K then $m_0m_1 = m_0'm_1'$ $\Rightarrow$ $m_0 = m_0'$ and $m_1 = m_1'$; $m_0m_1 \neq \{m_0'\}_{k'}$; $m_0m_1 \notin K \cup T$; $m_0m_1 \neq$ hash(m_1'); $\{m_0\}_k \notin K \cup T$; $\{m_0\}_k \neq$ hash(m_1); hash(m_0) $\notin K \cup T$

Definition 2.1 $S \subseteq A$, W[S] is the minimal set satisfying the following conditions: (1) $\forall g \in S$, then $g \in W[S]$ (2) $\forall g \in W[S]$, $\forall h \in A$, then gh, hg $\in$ W[S]

Definition 2.2 $S \subseteq A$, H[S] is the minimal set satisfying the following conditions: $\forall g \in W[S]$, then hash (g)$\in$ H[S]. Because H[S] is the minimal set satisfying conditions, the mapping from W[S] to H[S] is surjection, and according to Axiom 1 it is assured that hash function is injection, thus there is bi-jection between W[S] and H[S] (hash is just one of them). Appoint that if there is only one element a in S, we can note W[{a}] and H[{a}] as W[a] and H[a]. Proposition 2.1 and 2.2 are the special cases of Proposition 1.6 and 1.5.

Proposition 2.1 $S \subseteq A$, $K' \subseteq K$, $\forall x \in H[S]$, if $gh \in I_{K'}[x]$, then g $\in$ $I_{K'}[x]$ or h $\in$ $I_{K'}[x]$.

Proposition 2.2 $S \subseteq A$, $k \in K$, $K' \subseteq K$, $\forall x \in H[S]$, if $\{h\}_k \in I_{K'}[x]$, then h $\in$ $I_{K'}[x]$.

Proposition 2.3 $\forall S \subseteq A$, $g \in A$, if $\exists x \in H[S]$, causing that hash (g)$\in$$I_K[x]$, then $g \in W[S]$.

Prove:

① Firstly we prove hash(g)=x. According to the definition of ideal, we can get the elements in $I_K[x]$ by three approaches: $x \in I_K[x]$; $\forall g \in A$, $h \in I_K[x]$, then hg, gh $\in$ $I_K[x]$; $\forall k \in K$, $h \in I_K[x]$, then $\{h\}_k \in I_K[x]$. According to Axiom 2, the range of hash function is not intersected with ranges of values from other two operations, so hash(g) never belongs to the latter two situations, then hash(g)=x.

② Prove $g \in W[S]$. Since hash(g)=x, hash(g) $\in$ H[S]. hash is bijection, then $\exists g' \in W[S]$ and hash(g')= hash(g), according to Axiom 1, then g'=g, thus $g \in W[S]$ □.

Proposition2.4 $S \subseteq K \cup T$, $\forall g, h \in A$, if $gh \in W[S]$, then $g \in W[S]$ or $h \in W[S]$.

Prove: Proof by contradiction. Assume that $gh \in W[S]$ and $g \notin W[S]$, $h \notin W[S]$. From definition 3.1, know that elements in W[S] can be obtained by two approaches. Since $S \subseteq K \cup T$, from Axiom 2, gh can not be obtained by the first approach; and from the assumption condition

978-1-935068-01-3 © 2009 SciRes.

above, gh can not be obtained by the second approach, then $gh \notin W[S]$, which is inconsistent with the assumption. Proposition is proved. □

Proposition 2.5 $K' \subseteq K$, $S_0=S$, $S_{i+1}=\{\{g\}_k: g \in W[S_i], k \in K'\}$, then $I_{K'}[S]=\cup_i W[S_i]$.

Prove: $\forall i$, $S_i \subseteq I_{K'}[S]$, so $\cup_i W[S_i] \subseteq I_{K'}[S]$. And obviously $\cup_i W[S_i]$ is close under the operations of connection and encryption, then $I_{K'}[S] \subseteq \cup_i W[S_i]$. Proposition is proved. □

Proposition2.6 $\forall S \subseteq A$, $g \in W[S]$, then $W[g] \subseteq W[S]$.

Prove: $\forall x \in W[g]$, if $x=g$ then $x \in W[S]$; Otherwise, $x=gh$ or hg, also $g \in W[S]$, then $x \in W[S]$. □

The capability of penetrator is also intensive in the intensive strand space model, if penetrator is aware of the data related to hash, the hash value will be known (this is in line with the actual situation, commonly the definition of hash function is public). We need to add a penetrator strand: H. <-g, +hash (g)>.

Definition 2.3 C is a bundle on A, $\forall k \in K$, and there is not any regular node n which makes that $uns\_term(n) \in I_K[k]$, and $\forall h \in A$, $\forall k^{-1} \in K_p$, and there is not existing any regular node n which makes that $uns\_term(n) \in W[\{h\}_k]$, then we say that C is normal. Bundle C is normal, if all the regular nodes among didn't transcend any private key in the protocol and didn't use the data encrypted by any private key which is destroyed.

Lemma 2.1 $S \subseteq K \cup T$, bundle C is normal, then except M and K, other penetrator strand can not include the access point of W[S].

Lemma 2.2 $S \subseteq A$, if W[S] can not find an access point in bundle C then $\forall x \in H[S]$, no penetrator node can be the access point for $I_K[x]$ in C.

3 IKE Protocol Analysis Instance

In order to make up the inborn deficiencies of IP protocol in the network security, it is proposed IPSec protocol set, one of them is IKE which is responsible for authentication and key exchange, and it is the important guarantee to that the whole protocol can be smoothly implemented. Its second phase is consultation SA for the specific service on the basis of the first stage; The first phase is set up the secure channel namely SA (Security Association), including two modes (main mode / active mode) and four types of authentication methods (digital signature / public key encryption / modified public key encryption / shared private key), in which the main mode is more complex than the active mode and must be implement. Here we will prove in detail the security protocol on the foundation of the main mode of public key system.

Initiator (I)		Responder(R)
HDR, SA_i	→	
	←	HDR, SA_r
HDR, KE_i, $\{I\}_{Pr}$, $\{N_i\}_{Pr}$	→	
	←	HDR, KE_r, $\{R\}_{Pi}$, $\{N_r\}_{Pi}$
HDR, $\{HASH\_I\}_k$	→	
	←	HDR, $\{HASH\_R\}_k$

We will introduce in brief about the content of protocol under the main mode based on public key system. In the protocol, HDR is the header information, which consists of mainly the cookies of a couple of initiator and responder. According to the SA payload, initiator put forward a set of alternative security options (like encryption algorithm represented by SA_i), then responder will choose an answer (SA_r) among them. KE_i, KE_r are the contents of Diffie-Hellman exchange, which is used to consult private key. N_i, N_r are the random number of exchange. I, R are both identities, and Pi and Pr are respectively the public keys of I and R. At the end, what are exchanged are hash values encrypted by the consulted algorithm and the exchanged private key, in which the calculation of HASH_I and HASH_R are following, where the prf and hash can be comprehended simply as hash function.

$SKEYID = prf(hash(N_i|N_r), C_i|C_r)$

$HASH\_I = prf(SKEYID, KE_i|KE_r|C_i|C_r|SA_i| I)$

$HASH\_R= prf(SKEYID, KE_r|KE_i|C_r|C_i|SA_i|R)$

Figure1 are the implementation of the protocol in the first phase of IKE based on the main mode of public key system. We will make a minor amendment of the exchange of data to simplify the data in the HDR for a pair of Cookie. Because here the nodes are regular and can make a distinction according to the Cookie previous received before for data latter, we can omit the Cookie after the third step. Taking into account the HASH_I and HASH_R are the proof for owning all data, which will be expressed in the form of a hash function respectively.

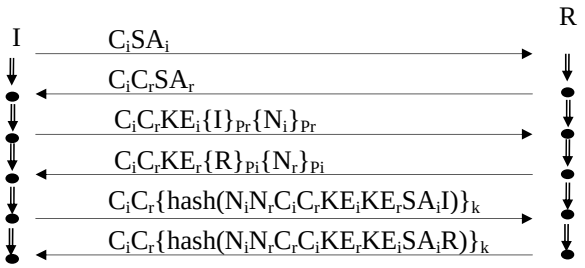

Figure 1. Part of IKE protocol strand graphic description

Definition 3.1 Set of strand with the trace $<+C_iSA_i$, $-C_iC_rSA_r$, $+KE_i\{I\}_{Pr}\{N_i\}_{Pr}$, $-KE_r\{R\}_{Pi}\{N_r\}_{Pi}$, $+\{hash(N_iN_rC_iC_rKE_iKE_rSA_iI)\}_k$, $-\{hash(N_iN_rC_rC_iKE_rKE_iSA_iR)\}_k>$ is designated by $Init(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$, in which $C_i,C_r,SA_i,SA_r,KE_i,KE_r \in T\backslash\{T_{name} \cup T_{nonce}\}$, $I,R \in T_{name}$, $N_i,N_r \in T_{nonce}$. I is the entity related to a strand in that set.

Definition 3. 2 Set of strand with trace $<-C_iSA_i$, $+C_iC_rSA_r$, $-KE_i\{I\}_{Pr}\{N_i\}_{Pr}$,$+KE_r\{R\}_{Pi}\{N_r\}_{Pi}$,$-\{hash(N_iN_rC_iC_rKE_iKE_rSA_iI)\}_k$, $+\{hash(N_iN_rC_rC_iKE_rKE_iSA_iR)\}_k>$ is designated by $Resp(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$, in which $C_i,C_r,SA_i,SA_r,KE_i,KE_r \in T\backslash\{T_{name} \cup T_{nonce}\}$, $I,R \in T_{name}$, $N_i,N_r \in T_{nonce}$. R is the entity related to a strand in that set. Sign * represents that can be replaced by any value, like that $Init(I,*,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$ represents that the one to communicate with entity I is not sure. The strand space of that protocol is $\Sigma = Init \cup Resp \cup P$, in which P represents all possible penetrator strand. The propositions following describe: No matter N_i or N_r is not used before.

Proposition 3. 1 If $s \in Init(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$, then node <s,3> is the access point of $I_K[N_i]$.

Prove: Because $C_i,C_r,SA_i,SA_r \in T\backslash\{T_{name} \cup T_{nonce}\}$, and $N_i \in T_{nonce}$, there is no perm of $I_K[N_i]$ existing in nodes before <s,3>, and $\{N_i\}_{Pr} \in I_K[N_i]$, so node <s,3> is the access point of $I_K[N_i]$. □

Proposition 3 .2 If $s \in Resp(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$ and N_i and N_r are different, then node <s,4> is the access point of $I_K[N_r]$.

Theorem 3.1 There is bundle C in Σ, I and R are different, and there is the only access point in C for $I_K[N_i]$. If the C-height is 6, and $s \in Init(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$, then $r \in Resp(I,R,C_i,C_r,SA_i,*,KE_i,KE_r,N_i,N_r)$ and the C-height is 6 at least.

Theorem 3.2 C is a bundle in Σ, and I is different with R, N_i is different with N_r also; $I_K[N_r]$ is the only access point in C. If $r \in Resp(I,R,C_i,C_r,SA_i,SA_r,KE_i,KE_r,N_i,N_r)$, and the height of C- is 6, there exists $s \in Init(I,R,C_i,C_r,SA_i,*,KE_i,KE_r,N_i,N_r)$, and the height of C- is 5 at least.

Theorem 3.3 C is a bundle in Σ, $Pr^{-1} \notin K_p$, then the access point of W [I] can not be the penetrator node except M.

Prove: Let $S=\{I\}$. $Pr^{-1} \notin K_p$, and according to Definition 3.1 and 3.2, C is normal. So according to Lemma 2.1, the access point of W[I] can not be penetrator node except M and K. Also $I \in T_{name}$, it is impossible that it is accessed from node K.

□

Theorem 3.4 C is a bundle in Σ, $Pi^{-1} \notin K_p$, then the access point of W[R] can not be the penetrator node except M.

These two theorems indicate the penetrator can only rely on guessing to get the identity of the initiator or the respondents.

4 Conclusions

Strand space model based on algebraic methods, which has the credibility and accuracy to prove security protocol and are relatively easy, and often finished manually. We extend and improve the strand space model by adding the description of the hash function, meanwhile it gives the associated nature definition and lemma, according to the improvement and intense of the new strand space model, we can analyze and validate with the protocol with the hash function. Take IKE protocol as an example, we chose one of a sub-protocol for a complete analysis, and demonstrate how to use the intensive model and our lemma-verified security protocol to shows that the new verification capacity of an intensive model. The current protocols and the network communication security protocols in particular are more and more complicated (such as IKE), which is usually composed of multiple sub-protocols, and commonly in practical applications, there are usually a number of protocols existing at the

978-1-935068-01-3 © 2009 SciRes.

same time, so how to prove security of the protocol in a mixed protocol space will be our next step.

References(参考文献)

[1] W. Diffie, P.C. van Oorschot, and M.J. Wiener. Authentication and Authenticated Key Exchanges. Designs, Codes and Cryptography, 2:107--125, 1992.
[2] D. Harkins, and D. Carrel. The Internet Key Exchange (IKE), RFC 2409, November 1998.
[3] Catherine Meadows. The NRL Protocol Analyzer: An overview. Journal of Logic Programming, 26(2):113--131, 1996.
[4] Roger Needham and Michael Schroeder. Using Encryption for Authentication in Large Networks of Computers. Communications of the ACM, 21(12), December 1978.
[5] Lawrence C. Paulson. The inductive approach to verifying cryptographic protocols. Journal of Computer Security, 6:85--128, 1998.
[6] F. Javier Thayer Fabrega, Jonathan C. Herzog, and Joshua D. Guttman. Honest ideals on strand spaces. In Proc. 11th IEEE Computer Security Foundations Workshop. IEEE Computer Society Press, June 1998.
[7] J. Thayer, J. C. Herzog, and J. D. Guttman. Strand spaces: Proving security protocols correct. Journal of Computer Security, 7, 1999.
[8] R. Thayer, N. Doraswamy, and R. Glenn. IP Security Document Roadmap, RFC 2411, Nov 1998.
[9] J. Zhou. Fixing a security flaw in IKE protocols. Electronics Letters, 35(13):1072--1073, June 19