A Note on "Three-Factor Anonymous Authentication and Key Agreement Based on Fuzzy Biological Extraction for Industrial Internet of Things"

Zhengjun Cao, Lihua Liu

Abstract. We show that the key agreement scheme [IEEE Trans. Serv. Comput. 16(4): 3000-3013, 2023] fails to keep user anonymity, not as claimed. The scheme simply acknowledges that user anonymity is equivalent to preventing user's identity from being recovered. But the true anonymity means that the adversary cannot attribute different sessions to target users. It relates to entity-distinguishable, not just identity-revealable. To the best of our knowledge, it is the first time to clarify the explicit signification of user anonymity.

Keywords: Key agreement, anonymity, mutual authentication, entity-distinguishable, identity-revealable.

1 Introduction

Recently, Xu *et al.* [1] have presented a mutual authentication and key agreement protocol in Industrial Internet of Things (IIoT) environment. It is designed to meet many security requirements, such as mutual authentication, session key establishment, user anonymity, forward secrecy, resistance to replay attack, man-in-the-middle attack, modification attack, DoS attacks, etc.

In the proposed scenario, there are different entities including trusted authority (TA), user (U_k), control node (CN_i), and smart sensor device (SD_j). The scheme consists of five phases: Initialization, Registration, Login and authentication, User join/revocation and session key update phase, Smart sensor devices join phase. TA picks a prime q to generate public parameters $F_q, E/F_q, G, P, p$ for elliptic curve domain, where P is a base point. Set $pr_{TA} \in Z_p^*$ as its secret key, and $Pub_{TA} = pr_{TA} \cdot P$ as its public key. Let $h(\cdot)$ be a hash function, $Gen(\cdot)$ be a probabilistic generation function and $Rep(\cdot)$ be a reproduction function of Fuzzy Extractor. The scheme can be briefly depicted as follows (see Table 1). In this note, we show that the scheme fails to keep user anonymity, not as claimed.

Z. Cao is with Department of Mathematics, Shanghai University, Shanghai, 200444, China.
L. Liu is with Department of Mathematics, Shanghai Maritime University, Shanghai, 201306, China.
Email: liulh@shmtu.edu.cn

Table 1: The Xu *et al.*'s key agreement scheme

CN_i	TA	SD_j
$\xRightarrow[\text{[secure channel]}]{ID_{CN_i}}$	Pick $pr_{CN_i}, r_{CN_i} \in Z_p^*$. Compute $Pub_{CN_i} = pr_{CN_i} \cdot P$, $R_{CN_i} = r_{CN_i} \cdot P,\ RID_{CN_i} = H(ID_{CN_i} \| pr_{TA})$, $Cert_{CN_i} = pr_{TA} + h(RID_{CN_i} \| Pub_{TA} \| Pub_{CN_i}) \cdot r_{CN_i}$.	
Store the parameters RID_{CN_i}, pr_{CN_i}, $R_{CN_i}, Cert_{CN_i}$.	$\xLeftarrow{\{RID_{CN_i}, pr_{CN_i}, Pub_{CN_i}, R_{CN_i}, Cert_{CN_i}\}}$ Pick $pr_{SD_j}, r_{SD_j} \in Z_p^*$. Compute $Pub_{SD_j} = pr_{SD_j} \cdot P,\ R_{SD_j} = r_{SD_j} \cdot P$, $RID_{SD_j} = h(ID_{SD_j} \| pr_{TA})$, $Cert_{SD_j} =$ $pr_{TA} + h(RID_{SD_j} \| Pub_{TA} \| Pub_{CN_i} \| Pub_{SD_j}) \cdot r_{SD_j}$, $s_{CN_i,SD_j} = h(RID_{CN_i} \| RID_{SD_j} \| r_{CN_i} \| r_{SD_j} \| ts_{SD_j})$, where ts_{SD_j} is a timestamp.	$\xLeftarrow{ID_{SD_j}}$
Store the shared key s_{CN_i,SD_j} into the memory.	$\xLeftarrow{s_{CN_i,SD_j}}$ $\xRightarrow[RID_{CN_i}, RID_{SD_j}, RID_{CN_i}, pr_{SD_j},]{Pub_{SD_j}, R_{SD_j}, Cert_{SD_j}, s_{CN_i,SD_j}}$	Store the parameters RID_{SD_j}, $pr_{SD_j}, R_{SD_j}, Cert_{SD_j}, s_{CN_i,SD_j}$.
CN_i	TA	U_k
	Pick $r_{U_k} \in Z_p^*$. Compute $R_{U_k} = r_{U_k} \cdot P$, $pid_{U_k} = h(ID_{U_k} \| pr_{TA})$, $Cert_{U_k} = r_{U_k} + h(pid_{U_k} \| R_{U_k}) \cdot pr_{TA}$.	$\xLeftarrow{ID_{U_k}}$ Input identity ID_{U_k}, password PW_{U_k}. Imprint the biometric BIO_{U_k}. Compute $Gen(BIO_{U_k}) = (\sigma_{U_k}, \tau_{U_k})$,
Store the user's pseudo identity pid_{U_k} and accession number R_{U_k}.	$\xLeftarrow[\text{[secure channel]}]{\{pid_{U_k}, R_{U_k}\}}$ $\xRightarrow[\text{[secure channel]}]{\{pid_{U_k}, R_{U_k}, Cert_{U_k}\}}$	$L_{U_k} = h(ID_{U_k} \| \sigma_{U_k} \| PW_{U_k})$. Check if $Cert_{U_k} \cdot P = R_{U_k} + h(pid_{U_k} \| R_{U_k}) \cdot Pub_{TA}$. Pick $pr_{U_k} \in Z_p^*$ to set $Pub_{U_k} = pr_{U_k} \cdot P$. Compute $W_{U_k} = h(Cert_{U_k} \| pid_{U_k} \| R_{U_k} \| pr_{U_k})$. Store $pid_{U_k}, R_{U_k}, Cert_{U_k}, W_{U_k}, L_{U_k}, \tau_{U_k}$.
U_k: $\{pid_{U_k}, R_{U_k}, Cert_{U_k}, W_{U_k}, L_{U_k}, \tau_{U_k}\}$	CN_i: $\{pid_{U_k}, R_{U_k}, RID_{CN_i}, pr_{CN_i}, R_{CN_i}, Cert_{CN_i}\}$	SD_j: $\{RID_{SD_j}, pr_{SD_j}, R_{SD_j}, Cert_{SD_j}\}$
Input $ID_{U_k}, PW_{U_k}, BIO_{U_k}$. Compute $\sigma_{U_k} = Rep(BIO_{U_k}, \tau_{U_k})$. Check $L_{U_k} = h(ID_{U_k}, \| \sigma_{U_k} \| PW_{U_k})$. If so, pick the timestamp ts_1, $a \in Z_p^*$. Compute $A = a \cdot P$, $PID_{U_k} = h(pid_{U_k} \| ts_1)$. $Auth_{U_k} = (a + Cert_{U_k}) \cdot Pub_{SD_j}$, $Ver_{U_k} = h(ts_1 \| PID_{U_k} \| A \| R_{U_k})$. $\xrightarrow[\text{[open channel]}]{M_1=\{ts_1,\ PID_{U_k},\ A,\ R_{U_k},\ Ver_{U_k}\}}$	Check the timestamp ts_1. Use R_{U_k} to retrieve pid_{U_k}. Check $PID_{U_k} = h(pid_{U_k} \| ts_1)$ and $Ver_{U_k} = h(ts_1 \| PID_{U_k} \| A \| R_{U_k})$. If so, compute $Auth_{CN_i} = (pr_{CN_i} + Cert_{CN_i}) \cdot Pub_{SD_j}$, $X_i = h(Auth_{CN_i} \| s_{CN_i,SD_j})$, $Ver_{CN_i} =$ $h(PID_{U_k} \| R_{U_k} \| RID_{CN_i} \| R_{CN_i} \| A \| ts_1 \| ts_2 \| X_i)$.	Check the timestamp ts_2. Check $Ver_{CN_i} =$ $h(PID_{U_k} \| R_{U_k} \| RID_{CN_i} \| R_{CN_i} \| A \| ts_1 \| ts_2 \| X_i)$. If so, compute $Auth'_{CN_i} = Pub_{CN_i} + Pub_{TA} +$ $h(RID_{CN_i} \| Pub_{TA} \| Pub_{CN_i}) \cdot R_{CN_i} \cdot pr_{SD_j}$. Check $X_i = h(Auth'_{CN_i} \| s_{CN_i,SD_j})$. If so, pick $b \in Z_p^*$, compute $B = b \cdot P$, $k_{SD_j} = b \cdot A$,
Check the timestamp ts_3. Check if $Ver_{SD_j} = h(B \| RID_{SD_j} \| R_{SD_j} \| ts_3 \| Y_j)$. If so, compute $k_{U_k} = a \cdot B$, $s_{U_k,SD_j} = h(k_{U_k} \| Auth_{U_k} \| ts_1 \| ts_3)$, $Auth_{SD_j}^{new'} = Pub_{SD_j} + Pub_{TA} + h(RID_{SD_j}$ $\| Pub_{TA} \| Pub_{CN_i} \| Pub_{SD_j}) \cdot R_{SD_j} \cdot pr_{U_k}$. Check if $Y_j = h(Auth_{SD_j}^{new'} \| s_{U_k,SD_j})$.	$\xrightarrow[\text{[open channel]}]{M_2=\{PID_{U_k},\ R_{U_k},\ RID_{CN_i},\ A,\ ts_1,\ ts_2,\ X_i,\ Ver_{CN_i}\}}$ $\xleftarrow[\text{[open channel]}]{M_3=\{B,\ RID_{SD_j},\ ts_3,\ Y_j,\ Ver_{SD_j}\}}$	$Auth_{SD_j} = pr_{SD_j}(A + R_{U_k} + h(PID_{U_k} \| R_{U_k}) \cdot Pub_{TA})$, $s_{U_k,SD_j} = h(k_{SD_j} \| Auth_{SD_j} \| ts_1 \| ts_3)$, $Auth_{SD_j}^{new} = (pr_{SD_j} + Cert_{SD_j}) Pub_{U_k}$, $Y_j = h(Auth_{SD_j}^{new} \| s_{U_k,SD_j})$, $Ver_{SD_j} = h(B \| RID_{SD_j} \| R_{SD_j} \| ts_3 \| Y_j)$.

2 The signification of user anonymity

Anonymity is a security requirement adopted by many cryptographic protocols. But we find its signification is often misunderstood. We want to stress that the true user anonymity means that the adversary cannot attribute different sessions to target users. In other words, it actually relates to entity-distinguishable, not just identity-revealable. To illustrate the explicit signification of anonymity, we refer to Fig.1.

In Fig.a, the user's identity ID_{U_k} uniquely corresponds to the pseudo-identity pid_{U_k}, which uniquely corresponds to the accession number R_{U_k}. Thus, different sessions (launched by this entity) can be attributed to the entity. In this case, *the unique accession number can be eventually used to recognize this entity.* But in Fig.b, pid_{U_k} corresponds to different random accession numbers $R_{U_k}^{(1)}, \cdots, R_{U_k}^{(n)}$. Therefore, the adversary cannot attribute different sessions to the entity, even though these sessions are launched by this entity.

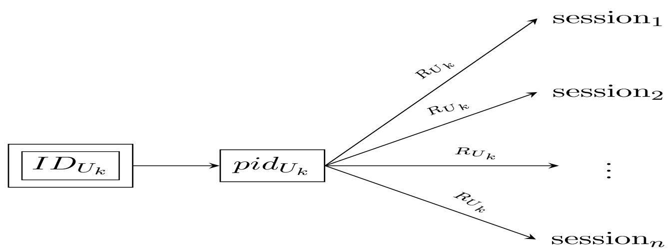

Fig.a: The false anonymity
(with the same accession number R_{U_k})

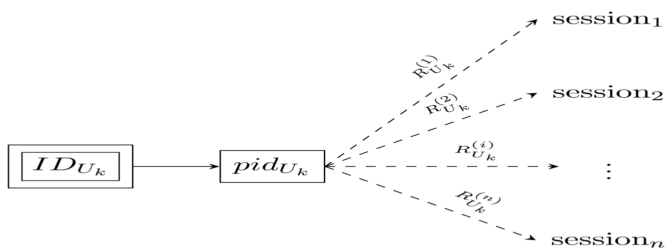

Fig.b: The true anonymity
(with different accession number $R_{U_k}^{(i)}$)

Figure 1: False anonymity versus true anonymity

3 The loss of user anonymity

The identity of a person or thing is the characteristics that distinguish it from others. In the scheme, the user U_k's *real identity* ID_{U_k} could be a regular string of some meanings, while the *pseudo-identity* pid_{U_k} is a random string, i.e.,

$$pid_{U_k} = h(ID_{U_k} \| pr_{TA}) \tag{1}$$

Since a real identity uniquely corresponds to a pseudo-identity (due to the collision-free property of the hash function h), one should prevent both identifiers from exposure. So, the user U_k needs to generate the *session pseudo-identity*

$$PID_{U_k} = h(pid_{U_k} \| ts_1) \tag{2}$$

where ts_1 is the current timestamp. Since the value of PID_{U_k} randomly varies in different sessions, the adversary cannot attribute different sessions to a target user. Based on this observation, the scheme is claimed to be of user anonymity (see §V, A, [1]). But we find the claim is false.

As we know, the controller node CN_i serves for a lot of users. By the received message

$$M_1 = \{ts_1, PID_{U_k}, A, R_{U_k}, Ver_{U_k}\} \tag{3}$$

which is sent via a public channel, the controller node should retrieve the target user's pseudo-identity pid_{U_k} from its local database. To do this, the node needs to use the accession number R_{U_k} to search out the long-term pseudo identity pid_{U_k}. Then compute the session pseudo identity PID_{U_k} and check its consistency.

The accession number R_{U_k} is a long-term parameter, which is issued by the trust authority in the user join phase. An accession number uniquely corresponds to a legitimate user. The adversary can obtain R_{U_k} from the captured message M_1. So, the accession number can be used to recognize the target user even though the adversary cannot use it to reveal the strings ID_{U_k} and pid_{U_k}.

To fix this flaw, one should specify a mechanism to randomly update the shared accession number R_{U_k} in each session, both for the user U_k and controller node CN_i. We refer to Ref.[2] for a possible updating mechanism, in which the shared one-time temporary identity tid_w between the sensor and controller node is randomly updated as tid_w^{new}.

By the way, the scheme also fails to keep controller node anonymity and smart device anonymity. In fact, the message

$$M_2 = \{PID_{U_k}, R_{U_k}, RID_{CN_i}, A, ts_1, ts_2, X_i, Ver_{CN_i}\}$$

contains RID_{CN_i}, where $RID_{CN_i} = H(ID_{CN_i} \| pr_{TA})$. The adversary can use the long-term random pseudo identity RID_{CN_i} to recognize the target controller node. Likewise, the message $M_3 = \{B,\ RID_{SD_j},\ ts_3,\ Y_j,\ Ver_{SD_j}\}$ contains RID_{SD_j}, where $RID_{SD_j} = h(ID_{SD_j} \| pr_{TA})$ which can also be used to recognize the target smart device.

4 Conclusion

We show that the Xu *et al.*'s key agreement scheme is flawed. We hope the findings in this note could be helpful for the future work on designing such key agreement schemes.

References

[1] H. Xu, et al.: Three-Factor Anonymous Authentication and Key Agreement Based on Fuzzy Biological Extraction for Industrial Internet of Things. *IEEE Trans. Serv. Comput.* 16(4): 3000-3013 (2023)

[2] M. Zia, et al.: A Provably Secure Lightweight Key Agreement Protocol for Wireless Body Area Networks in Healthcare System. *IEEE Trans. Ind. Informatics* 19(2): 1683-1690 (2023)