PASSWORD AUTHENTICATED KEY EXCHANGE: FROM TWO PARTY METHODS TO GROUP SCHEMES

STEPHEN MELCZER, TARAS MYCHASKIW AND YI ZHANG

Abstract. This project investigates password authenticated key exchange methods (PAKEs), and variants involving many agents: group password authenticated key exchange methods (GPAKEs). In particular, after surveying classical information on two party PAKEs we detail the so-called 'fairy ring dance' method recently described by Hao et. al. [13] which allows for the extension of two party password authenticated key exchange methods (PAKEs) with key confirmation to a group setting with an arbitrary number of users without increasing round complexity (the computational complexity, of course, increases with the number of users). This paper presents two new GPAKEs constructed through these means, based on the Dragonfly and PAK/PPK two party protocols, and includes timings comparing them to previous GPAKEs of Hao et. al. [13].

1. Introduction

Since their introduction in the 1990s, password-authenticated key exchange (PAKE) methods – also known as password-authenticated key agreement methods – have become popular for their ability to allow agents sharing a (typically low entropy) password to securely establish shared cryptographic keys (see Bellovin and Merritt [5] or Jablon [15] for early examples, and Hao and Ryan [12] for a more recent paper). Although they have been around for decades, most research on PAKEs has focused on key establishment between two parties. For our project, we have studied the problem of establishing Group PAKEs (GPAKEs) – that is, using a low entropy password shared between many agents to set up cryptographic keys. This has modern applications with the rise of the so-called 'Internet of Things', where many consumer devices connected through a local Internet connection wish to securely communicate (such schemes would allow, for instance, secure communication between a smart television, DVD player, and cable box after their owner inputs a short shared password into each upon purchase).

The major issue in designing an efficient GPAKE is to minimize the number of rounds of communication between the agents involved, as the latency of such a protocol is determined by the slowest responder in each round. A recent pre-print of Hao et. al. [13] proposes a construction which allows for the extension of any secure two-party PAKE with key confirmation to a multi-party PAKE, without adding any extra rounds of communication (if the underlying two-party PAKE only allows for key authentication, then the associated GPAKE has one extra round of communication). The authors continue on to give two explicit schemes following from this template: SPEKE+ (using two rounds of communication, adapted from the SPEKE [15] protocol) and J-PAKE+ (using three rounds of communication, adapted from the J-PAKE [12] protocol).

The structure of this document is as follows: Section 2 begins by giving a survey of classical two-party PAKEs – including explicit descriptions of the PAKEs which will be extended into the group setting. Section 3 starts with a description of the theoretical methodology developed by Hao et. al. [13] to extend two-party PAKEs into a group setting. After this general background, we give two explicit GPAKEs (SPEKE+ and J-PAKE+) constructed by Hao et. al. using this methodology, followed by two explicit GPAKEs which we have derived through the same means (a group variant of the IEEE 802.11-2012 standard Dragonfly protocol [14], and a variant of the PAK/PPK protocol [6]). Security properties of these GPAKEs follow from the security properties of the underlying two party PAKES, in a manner described by Hao et. al. [13] and in Section 3 of this document. In Section 4 we test the practical efficacy of our new methods against the Java implementations of SPEKE+ and J-PAKE+ given by Hao et. al. [13]. Section 5 concludes with an

overview of these results and possible directions for future work.

The main original contributions found in this project come from the two new group PAKEs we have constructed – see sub-Sections 2.3 and 2.4 – and timings which compare these methods against previous Java implementations of SPEKE+ and J-PAKE+ (the code for this project is available at https://github.com/twentylemon/gpake). We also survey the relevant background material on Zero Knowledge Proofs and classical PAKEs missing from Hao et. al. [13] (which had constrained space as a conference abstract), and fixed some minor Java implementation oversights which could cause the timings in that paper to be slightly inaccurate.

2. Two Party Password-Authenticated Key Exchange (PAKE) [1]

The genesis of password-authenticated key exchange is widely credited to the 1992 work of Bellovin and Merrit [5], whose protocol – known as Encrypted Key Exchange, or EKE, for short – came to be known as the first PAKE (previous password based protocols, like the one proposed in 1989 by Lomas et.al. [16], contained key features of PAKEs such as the offline dictionary attack resistance detailed below, although they still relied on one party having another's public key). All PAKEs aim for two main goals: to require their users to provide a zero knowledge proof of a short password known to both parties *a priori* (that is, before the protocol has begun) and to leverage knowledge of this password to facilitate an authenticated key exchange. As password are assumed to be low entropy – for instance, they are often treated as human memorable passwords (typically assumed to be approximately 20-30 bits of entropy) – if the passwords themselves were broadcast they would need to be protected, for instance using SSL. This would require Public Key Infrastructure, such as a Trusted Authority / Certificate Authority to maintain public keys, which can be expensive. The ability to work around such infrastructure is often the point of PAKE protocols, which essentially use pre-established shared knowledge (of the common password) as an alternative to Trusted Authorities.

Indeed, it is somewhat miraculous that PAKEs – which transform a low entropy shared secret into a much larger and more complicated shared key – exist at all. Although the EKE protocol of Bellovin and Merrit was later shown to have weaknesses (see Jaspan [15], for example) its great contribution was to show that such schemes can be achieved. Due to its historical significance, we outline the Diffie-Hellman variant of the method here (an RSA variant, also by Bellovin and Merrit, was later shown to be insecure). Given a symmetric encryption function $[\cdot]_\pi$ which uses a password π shared by agents Alice and Bob as a key, the algorithm does the following:

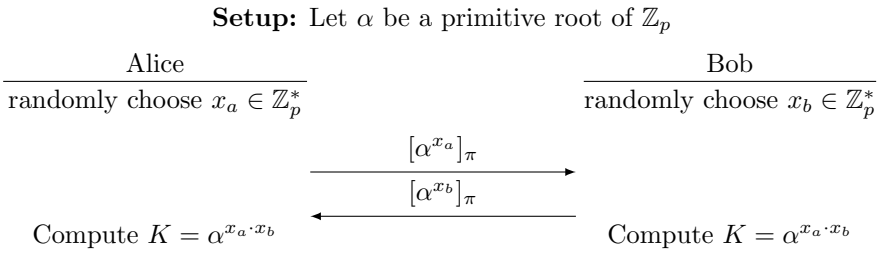


Figure 1. The flow diagram for EKE

At the end, both Alice and Bob share the key $K = \alpha^{x_a \cdot x_b}$. The weaknesses of the algorithm stems from the issues discussed above: as the password has low entropy in order for the scheme to be secure the input into $[\cdot]_\pi$ must essentially look like a random number. But a 1024 bit number modulo p is not random, and a passive attacker can try candidate passwords π' to attempt to decipher $[\alpha^{x_a}]_\pi$ and immediately rule out

[1] The background information in this section, and details about SPEKE and J-PAKE, are mainly based on the presentation in Hao and Ryan [12]. The information on the Dragonfly protocol was taken from Harkins [14] and Clarke and Hao [9]. The sub-section about PAK/PPK is based on the work of Boyko et. al. [6].

any passwords giving a result in the range $[p, 2^{1024} - 1]$.

Although EKE has this, and other, weaknesses, it was extremely influential and its general characteristics are reflected in many of the more advanced protocols we outline below (there is also a minor variant known as EKE2, which was shown to be secure by Bellare, Pointcheval, and Rogaway [4]). Before giving these methods, we must outline what constitutes a good measure of security for a PAKE. To begin, a secure two party PAKE satisfies each of the following properties coming from the security of general key exchange protocols:

(Offline dictionary attack resistance)
The PAKE does not leak any information to a passive or active attacker which can be used by the attacker to determine the password through a brute force search (the protocol cannot reveal a hash of the password, for instance).

(Forward secrecy for established keys)
If the password is disclosed, past session keys cannot be computed by an attacker. This implies that a *passive* attacker who knows the password cannot learn a session key by observing communication between Alice and Bob (of course, an active attacker could establish a shared key with one of the participants as he would have access to all of their secret information).

(Known session security)
Even if an attacker learns all session specific secrets of a protocol in progress, these secrets do not reveal any information about other established sessions.

(Online dictionary attack resistance)
An active attacker can only test one password per protocol execution (this is the best that we can reasonably assume, as any attacker can randomly guess a password and run the protocol – at some point the key must be confirmed, either explicitly through the PAKE or when the key is used in some other protocol, and the attacker will know whether or not his guess was correct). This is sometimes relaxed (for instance, in the proof of SPEKE given by MacKenzie [17]) to restricting the attacker to at most a small constant number of tests per protocol execution.

In the modern literature, a full proof of security essentially requires showing that an attacker can only gain information about established keys or a shared password if he is active, and that even an active attacker can gain extremely little information (for instance, can only guess one password per protocol execution). The formal model of Bellare, Pointcheval, and Rogaway [4] is commonly used as a standard, and three of the four PAKEs discussed later in this section have been proven secure – under various assumptions, see Figure 1 – in this model (the fourth, which is the Dragonfly protocol of sub-Section 2.3, we include despite a formal proof of security as it is an IEEE 802.11-2012 standard). We refer an interested reader to the work of Abdalla et. al. [2] and MacKenzie [18] for in-depth discussions of PAKE security and attack models.

	Rounds / Flows	Assumptions	Communication[2]	Time
J-PAKE w/ Schnorr	2 / 4	ROM, AAM, DSDH	$12 \times G + 6 \times \mathbb{Z}_p$	28 exponentiations
SPEKE	1 / 2	ROM, DIDH	$2 \times G$	8 exponentiations
PPK	2 / 2	ROM, DDH	$2 \times G$	6 exponentiations

TABLE 1. Table comparing the security assumptions needed for the provably secure methods (in the BPR model) discussed here. 'Communication' and 'Time' refer to the complexity of the algorithms, taken from Abdalla et. al. [2]. The assumptions are described below.

Some PAKEs satisfy the additional requirement that an attacker not be allowed to impersonate other users to some fixed target after obtaining (through illicit means) password verification files for those users which were stored by the target. The schemes with this additional property are known as augmented PAKEs, although some (for instance, Hao et. al. [13]) have argued that such a requirement is not useful as the low entropy of the password means that it will soon be discovered through an offline dictionary attack on the verification files. Nevertheless, augmented variants exist for a number of PAKEs (for example Augmented-EKE for EKE, B-SPEKE for SPEKE and PAK-X for PAK/PPK).

2.1. **SPEKE.** A more advanced PAKE which we consider is the Simple Password Exponential Key Exchange (SPEKE) protocol, designed by Jablon [15] in 1996. SPEKE tries to work around the deficiencies of EKE by using the shared password π of the two participants to change the generator of a Diffie-Hellman like scheme. The protocol runs as follows:

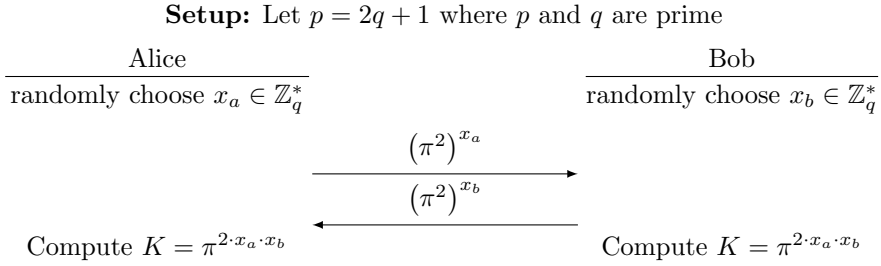


FIGURE 2. The flow diagram for SPEKE

Note that the password is squared so that the exponentiations in the protocol can occur in a subgroup of prime order q (the participants must check that that $\pi^2 \not\equiv \pm 1 \bmod p$ – if this is not the case, all work is carried out in the order 2 subgroup of $\mathbb{Z}_p^*$ which renders the protocol insecure, and a new password or prime p must be chosen).

There are drawbacks to using the password directly: mainly that an attacker can guess multiple passwords in one execution of the protocol (as multiple passwords may have the same square mod p) and that the size of the subgroup in which the protocols occur is large (if p is a 1024-bit prime, for example, then q is a 1023-bit prime). Although it is troubling to allow an active attacker multiple guesses at the password, for practical purposes as long as they are limited to a small constant number of guesses per protocol execution the security of the protocol can be safely assumed. Indeed, a variant of the basic protocol presented in Figure 2 where a hash of the password is squared (as defined by the IEEE P1363.2 standard regarding SPEKE[3]) was later proven secure by MacKenzie [17] in a common formal model (proposed by Boyko, MacKenzie and Patel [6]) under the assumptions of the random-oracle model and the hardness of the decision inverted-additive Diffie-Hellman (DIDH) problem[4]. Here, the notion of 'secure' is relaxed to allow an active attacker to rule out a (small) constant number of guesses per protocol execution.

[2] G refers to the sending of an element from the cyclic group where each algorithm takes place – which does not necessarily have to be a subgroup of the units of a finite field, although our implementations and descriptions always use such a G – and $\mathbb{Z}_p$ refers to the sending of a member of the finite field.

[3] Our implementation of SPEKE uses this variant

[4] This somewhat non-standard Diffie-Hellman assumption asks one to distinguish between an element $g^{(x+y)^{-1}}$ and a random group element, given the elements $X = g^{x^{-1}}$ and $Y = g^{y^{-1}}$. It has been shown that if the typical computational Diffie-Hellman problem (CDH) is hard, so is the computational inverted-additive Diffie-Hellman problem. Furthermore, if the Decision Square Diffie-Hellman problem is hard (which is assumed in the security proof of the J-PAKE protocol, discussed below) then the DIDH problem is hard. See Figure 2 of Abdalla et. al. [2] for a comparison of all the Diffie-Hellman type assumptions used in the protocols presented here.

2.2. **J-PAKE.** In 2010, Hao and Ryan [12] proposed the Password Authenticated Key Exchange by Juggling (J-PAKE) protocol, at least in part to get around the deficiencies of the SPEKE method described in the previous section.[5] J-PAKE, which is used in Firefox and (as an optional protocol) in OpenSSL, among others, is quite straightforward and uses the shared password to make a nice simplification of randomly chosen Diffie-Hellman like exponentiations.

The protocol as specified relies on a Zero Knowledge Proof (ZKP) of an exponent: that is, a protocol such that a sender can transmit $X = g^x$ and a message which allows a receiver to determine almost certainly that the sender knows x, without revealing any knowledge of x (the element X is assumed to be a member of a group in which the computational Diffie-Hellman problem is hard). Our implementation uses the common Schnorr non-interactive ZKP: roughly, the sender transmits their ID, sID, along with the values

$$V = g^v \qquad \text{and} \qquad r = v - xh$$

where $v \in_R \mathbb{Z}_q$ and $h := H(g||V||X||sID)$ for a suitable hash function H. The receiver checks that X is in the proper group, that h is the correct hash, and that $V = g^r \cdot X^h$. This choice of ZKP is also used by Hao et. al. [13], and we refer the reader to that paper for more details.

We are now ready to describe the protocol. Let Q be a subgroup of $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ with prime order q, g be a generator of this subgroup, and $\pi \in \mathbb{Z}_q^*$ be the shared password between Alice and Bob. J-PAKE consists of a setup round followed by two rounds of communication:

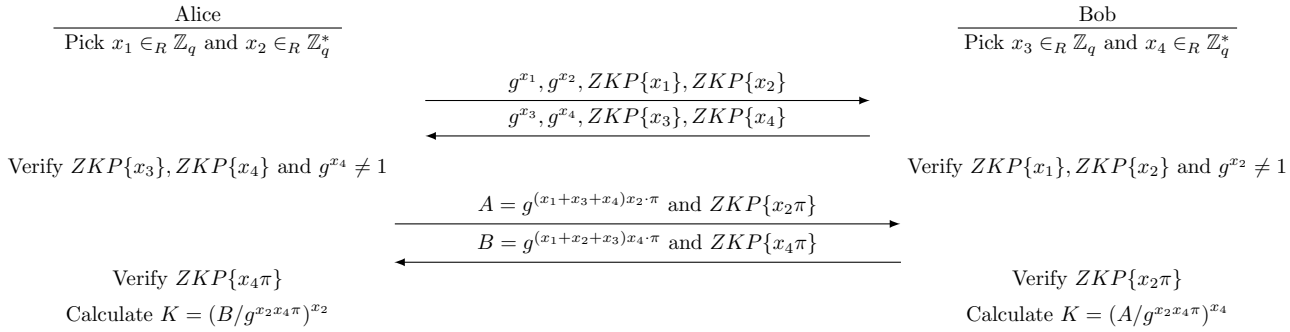


FIGURE 3. The J-PAKE protocol.

As noted in the figure, both Alice and Bob are able to determine

$$K = \underbrace{(B/g^{x_2 x_4 \pi})^{x_2}}_{\text{Computable by Alice}} = g^{(x_1+x_3)x_2 x_4 \pi} = \underbrace{(A/g^{x_2 x_4 \pi})^{x_4}}_{\text{Computable by Bob}} .$$

The shared session key is then taken to be $\kappa = H(K)$, where H is a suitable hash function. We note that J-PAKE (like SPEKE) admits only key authentication: each of Alice and Bob know that the only people who can calculate the shared key are themselves; key confirmation can additionally be performed if desired (which will increase the number of rounds of communication by one).

In their original paper, Hao and Ryan [12] gave proofs that J-PAKE satisfies the four security properties (offline dictionary attack resistance, forward secrecy for established keys, known session security, and online dictionary attack resistance) discussed at the beginning at this section. Although these proofs were straightforward, they did not take place in the framework of an established and commonly used formal model for PAKE security, and relied on unstated assumptions about an adversary's range of potentially attack techniques. As Feng Hao, one of the J-PAKE authors, wrote in a blog post[6]: "Some researchers might like to take it from here and add more 'formalism' into the paper. I'm sure that will be a valuable addition in

[5] In addition to the security flaws, such as allowing multiple guesses of the password per execution, SPEKE is also patented by Phoenix Technologies, while J-PAKE proudly presents its freedom from patents.

[6] Accessible at https://www.lightbluetouchpaper.org/2008/05/29/j-pake/#comment-9550

future work." In 2015, the security of J-PAKE was proven in the model of Bellare, Pointcheval, and Rogaway by Abdalla et. al. [2] under the assumptions of the random-oracle model, the algebraic adversary model (AAM)[7] and the hardness of the Decision Square Diffie-Hellman problem (DSDH). DSHD is the problem of determining the group element g^{x^2} from a random element, given access to g^x. Hardness of DSDH implies the hardness of the standard decision Diffie-Hellman (DDH) problem, and it is currently unknown whether or not it is harder (i.e., whether there is a separation in the complexity classes).

2.3. **Dragonfly.** Dragonfly – another PAKE, created by Harkins [14] – is also based on discrete logarithm cryptography. This general setup means one can work either in a finite field or use elliptic curves, and like all our implementations for this project we describe and use the finite field version. Similar to J-PAKE (and unlike SPEKE), there are no assumptions on the order of the underlying group (i.e., its order does not have to be of the form $p = 2q - 1$, for q prime).

Let p be a large prime. We will let Q denote a cyclic subgroup of $\mathbb{Z}_p^*$ with prime order q – hence $q|(p-1)$. In addition to p and q, a hash function H is also agreed upon, and it assumed that Alice and Bob share the password $\pi \in Q$. The protocol specification maps the password arbitrarily (but deterministically) to the element π, and includes some example algorithms to perform the actual mapping. These examples are omitted here. Due to a (slightly) increased complexity over the protocols discussed so far, we begin with a text description:

(Round 1) Alice chooses two random values $r_A, m_A \in_R \mathbb{Z}_q^*$ and computes $s_A = r_A + m_A \mod q$ along with the element $E_A = \pi^{-m_A} \mod p$. If $s_A < 2$ (to avoid a small subgroup attack), she repeats this step. She sends s_A and E_A to Bob.

Bob chooses two random values $r_B, m_B \in_R \mathbb{Z}_q^*$. He computes $s_B = r_B + m_B \mod q$ and the element $E_B = \pi^{-m_B} \mod p$. If $s_B < 2$, he repeats this step over. He sends s_A and E_A to Alice.

Each member verifies that one of $E_A \neq E_B$ or $s_A \neq s_B$ is true to avoid a reflection attack. [8]

(Round 2) Alice computes the shared secret $ss = (\pi^{s_B} E_B)^{r_A} = \pi^{r_A r_B} \mod p$.
Alice sends $A = H(ss||E_A||s_A||E_B||s_B)$ to Bob.

Bob computes the shared secret $ss = (\pi^{s_A} E_A)^{r_B} = \pi^{r_A r_B} \mod p$.
Bob sends $B = H(ss||E_B||s_B||E_A||s_A)$ to Alice.

Alice and Bob both confirm the received hash values are correct and compute the shared key

$$K = H(ss||E_A \times E_B||(s_A + s_B) \mod q).$$

This protocol is illustrated in Figure 4. We note that despite its status as an IEEE 802.11-2012 standard no security proof of Dragonfly has yet been derived (the document of Harkins [14] states that it possesses security features such as offline dictionary attack resistance, but does not prove anything). In fact, Round 1 of the protocol was modified in the most recent update to the Dragonfly protocol after a small subgroup attack was discovered by Clarke and Hao [9]; the protocol now works around this by checking s_A and s_B in Round 1, and repeating the step until the values generated are safe. The chief benefit of Dragonfly is its speed: as the results of Section 4 show, Dragonfly+ is the fastest group protocol we tested.

[7] The AAM, originated by Dolev and Yao [10] states that an adversary can only perform operations in the underlying group of the protocol, on known messages (for instance, the attacker cannot modify the bits of messages or guess keys)

[8] Otherwise an attacker could be accepted as the variables A and B would be equal in Round 2. Note, however, that the attacker would still be unable to compute the key without knowledge of the password.

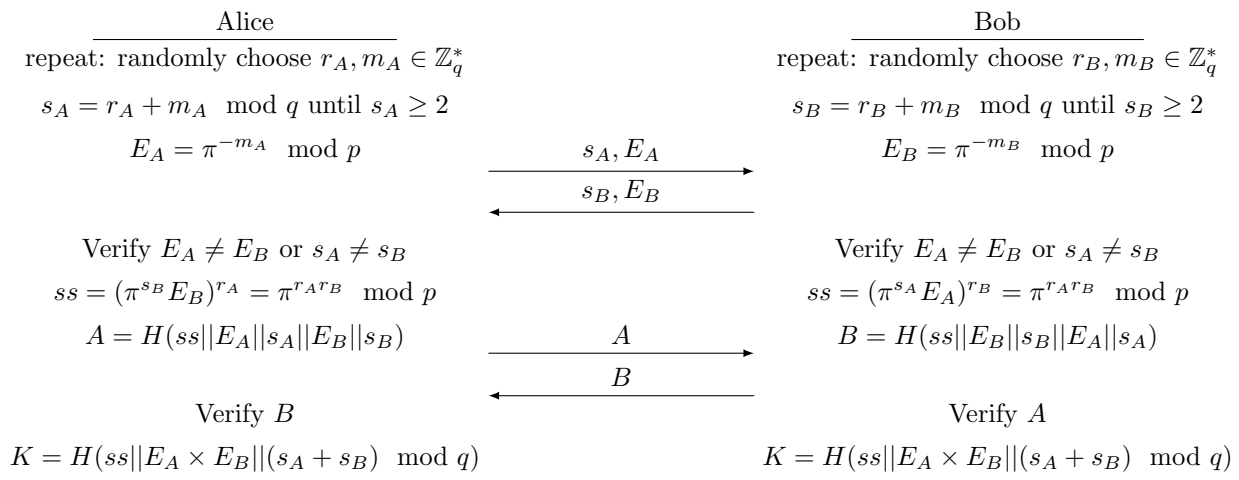


FIGURE 4. Flow diagram of the Dragonfly protocol.

2.4. **PAK/PPK.** The PAK/PPK protocols were first introduced by Boyko, MacKenzie and Patel in [6] in 2000 as Diffie-Hellman based provably secure PAKEs, with the PAK protocol admitting key confirmation and PPK admitting only key authentication (although PPK uses one less round of communication). An augmented variant of PAK, called PAK-X, was also introduced in the same paper.

Let π be the password shared by Alice and Bob and p and q be primes with $p = rq + 1$, where q does not divide r. Furthermore, let g be a generator of a subgroup of $\mathbb{Z}_p^*$ of size q where the Decision Diffie-Hellman (DDH) problem is infeasible. Finally, we take H_1, H_{2a}, H_{2b}, H_3 to be independent random hash functions. The PAK and PPK protocols are described as in Figures 5 and 6.

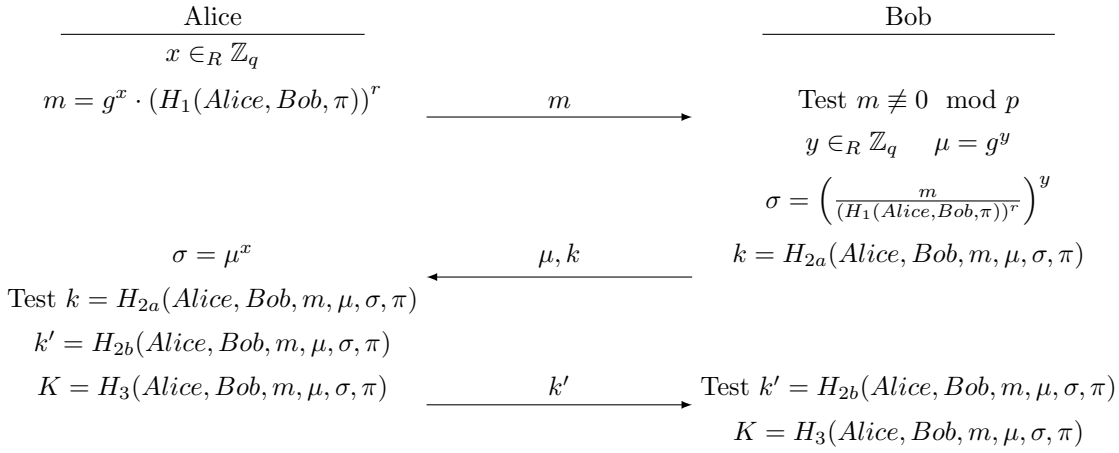


FIGURE 5. The PAK protocol.

In their original paper, Boyko, MacKenzie and Patel [6] developed a new formal model for PAKE security, in which they proved that PAK and PPK are secure under the assumptions of the random-oracle model and the hardness of the DDH is intractable. This newly proposed model was well designed, and security proofs of other PAKE protocols have been tailored to it (for instance, the security proof of SPEKE given by MacKenzie[17] and mentioned in Section 2.1).

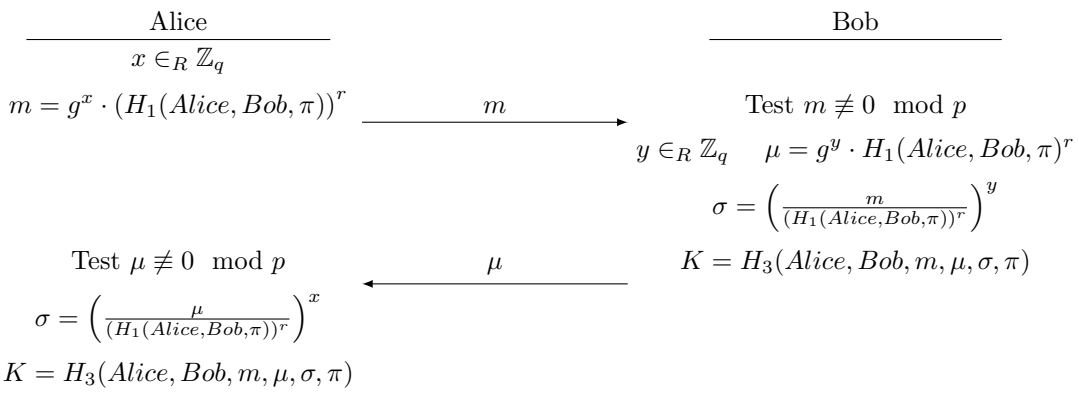


FIGURE 6. The PPK protocol.

3. Group Password-Authenticated Key Exchange (GPAKE)

We now turn our attention to the group setting, where n agents all share knowledge of a common password π and wish to establish a shared secure group key K (to be explicit, knowledge of the shared password is the only thing that makes an agent an authentic group member). Although there has been some previous work on this topic, much of it – see, for instance, Dutta and Barua [11] – required $O(n)$ rounds of communication to establish the shared group key, which is an issue as the protocol can easily be disrupted by one slow participant (for example, setting up a group key with 10 participants would require around 10 rounds of communication, meaning the group would have to wait for the slowest responder 10 times)[9]. In 2006, Abdalla et. al. [3] gave a GPAKE using only 4 rounds of communication (independent of n), a significant improvement. Furthermore, in 2015 Hao et. al. [13] outlined a general method – which they called the 'fairy-ring dance', due to a similarity between its structure of establishing pairwise keys and a traditional Scottish country dance – of taking a PAKE using finite fields and transforming it into a GPAKE which has at most one additional round of communication (so the number of rounds of communication will again be independent of the number of members in the group)[10]. The paper of Hao et. al. continued on to give two explicit GPAKEs using this construction based on the SPEKE and J-PAKE protocols detailed in the previous section. We describe their algorithms, together with two new GPAKEs we have created using this general construction, based on the Dragonfly and PPK protocols described above.

Before describing these GPAKEs, we give a general overview of how the generic construction works (using a bit more detail than in the original paper of Hao et. al. [13], which puts a larger focus on the explicit methods SPEKE+ and J-PAKE+). Essentially, the construction proceeds as follows:

- Each pair of members P_i, P_j in the group will compute a two-party pairwise key K_{ij} using some two-party PAKE protocol. All information is broadcast simultaneously by each group member
- At the same time, each participant P_i chooses another random element y_i of the finite field in which the PAKE operations are performed, and broadcasts g^{y_i} together with a Zero Knowledge Proof (for instance, the Schnorr ZKP detailed in the previous section) ZKP$\{y_i\}$ that P_i knows the exponent y_i. Each participant P_i then computes $g^{z_i} = g^{y_{i+1}}/g^{y_{i-1}}$, where indices are taken modulo n.

[9]To be fair, the construction of Hao et. al. has security properties which are proven in an informal attack model while the scheme of Dutta and Barua is proven secure in the widely accepted BPR model of Bellare et. al. [4] under the CDH assumption in the random oracle model. Also, the greatly decreased number of rounds comes with more computational work for each group member.

[10]The resulting GPAKE will always have at least two rounds of communication.

- Each P_i broadcasts $(g^{z_i})^{y_i}$ and a Zero Knowledge Proof ZKP$\{\tilde{y_i}\}$ that the discrete logarithm of $(g^{z_i})^{y_i}$ with respect to the base g^{z_i} is equal to the discrete logarithm of g^{y_i} with respect to the base g. For this we use the well known Chaum-Pedersen ZKP. Essentially, for Alice to prove to Bob that $\log_a(e_1) = \log_b(e_2)$ is some common value $x \in \mathbb{Z}_q$ – when Bob knows a, b, e_1, e_2 – Alice picks some random $s \in_R \mathbb{Z}_q$ and sends Bob a^s and b^s. Bob then sends Alice a challenge $c \in \mathbb{Z}_q^*$, to which Alice must respond with $t = s - cx \bmod q$. Bob will accept if $a^s = a^t e_1^c$ and $b^s = b^t e_2^c$. See the paper of Chaum and Pedersen [8] for details and information about the scheme's security.

- Each member P_i computes, for all $j \neq i$,
$$\kappa_{ij}^{MAC} = H(K_{ij}||\text{`MAC'}) \qquad \kappa_{ij}^{KC} = H(K_{ij}||\text{`KC'})$$
where H is a suitable hash function, and broadcasts
$$t_{ij}^{MAC} = HMAC(\kappa_{ij}^{MAC}, g^{y_i}||\text{ZKP}\{y_i\}||(g^{z_i})^{y_i}||\text{ZKP}\{\tilde{y_i}\})$$
$$t_{ij}^{KC} = HMAC(\kappa_{ij}^{KC}, \text{“}KC''||i||j||A_{ij}),$$
where HMAC is a suitable keyed-hash message authentication code – that is, a MAC which is based on a hash function – and A_{ij} is a concatentation of all information required to compute the shared key for the pairwise PAKE between P_i and P_j (this is elaborated on in the four examples below).

- If agent P_i is able to verify all Zero Knowledge Proofs sent by all other group members, and can verify the message authentication codes t_{ji}^{MAC} and t_{ji}^{KC} for all $j \neq i$, then P_i accepts and calculates the shared group key

(1) $$K = (g^{y_{i-1}})^{n\cdot y_i}(g^{z_i y_i})^{n-1}(g^{z_{i+1}y_{i+1}})^{n-2}\cdots(g^{z_{i-2}y_{i-1}})^{n-1} = g^{y_1\cdot y_2 + y_2\cdot y_3 + \cdots + y_n\cdot y_1},$$

where again the indices in the above equation are taken modulo n. The steps used to compute the key (namely, the construction of the group key using the elements g^{y_i} and g^{z_i}) are based on the Burmester-Desmedt cyclic key computation technique, which was used by Burmester and Desmedt [7] to construct a shared secure key amongst a group of agents using Public Key Infrastructure. In their article, Burmester and Desmedt show the security of the scheme under a realistic attack model assuming the intractability of the DDH.

In their paper, Hao et. al. [13] give arguments showing the security of the resulting GPAKE, in an informal attack model where α of the n group participants are legitimate group members and $n-\alpha$ members are active attackers which can send/receive/modify messages as usual (there may also be passive attackers watching the exchange take place). Their explanation shows that – under assumptions such as the hardness of the DDH, the security of the Schnorr and Chaum-Pedersen ZKPs, and the security of the Burmester-Desmedt group key agreement protocol – when the underlying PAKE has each of the four security properties discussed in Section 2 (offline dictionary attack resistance, forward secrecy for established keys, known session security, and online dictionary attack resistance) then so will the resulting GPAKE, where online dictionary attack resistance now means that at most $\alpha \cdot (n - \alpha)$ passwords may be guessed in one run of the protocol (each active attacker can try one guess of the password with each authentic password holder). Their argument for security roughly takes the following form: the shared key is constructed from the Burmester-Desmedt group key agreement protocol, which is secure when Public Key Infrastructure is available. In order to verify participants as authentic password holders without PKI – to prevent man in the middle attacks, for instance – a pairwise PAKE is established between each pair of members. Assuming the underlying PAKE is secure, this will correctly identify attackers. For the full argument, we refer the reader to Hao et. al. [13].

We note also that the verification of t_{ij}^{MAC} by each group member results in key confirmation for the GPAKE scheme, in the sense that any group member will know that all other group members have all information needed to calculate the shared key (verification of t_{ij}^{KC} results in key confirmation in each pairwise scheme). We now give four explicit schemes derived by this method, beginning with the two GPAKEs of Hao et. al.

3.1. **SPEKE+ and J-PAKE+.** The construction by Hao et. al. of the explicit GPAKE SPEKE+ follows the general outline above very closely. Let $p = 2q - 1$ where p and q are prime, and let g be a fixed generator of the subgroup $\mathbb{Z}_q$ of $\mathbb{Z}_p^*$ with order q. To the password π we associate a group element $g_\pi = H(\pi)^2 \bmod p$, where H is a hash function. SPEKE+ is run as follows:

(Round 1) Every participant P_i selects $x_i \in_R \mathbb{Z}_q^*$ and $y_i \in_R \mathbb{Z}_q$ and broadcasts, for all $i \neq j$, the values

$$g_\pi^{x_i}, \qquad g^{y_i}, \qquad \mathrm{ZKP}\{y_i\}.$$

Define $z_i = y_{i+1}/y_{i-1}$ (with cyclic index i). Each P_i is able to compute $g^{z_i} = g^{y_{i+1}}/g^{y_{i-1}}$ and checks that:

- $g^{z_i} \neq 1 \mod p$ for $i = 1, \ldots, n$;
- $g_\pi^{x_j} \notin \{1, p-1\}$ for $j \neq i$;
- the $\mathrm{ZKP}\{y_j\}$ for all $j \neq i$ are valid.

(Round 2) Every participant P_i broadcasts $(g^{z_i})^{y_i}$ and a zero knowledge proof $\mathrm{ZKP}\{\tilde{y_i}\}$ proving the equality of the discrete logarithm of $(g^{z_i})^{y_i}$ to the base g^{z_i} and the discrete logarithm g^{y_i} to the base g. Each member computes the pairwise SPEKE keys

$$K_{ij} = g_\pi^{x_i x_j},$$

and the authentication and confirmation keys

$$\kappa^{\mathrm{MAC}} = H(K_{ij}, \text{“MAC”}) \qquad \kappa^{\mathrm{KC}} = H(K_{ij}, \text{“KC”}).$$

Each member additionally broadcasts

$$t_{ij}^{MAC} = HMAC(\kappa_{ij}^{MAC}, g^{y_i} || \mathrm{ZKP}\{y_i\} || (g^{z_i})^{y_i} || \mathrm{ZKP}\{\tilde{y_i}\})$$
$$t_{ij}^{KC} = HMAC(\kappa_{ij}^{KC}, \text{“}KC''||i||j||g_\pi^{x_i}||g_\pi^{x_j}).$$

Finally, all members confirm:

- the received $\mathrm{ZKP}\{\tilde{y_j}\}$ for $j \neq i$ are valid;
- the received key confirmation strings t_{ji}^{KC} for $j \neq i$ are valid;
- the received message authentication tags t_{ji}^{MAC} for $j \neq i$ are valid.

and establish the group key via Equation (1), according to the Burmester-Desmedt group key agreement protocol.

The construction of J-PAKE+ is similar. Let $p = rq - 1$ where p and q are prime, and let g be a fixed generator of the subgroup $\mathbb{Z}_q$ of $\mathbb{Z}_p^*$ with order q. Then the J-PAKE+ protocol consists of the following steps:

(Round 1) Every participant P_i selects $a_{ij} \in_R \mathbb{Z}_q$ and $b_{ij} \in_R \mathbb{Z}_q^*$ for all $j \neq i$.Additionally, each member selects $y_i \in_R \mathbb{Z}_q$ and broadcasts the values

$$g^{a_{ij}}, \qquad g^{b_{ij}}, \qquad g^{y_i} \bmod p, \qquad \mathrm{ZKP}\{a_{ij}\}, \qquad \mathrm{ZKP}\{b_{ij}\}, \qquad \mathrm{ZKP}\{y_i\}.$$

Define $z_i = y_{i+1}/y_{i-1}$ (with cyclic index i). Each member is able to compute $g^{z_i} = g^{y_{i+1}}/g^{y_{i-1}}$, and check:

- $g^{z_i} \neq 1 \mod p$ for $i = 1, \ldots, n$;
- $g^{b_{ji}} \neq 1$ for $j \neq i$;
- the $\mathrm{ZKP}\{a_{ji}\}$, $\mathrm{ZKP}\{b_{ji}\}$, and $\mathrm{ZKP}\{y_j\}$, are valid for all $j \neq i$.

(Round 2) Every participant P_i can compute and broadcast, for each $j \neq i$,

$$\beta_{ij} = (g^{a_{ij} + a_{ji} + b_{ji}})^{b_{ij} \cdot s}, \qquad \mathrm{ZKP}\{b_{ji} \cdots s\}.$$

Each participant P_i verifies ZKP$\{b_{ji} \cdots s\}$ for all $j \neq i$.

(Round 3) Every participant P_i broadcasts $(g^{z_i})^{y_i}$ and a zero knowledge proof ZKP$\{\tilde{y_i}\}$ proving the equality of the discrete logarithm of $(g^{z_i})^{y_i}$ to the base g^{z_i} and the discrete logarithm g^{y_i} to the base g. Each member computes the pairwise J-PAKE keys

$$K_{ij} = (\beta_{ji}/g^{b_{ij}b_{ji}s})^{b_{ij}}$$

and the authentication and confirmation keys

$$\kappa^{\text{MAC}} = H(K_{ij}, \text{“MAC”}) \qquad \kappa^{\text{KC}} = H(K_{ij}, \text{“KC”}).$$

Each member additionally broadcasts

$$t_{ij}^{MAC} = HMAC(\kappa_{ij}^{MAC}, g^{y_i}||\text{ZKP}\{y_i\}||(g^{z_i})^{y_i}||\text{ZKP}\{\tilde{y_i}\})$$
$$t_{ij}^{KC} = HMAC(\kappa_{ij}^{KC}, \text{“}KC''||i||j||E_{ij}||E_{ji}).$$

Finally, all members confirm:
- the received ZKP$\{\tilde{y_j}\}$ for $j \neq i$ are valid;
- the received key confirmation strings t_{ji}^{KC} for $j \neq i$ are valid;
- the received message authentication tags t_{ji}^{MAC} for $j \neq i$ are valid.

and establish the group key via Equation (1), according to the Burmester-Desmedt group key agreement protocol.

3.2. **Dragonfly+.** We now present our group extension of the Dragonfly protocol using the general construction. The resulting GPAKE follows the Dragonfly protocol closely, with only minor modifications to the first round of communication and the addition of a final round to establish the group key. The setup is the same as Dragonfly (see Section 2.3), except an explicit generator g of the subgroup Q is required. The Dragonfly+ protocol is executed as follows:

(Round 1) Every participant P_i selects $r_{ij}, m_{ij} \in_R \mathbb{Z}_q^*$ for all $j \neq i$ and computes $s_{ij} = r_{ij} + m_{ij} \mod q$ along with the element

$$E_{ij} = \pi^{-m_{ij}} \mod p.$$

If any $s_{ij} < 2$, r_{ij} and m_{ij} must be re-established. Additionally, each member selects $y_i \in_R \mathbb{Z}_q$ and broadcasts the values

$$s_{ij}, \qquad E_{ij}, \qquad g^{y_i} \bmod p, \qquad \text{ZKP}\{y_i\}.$$

Define $z_i = y_{i+1}/y_{i-1}$ (with cyclic index i). Each member is able to compute $g^{z_i} = g^{y_{i+1}}/g^{y_{i-1}}$, and check:
- $g^{z_i} \neq 1 \mod p$ for $i = 1, \ldots, n$;
- at least one of $E_{ij} \neq E_{ji}$ or $s_{ij} \neq s_{ji}$ is true for all $j \neq i$;
- the received ZKP$\{y_j\}$ for all $j \neq i$ is valid.

(Round 2) Every participant P_i can compute the pairwise shared secrets $ss_{ij} = (\pi^{s_{ji}}E_{ji})^{r_{ij}} = \pi^{r_{ij}r_{ji}} \mod p$. They broadcast $A_{ij} = H(ss_{ij}||E_{ij}||s_{ij}||E_{ji}||s_{ji})$.

Each participant confirms the hashes are correct.

(Round 3) Every participant P_i broadcasts $(g^{z_i})^{y_i}$ and a zero knowledge proof ZKP$\{\tilde{y_i}\}$ proving the equality of the discrete logarithm of $(g^{z_i})^{y_i}$ to the base g^{z_i} and the discrete logarithm g^{y_i} to the base g. Each member computes the pairwise Dragonfly keys

$$K_{ij} = H(ss_{ij}||E_{ij} \times E_{ji}||(s_{ij} + s_{ji}) \mod q)$$

and the authentication and confirmation keys

$$\kappa^{\text{MAC}} = H(K_{ij}, \text{“MAC”}) \qquad \kappa^{\text{KC}} = H(K_{ij}, \text{“KC”}).$$

Each member additionally broadcasts

$$t_{ij}^{MAC} = HMAC(\kappa_{ij}^{MAC}, g^{y_i}||\text{ZKP}\{y_i\}||(g^{z_i})^{y_i}||\text{ZKP}\{\tilde{y_i}\})$$
$$t_{ij}^{KC} = HMAC(\kappa_{ij}^{KC}, \text{“}KC''||i||j||E_{ij}||E_{ji}).$$

Finally, all members confirm:
- the received ZKP$\{\tilde{y_j}\}$ for $j \neq i$ are valid;
- the received key confirmation strings t_{ji}^{KC} for $j \neq i$ are valid;
- the received message authentication tags t_{ji}^{MAC} for $j \neq i$ are valid.

and establish the group key via Equation (1), according to the Burmester-Desmedt group key agreement protocol.

Note that for simplicity in our implementation of Dragonfly+ we mapped the password directly to g, the generator of Q. This specific strategy is *not* secure since the group element is not password dependent – however, our implementation is simply used for comparing latency in the different GPAKEs, which will be relatively unaffected by the choice of generator the password is mapped to. In a secure implementation the password should be mapped in a manner similar to Harkins [14].

3.3. **PPK+.** In this section we present the Group PAKE extension of the PPK protocol, the simpler of the PAK/PPK suite. The PPK+ protocol, which has the same setup as the PPK protocol of Section 2.4, is described as follows:

(Round 1) Every participant P_i selects $x_i \in_R \mathbb{Z}_q$ and $y_i \in_R \mathbb{Z}_q$ and broadcasts, for all $i \neq j$, the values

$$m_{ij} = g^{x_i} \cdot (H_1(i, j, \pi))^r, \qquad g^{y_i}, \qquad \text{ZKP}\{y_i\}.$$

Define $z_i = y_{i+1}/y_{i-1}$ (with cyclic index i). Each P_i is able to compute $g^{z_i} = g^{y_{i+1}}/g^{y_{i-1}}$ and

$$\sigma_{ij} = \left(\frac{m_{ji}}{H_1(j, i, \pi)^r}\right)^{x_i} = g^{x_i x_j},$$

and check:
- $g^{z_i} \neq 1 \mod p$ for $i = 1, \ldots, n$;
- $m_{ij} \neq 0$ for $j \neq i$;
- the ZKP$\{y_j\}$ for all $j \neq i$ are valid.

(Round 2) Every participant P_i broadcasts $(g^{z_i})^{y_i}$ and a zero knowledge proof ZKP$\{\tilde{y_i}\}$ proving the equality of the discrete logarithm of $(g^{z_i})^{y_i}$ to the base g^{z_i} and the discrete logarithm g^{y_i} to the base g. Each member computes the pairwise PPK keys

$$K_{ij} = H_3(i, j, m_{ij}, m_{ji}, \sigma_{ij}, \pi),$$

and the authentication and confirmation keys

$$\kappa^{\text{MAC}} = H(K_{ij}, \text{“MAC”}) \qquad \kappa^{\text{KC}} = H(K_{ij}, \text{“KC”}).$$

Each member additionally broadcasts

$$t_{ij}^{MAC} = HMAC(\kappa_{ij}^{MAC}, g^{y_i}||\text{ZKP}\{y_i\}||(g^{z_i})^{y_i}||\text{ZKP}\{\tilde{y_i}\})$$
$$t_{ij}^{KC} = HMAC(\kappa_{ij}^{KC}, \text{“}KC''||i||j||E_{ij}||E_{ji}).$$

Finally, all members confirm:

- the received $\text{ZKP}\{\tilde{y_j}\}$ for $j \neq i$ are valid;
- the received key confirmation strings t_{ji}^{KC} for $j \neq i$ are valid;
- the received message authentication tags t_{ji}^{MAC} for $j \neq i$ are valid.

and establish the group key via Equation (1), according to the Burmester-Desmedt group key agreement protocol.

We now go over a few details of our implementation of the protocol. First, the supposedly independent random hash functions H_1 and H_3 are both the *SHA-256* hash function shifted by two different constants. As with the modification for Dragonfly this is not secure, however our implementation is simply used for comparing latency in the different GPAKEs – this would be fixed before using the code for any other purpose. Additionally, as the formula for calculating the raw pairwise keys K_{ij} at the end of Round 1 above is not symmetric between i and j, we calculate K_{ij} where $i < j$ (practically, the participants can be ordered in some arbitrary way, such as lexicographically by their unique identifiers or by the time they accepted to enter the group protocol), then both parties P_i and P_j still calculate the same pairwise key K_{ij}.

4. Implementation Results

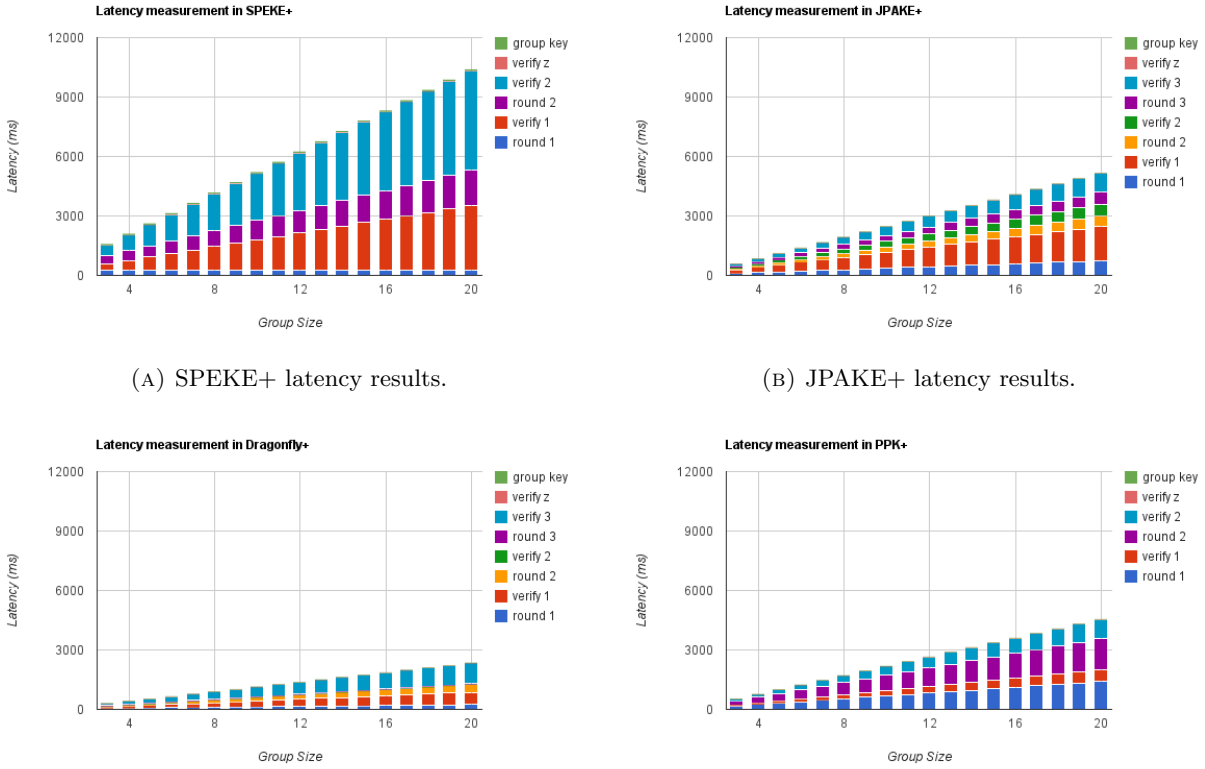
All of the protocols were implemented in Java 6 on a server (3GHz AMD processor, 6GB of RAM) running Ubuntu 12.04. These benchmarks measured latency, the amount of work each device would have to do in the group *excluding communication*. In each test, 2048-bit primes p are used; the values for SPEKE+ and J-PAKE+ were used as in Hao et. al. [13]. Values for PPK+ were taken from J-PAKE+, and values for Dragonfly+ were taken from the NIST cryptographic toolkit [1].

For groups sizes from 3 to 20 participants, we measure the latency of computation at each round of the protocol. The measurement was done by repeating the same experiment 100 times and taking the average values. The results are summarized in Figure 7, where it is observed that the total latency for each member increases linearly with the size of group, regardless of the protocol. Furthermore, we see that SPEKE+ is the slowest protocol, due in part to its need of a ‘safe’ prime of the form $q = 2p - 1$. This means that when p is a 2048-bit prime, q is a 2047-bit prime, which is much larger than what is needed for the other algorithms (typically, with primes in the hundreds of digits the assumptions needed for our PAKE security are considered to be safe). Although Dragonfly+ is considerably the fastest GPAKE here, it uses 3 rounds of communication – compared to only 2 for J-PAKE+ and PPK+ – and the underlying PAKE protocol does not have a security proof. The PPK+ protocol which we constructed, on the other hand, is built from a two party PAKE with a stronger security proof than J-PAKE and runs faster than J-PAKE+. Thus, this seems like a good candidate for future research.

We now conclude by summing up the work presented here and giving some directions for future work in this area.

5. Conclusion

This project has investigated password authenticated key exchange methods (PAKEs), both in classical two party settings and in the context of group key establishment. After surveying some literature on PAKEs, we outlined a general construction recently proposed by Hao et. al. [13] to convert a two party PAKE into



(A) SPEKE+ latency results.

(B) JPAKE+ latency results.

(C) Dragonfly+ latency results.

(D) PPK+ latency results.

FIGURE 7. Latency measurement results for the GPAKE extensions of each of the PAKE protocols.

a group PAKE (with key confirmation) which adds at most one additional round of communication. After detailing two explicit GPAKEs derived by Hao et. al. – SPEKE+ and J-PAKE+ – we constructed two new GPAKEs – Dragonfly+ and PPK+ – based on well known two party PAKE protocols. These four schemes were implemented and timings were shown comparing the latency of each method.

Although the work of Hao et. al. [13] includes informal proofs of desired security properties which are inherited from an underlying PAKE to the GPAKE which results from their general construction, the authors do not show security in a formal model (in the sense of Bellare, Pointcheval, and Rogaway [4] for two party PAKEs). Future work could (and should!) look into formal models for GPAKEs and provide a proof that the GPAKE resulting from any PAKE which is secure in a formal model (like the BPR model) is secure in a formal GPAKE attack model. This work will become more important as the number of Internet compatible devices continues to grow, increasing the need for group key establishment schemes.

REFERENCES

[1] NIST cryptographic toolkit, July 2014. http://csrc.nist.gov/groups/ST/toolkit/.

[2] M. Abdalla, F. Benhamouda, and P. MacKenzie. Security of the j-pake password-authenticated key exchange protocol. In *SP*, 2015.

[3] M. Abdalla, E. Bresson, O. Chevassut, and D. Pointcheval. Password-based group key exchange in a constant number of rounds. *PKC'06, LNCS 3958*, pages 427–442, 2006.

[4] M. Bellare, D. Pointcheval, and P. Rogaway. Authenticated key exchange secure against dictionary attacks. In *LNCS, B. Preneel, Ed., vol. 1807. Springer*, pages 139 – 155. EUROCRYPT 2000, May 2000.

[5] S. M. Bellovin and M. Merritt. Encrypted key exchange: Password-based protocols secure against dictionary attacks. In *1992 IEEE Computer Society Symposium on Research in Security and Privacy Proceedings*, pages 72–84. IEEE, 1992.

[6] V. Boyko, P. MacKenzie, and S. Patel. Provably secure password-authenticated key exchange using diffie-hellman. In *LNCS 1807,Springer-Verlag, Berlin*, pages 156–171. Eurocrypt 2000, 2000.
[7] M. Burmester and Y. Desmedt. A secure and efficient conference key distribution system. *EUROCRYPT'95, LNCS 950*, pages 275–286, 1995.
[8] D. Chaum and T.P. Pedersen. Wallet databases with observers. *Advances in Cryptology – Crypto '92*, pages 89–105, 1992.
[9] Dylan Clarke and Feng Hao. Cryptanalysis of the dragonfly key exchange protocol. In *IET Information Security*, pages 283–289, 2014.
[10] D. Dolev and A. C. Yao. On the security of public key protocols. *IEEE trans. on Information Theory*, pages 198–208, 1983.
[11] R. Dutta and R. Barua. Password-based encrypted group key agreement. *International Journal of Network Security*, 3, 2006.
[12] F. Hao and P. Ryan. J-PAKE: Authenticated key exchange without PKI. *Transactions on Computational Science XI, Lecture Notes in Computer Science*, 6480:192–206, 2010.
[13] F. Hao, X. Yi, L. Chen, and S. F. Shahandashti. The fairy-ring dance: Password authenticated key exchange in a group. Cryptology ePrint Archive, Report 2015/080, 2015. http://eprint.iacr.org/2015/080.
[14] D. Harkins. Dragonfly key exchange – internet research task force internet draft, 2015. http://datatracker.ietf.org/doc/draft-irtf-cfrg-dragonfly/.
[15] B. Jaspan. Dual-workfactor encrypted key exchange: efficiently preventing password chaining and dictionary attacks. In *Proceedings of the Sixth Annual USENIX Security Conference*, pages 43–50, July 1996.
[16] T. M. A. Lomas, L. Gong, J. H. Saltzer, and R. M. Needham. Reducing risks from poorly chosen keys. In *ACM Operating Systems Review, 23(5)*, pages 14–18. Proceedings of the 12th ACM Symposium on Operating System Principles, Dec 1989.
[17] P. MacKenzie. On the security of the speke password-authenticated key exchange protocol. Cryptology ePrint Archive, Report 2001/057, 2001. http://eprint.iacr.org/2001/057.
[18] Philip MacKenzie. The pak suite: Protocols for password-authenticated key exchange. In *IEEE P1363.2*, 2002.