The NoiseSocket Protocol

Alexey Ermishkin Trevor Perrin

Revision 2draft, 2018-05-01, official/unstable

Contents

Abstract

NoiseSocket provides an encoding layer for the Noise Protocol Framework.

NoiseSocket can encode Noise messages and associated negotiation data into a form suitable for transmission over reliable, stream-based protocols such as TCP.

1. Overview

The Noise Protocol Framework [1] describes **Noise protocols**. A Noise protocol sends a fixed sequence of handshake messages based on a fixed set of cryptographic choices. In some situations the responder needs flexibility to accept or reject the initiator's Noise protocol choice, or make its own choice based on options offered by the initiator.

The **NoiseSocket** framework allows **compound protocols** where the initiator and responder negotiate a particular Noise protocol. This is a two-step process:

- The initiator begins executing an initial Noise protocol and sends an initial Noise handshake message. As a preamble to this message, the initiator can send some **negotiation data** which indicates the initial Noise protocol and can advertise support for other Noise protocols.
- The responder can **accept** the initiator's choice of initial Noise protocol; **switch** to a different Noise protocol; **request retry** with a different Noise protocol; or **reject** the handshake entirely. The responder indicates this choice by sending some negotiation data back to the initiator, or closing the connection.

NoiseSocket doesn't specify the contents of negotiation data, since different applications will encode and advertise protocol support in different ways. NoiseSocket just defines a message format to transport this data.

NoiseSocket handles two other low-level issues:

- NoiseSocket defines length fields for all messages, so NoiseSocket messages can be used with stream-based protocols like TCP.
- NoiseSocket defines padding fields which are included in ciphertexts so that applications can pad their messages to avoid revealing plaintext lengths to an eavesdropper.

2. Message Formats

A NoiseSocket protocol begins with a **handshake phase**. During the handshake phase each NoiseSocket message contains a single **handshake message** from some underlying Noise protocol, plus optional negotiation data.

After the handshake completes, NoiseSocket enters the **transport phase** where each NoiseSocket message contains a **transport message** from some underlying Noise protocol.

All transport messages and some handshake messages contain an encrypted Noise **payload**. Each encrypted payload contains a plaintext with a **body** (its actual contents) followed by **padding**.

The following sections describe the format for NoiseSocket handshake and transport messages, and encrypted payloads.

2.1. Handshake messages

All NoiseSocket handshake messages have the same structure:

- `negotiation_data_len` (2 bytes)
- `negotiation_data`
- `noise_message_len` (2 bytes)
- `noise_message`

The `negotiation_data_len` and `noise_message_len` fields are 2-byte unsigned integers, encoded in big-endian, that store the number of bytes for the following `negotiation_data` and `noise_message` fields.

2.2. Transport messages

All NoiseSocket transport messages have the same structure:

- `noise_message_len` (2 bytes)
- `noise_message`

The `noise_message_len` field is a 2-byte unsigned integer, encoded in big-endian, that stores the number of bytes for the following `noise_message` field.

2.3. Encrypted payloads

Each Noise transport message consists of a single encrypted payload. Each Noise handshake message might contain a a single encrypted payload (or might contain a cleartext payload). When these payloads are decrypted, the plaintext will have the following structure:

- `body_len` (2 bytes)
- `body`
- `padding`

The `body_len` field is a 2-byte unsigned integer, encoded in big-endian, that stores the number of bytes for the following `body` field. Following the `body` field the remainder of the decrypted plaintext will be padding bytes, which may contain arbitrary data and must be ignored by the recipient.

3. Negotiation

The initiator will choose the initial Noise protocol, and will indicate this to the responder using the `negotiation_data` field. Upon receiving an initial NoiseSocket message, the responder has five options:

- **Accept**: The responder accepts the initial Noise protocol. If this is an interactive protocol, the responder sends a NoiseSocket handshake message containing the next handshake message in the initial Noise protocol. The `negotiation_data` field of this response message must be empty.
- **Switch**: The responder sends a NoiseSocket handshake message containing a handshake message from a new Noise protocol (e.g. a fallback protocol), different from the initial Noise protocol. The `negotiation_data` field must be non-empty. The `noise_message` field must be non-empty.
- **Request Retry**: The responder requests the initiator to send a NoiseSocket handshake message containing a handshake message from a new Noise protocol, different from the initial Noise protocol. The `negotiation_data` field must be non-empty. The `noise_message` field must be empty.
- **Explicit Reject**: The responder sends a single NoiseSocket handshake message. The `negotiation_data` field must be non-empty and contain an error message. The `noise_message` field must be empty. After sending this message, the responder closes the network connection.
- **Silent Reject**: The responder closes the network connection.

The following table indicates the cases where the response `negotiation_data` and `noise_message` fields are non-empty.

	Negotiation	Message
Accept	-	Yes
Switch	Yes	Yes
Request Retry	Yes	-
Explicit Reject	Yes	-

The initiator's first `negotiation_data` field must indicate the initial Noise protocol and what other Noise protocols the initiator can support for the switch and retry cases. How this is encoded is up to the application.

If the responder's first `negotiation_data` field is empty, then the initial protocol was accepted. If the field is non-empty, it must encode values that distinguish betwen the "switch", "retry", and "reject" cases. In the first two cases, the `negotiation_data` must encode the Noise protocol the initiator should switch to or retry with. In the last case, the `negotiation_data` must encode an error message.

When the initiator receives the first NoiseSocket response message, and for all later handshake messages received by both parties, the only options are silent rejection, explicit rejection, or acceptance.

Example negotiation flows:

- It's easy for the responder to change symmetric crypto options by switching to a different protocol. For example, if the initial Noise protocol is `Noise_XX_25519_AESGCM_SHA256`, the responder can switch to `Noise_XX+fallback_25519_ChaChaPoly_BLAKE2s`. This reuses the ephemeral public key from the initiator's initial message.
- If the initiator attempts 0-RTT encryption that the responder fails to decrypt, the responder can also switch to a fallback protocol. For example, if the initial Noise protocol is `Noise_IK_25519_AESGCM_SHA256`, the responder can fallback to `Noise_XX+fallback_25519_AESGCM_SHA256`. This reuses the ephemeral public key from the initiator's initial message.
- If the responder wants to use a DH function that the initiator supports but did not send an ephemeral public key for, in the initial message, then the responder might need to request a retry. For example, if the initial Noise protocol is `Noise_XX_25519_AESGCM_SHA256`, the responder can request retry with `Noise_XX_448_AESGCM_SHA256`, causing the initiator to respond with a NoiseSocket message containing the initial message from the `Noise_XX` pattern with a Curve448 ephemeral public key.

4. Prologue

Noise protocols take a **prologue** input. The prologue is cryptographically authenticated to make sure both parties have the same view of it.

The prologue for the initial Noise protocol is set to the UTF-8 string "NoiseSocketInit1" followed by all bytes transmitted in the NoiseSocket protocol prior to the `noise_message_len`. This consists of the following values concatenated together:

- The UTF-8 string "NoiseSocketInit1"
- The initial message's `negotiation_data_len`
- The initial message's `negotiation_data`

If the responder switches the Noise protocol, the prologue is set to the UTF-8 string "NoiseSocketInit2" followed by all bytes received and transmitted in the NoiseSocket protocol prior to the `noise_message_len` in the response message. This consists of the following values concatenated together:

- The UTF-8 string "NoiseSocketInit2"
- The initial message's `negotiation_data_len`
- The initial message's `negotiation_data`
- The initial message's `noise_message_len`
- The initial message's `noise_message`
- The responding message's `negotiation_data_len`
- The responding message's `negotiation_data`

If the responder requests retry with a different Noise protocol, the prologue is set to the UTF-8 string "NoiseSocketInit3" followed by all bytes received and transmitted in the NoiseSocket protocol prior to the `noise_message_len` in the retry message. This consists of the following values concatenated together:

- The UTF-8 string "NoiseSocketInit3"
- The initial message's `negotiation_data_len`
- The initial message's `negotiation_data`
- The initial message's `noise_message_len`
- The initial message's `noise_message`
- The responding message's `negotiation_data_len`
- The responding message's `negotiation_data`
- The responding message's `noise_message_len` (i.e. two bytes of zeros)
- The retry message's `negotiation_data_len`
- The retry message's `negotiation_data`

Finally, the application using NoiseSocket may append an arbitrary **application prologue** byte sequence following the above data.

5. IPR

The NoiseSocket specification (this document) is hereby placed in the public domain.

6. Acknowledgements

Thanks to Gerardo di Giacomo, Nemanja Mijailovic, Rhys Weatherley, Christopher Wood, and Justin Cormack for helpful discussion.

7. References

[1] T. Perrin, "The Noise Protocol Framework." 2017. https://noiseprotocol.org