AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

WYDZIAŁ INFORMATYKI, ELEKTRONIKI I TELEKOMUNIKACJI

KATEDRA TELEKOMUNIKACJI

Praca dyplomowa inżynierska

Opracowanie biblioteki programistycznej do bezpiecznego uwierzytelniania urządzeń AVR.

Development of libraries for authentication of AVR devices.

Autor: *Kacper Żuk*
Kierunek studiów: *Teleinformatyka*
Opiekun pracy: *dr inż. Jarosław Bułat*

Kraków, 2016

Uprzedzony o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystycznego wykonania albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.", a także uprzedzony o odpowiedzialności dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t.j. Dz. U. z 2012 r. poz. 572, z późn. zm.): „Za naruszenie przepisów obowiązujących w uczelni oraz za czyny uchybiające godności studenta student ponosi odpowiedzialność dyscyplinarną przed komisją dyscyplinarną albo przed sądem koleżeńskim samorządu studenckiego, zwanym dalej «sądem koleżeńskim».", oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i że nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.

Spis treści

Spis skrótów

AES	Advanced Encryption Standard.
API	Application Programming Interface.
CBC	Cipher Block Chaining.
CBC-MAC	Cipher Block Chaining Message Authentication Code.
DDoS	Distributed Denial of Service.
ECBC-MAC	Encrypt-last-block Cipher Block Chaining Message Authentication Code.
ECC	Elliptic Curve Cryptography.
ECDH	Elliptic Curve Diffie-Hellman.
IoT	Internet of Things.
NIST	National Institute of Standards and Technology.
PKCS	Public Key Cryptography Standards.
RFC	Request For Comments.
SHA	Secure Hash Algorithm.
SRAM	Static Random Access Memory.

1. Wstęp

Rys. 1.1. Relatywna liczba wyszukiwań frazy „Arduino” w ostatnich pięciu latach.
Źródło: Google Trends

AVR to rodzina mikroprocesorów opracowana i rozwijana przez firmę Atmel. Oparta o nią jest m. in. platforma Arduino, która – jak przedstawiono na Rys. 1.1 – z roku na rok zyskuje popularność. Platforma Arduino zaprojektowana została z myślą o osobach, które niekoniecznie posiadają formalne wykształcenie inżynierskie [1]. Jest ona też często używana do prototypowania urządzeń, wpisujących się w koncepcję *Internetu Rzeczy (ang. Internet of Things (IoT))*.

Urządzenia wbudowane podłączone do Internetu są szczególnie narażone na ataki. W 2016 roku podatne urządzenia wbudowane zostały wykorzystane do przeprowadzenia masywnych ataków typu Distributed Denial of Service (DDoS) [2].

1.1. Cele pracy

Istotne jest więc dostarczenie narzędzi, które pozwalają nie tylko na szybkie prototypowanie, ale które pozwolą także zachować odpowiedni poziom bezpieczeństwa. Należy pamiętać przede wszystkim o tym, że urządzenia *IoT* są tworzone także przez ludzi bez formalnego wykształcenia inżynierskiego.

W niniejszej pracy przedstawiono protokół bezpiecznej komunikacji oraz bibliotekę programistyczną na urządzenia AVR zaprojektowane z myślą o prostocie obsługi. Wybrane zostały zestawy algorytmów, które zapewniają niezbędny poziom bezpieczeństwa. Ich złożoność została ukryta za prostym interfejsem programistycznym *(ang. IoT)*, który nie pozwala na wprowadzenie błędów zmniejszających bezpieczeństwo. Zaproponowane rozwiązanie zapewnia poufność, autentyczność oraz integralność przesyłanych danych.

1.2. Zawartość pracy

W rozdziale 2 scharakteryzowana jest platforma sprzętowa AVR, ze szczególnym uwzględnieniem jej ograniczeń. Następnie w rozdziale 3 przedstawione zostały różne metody uwierzytelniania i uzasadniony został wybór konkretnych rozwiązań. Implementacja została szczegółowo opisana w rozdziale 4. Całość rozwiązania została zwalidowana poprzez porównanie z implementacją na inną platformę, co opisano w rozdziale 5. W rozdziale 6 podsumowano całe rozwiązanie oraz przedstawiono jego ograniczenia i słabe strony.

Całość kodu źródłowego dostępna jest w serwisie GitHub[1].

[1] *https://github.com/kacperzuk/seconn*

2. Charakterystyka platformy sprzętowej

Mikropocesory Atmel AVR są w większości 8-bitowe i na takich skupia się praca. Rodzina AVR jest szeroka, od ATtiny4 z 32 bajtami Static Random Access Memory (SRAM) [3] do ATxmega384C3 z 32 kilobajtami SRAM [4]. W pracy wykorzystywany był model ATmega32u4 z 2,5 kilobajta SRAM [5].

SRAM jest głównym ograniczeniem, ponieważ 32 bajty nie są wystaczające do przeprowadzania operacji, przy których sam klucz zajmuje 16 lub 32 bajty. Należy też pamiętać, że obsługa bezpiecznego połączenia nie może zajmować całości pamięci. Część należy przeznaczyć na obsługę peryferiów oraz właściwą logikę programu.

Istotnym elementem jest też wielkość domyślnych buforów. *Arduino* w modułach *Serial* oraz *SoftwareSerial* domyślnie używa 16- lub 64-bajtowego (w zależności od ilości dostępnej pamięci) buforu na przychodzące dane[1]. Przy wiadomościach dłuższych niż 64 bajty oznacza to, że zbyt długie przetwarzanie jednej wiadomości spowoduje błędne odebranie następnej, jeżeli zostanie ona za szybko wysłana.

FIXME tutaj jeszcze cos o taktowaniu procesora.

[1] https://github.com/arduino/Arduino/blob/master/hardware/arduino/avr/cores/arduino/HardwareSerial.h

3. Metody uwierzytelniania

W zależności od potrzeb i ograniczeń stosuje się różne metody uwierzytelniania podmiotów w komunikacji. Wyróżnić należy uwierzytelnianie przy pomocy kryptografii asymetrycznej, w której używana jest para matematycznie związanych ze sobą kluczy, oraz uwierzytelnianie przy pomocy kryptografii symetrycznej, w której używany jest jeden, współdzielony, tajny klucz.

Klucze w przypadku kryptografii asymetrycznej muszą posiadać konkretne właściwości. W przypadku algorytmu RSA bezpieczeństwo polega na trudności w faktoryzowaniu dużych liczb, co wymaga stosowania kluczy co najmniej 2048 bitowych [6]. Klucze w przypadku kryptografii symetrycznej nie muszą mieć konkretnych właściwości poza ich nieprzewidywalnością.

Ważną różnicą jest też wydajność. Kryptografia asymetryczna jest dużo bardziej złożona obliczeniowo od symetrycznej [7]. Jest to szczególnie istotne na ograniczonych sprzętowo systemach wbudowanych. Przewagą kryptografii asymetrycznej jest jednak brak konieczności ustalenia wspólnego klucza przed rozpoczęciem komunikacji, jak ma to miejsce w przypadku kryptografii symetrycznej.

Zalecanym rozwiązaniem jest najpierw ustalenie wspólnego, tajnego klucza przy użyciu kryptografii asymetrycznej, a następnie użycie tego klucza do kryptografii symetrycznej [7].

3.1. Kryptografia asymetryczna

Przy wyborze algorytmu dla potrzeb pracy istotne były:

– jakość implementacji dostępnych na mikroprocesory AVR,

– złożoność obliczeniowa (niższa jest lepsza),

– długość klucza wymagana do zapewnienia niezbędnego poziomu bezpieczeństwa.

Biblioteka *AVR-Crypto-Lib* dostarcza implementację algorytmów RSA oraz DSA[1]. Biblioteka *Emsign* dostarcza implementację RSA, lecz tylko z 64 bitowym kluczem[2], co nie jest wystarczające dla zapewnienia bezpieczeństwa. Komercyjna biblioteka *LightCrypt-AVR8-ECC* oraz biblioteka *micro-ecc* dostarczają implementację kryptografii opartej o krzywe eliptyczne[3]. Brak jest na rynku implementacji

[1] *https://trac.cryptolib.org/avr-crypto-lib/browser*

[2] *http://www.emsign.nl/*

[3] *http://industrial.crypto.cmmsigma.eu/lightcrypt_avr8/lc_avr8_ecc.pl.html*

innych algorytmów klucza publicznego. Dostępność implementacji ogranicza wybór algorytmu do RSA, DSA oraz krzywych eliptycznych.

Następnym kryterium jest złożoność obliczeniowa. W analizie przeprowadzonej przez pracowników *Sun Microsystems Laboratories* wykazano, że na mikroprocesorach AVR algorytmy oparte o krzywe eliptyczne są o rząd wielkości szybsze od algorytmu RSA [8].

Krzywe eliptyczne wymagają najkrótszych kluczy. Rekomendacje National Institute of Standards and Technology (NIST) [6] podają, że 256-bitowy klucz Elliptic Curve Cryptography (ECC) zapewnia bezpieczeństwo porównywalne do 3072-bitowego klucza RSA lub DSA.

W związku z przewagą krzywych eliptycznych przy zadanych założeniach do ustalenia wspólnego klucza wybrano algorytm Elliptic Curve Diffie-Hellman (ECDH). Wadą tego rozwiązania jest niezmienność klucza ustalanego tą metodą. Powoduje to brak utajnienia przekazywania *(ang. forward secrecy)*.

3.2. Kryptografia symetryczna

ECBC-MAC OMAC CCM HMAC

4. Implementacja

Implementację wykonano z następującymi założeniami:

- pełna funkcjonalność przy jak najmniejszych wymaganiach sprzętowych
- niezależność od warstwy sieciowej
- zapewnienie uwierzytelniania i szyfrowania wiadomości

4.1. Protokół komunikacji

W protokole wymieniane są rekordy o następującej budowie:

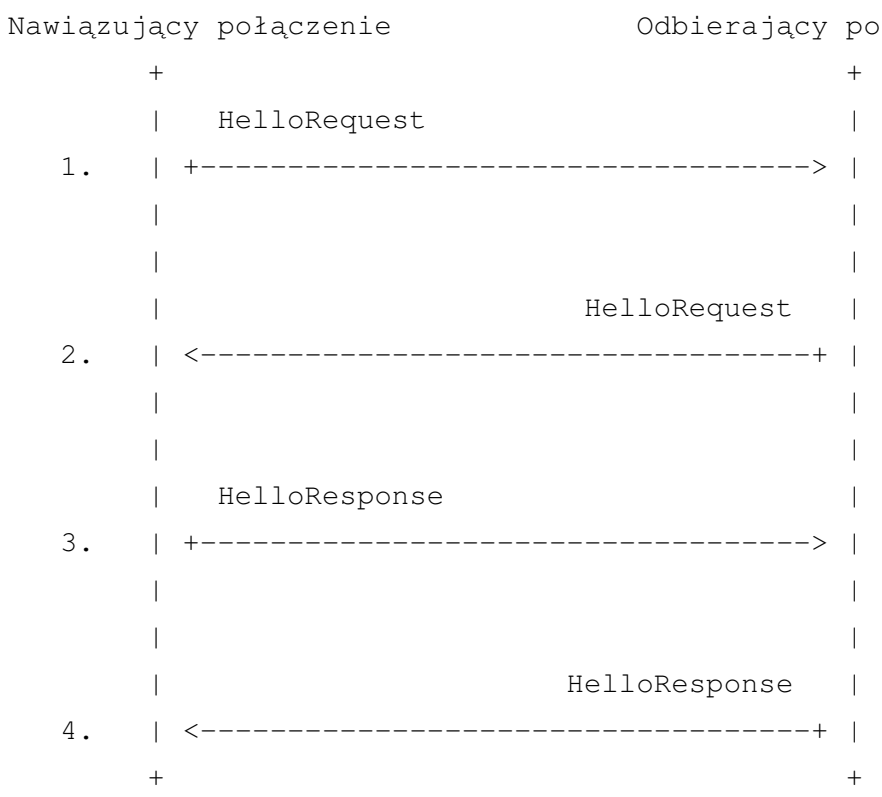
- piersze dwa bajty definiują wersję protokołu i mają wartość 0x00 0x01,
- następny bajt definiuje typ rekordu,
- następne dwa bajty definiują długość zawartości rekordu (najbardziej znaczący bajt jako pierwsze),
- zawartość rekordu.

Zdefiniowane są następujące typy rekordów.

- HelloRequest o wartości 0x00. Długość zawartości tego rekordu to zawsze 64 bajty. Zawartość tego rekordu nie jest zaszyfrowana ani uwierzytelniona.
- HelloResponse o wartości 0x01. Długość zawartości tego rekordu to zawsze 128 bajtów. Zawartość tego rekordu jest zaszyfrowana oraz uwierzytelniona.
- EncryptedData o wartości 0x02. Długość nie jest zdefiniowana. Zawartość tego rekordu jest zaszyfrowana oraz uwierzytelniona

Odbiorca rekordu powinien zweryfikować:

- zgodność wersji protokołu – wymagane bajty 0x00 oraz 0x01,
- prawidłowość bajtu określającego typ rekordu – wymagana wartość 0x00, 0x01 lub 0x02,



Rys. 4.1. Kolejność wymiany rekordów w procesie nawiązywania połączenia

– zgodność zadeklarowanej długości zawartości rekordu z typem rekordu,

– w przypadku typów HelloResponse oraz EncryptedData – prawidłowość kodu uwierzytelniającego.

W przypadku niezgodności któregokolwiek elementu rekord powinien zostać zignorowany.

4.1.1. Nawiązywanie połączenia

Kolejność przesyłania rekordów w celu nawiązania połączenia przedstawiona jest na Rys. 4.1. Rekordy HelloRequest zawierają klucz publiczny węzła, który je wysyła. Węzeł, który odbiera HelloRequest używa swojego klucza publicznego, oraz klucza publicznego z odebranego rekordu do ustalenia sekretnego klucza.

Po ustaleniu wspólnego klucza węzły mogą wysłać rekord HelloResponse, który zawiera zaszyfrowany i uwierzytelniony klucz publiczny węzła wysyłającego rekord. Jeżeli węzeł odbierający rekord jest w stanie potwierdzić, że rekord jest prawidłowo uwierzytelniony, a zdeszyfrowany klucz publiczny pokrywa się z kluczem przesłanym wcześniej w rekordzie HelloRequest, połączenie uznawane jest za nawiązane. Po nawiązaniu połączenia wymieniane mogą być tylko rekordy typu EncryptedData.

4.2. Generowanie współdzielonego klucza

Każdy z węzłów po odebraniu rekordu HelloRequest używa odebranego klucza publicznego oraz swojego klucza publicznego do ustalenia wspólnego sekretu przy użyciu algorytmu ECDH oraz proponowanej przez NIST krzywej eliptycznej P-256 [9] (w RFC 5480 nazwaną krzywą secp256r1 [10]).

Z sekretu będącego wynikiem algorytmu ECDH liczony jest skrót przy użyciu algorytmu SHA-256. Następnie jest on dzielony na dwie części po 128-bitów. Pierwsza część staje się współdzielonym kluczem używanym do szyfrowania, druga część staje się współdzielonym kluczem używanym do uwierzytelniania.

Implementacja algorytmu ECDH z krzywą eliptyczną P-256 pochodzi z biblioteki *micro-ecc*. Jest to jedyna darmowa biblioteka implementująca ECDH.

4.3. Szyfrowanie i deszyfrowanie wiadomości

Szyfrowanie wiadomości odbywa się za pomocą szyfru blokowego AES ze 128-bitowym kluczem używanym w trybie Cipher Block Chaining (CBC). Wektor inicjalizacyjny jest losowy i dołączany przed szyfrogramem. Tekst jawny jest dopełniany do pełnego bloku według algorytmu zdefiniowanego w PKCS#7 [11].

Właściwe kroki potrzebne do zaszyfrowania wiadomości:

1. dopełnienie tekstu jawnego do pełnego bloku
   - jeżeli długość tekstu jawnego jest wielokrotnością długości bloku, do tekstu jawnego doklejone musi być 128 bajtów o wartości 128.
   - w przeciwnym wypadku, gdy wymagane jest dopełnienie N bajtów, do tekstu jawnego doklejone musi być N bajtów o wartości N.
2. zaszyfrowanie dopełnionego tekstu jawnego w trybie CBC z losowym wektorem inicjalizacyjnym
3. doklejenie przed szyfrogramem wektora inicjalizacyjnego

Właściwe kroki potrzebne do odszyfrowania wiadomości:

1. oddzielenie wektora inicjalizacyjnego od szyfrogramu
2. zdeszyfrowanie szyfrogramu w trybu CBC korzystając z oddzielonego wektora inicjalizacyjnego
3. pobranie wartości ostatniego bajtu zdeszyfrowanego ciągu
   - wartość ta nazywana jest dalej N
4. zweryfikowanie poprawności dopełnienia

- ostatnie N bajtów musi mieć wartość N
- jeżeli dopełnienie jest nieprawidłowe, cały rekord jest ignorowany

5. usunięcie ostatnich N bajtów

Implementacja algorytmu AES pochodzi z biblioteki *AVR-Crypto-Lib*. Jest to najlepiej udokumentowana, darmowa biblioteka implementująca algorytm AES. Implementacja trybu CBC oraz algorytmu dopełniania została zrealizowana w ramach pracy.

4.4. Uwierzytelnienie wiadomości

Szyfrowanie wiadomości odbywa się za pomocą szyfru blokowego AES ze 128-bitowym kluczem używanym w trybie Encrypt-last-block Cipher Block Chaining Message Authentication Code (ECBC-MAC). Wektor inicjalizacyjny wypełniony jest zerami i nie jest przesyłany. Uwierzytelniany jest kompletny szyfrogram wraz z wektorem inicjalizacyjnym użytym do szyfrowania, a nie tekst jawny. Długość szyfrogramu wraz z wektorem inicjalizacyjnym zawsze będzie wielokrotnością długości bloku, a więc nie jest stosowane dopełnianie.

Tryb ECBC-MAC to tryb Cipher Block Chaining Message Authentication Code (CBC-MAC), którego wynik jest dodatkowo szyfrowany innym kluczem niż ten użyty do CBC-MAC. W tej pracy do CBC-MAC użyty jest klucz służący do uwierzytelniania, a wynik CBC-MAC jest zaszyfrowany używając klucza służącego do szyfrowania.

Właściwe kroki potrzebne do obliczenia kodu uwierzytelniającego:

1. obliczenie ostatniego bloku będącego wynikiem zaszyfrowania szyfrogramu w trybie CBC z wektorem inicjalizacyjnym wypełnionym zerami przy użyciu klucza przeznaczonego do uwierzytelniania
2. zaszyfrowanie bloku korzystając z AES i klucza przeznaczonego do szyfrowania

Węzeł wysyłający dokleja kod uwierzytelniający przed szyfrogramem. Węzeł odbierający oddziela otrzymany kod od szyfrogramu, oblicza kod uwierzytelniający dla danego szyfrogramu i porównuje, czy zgadza się on z kodem otrzymanym. Jeżeli kod obliczony różni się od kodu otrzymanego, cały rekord jest ignorowany.

Implementacja trybu ECBC-MAC została zrealizowana w ramach pracy.

5. Walidacja

Informacja o bibliotece dla Javy i przykladowych implementacjach.

6. Podsumowanie

?

Spis rysunków

Bibliografia

[1] Massimo Banzi and Michael Shiloh. *Getting Started with Arduino: The Open Source Electronics Prototyping Platform*. Sebastopol: Maker Media, Inc., 2014.

[2] Martin McKeay i in. *Q3 2016 State of the Internet Security Report*. Spraw. tech. Akamai Technologies, Inc., 2016.

[3] *ATtiny4 / ATtiny5 / ATtiny9 / ATtiny10 - Datasheet Summary*. Atmel. 2016. URL: *http://www.atmel.com/Images/Atmel-8127-AVR-8-bit-Microcontroller-ATtiny4-ATtiny5-ATtiny9-ATtiny10_Datasheet-Summary.pdf* (dostęp dnia 2016-12-06).

[4] *ATxmega384C3 - Datasheet*. Atmel. 2016. URL: *http://www.atmel.com/Images/Atmel-8361-8-and-16-bit-AVR-XMEGA-Microcontrollers-ATxmega384C3_Datasheet.pdf* (dostęp dnia 2016-12-06).

[5] *ATmega16U4/ATmega32U4 - Datasheet*. Atmel. 2016. URL: *http://www.atmel.com/Images/Atmel-7766-8-bit-AVR-ATmega16U4-32U4_Datasheet.pdf* (dostęp dnia 2016-12-06).

[6] *Recommendation for Key Management Part 1: General*. National Institute of Standards and Technology, U.S. Department of Commerce. 2016. URL: *http://dx.doi.org/10.6028/NIST.SP.800-57pt1r4* (dostęp dnia 2016-12-07).

[7] Abdullah Al Hasib i Abul Ahsan Md Mahmudul Haque. „A comparative study of the performance and security issues of AES and RSA cryptography". W: *Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on*. T. 2. IEEE. 2008, s. 505–510.

[8] Nils Gura i in. „Comparing Elliptic Curve Cryptography and RSA on 8-bit CPUs". W: *Cryptographic Hardware and Embedded Systems - CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings*. Wyed. Marc Joye i Jean-Jacques Quisquater. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, s. 119–132. ISBN: 978-3-540-28632-5. DOI: *10.1007/978-3-540-28632-5_9*.

[9] Cameron F Kerry. „Digital Signature Standard (DSS)". W: *National Institute of Standards and Technology* (2013).

[10] Sean Turner i in. „Elliptic Curve Cryptography Subject Public Key Information". W: (2009).

[11] Burt Kaliski. „Pkcs# 7: Cryptographic message syntax version 1.5". W: (1998).