Dane osobowe w eduroam

Na podstawie opracowania dr Tomasza Wolniewicza, UCI UMK Toruń

1 Terminologia

a) Użytkownik – osoba fizyczna uprawniona do korzystania z eduroam
b) Instytucja macierzysta – instytucja odpowiedzialna za wydanie poświadczenia o uprawnieniu użytkownika do korzystania z eduroam.
c) Instytucja udostępniająca sieć – instytucja udostępniająca własną sieć użytkownikom eduroam.
d) Instytucja pośrednicząca – instytucja utrzymująca serwer pośredniczący eduroam, są to instytucje utrzymujące światowe, krajowe lub regionalne węzły eduroam we wszystkich państwach włączonych w strukturę eduroam.

Uwaga. W eduroam zazwyczaj instytucja występuje zarówno w roli instytucji macierzystej, jak i instytucji udostępniającej sieć. Te role mają jednak zastosowanie w stosunku do różnych grup użytkowników i muszą być prawidłowo rozróżniane na potrzeby opisu procesów przetwarzania danych osobowych.

2 Ogólny opis przetwarzania danych osobowych

W ramach eduroam można wyróżnić trzy procesy przetwarzania danych. Uzasadnienie przetwarzaniach konkretnych atrybutów użytkownika jest dodatkowo opisane w dalszych sekcjach niniejszego dokumentu.

2.1 Proces uzyskania dostępu do sieci

eduroam umożliwia gościnne korzystanie z sieci instytucji udostępniającej sieć w oparciu o poświadczenie instytucji macierzystej użytkownika. Instytucja macierzysta przed wydaniem poświadczenia jest zobowiązana do potwierdzenie tożsamości oraz uprawnień użytkownika. Instytucja udostępniająca sieć ponosi ryzyko związane z nieuprawnionymi działaniami użytkowników korzystających z dostępu gościnnego. Nieuprawnione działania użytkownika mogą nieść za sobą różne konsekwencje poczynając od zwrócenia użytkownikowi uwagi, poprzez odebranie użytkownikowi uprawnień do korzystania z sieci, aż do współpracy z uprawnionymi organami ścigania.

Należy zaznaczyć, że w typowej sytuacji, tzn. kiedy urządzenie użytkownika jest prawidłowo skonfigurowane, instytucja udostępniająca sieć nie posiada informacji, która pozwoliłaby jej na powiązanie dostępu sieciowego z konkretną osobą fizyczną. Takie powiązanie jest możliwe wyłącznie poprzez analizę połączonych danych posiadanych przez zarówno prze instytucję udostępniającą sieć, jak i instytucję macierzystą. Zgodnie z prawem ochrony danych osobowych oraz regulaminem eduroam, instytucja udostępniająca sieć nie jest uprawniona do otrzymania z instytucji macierzystej użytkownika danych, które pozwoliłyby jej na zidentyfikowanie osoby fizycznej, taki proces może być realizowany wyłączenie poprzez organy uprawnione przez prawo.

Proces uwierzytelnienia w eduroam jest złożony i ma na niego wpływ szereg elementów. Dane przekazywane w procesie uwierzytelnienia są niezbędne do analizy problemów zgłaszanych przez użytkowników, są przechowywane i mogą być przetwarzane przez instytucje prowadzące serwery pośredniczące. To, jakie serwery pośredniczące biorą udział w procesie uwierzytelnienia zależy od miejsca, w którym znajduje się użytkownik.

2.2 Proces przetwarzania danych do celów statystycznych

Dane uwierzytelnień są również przetwarzane w celach statystycznych na potrzeby skuteczności działania projektu eduroam. Dane statystyczne są przetwarzane lokalnie, przekazywane do centralnego serwera Polski, a również dalej do serwera na poziomie europejskim.

2.3 Proces zbierania danych o trwającym połączeniu

Urządzenia sieciowe z reguły posiadają możliwości szczegółowego raportowania sieci. Zbieranie tych informacji pozwala na zauważanie różnego rodzaju nieprawidłowości, a również analizę wcześniej zgłoszonych problemów. Zwłaszcza w sieciach bezprzewodowych, wadliwe urządzenie użytkownika może skutecznie zakłócić transmisję innych użytkowników. Zbieranie takich danych jest zatem uzasadnione dbałością o jakość sieci. Zgodnie z zapisami polskiego regulaminu eduroam oraz regulaminu europejskiego, dane dotyczące zużycia sieci muszą być blokowane przez serwery pośredniczące, z wyjątkiem sytuacji, gdy zainteresowane instytucje zawierają bezpośrednie porozumienia.

3 Podstawa formalna przetwarzania danych

a) Regulamin usługi eduroam w Polsce [eduroam-pl]
b) Zasady działania europejskiej konfederacji eduroam [eduroam-org]
c) Zasady współpracy regionalnych konfederacji eduroam [eduroam-complience]
d) Deklaracja przystąpienia Polski do europejskiej konfederacji eduroam [deklaracja-pl]
e) Deklaracje polskich instytucji korzystających z eduroam [deklaracja]
f) Regulaminy zagranicznych federacji eduroam

4 Okres retencji danych

a) Instytucje macierzyste – minimum 6 miesięcy na podstawie regulaminu eduroam, dłuższy okres retencji może wynikać z wewnętrznych regulaminów instytucji
b) Instytucja udostępniająca sieć – zalecane minimum 6 miesięcy
c) Instytucja pośrednicząca – minimum 6 miesięcy

Uwaga. Dane osobowe w instytucjach udostępniających sieć i instytucjach pośredniczących są wysoce zanonimizowane, dzięki czemu są zminimalizowane zagrożenia wynikające z ich przechowywania i przetwarzania.

5 Obowiązek informacyjny

Informacje na temat eduroam, w szczególności regulamin, opisy działania systemu itp. są dostępne na stronach https://eduroam.pl oraz https://eduroam.org.

6 Dane użytkownika w systemie eduroam

6.1 Dane niezbędne w procesie uwierzytelnienia

Typ	gdzie
adres interfejsu sieciowego (MAC)	M,P,S
identyfikator podłączenia, a w przypadku korzystania z certyfikatów indywidualnych, również cały certyfikat – zazwyczaj dana anonimowa lub spseudonimizowana	M,P,S
identyfikator użytkownika (pozwalający na jego jednoznaczną identyfikację)	M
identyfikator instytucji macierzystej	M,P,S
identyfikator spseudonimizowany CUI	M,P,S
czas nawiązania połączenia	M,P,S

przydzielony adres internetowy	S
lokalizacja użytkownika na terenie sieci (identyfikator punktu bezprzewodowego obsługującego użytkownika)	S,P*,M*
hasło użytkownika lub NT-hash (o ile są stosowane)	M

*Identyfikator może zawierać adres internetowy punktu dostępowego, ale ta informacja nie powinna być przekazywana poza instytucję udostępniającą sieć.

Oznaczenia:
M – instytucja macierzysta
P – instytucja pośrednicząca
S – instytucja udostępniająca sieć

Jeżeli instytucja udostępniająca sieć prowadzi analizę ruchu generowanego przez użytkowników sieci, to istnieje możliwość powiązania takiego ruchu z parametrami wymienionymi powyżej, wykracza to jednak poza zakres niniejszego dokumentu.

6.2 Dane przekazywane dla potrzeb statystycznych

System eduroam jest finansowany w zdecydowanej większości ze środków publicznych (projekt GEANT, dofinansowania krajowych akademickich sieci komputerowych, środki uczelni publicznych). Na potrzeby raportowania skuteczności projektu eduroam zbierane są informacje o każdym wykonanym połączeniu. Ponieważ obecnie wszystkie połączenie gościnne są realizowane za pośrednictwem serwerów pośredniczących, dane statystyczne są zbierane tylko z tych serwerów.

Dane na potrzeby statystyczne są zachowywane w logach serwera pośredniczącego oraz przekazywane do serwera krajowego, który następnie przekazuje je do serwera europejskiego.

Przekazywane dane zawierają: czas uwierzytelnienia, identyfikator instytucji udostępniającej sieć, identyfikator instytucji macierzystej, identyfikator urządzenia.

7 Uproszczony opis procesu uwierzytelnienia

Urządzenie użytkownika, odpowiadając na sygnał sieci, rozpoczyna transmisję, w ramach której wstępnie przekazuje minimalny zakres informacji dostępowych. Te informacje są przekazywane przez infrastrukturę eduroam do instytucji macierzystej użytkownika. Serwer instytucji macierzystej użytkownika nawiązuje bezpośrednie, zabezpieczone przed podsłuchem połączenie z urządzeniem użytkownika i poprzez to połączenie następuje wymiana bardziej szczegółowej informacji uwierzytelniającej. Po poprawnym uwierzytelnieniu użytkownika serwer instytucji macierzystej przekazuje serwerowi instytucji udostępniającej sieć zgodę na podłączenie użytkownika.

Dane są przekazywane najkrótszą możliwą ścieżką pomiędzy instytucją udostępniającą sieć, a instytucją macierzystą.

Rozbudowany opis zasad działania eduroam jest zawarty w dokumencie RFC [rfc7593] oraz wielu polskich opracowaniach dostępnych na stronie https://eduroam.pl .

8 Uzasadnienie i analiza ryzyka przetwarzania konkretnych danych osobowych

8.1 Adres interfejsu sieciowego MAC

Adres sieciowy przypisany przez urządzenie użytkownika jest zazwyczaj związany z tym urządzeniem w trwały sposób. Na podstawie adresu często można ustalić producenta urządzenia. Adres MAC musi być przekazywany w ramach protokołu RADIUS, jest zatem niezbędnym składnikiem całego procesu uwierzytelnienia w eduroam. Adres MAC pozwala rozróżniać urządzenia jest więc przydatny również do celów statystycznych. Z uwagi na swój stały charakter, adres MAC jest trwałym śladem użytkownika pozostawianym we wszystkich sieciach bezprzewodowych do których się łączył, może być zatem użyty do celów lokalizacji użytkownika i z tego powodu powinien być uważany za daną wymagającą szczególnej ochrony. Należy zwrócić uwagę, że adres MAC jest niezbędnym elementem na drodze ustalenia tożsamości użytkownika w sytuacjach kiedy jest to niezbędne, niemożliwa jest zatem pseudonimizacja tej informacji i usuwanie informacji źródłowej.

8.2 Identyfikator połączenia

W sytuacjach, kiedy stosowane jest uwierzytelnienie typu użytkownik/hasło, identyfikator użytkownika powinien mieć postać id@domena.macierzysta, gdzie id powinien być wspólny dla wszystkich identyfikatorów instytucji, może być również pustym ciągiem znaków. W taki przypadku identyfikator nie powinien być kwalifikowany jako dana osobowa. Istnieje jednak potencjalne ryzyko, że nieprawidłowo skonfigurowane urządzenie końcowe użytkownika udostępnia w tym miejscu rzeczywisty, jednoznaczny identyfikator użytkownika. Taka sytuacja jest efektem błędu po stronie użytkownika, tym niemniej z uwagi na to zagrożenie, w stosunku do tego identyfikatora powinno się stosować ochronę, jak dla danych osobowych. Identyfikator połączenia jest widoczny na całej drodze uwierzytelnienia.

W sytuacji, kiedy zamiast pary identyfikator/hasło stosowany jest osobisty certyfikat użytkownika, w roli identyfikatora stosuje się pole jedno z pól certyfikatu, a cały certyfikat jest przekazywany bez zabezpieczenia poprzez całą infrastrukturę eduroam. Zalecane jest, aby w tym przypadku stosować certyfikaty pseudonimizowane, tak by nie zdradzać istotnych danych użytkownika. Stosowanie certyfikatów zawierających imię i nazwisko oraz adres email użytkownika jest zdecydowanie niezalecane.

8.3 Identyfikator użytkownika

Identyfikator musi być znany instytucji macierzystej i tylko jej.

W sytuacji, kiedy stosuje się uwierzytelnienie oparte o identyfikator/hasło, identyfikator jest przekazywany szyfrowanym tunelem bezpośrednio między urządzeniem użytkownika a serwerem instytucji macierzystej, nie jest zatem ujawniany ani w instytucji udostępniającej sieć, ani w instytucjach pośredniczących.

Bezpieczeństwo przekazywanych informacji zależy od prawidłowości konfiguracji urządzenia użytkownika (patrz niżej).

W sytuacji, kiedy zamiast pary identyfikator/hasło stosowany jest osobisty certyfikat użytkownika, należy stosować uwagi opisane w części „identyfikator połączenia”.

8.4 Hasło użytkownika lub NT-hash

Hasło musi być znane wyłącznie instytucji macierzystej. Jest przekazywane szyfrowanym tunelem bezpośrednio między urządzeniem użytkownika a serwerem instytucji macierzystej, nie jest zatem ujawniane ani w instytucji udostępniającej sieć, ani w instytucjach pośredniczących.

Hasło jest porównywane z danymi posiadanymi przez instytucję macierzystą, nie musi i nie powinno być zapisywane w jakichkolwiek logach.

Bezpieczeństwo przekazywanych informacji zależy od prawidłowości konfiguracji urządzenia użytkownika (patrz niżej).

8.5 Identyfikator instytucji macierzystej (realm)

Jest to identyfikator będący częścią identyfikatora połącznia i będący nazwą domenową domeny będącej własnością instytucji macierzystej. Identyfikator ten pozwala zatem określić instytucję macierzystą. Identyfikator jest niezbędny w procesie uwierzytelnienia, korzysta się z niego również w celach statystycznych

8.6 Czas nawiązania połączenia

Przechowywanie dokładnego czasu połączeń jest wymogiem regulaminu eduroam i jest niezbędne przy identyfikacji problemów.

8.7 Identyfikator Operator-Name instytucji udostępniającej sieć

Identyfikator będący nazwą domenową domeny będącej własnością instytucji udostępniającej sieć. Identyfikator ten może, ale nie musi być stosowany. Podstawowym celem jego stosowania jest znaczące ułatwienie procesu analizowania problemów uwierzytelniania użytkowników. Ten identyfikator jest również niezbędny w procesie zapytania o spseudonimizowany identyfikator CUI (patrz poniżej). Operator-Name pozwala na powiązanie danych użytkownika z miejscem dostępu do sieci, pozwala zatem na zgrubną lokalizację użytkownika.

8.8 Identyfikator spseudonimizowany CUI

Identyfikator ma na celu umożliwienie instytucji udostępniającej sieć powiązanie wszystkich uwierzytelnień związanych z jednym użytkownikiem. Przykładem takiej sytuacji jest nadużywanie sieci przez zaawansowanego użytkownika i zmienianie przez niego adresów MAC w celu uniknięcia blokad zakładanych adres MAC. Dysponowaniem identyfikatorem CUI pozwala instytucji udostępniającej sieć na założenie blokady na wszystkie uwierzytelnienia takie użytkownika.

Identyfikator CUI jest generowany przez instytucję macierzystą i wysyłany w pakiecie potwierdzającym uprawnienie do sieci. Identyfikator CUI jest tworzony w sposób zapewniający maksymalną ochronę prywatności użytkownika. Zgodnie z przyjętymi w eduroam zasadami, identyfikator jest wysyłany w odpowiedzi na prośbę instytucji udostępniającej sieć zawartą w pakiecie uwierzytelniającym. Warunkiem wygenerowania CUI jest przekazanie przez instytucję udostępniającą sieć identyfikatora Operator-Name. Instytucja macierzysta generuje identyfikator CUI tak, by był on inny dla każdej wartości Operator-Name (czyli dla każdej instytucji udostępniającej sieć).

8.9 Identyfikator (adres internetowy) punktu dostępowego w instytucji udostępniającej sieć

Adres internetowy punktu dostępowego, z którego korzysta użytkownik, może być przesyłany jako atrybut NAS-IP-Address. W wielu przypadkach sieci stosują tzw. adresy prywatne, które nie pozwalają na zidentyfikowanie sieci, z której przyszło zapytanie, może się jednak zdarzyć, że stosowany jest

adres pozwalający na zidentyfikowanie miejsca. Zaleca się, aby w przypadku stosowania publicznych adresów IP nie wysyłać atrybutu NAS-IP-Address, stosując zamiast tego atrybut NAS-Identifier.

9 Bezpieczeństwo informacji przekazywanych w szyfrowanym kanale

Na poprawnie skonfigurowanym urządzeniu użytkownika zapisane są: certyfikat urzędu certyfikującego, który wystawił certyfikat serwera instytucji macierzystej, nazwa serwera, zgodna z zapisaną w certyfikacie serwera.

W początkowej fazie ustanawiania kanału szyfrowanego między urządzeniem użytkownika, a serwerem instytucji macierzystej, urządzenie użytkownika powinno zweryfikować prawidłowość certyfikatu, którym przedstawia się serwer. Jeżeli certyfikat jest zgodny z parametrami zapisanymi w urządzeniu, to połączenie może dojść do skutku, jeżeli nie, to powinno być zerwane. Dzięki temu, urządzenie użytkownika nie przekaże nieznanemu serwerowi istotnych danych użytkownika, takich jak identyfikator i hasło.

Światowy koordynator eduroam – GEANT – udostępnia narzędzie wspomagające prawidłową konfigurację urządzeń użytkowników pod adresem https://cat.eduroam.org. Stosowanie tych konfiguratorów gwarantuje bezpieczeństwo użytkowników.

10 Źródła

- [eduroam-pl] http://www.eduroam.pl/Dokumenty/Regulamin_eduroam_v2_18_10_2013.pdf
- [eduroam-org] http://www.eduroam.pl/Dokumenty/GN3-12-192_eduroam-policy-service-definition_ver28_26072012.pdf
- [eduroam-complience] https://www.eduroam.org/wp-content/uploads/2016/05/eduroam_Compliance_Statement_v1_0.pdf
- [deklaracja-pl] http://www.eduroam.pl/Dokumenty/eduroam_policy_v2_pl.pdf
- [deklaracja] http://www.eduroam.pl/Dokumenty/Deklaracja_eduroam_v2_18_10_2013.docx
- [rfc7593] https://tools.ietf.org/html/rfc7593