UAI - Laboratorium 6 - VoIP

Stanisław Kwiatkowski, Bartosz Ziemba, Kamil Stachowicz

1. Wprowadzenie

Niniejsze ćwiczenie ma na celu zaprezentować analizę działania systemu Voice over IP (VoIP), przy wykorzystaniu aplikacji Twinkle i narzędzia do analizy sieci Wireshark. Celem tego ćwiczenia jest zrozumienie mechanizmów komunikacji pomiędzy klientem SIP (Session Initiation Protocol), jakim jest Twinkle, a serwerem VoIP. Ćwiczenie obejmuje połączenie z serwerem, zmianę ustawień dotyczących kodeku głosowego oraz transportu DTMF (Dual-Tone Multi-Frequency) w Twinkle, a następnie przeprowadzenie analizy przepływu danych między klientem a serwerem. Przechwycone dane będą następnie zapisane i analizowane. Ćwiczenie będzie zawierało szczegółowy przegląd sekwencji i zawartości wiadomości protokołu SIP, analizę danych RTP oraz informacji o sesjach opisanych przy pomocy języka SDP.

2. Sekwencja wiadomości protokołu SIP wymienionych z serwerem

Sekwencja wiadomości protokołu SIP wymienionych z serwerem w czasie pierwszej rozmowy:

Rys. 1: Graf komunikacji 1 rozmowy

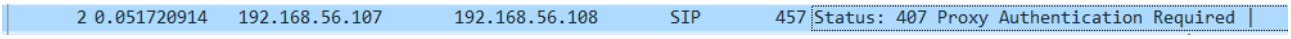
Sekwencja wiadomości protokołu SIP wymienionych z serwerem w czasie drugiej rozmowy:

Rys. 2: Graf komunikacji 2 rozmowy

3. Pierwsza Sesja z serwerem

3.1. Odpowiedz 407 Proxy Authentication

Odpowiedź ta jest prośbą serwera o dostarczenie poprawnych danych do uwierzytelniania, jeśli chce dalej kontynuować połączenie. Jest to potrzebne by zweryfikować użytkownika, zanim serwer udzieli mu dostępu do usługi.



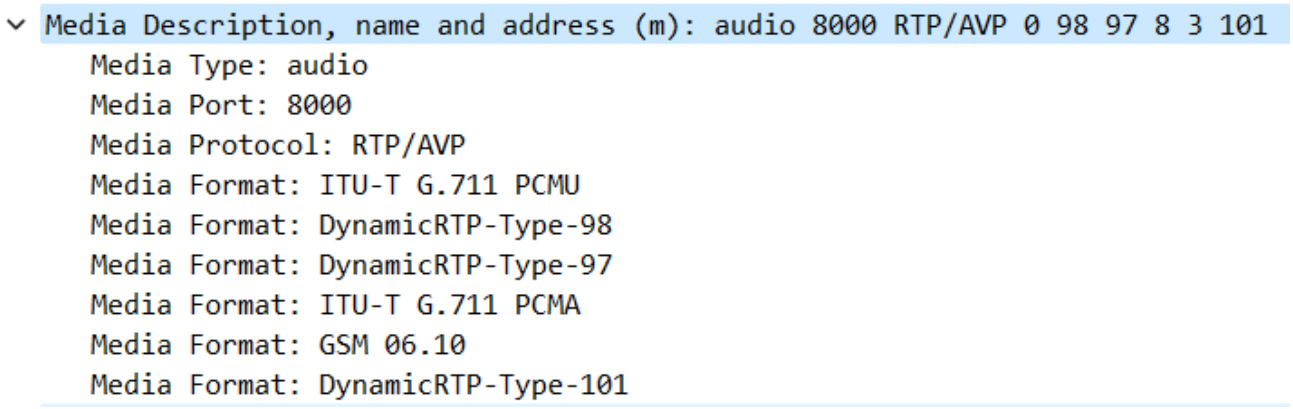
Rys. 3: 407 Proxy Authentication

3.2. Zawartość linii Media Description w wiadomościach SDP

3.2.1. Wiadomość INVITE

Wiadomość SIP (Session Initiation Protocol) "INVITE" jest używana do inicjowania sesji komunikacyjnej.

- **Media Type:** Typ medium do przesyłania w sesji, który w tym przypadku to "audio", oznaczające transmisję dźwięku.
- **Media Port:** Numer portu, na którym serwer oczekuje na nawiązanie połączenia RTP. W tym przypadku to port 8000.
- **Media Protocol:** Protokół do przesyłania danych multimedialnych. Wykorzystuje się tutaj RTP/AVP (Real-time Transport Protocol / Audio Video Profile), standardowy protokół do przesyłania danych audio i wideo w czasie rzeczywistym.
- **Media Format:**Formaty, które będą używane do przesyłania danych multimedialnych w tej sesji
  - "ITU-T G.711 PCMU" i "ITU-T G.711 PCMA" to formaty kodowania dźwięku, które korzystają z kodowania PCM (Pulse Code Modulation) z odpowiednio μ-Law i A-Law, zapewniające wysoką jakość dźwięku.
  - "GSM 06.10" to standard kompresji dźwięku używany w sieciach GSM.
  - "DynamicRTP-Type-97", "DynamicRTP-Type-98" i "DynamicRTP-Type-101" to dynamiczne typy ładunku RTP, których dokładne formaty są zdefiniowane dynamicznie w danej sesji.



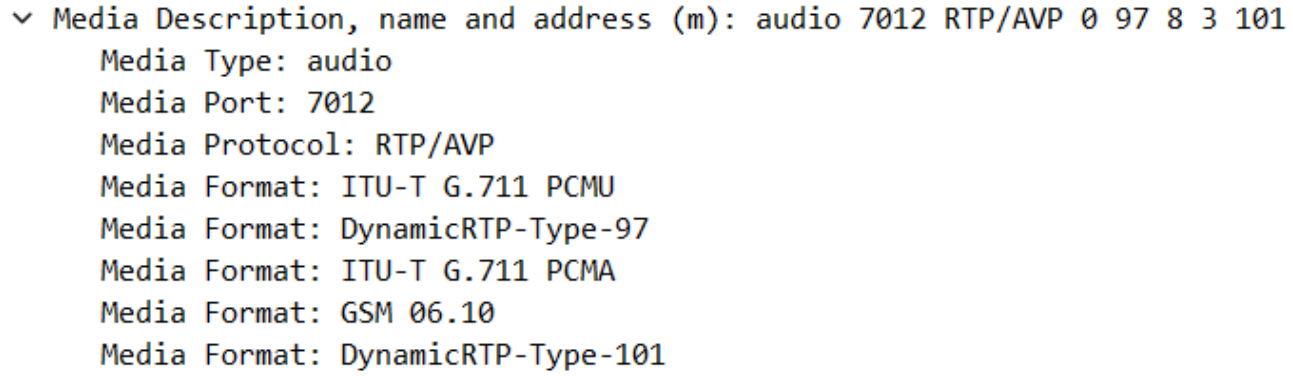
Rys. 4: Media Description wiadomości INVITE

Ogólnie rzecz biorąc, te pola określają, jakie media będą używane w sesji, jaki protokół będzie używany do transmisji tych mediów, i na jakim porcie serwer oczekuje na nawiązanie połączenia. Te szczegóły są niezbędne, aby obie strony mogły prawidłowo skonfigurować swoje systemy do komunikacji.

3.2.2. Wiadomość 200 OK

Wiadomość SIP (Session Initiation Protocol) "200 OK" jest wiadomością odpowiedzi wysyłaną przez serwer, wskazującą, że żądanie od klienta zostało pomyślnie przetworzone. Przekazujesz informacje z sekcji opisującej medium, która zawiera szczegóły o strumieniu audio, który ma być ustanowiony.

- **Media Type:** Typ medium, które ma być używane w sesji. W tym przypadku jest to "audio", co oznacza, że sesja będzie transmitować dźwięk.
- **Media Port:** Numer portu, na którym serwer oczekuje na nawiązanie połączenia RTP (Real-time Transport Protocol) dla strumienia audio. W tym przypadku portem jest 7012.
- **Media Protocol:** Protokół, który będzie używany do przesyłania danych multimedialnych. Tutaj używany jest RTP/AVP (Real-time Transport Protocol / Audio Video Profile), który jest standardowym protokołem do przesyłania audio i wideo w czasie rzeczywistym.
- **Media Format:**Formaty, w jakich dane multimedialne będą przesyłane przez sesję.
  - "ITU-T G.711 PCMU" i "ITU-T G.711 PCMA" to formaty kodowania dźwięku, które korzystają z kodowania PCM (Pulse Code Modulation) z odpowiednio µ-Law i A-Law, zapewniając wysoką jakość dźwięku.
  - "GSM 06.10" to standard kompresji dźwięku stosowany w sieciach GSM.
  - "DynamicRTP-Type-97" i "DynamicRTP-Type-101" to dynamiczne typy ładunku RTP, co oznacza, że ich dokładne formaty są zdefiniowane dynamicznie dla danej sesji.



Rys. 5: Media Description wiadomości 200 OK

3.3. Kodek mowy

Do przeprowadzenia pierwszego połączenia użyty został kodek G.771, informacje tą możemy odczytać z nagłówka pakietu wyświetlanego w wiresharku oraz informacji o payloadzie.

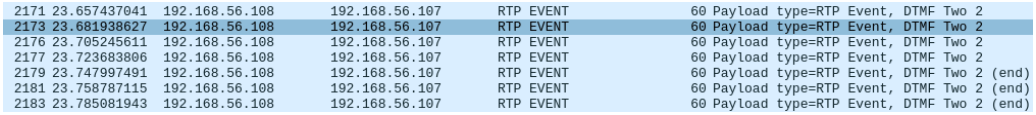
Rys. 6: Kodek mowy

3.4. Pole Mark

Marker bit jest swojego rodzaju wskazówką, może oznaczać początek strumienia głosowego, bądź też inne ważne wydarzenia w trakcie jego trwania. W naszym wypadku jest to na przykład wybranie klawisza numerycznego w kodzie tonowym DTMF. Ogólnie rzecz biorąc, pole Marker może być używane do przekazywania informacji o znaczących momentach w strumieniu multimediów, takich jak zmiana sekwencji, początek nowego nagłówka lub zdarzenie o szczególnym znaczeniu dla aplikacji odbiorczej.

3.5. Przekazywanie informacji o kodzie tonowym DTMF

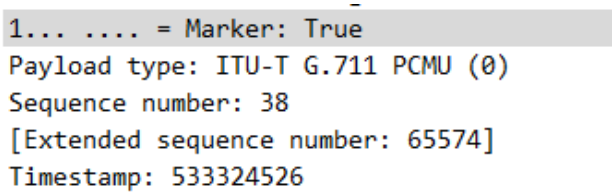
Informacja o kodzie tonowym przekazywana jest za pomocą protokołu RPC Event. Transmisja odbyła się przy pomocy 4 pakietów zawierających informacje odnośnie wciśniętego guzika (2) oraz 2 pakietów z flagą *end*.



Rys. 7: Protokół RPC Event

3.6. Numer sekwencji i Timestamp

Informacje o tych danych także możemy odczytać z nagłówka i jak widać pierwszy pakiet ma sequnce number równy 38 oraz timestamp = 533324526. W naszym protokole wartości inicjalizacyjne tych zmiennych są losowe, ponieważ zwiększa to odporność na ataki "ze znanym tekstem jawnym".

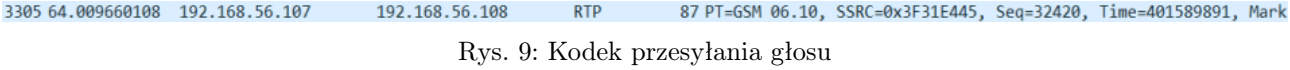


Rys. 8: Kodak mowy

4. Druga sesja z serwerem

4.1. Kodek przesyłania głosu

Do przeprowadzenia pierwszego połączenia użyty został kodek GSM 06.10, informacje tą możemy odczytać z nagłówka pakietu wyświetlanego w wiresharku oraz informacji o payloadzie.



Rys. 9: Kodek przesyłania głosu

4.2. Sposób przekazywania informacja o wciśniętym klawiszu numerycznym

W przypadku połączenia z protokołem transportowym SIP Info informacja o wciśniętym klawiszu przesyłana jest właśnie tym protokołem.

Rys. 10: Pakiet SIP Info

Po rozwinięciu i prześledzeniu pakietów protokołu SIP Info ujrzymy pewną wymianę wiadomości między hostem a klientem. Widać na nim, jak zostaje wysłana informacja o wciśniętym klawiszy 2 i odpowiedź hosta "OK".

```
Signal=2
Duration=100
SIP/2.0 200 OK
Via: SIP/2.0/UDP 192.168.56.108;rport=5060;branch=z9hG4bKigwjcjsn
Contact: <sip:801@192.168.56.107:5060>
To: <sip:801@192.168.56.107>;tag=7733756b
From: "Host machine lab voip" <sip:001@192.168.56.107>;tag=vycqy
Call-ID: rvgbyohdfjmplsq@iplvm
CSeq: 247 INFO
User-Agent: 3CXPhoneSystem 16.0.8.16 (16)
Content-Length: 0
```

Rys. 11: Wymiana wiadomości protokołem SIP Info

5. Podsumowanie

Podczas ćwiczenia przeprowadziliśmy dogłębną analizę sekwencji wiadomości protokołu SIP wymienionych z serwerem za pomocą programu Wireshark. Omówiliśmy rolę odpowiedzi 407 Proxy Authentication Required, która jest kluczowa dla uwierzytelnienia klienta przez serwer proxy w celu zapewnienia dodatkowego poziomu bezpieczeństwa. Przeanalizowaliśmy zawartość linii Media Description w wiadomościach SDP, które definiują typ media, port, protokół i format, co jest niezbędne do prawidłowego ustawienia sesji. Zbadaliśmy również, jaki kodek mowy został użyty w nawiązanej sesji, co ma znaczenie dla jakości i wydajności transmisji głosu. Dodatkowo, przeanalizowaliśmy pola pakietów RTP, zwracając uwagę na pole Mark, które jest używane do oznaczania pierwszego pakietu w strumieniu danych, co jest niezbędne dla prawidłowej synchronizacji odtwarzania. Zajęliśmy się również analizą przekazywania informacji o wciśniętym klawiszu numerycznym (kodzie tonowym DTMF) do serwera, co jest kluczowe dla przesyłania sygnałów sterujących. Na końcu, przeanalizowaliśmy pierwszy pakiet RTP przesłany przez serwer, koncentrując się na numerze kolejnym i znaczniku czasu, które nie zaczynają się od 0. Ten fakt wynika z zastosowania losowych wartości początkowych w celu zwiększenia bezpieczeństwa transmisji. Te same aspekty zostały przeanalizowane również w kontekście drugiej sesji, co pozwoliło nam zrozumieć, jak różne elementy protokołu SIP i RTP współpracują ze sobą, aby umożliwić komunikację głosową przez IP.