POLITECHNIKA WARSZAWSKA
Wydział Elektroniki i Technik Informacyjnych
Instytut Informatyki
Magisterskie Studia Uzupełniające – Informatyka

Rok akademicki 2003/2004

PRACA DYPLOMOWA MAGISTERSKA

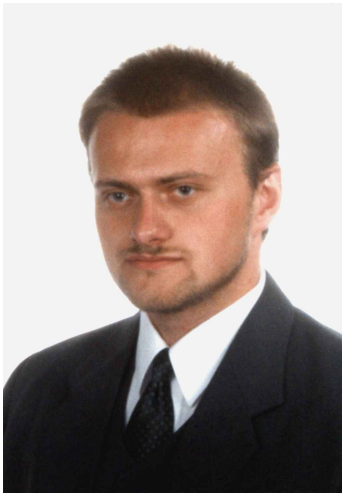
Marcin Stolarski

Problemy bezpieczeństwa w sieci Wi-Fi.

Opiekun pracy:
dr Jacek Wytrębowicz

Ocena.....

.....
Podpis Przewodniczącego
Komisji Egzaminu Dyplomowego



Kierunek: Informatyka

Imię i nazwisko: Marcin Stolarski

Data urodzenia 20.07.1976 r.

Data rozpoczęcia studiów: 1.10.2002 r.

Ż Y C I O R Y S

Urodziłem się 20 lipca 1976 roku w Warszawie. W latach 1983-1991 uczęszczałem do Szkoły Podstawowej Nr 164 im. Krajowej Rady Narodowej w Warszawie. W roku 1991 rozpocząłem naukę w III L.O. im. Gen. Sowińskiego w Warszawie w klasie o profilu matematyczno-fizycznym. W 1995 roku ukończyłem szkołę średnią zdając egzamin dojrzałości. W tym samym roku podjąłem naukę na dziennych studiach magisterskich na Politechnice Warszawskiej na wydziale Elektroniki i Technik Informacyjnych na kierunku Informatyka. W 1999 roku przerwałem studia dzienne i rozpocząłem naukę na Wieczorowych Studiach Zawodowych na Politechnice Warszawskiej na kierunku Informatyka oraz podjąłem pracę w firmie Poczta Polska jako Administrator systemów komputerowych. W roku 2002 ukończyłem studia zdobywając tytuł inżyniera, a następnie rozpocząłem studia magisterskie na tym samym kierunku.

.....

Podpis studenta

EGZAMIN DYPLOMOWY

Złożył egzamin dyplomowy w dn. .....

z wynikiem .....

Ogólny wynik studiów .....

Dodatkowe wnioski i uwagi Komisji .....

.....

.....

Streszczenie

Praca ta zajmie się szeroko rozumianym bezpieczeństwem sieci IEEE 802.11. Celem tej pracy jest analiza bezpieczeństwa sieci Wi-Fi oraz znalezienie metod wykrywania włamań do takiej sieci.

Pierwsza jej część dotyczy analizy mechanizmów bezpieczeństwa, jakie zastosowano w sieciach Wi-Fi. Wykazane jest, że bezpieczeństwo takiej sieci jest wysoce niewystarczające nawet w niekomercyjnych zastosowaniach. Podane są przykłady jak można włamać się do takich sieci oraz w jaki sposób można zwiększyć bezpieczeństwo, miedzy innymi za pomocą odpowiedniej polityki, wdrażania autorskich rozwiązań bądź za pomocą systemów IDS.

W drugiej części zaprezentowana jest próba nowatorskiego podejścia do tematu bezpieczeństwa w wyżej wymienionych sieciach, przy wykorzystaniu algorytmów eksploracji danych. Została przedstawiona aplikacja WiFi_Analysis do badania ruchu w sieci bezprzewodowej, napisana przez autora niniejszej pracy. Przy jej pomocy dokonano próby analizy na najniższych warstwach sieci, w celu wyszukania charakterystycznych zachowań sieci działającej prawidłowo oraz sieci, do której nastąpiło włamanie.

Podana jest interpretacja uzyskanych wyników. Niestety wynika z niej, że wyżej wymienione algorytmy nie pozwalają wykryć włamania do sieci. Pokazane są też zaobserwowane ciekawe zachowania sieci, na które dotychczas nie zwracano uwagi podczas badań.

Słowa kluczowe: Wi-Fi, 802.11, IDS, eksploracja danych, bezpieczeństwo sieciowe.

Security Issues at Wi-Fi networks.

Summary

The purpose of this work is to analyze the security of the Wi-Fi networks and to find new methods of intrusion detection.

The first part contains a detail analysis of security mechanisms built in Wi-Fi networks. It shows, that the security of such networks is extremely scarce even at commercial deployments. There are given some examples how one can break in such a network. Next I describe the rules one can use to improve security of these networks.

The second part of this work describes an application I have written for analysis of frame traffic in a Wi-Fi network. The application processes gathered data from MAC layer in order to find a characteristic behavior of the traffic using a data mining approach.

Finally I give interpretation of the results. Some interesting and not obvious behaviors of the frames traffic has been discovered. Unfortunately the selected approach do not lead to intrusion detection in a network.

Keywords: Wi-Fi, 802.11, IDS, data mining, network security.

Spis Treści

Wstęp

Lokalne sieci bezprzewodowe stają się coraz popularniejsze. Na półkach sklepowych bez problemu można znaleźć dość tanie urządzenia przeznaczone do użytku domowego, które łatwo się konfiguruje za pomocą przeglądarki WWW. Większość z nich wystarczy jedynie podłączyć by uruchomić sieć. Niestety powszechne jest nieautoryzowane dołączanie się do tych sieci.

Praca ta zajmie się szeroko rozumianym bezpieczeństwem w sieciach Wi-Fi. Jej celem jest analiza bezpieczeństwa sieci Wi-Fi, oraz próba zastosowania metod eksploracji danych do wykrywania włamań w takich sieciach.

Pierwsza część pracy dotyczy analizy mechanizmów bezpieczeństwa, jakie zastosowano w sieciach Wi-Fi. Wykazane jest, że bezpieczeństwo takiej sieci jest wysoce niewystarczające nawet do nie komercyjnych zastosowań. Podane są przykłady jak można włamać się do takich sieci oraz w jaki sposób można zwiększyć bezpieczeństwo, miedzy innymi za pomocą odpowiedniej polityki, wdrażania autorskich rozwiązań, bądź za pomocą systemów IDS.

W drugiej części jest pokazane nowatorskie podejście do tematu bezpieczeństwa przy wykorzystaniu algorytmów eksploracji danych. Została zaprezentowana, napisana przez autora niniejszej pracy, aplikacja WiFi_Analysis do badania ruchu w sieci bezprzewodowej. Za jej pomocą dokonano próby analizy ramek w warstwie MAC sieci. Poszukiwane były charakterystyczne zachowania sieci działającej prawidłowo oraz sieci, do której nastąpiło włamanie.

Dokonana interpretacja wyników, pokazuje, że wyżej zastosowane algorytmy nie pozwalają wykryć włamania do sieci. Pokazane są też zaobserwowane ciekawe zachowania sieci, na które dotychczas nie zwracano uwagi podczas badań.

Analiza problemu

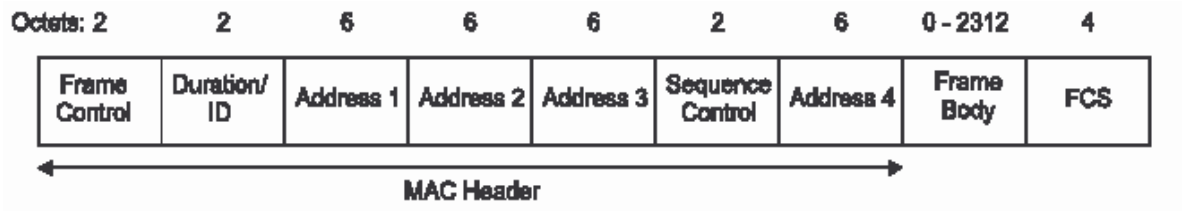
Gdy powstawały sieci WLAN, nie miały one dużej popularności ze względu na wolną transmisję oraz wysokie ceny urządzeń. Twórcy standardu zdefiniowali jedynie słabe mechanizmy autoryzacji oraz szyfrowania, aby nie podrażać i tak drogich urządzeń. Wraz ze standardem 802.11b pojawiła się większa prędkość transmisji, co wpłynęło na gwałtowne zainteresowanie się tymi sieciami przez rynek, co z kolei spowodowało gwałtowny spadek cen urządzeń. Kiedy sprzęt trafił pod strzechy szybko okazało się, że sieci Wi-Fi są niebezpieczne oraz że istnieją problemy w stosowaniu ich w rozwiązaniach biznesowych. Niestety na kolejne unormowania trzeba było poczekać.

Opis protokółu 802.11

Sieci WLAN pojawiły się już w latach 80-tych. Brak jednolitego standardu powodował wzajemną niekompatybilność rozwiązań różnych firm. W roku 1997 pojawił się pierwszy standard w postaci normy 802.11 [5]. Zakładała ona transmisję ramek z prędkościami 1 i 2 Mbps na wolnej częstotliwości 2.4 GHz. Jako mechanizmy bezpieczeństwa stosowano identyfikator SSID oraz szyfrowanie kluczem WEP. Z czasem opracowano poprawki do normy w postaci norm 802.11a oraz 802.11b, które pozwoliły na większe prędkości (54Mbps dla standardu „a", 11Mbps dla standardu „b") oraz na wykorzystanie innych częstotliwości (5.6 GHz dla standardu „a"). Potem pojawił się standard 802.11g, który dzięki zastosowaniu modulacji kwadraturowej ze standardu „a" umożliwił osiągnięcie prędkości 54Mb na częstotliwości 2.4GHz pozostając kompatybilnym ze standardem „b". Jak widać platforma sprzętowa rozwijała się, czego nie można powiedzieć o bezpieczeństwie. W roku 2000 pojawiły się pierwsze prace i artykuły mówiące o problemach i słabych punktach protokołu. W 2001 powstał ruch zwany „)(Warchalking" zajmujący się włamaniami do sieci radiowych. Mimo wyraźnych sygnałów ze świata [1] nadal nic nie zmieniono w kwestii bezpieczeństwa. Jedyna ochroną nadal są mechanizmy: SSID oraz WEP. W sierpniu roku 2004 pojawił się standardu 802.11i, w którym główny nacisk kładzie się na sprawy bezpieczeństwa. W miedzy czasie pojawiły się rozwiązania firmowe takie jak np. CISCO TKIP.

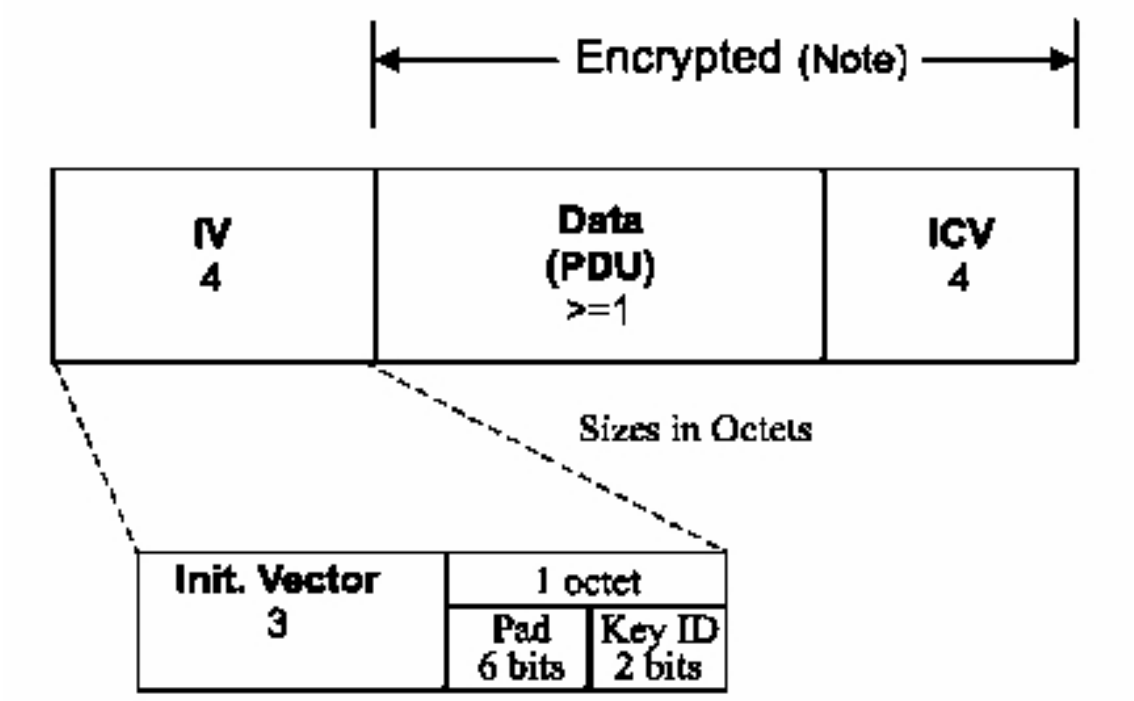
Budowa ramki

Na rysunku 1 przedstawiono budowę ramki protokołu 802.11. Składa się ona z nagłówka, typu, adresu MAC, numeru ramki, pola z danymi oraz sumy kontrolnej.



Rysunek 1. Budowa ramki 802.11.

W przypadku, kiedy zostanie włączone kodowanie WEP, pole z danymi jest podzielone na 3 sekcje (Rysunek 2). Pierwsza to nr klucza RC4, druga to pole z danymi, trzecia to suma kontrolna pola danych. Części druga i trzecia są kodowane algorytmem WEP.



Rysunek 2. Blok ramki 802.11 szyfrowany algorytmem WEP.

Bezpieczeństwo 802.11

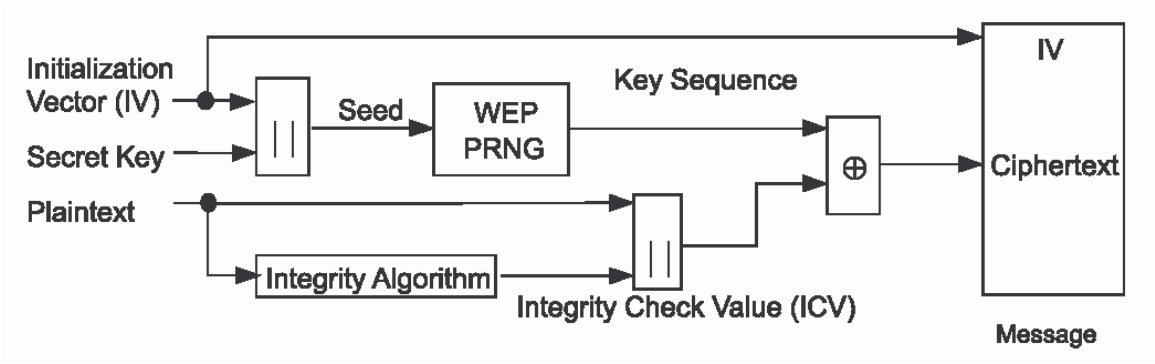
W ramach protokołu 802.11 stworzono kilka mechanizmów ochrony sieci: autentyfikację, SSID, WEP oraz MAC.

Podczas pierwszego połączenia stacji klienckiej najpierw musi nastąpić autoryzacja stron. Stosowane są dwa algorytmy. Typ 0 (open system - domyślny) polega jedynie na wymianie informacji kto z kim ma się połączyć i następuje uwierzytelnienie. W algorytmie typu 1 (shared-key) autoryzacja jest wykonywana przy wykorzystaniu kluczy WEP, tak więc obie strony muszą mieć taki sam klucz.

Kolejnym mechanizmem ochrony jest identyfikator SSID. Mechanizm ten początkowo był wykorzystywany do wydzielenia oddzielnego VLAN-u w sieciach Wi-Fi, aby uniemożliwić innym dostęp do danej podsieci. Każde urządzenie posiada identyczne hasło do sieci (SSID).

Aby mogło transmitować pakiety, musi do pakietu dołączyć SSID jawnym tekstem, co powoduje wydzielenie podsieci. Niestety, jeśli zostanie uruchomione oprogramowanie nasłuchujące, to identyfikator SSID zostanie bardzo szybko znaleziony. Dodatkowo w Acccess Point-ach jest usługa wysyłania ramek rozgłoszeniowych (broadcast) z identyfikatorem SSID, aby inne urządzenia mogły pobrać sobie identyfikator, w celu przyłączenia do sieci. Jak widać, jeśli medium jest powietrze, to każdy praktycznie może nasłuchiwać i wysyłać pakiety w pobliżu naszej sieci. Oznacza to, że może bez problemów uzyskać identyfikator SSID i dołączyć się do naszej sieci.

Mechanizm WEP służy do szyfrowania danych w pakietach sieci 802.11. Aby zaszyfrować dane (Rysunek 3 oraz 4), wykonujemy operacje XOR na danych (wraz z sumą kontrolną CRC) oraz na ciągu szyfrującym.

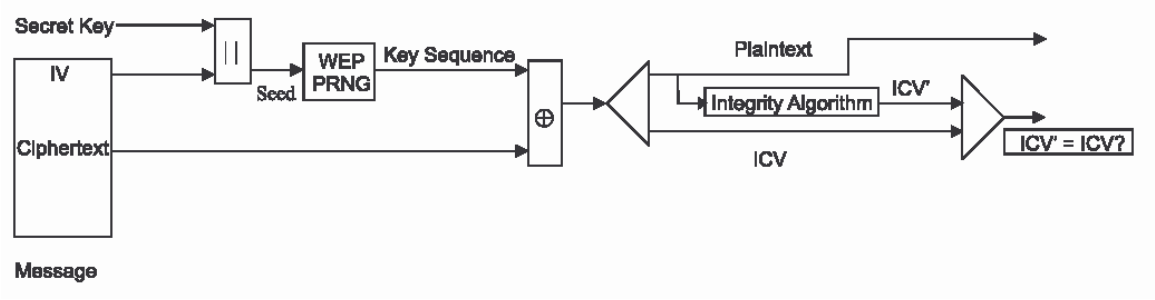


Rysunek 3. Szyfrowanie WEP.

	Blok danych	CRC
XOR	Ciąg szyfrujący RC4(V, K)	
Wektor inicjujący	Blok zaszyfrowany	

Rysunek 4. Szyfrowanie WEP.

Ciąg szyfrujący generowany jest przy wykorzystaniu algorytmu RC4, który na podstawie zmiennego 24b wektora inicjującego (V) i stałego klucza (K, 40b dla WEP 64b) generuje funkcja jednokierunkową pseudolosowy ciąg szyfrujący. Tak zaszyfrowany blok danych jest przesyłany wraz z wektorem inicjującym. Jeśli strona odbiorcza posiada identyczny klucz K (Rysunek 5 oraz 6), to na podstawie przysłanego wektora V jest w stanie wygenerować ciąg szyfrujący i po wykonaniu operacji XOR na ciągu szyfrującym oraz zaszyfrowanych danych otrzymuje oryginalne dane wraz z sumą kontrolną.



Rysunek 5. Odszyfrowanie WEP.

Wektor inicjujący	Blok zaszyfrowany	
XOR	Ciąg szyfrujący RC4(V, K)	
	Blok danych	CRC

Rysunek 6. Odszyfrowanie WEP.

Matematycznie można to przedstawić za pomocą równań:

P XOR RC4(V,K)=C
C XOR RC4(V,K)=P

Wektor V po każdej operacji jest zmieniany (zwykle inkrementowany) i ma długość 24b niezależnie od długości klucza K. Daje to 2^24 (16.777.216) kombinacji ciągu szyfrującego. Zgodnie z rachunkiem poniżej oznacza to, że dla sieci 11Mbps co około 5 godzin powtarza się ciąg szyfrujący.

11Mbps / (1500B [na pakiet] * 8b [na B])=91667 [pakietów na sekundę]

16777216 [kombinacji V]/ 91667 [pakietów na sekundę]=18302,41745 [sek. bez powtórki V]

18302,41745[sek bez powtórki V] / (60[sek na min] * 60[min na godz])=5,0840048[godz bez powtórki V]

Poniższe działania matematyczne ukazują nam słabość WEP-a, a mianowicie:

P XOR RC4(V,K)=C
C XOR RC4(V,K)=P

C1 = P1 XOR RC4(V, K)
C2 = P2 XOR RC4(V, K)
C1 XOR C2 = (P1 XOR RC4(V, K)) XOR (P2 XOR RC4(V, K)) = P1 XOR P2
(P2 XOR RC4(V, K)) = P1 XOR P2

Oznacza to, że jeśli zostaną przechwycone dwie ramki z identycznym ciągiem szyfrującym (wiemy to na podstawie V) i jest znana zawartość danych jednej z nich, to można bez większych problemów wyliczyć zawartość danych drugiej ramki. Do tego dochodzi fakt, że w polu danych przenoszone są ramki innych protokołów, które mają ściśle określone wartości na określonych polach.

Kolejnym problemem algorytmu WEP są tak zwane słabe klucze, które pozwalają złamać klucz główny WEP. Np. przy kodowaniu 64b, aby złamać klucz potrzeba przechwycić około 15 słabych kluczy z puli 1280 (1280 to 0.008% wszystkich kluczy). Na uwagę zasługuje fakt, że pula słabych kluczy rośnie wraz ze wzrostem długości klucza głównego i tak np. przy kluczu 128b pula ta stanowi 0.020% wszystkich kluczy.

Ostatnim mechanizmem ochrony jest filtrowanie adresów MAC kart sieciowych. Adresy dozwolone bądź zabronione można wpisywać na listy, nie mniej, ponieważ listy te nie są zbyt duże, nie pozwalają na budowę dużych mobilnych sieci. Poza tym adres MAC można zmienić w karcie za pomocą specjalnego oprogramowania, co pozwala się podszyć pod autoryzowane urządzenie.

Atak 802.11

Widząc słabość mechanizmów ochrony w sieciach bezprzewodowych, można pokusić się po parę scenariuszy ataku na taką sieć.

Jeśli sieć ma ustawiony tryb autoryzacji na 0, to wystarczy zostawić u siebie puste pole SSID i domyślny protokół autoryzacji wypełni to pole dając nam dostęp do sieci. Gdy w sieci jest uruchomiony serwer DHCP, to w zasadzie samo oprogramowanie karty radiowej podłącza nas do takiej sieci. Jeśli będzie ustawione wymaganie na podanie identyfikatora SSID wystarczy go podsłuchać od innych użytkowników sieci, ponieważ jest nadawany jawnym tekstem.

Jeśli sieć jest zabezpieczona kodowaniem WEP, możemy do sprawy podejść na wiele sposobów.

Pierwszy najprostszy to atak brutalny na klucz. Łapiemy ramkę, a następnie próbujemy ja zdekodować kolejno generowanymi kluczami. Jeśli zgadza się zakodowana suma kontrolna, to na kolejnych ramkach sprawdzamy czy tez się uda sprawdzić sumę i jeśli nam się to udaje to mamy klucz. Gdy następne ramki się nie dekodują, to szukamy dalej klucza aż do skutku. Na maszynie typu P4 2GHz dla klucza 40b taka operacja zajmie około 1 roku. Jeśli jednak klucz WEP zostanie wpisany nie jako fraza zapisana w postaci liczb o podstawie 16 (HEX), ale jako fraza ASCII to jego przedział zostanie zmniejszony do 2^21, co łamie się w przeciągu 10 sekund. Do tego, jeśli w ASCII jest napisana jakaś fraza słowna (np. nazwa firmy) możemy posiłkować się słownikiem (acz atak słownikowy należałoby stosować do dłuższych kluczy).

Kolejna formą ataku na WEP jest wpuszczenie w sieć znanego tekstu (np. poczty elektronicznej) i zebraniu zestawu wszystkich kluczy sesyjnych do dekodowania innych ramek. Metodę te można rozwijać przez poszukiwanie w szyfrogramach znanych fragmentów

protokołów komunikacyjnych takich jak http i łamania kolejnych kawałków kluczy sesyjnych. Wymaga to jednak posiadania pojemnych twardych dysków. Jeden zestaw kluczy to ok. 20MB.

Sieć bezprzewodową można też atakować wysyłając specjalnie spreparowane ramki. Metoda polega na przechwyceniu zakodowanej ramki, ingerencji w jej treść tak by np. zmienić adres jej przeznaczenia np. na własny komputer podłączony do Internetu, tak manipulując danymi, aby zgadzała się zakodowana suma CRC i jeśli zostanie odebrana ze strony internetu zdekodowana ramka, można sobie wyliczyć określony klucz sesyjny.

Kolejnym miejscem, w które można uderzyć to filtrowanie adresów MAC, przez podsłuchanie autoryzowanych adresów, a następnie podmienienie takiego adresu we własnej karcie radiowej.

Ostatnim możliwym atakiem na sieć radiową jest atak DoS, który może przybrać formę od zaawansowanej wpuszczając w sieć wadliwe pakiety lub blokując sieć przez nasycenie jej ramkami wysyłanymi ciągłym strumieniem z własnej karty, a kończąc na brutalnej, przesterowującej stopnie wejściowe odbiorników radiowych silnym polem radiowym pochodzącym np. z otwartej kuchenki mikrofalowej.

Poniżej pokazane są dwa przykładowe ataki na sieć 802.11 [1]. Pierwszy (Rysunek 7) pokazuje, jak możemy sprawdzać zawartość ramek nie posiadając klucza, a drugi (Rysunek 8) jak można manipulować zawartością ramki tak by zachować zgodność z sumą kontrolną CRC.

P „a”	01100001	P „b”	01100010
K „n”	01101110	K „n”	01101110
XOR ”a”	00001111	XOR ”b	00001100
XOR ”a”	00001111	P „a”	01100001
XOR ”b”	00001100	P „b”	01100010
XOR ”XaXb”	00000011	XOR ”PaPb”	00000011

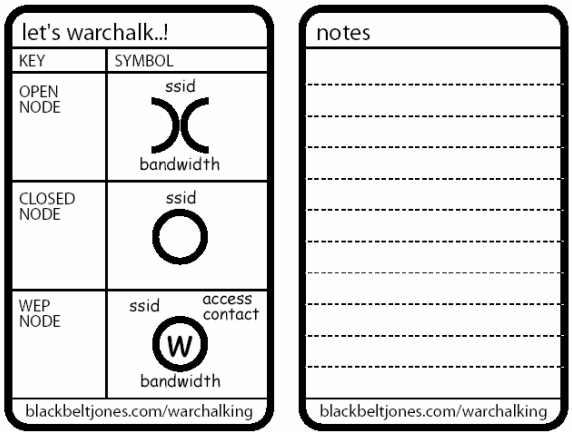
Rysunek 7. Atak na WEP. P-tekst jawny, K-klucz, XOR-operacja matematyczna XOR.

Sieć			Haker
	Data	CRC-8	
P „b”	01100001	00101001	
K „n”	01101110	01101110	
XOR ”b	00001100	01000111	
			Data CRC-8
			XOR ”b” 00001100 01000111
			Zmiana 00000011 00001001
			XOR XOR ”b”
			00001111 01001110
	Data	CRC-8	
XOR XOR ”b”			
	00001111	01001110	
K „n”	01101110	01101110	
Wynik ”a”			
	01100001	00100000	

Rysunek 8. Atak na CRC. P-tekst jawny, K-klucz, XOR-operacja matematyczna XOR, Zmiana-zamiana bitow przez hakera.

Atak w praktyce

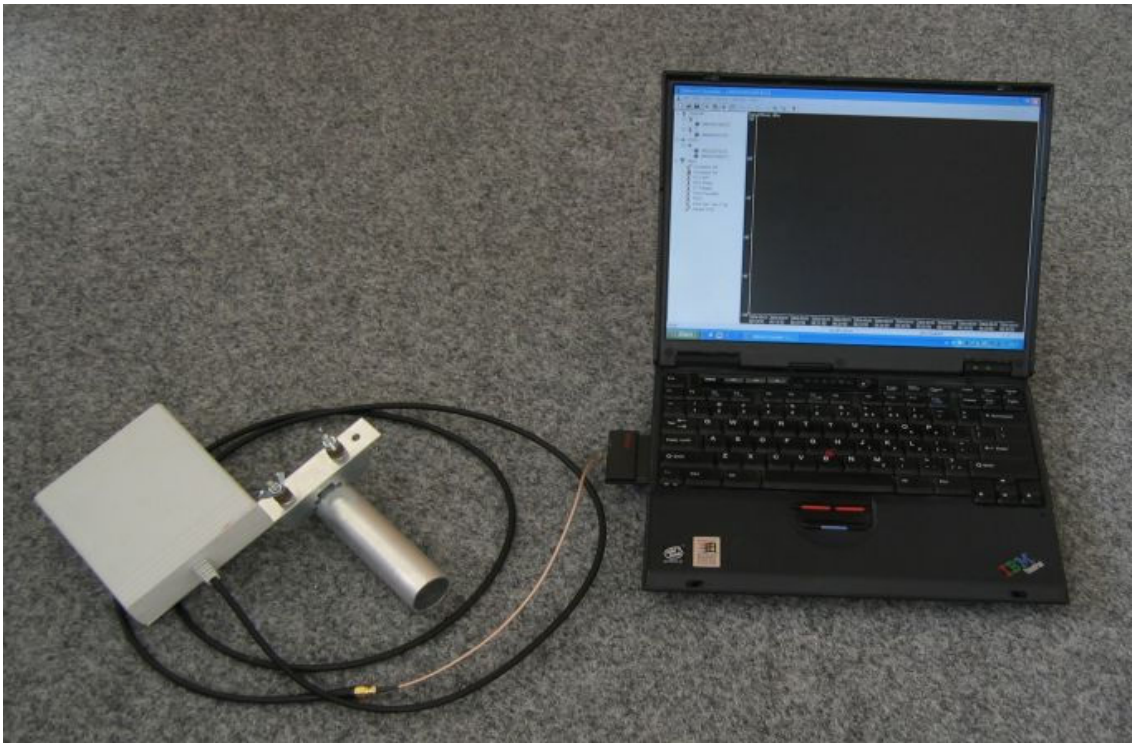
Przed atakiem na daną sieć można się rozejrzeć czy przypadkiem ktoś już tego nie zrobił wcześniej i nie zostawił stosownych informacji na murze czy też płocie posesji. Na rysunku 9 pokazana jest karta z notacją znaków.



Rysunek 9. Karta warchalk.

Aby przeprowadzić atak w praktyce należy zaopatrzyć się w (Rysunek 10):

- komputer (najlepiej przenośny)
- kartę radiową
- antenę
- przewód antenowy nisko stratny
- ewentualne przejściówki i złączki (karty radiowe mają różne złącza antenowe)
- odpowiednie oprogramowanie



Rysunek 10. Sprzęt do podsłuchu sieci Wi-Fi.

Przy wyborze anteny należy zastanowić się skąd i w jakich warunkach sieć będzie podsłuchiwana. Są do wyboru anteny dookólne oraz kierunkowe (YAGI, panele {Rysunek 11}, talerze).

Rysunek 11. Panelowa antena Wi-Fi.

Kolejną sprawą jest wybór karty radiowej. Pojawia się tutaj zagadnienie, które co prawda zanika, ale jest jeszcze obecne i generuje różne problemy. Chodzi o to, że w zależności od środowiska (Windows, Linux) programiści stosownych aplikacji preferują dwa różne chipsety kart. W Linux oprogramowanie było tworzone pod chipset PRISM2, a w Windows pod chipset HERMES. Stąd też wybierając środowisko pracy, należałoby pod system zakupić stosowną kartę. Poniżej w tabelach 1 oraz 2 są sprawdzone rozwiązania rynkowe w zależności od chipset-u.

Tabela 1. Karty Wi-Fi zbudowane na chipsecie HERMES.

Chipset HERMES
• ORINOCO (Lucent PC) Card (FC-WD-11Ch, ETS-EU-13Ch) • Dell TrueMobile 1150 • Avaya Wirless PC Card • Compaq WL110 • Enterasysy Roamabout • Elsa Airlancer MC-11 • Arterm CC-WL11 • IBM High Rate Wirles LAN • Buaffalo WLI-PCM-L11 • 1stWave 1ST-PC-DSS11IS, DSS11IG, DSS11ES, DSS11EG • D-Link DWL-660 • Compex WL11A+ • Compex WL11B+ (Hermes II)

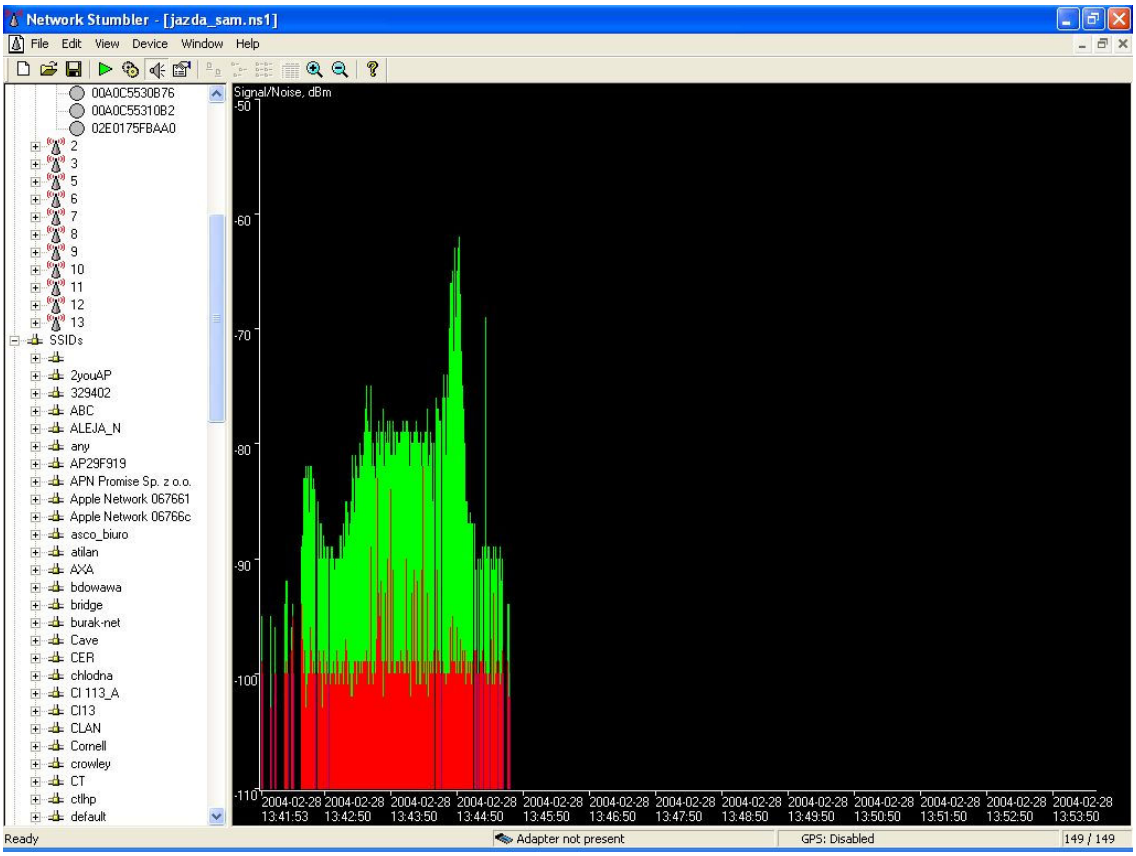
Tabela 2. KArty Wi-Fi zbudowane na chipsecie PRISM2.

Chipset PRISM2
• Addtron AWP-100 • Ambicom WL100B-PC • Bromax Freeport • Compaq WL100 • D-Link DWL-650 • GemTek WL-211 • Intalk/Nokia WL201 • Linksys WPC11 • Samsung SWL2000-N • SMC 2632W • Teletronic WL1000 • YDI Diamond • Z-Com XI300 • Zoom Telephonics ZoomAir 4100

Aktualnie popularne oprogramowanie posiada stosowne poprawki umożliwiające prace na obydwu kartach, nie mniej chipset hermes zasługuje na szczególną uwagę, ponieważ podczas skanowania sieci nasycenie pola podaje w decybelach, a nie w bliżej nieokreślonych jednostkach porównawczych. Liderem na rynku jest firma Orinoco (dawniej Avaya – Rysunek 12) i najbardziej polecaną kartą jest karta Orinoco Gold Clasic (Gold oznacza możliwość kodowania WEP do 128b, a Clasic możliwość pracy w Access Poincie firmy Orinoco). Charakteryzuje się bardzo dużą stabilnością oraz popularnością wśród programistów. Niestety jest też bardzo droga (ok. 350 zł) i dość trudno dostępna w wersji europejskiej (13 kanałów). Można się też posiłkować tak zwanym klonem Orinoco, czyli kartą, która elektronicznie jest kartą identyczną z Orinoco, a różni się jedynie frameware-m oraz sterownikami. Okazuje się , że klony często można uruchomić ze sterownikami Orinoco bądź Agree, które dają im pełną funkcjonalność Orinoco przy niższej cenie. (Np. Compex WL11B+ kosztuje ok. 130 zł)

Rysunek 12. Karta radiowa Wi-Fi.

Kolejna rzeczą, którą będziemy potrzebować jest oprogramowanie. Spośród wielu wyróżnić należy darmowe oprogramowanie pod Windows NetStumbler oraz pod Linux AirSnort oraz Wepcrack.

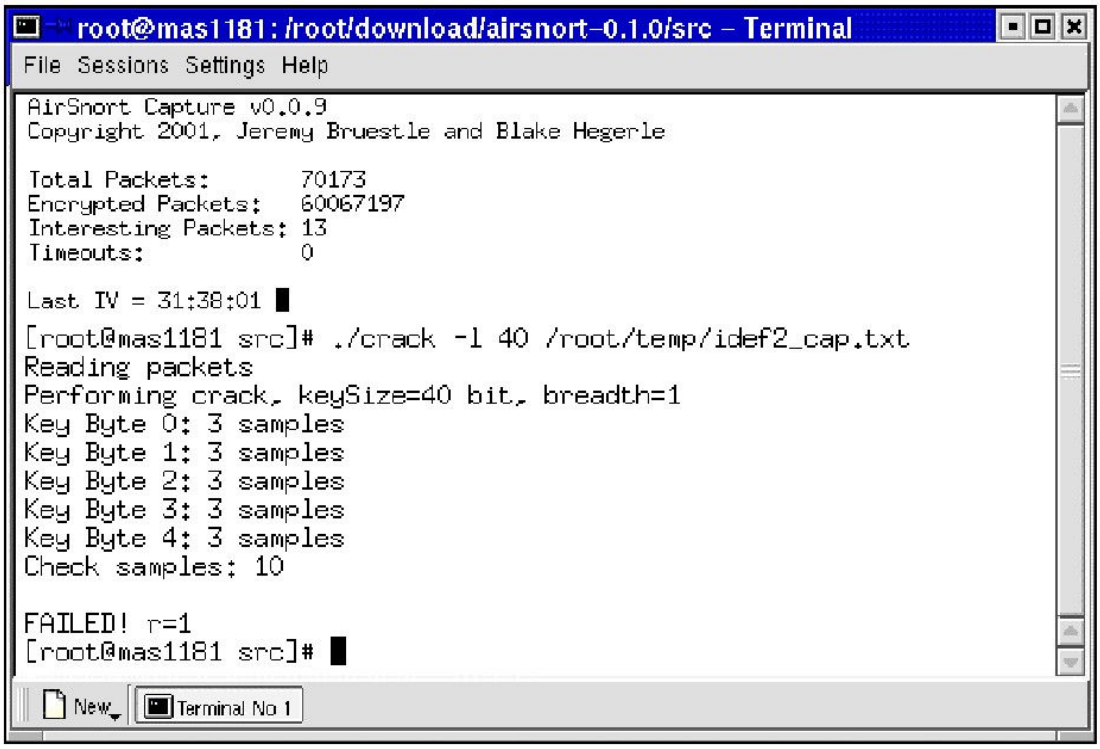


Rysunek 13. Aplikacja NetStumbler.

Program NetStumbler (Rysunek 13) należy zaliczyć do zaawansowanych skanerów. Pokazuje, na jakich kanałach są określone sieci (po SSID) i jakie urządzenia (po adresach MAC). Dodatkowo informuje, czy dana sieć jest kodowana algorytmem WEP. Skaner ten jest chyba najczęściej wykorzystywaną aplikacją przez instalatorów urządzeń, ponieważ pokazuje wykres stosunku sygnału (kolor zielony) do szumu (kolor czerwony), co pozwala określić jakość i stabilność linku. Dodatkowym atutem oprogramowania jest możliwość podłączenia do niego GPS-a, który podaje swoją dokładną pozycję geograficzną. Jeśli byśmy taki zestaw (laptop z NetStumblerem, kartą radiową, GPS-em i anteną samochodową) umieścili w samochodzie i przejechali się po mieście, dostajemy mapę lokalizacji sieci radiowych w danej okolicy. Jeśli sieć jest niechroniona (a tak jest w około 75% przypadków), pozostaje wpisanie stosowych parametrów do sterownika karty sieciowej.

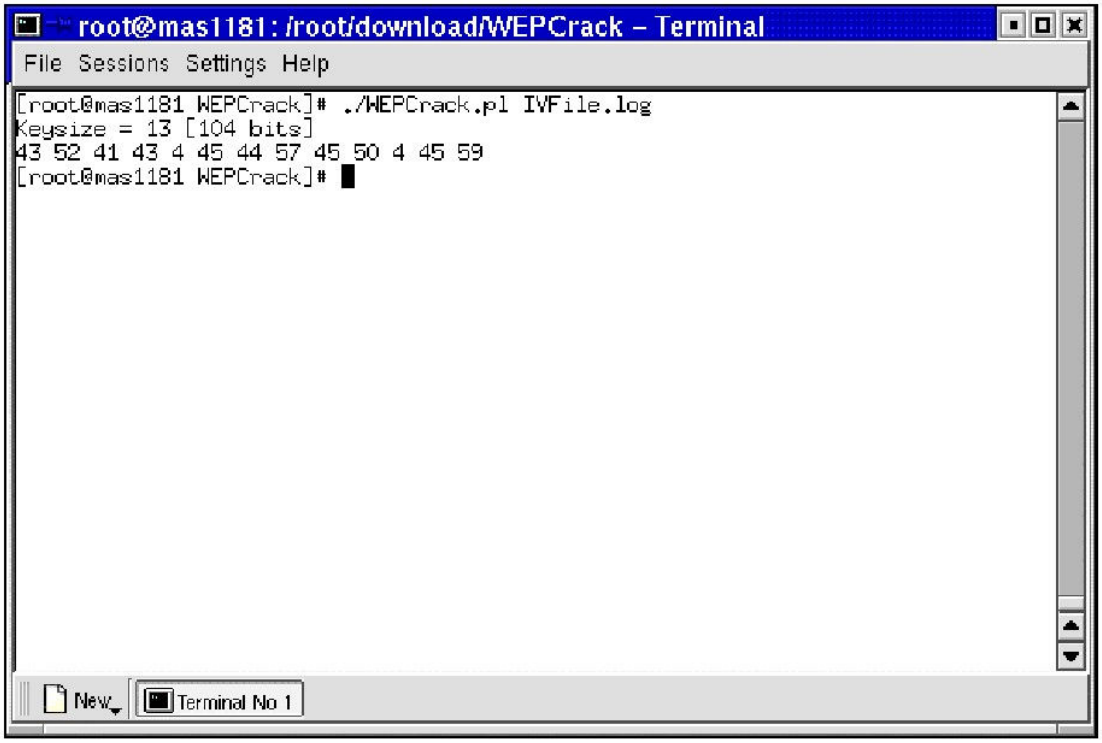
Jeśli jednak sieć jest chroniona kodowaniem WEP, możliwy jest atak statystyczny na słabe klucze. Taki atak może przeprowadzić program AirSnort bądź Wepcrack.

Program AirSnort (Rysunek 14) to w zasadzie aktywny skaner, który zbiera ramki 802.11, a następnie poszukując w nich słabych kluczy stara się złamać klucz WEP.



Rysunek 14. Aplikacja AirSnort.

Program Wepcrack (Rysunek 15) to skrypt w Perlu poszukujący klucza WEP na podstawie słabych kluczy. Aby można było z niego skorzystać należy najpierw zapisać na dysku stosowną ilość ruchu za pomocą snifera. Poniżej pokazane jest, jak Wepcrack złamał klucz WEP 128.



Rysunek 15. Aplikacja WEPCrack.

Jak widać, sieci oparte o standard 802.11 nie są bezpieczne. Aby się do nich włamać potrzebny jest niezbyt drogi sprzęt i odrobina wolnego czasu. Aby ta odrobina była jak najdłuższa, zamieszczono poniżej parę sugestii, którymi należy się kierować budując sieci bezprzewodowe. Pamiętać jednak należy o tym, że sieć radiową należy traktować jako słabo zabezpieczoną i jeśli zamierzamy w niej umieścić jakieś cenne dane, musimy posiłkować się dodatkowymi zabezpieczeniami takimi jak szyfrowane tunele SSH czy IPSec.

Przykazania administratora WLAN

Aby ułatwić budowanie bezpiecznych (w miarę ich możliwości) sieci Wi-Fi powstał zbór reguł, którymi należy się kierować.

- Nie należy bezgranicznie ufać bezpieczeństwu WEP-a (może zostać bardzo szybko złamany).
- Należy separować sieci Wi-Fi (stosując tunele VPN).
- Nie należy używać opisowych nazw SSID oraz urządzeń Wi-Fi (są one łamane za pomocą słownika).
- Jeśli to możliwe, to należy blokować MAC adresy (zabezpiecza to przed amatorami).
- Należy zmieniać klucze WEP możliwie często (może komuś już nie będzie się chciało po raz kolejny łamać WEP-a skoro 75% sieci w sąsiedztwie nie stosuje WEP-a).
- Należy wyłączyć bacon (aby nie rozsyłać zaproszenia innym by się włamali).
- Powinno się instalować AP centralnie (AP ma ograniczony zasięg, czym dalej jest od ogrodzenia tym trudniej jest za ogrodzeniem uzyskać użyteczny sygnał do podsłuchiwania sieci).
- Należy zmieniać domyślne hasła i nr IP (hakerzy maja listy z takimi hasłami i będą próbować zmienić konfigurację sieci).
- Należy unikać słabych kluczy WEP (powinno się wpisywać losowe kody HEX jako klucz, unikać zapisu ASCII, a w szczególności unikać jako klucza haseł słownikowych).
- Nie powinno się używać w WLAN serwera DHCP (jeśli jest dostępny tryb autoryzacji 0 i serwer DHCP, to każdy kto włączy komputer z kartą Wi-Fi zaloguje się automatycznie do sieci).
- Należy usuwać z sieci obce urządzenia Wi-Fi (nie pozwalać pracownikom uruchamiać własnych komponentów Wi-Fi. Zwykle uruchamiają je w konfiguracji domyślnej, która dla ułatwienia ma wyłączone wszystkie mechanizmy ochrony).
- Jeśli to możliwe należy włączyć dodatkowe zabezpieczenia producentów (zwykle oprogramowanie hakerskie nie potrafi sobie z nimi poradzić, a poza tym takie rozwiązania są skuteczne, gdyż sięgają do przyszłych standardów).

Przyszłość

Dzisiejszy standard 802.11 nie zapewnia mocnego bezpieczeństwa. Producenci na rynku stosują nowsze rozwiązania (niestety niekompatybilne miedzy różnymi producentami).

Największą bolączką WEP-a jest czas życia kluczy oraz ich dystrybucja. Ponieważ są to klucze statyczne i aby je zmienić administrator musi wpisywać ręcznie nowe klucze na wszystkich urządzeniach danej sieci, daje to czas na łamanie tych kluczy. Z pomocą przychodzi tu protokół TKIP, który do kodowania pakietów nie używa statycznego klucza WEP, ale krótkoterminowe klucze sesyjne, nie dając czasu na ich złamanie. Gdyby protokół TKIP stał się kolejnym rozszerzeniem standardu 802.11, sieci bezprzewodowe były by bardzo trudne do złamania. Dodatkowo taka zmiana nie pociągnęłaby potrzeby wymiany urządzeń, a jedynie ich oprogramowania.

Kolejną rzeczą, jaką się proponuje, jest wymiana kodowania WEP na AES. Niestety AES przez to, że jest znacznie bardziej obciążający obliczeniowo, wymaga wymiany sprzętu, co niestety w związku z faktem, że sporo tego typu sprzętu już jest na rynku, nie wróży mu dobrze na przyszłość.

Oba powyższe mechanizmy są zawarte w nowym standardzie bezprzewodowym 802.11i zwanym też WEP2. Ale standard ten został zatwierdzony dopiero w lipcu 2004 roku i upłynie trochę czasu zanim pojawią się urządzenia zgodne z tą normą. Tymczasem sieci Wi-Fi powstają teraz. Czołowi operatorzy telefonii komórkowej wprowadzają dostęp do internetu przy wykorzystaniu Wi-Fi i myślę, że ciężko będzie ten standard wprowadzać, ponieważ ze względu na kodowanie AES będzie potrzeba wymiana sprzętu. Ewentualna emulacja AES-a za pomocą oprogramowania będzie się wiązała z drastycznym obniżeniem wydajności sieci (kiedy kodowanie WEP nie było popularne i traktowano je jako opcja w AP to realizowano je programowo, co powodowało że przepustowość sieci spadała z 11Mb do 2Mb). Mówiąc krótko, standard ten pojawia się za późno i jest niekompatybilny ze starym sprzętem. Dodatkowo pojawiają się artykuły na podstawie wersji roboczej standardu, że tutaj też są problemy z bezpieczeństwem oraz Atakami DoS (proponuje się, aby reakcją AP na atak było jego tymczasowe milknięcie, jeśli atak będzie polegał na zalaniu AP błędnymi ramkami spowoduje to wstrzymanie pracy AP). Zastrzeżenia są też do zatwierdzonej normy, tj. pojawiły się wzmianki w sieci, że jeśli wykorzystujemy WEP-a bez serwera RADIUS-a to jest on podatny na atak brute-force. Zapotrzebowanie rynku na bezpieczne sieci radiowe jest jednak duże i widać już pierwsze ruchy firm w celu przynajmniej częściowego wdrożenia tego standardu, a mianowicie autoryzacji WPA, a w ramach niej szyfrowania TKIP. Pojawiła się nawet lista sprzętu zgodnego w WPA. Standard WPA składa się z mechanizmów 802.11x (omówiony w następnym akapicie), TKIP, EAP (nowy sposób na autoryzacji AP-APC) oraz MIC (nowy sposób sprawdzania integralności pakietów, ponieważ CRC32 nie jest bezpieczny). Jako ciekawostkę można podać, że wymienione mechanizmy już dość długo można znaleźć w urządzeniach CISCO.

Standard 802.11x zajmuje się transportem danych związanym z autoryzacją klientów w AP w oparciu o centralną bazę danych klientów opartą o serwer Radius-a. Jego uniwersalność

pozwala użyć go także do innych celów np. do transportu kluczy sesyjnych TKIP. Działanie tego mechanizmu autoryzacji w AP w oparciu o 802.11x jest proste:

- klient prosi o autoryzację i podaje swoją nazwę
- AP wysyła do klienta ciąg znaków i prosi o zakodowanie algorytmem jednokierunkowym przy wykorzystaniu klucza, jaki jest trzymany na serwerze Radius-a
- AP podobne zapytanie wysyła do serwera Radius-a dołączając nazwę klienta.
- kiedy AP otrzyma odpowiedzi z obu stron i są one identyczne, autoryzuje klienta

Standard 802.11x jest odpowiedzią na słabą autoryzację 802.11 oraz pozwala na centralne zarządzanie bazą danych o klientach. Standard ten jest bardzo chętnie wdrażany przez producentów i obecnie każde nowe urządzenie jest z nim zgodne.

Wnioski analizy

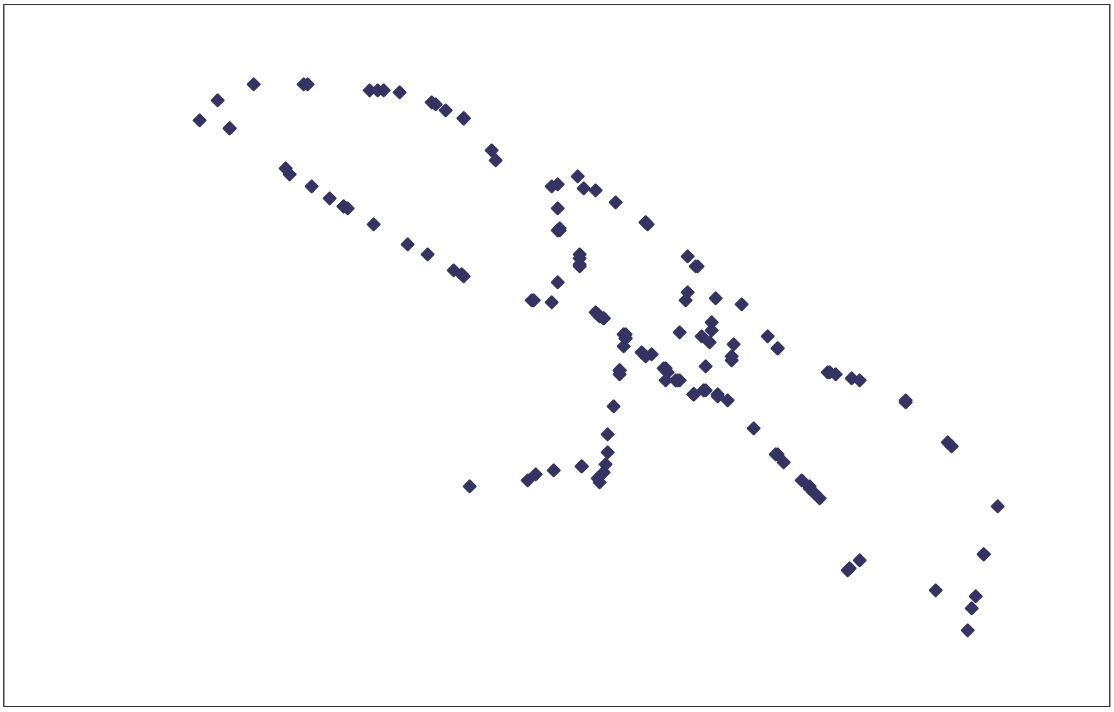
Standard 802.11 jest niebezpieczny. Trzeba pamiętać, że nawet kodowanie WEP nie zapewnia bezpieczeństwa, a jedynie szyfrowane tunele mogą rozwiązać ten problem. Można też posiłkować się rozwiązaniami autorskimi firm zapewniające odpowiedni poziom bezpieczeństwa, ale należy pamiętać, że o ile rozwiązania te wydają się skuteczne, to nie są kompatybilne z urządzeniami innych firm oraz, że nie zostały one gruntownie przetestowane na „światowym" poziomie. Ciekawie przedstawia się sprzęt zgodny z WPA (fragmentem standardu 802.11i w wersji draft). Można powiedzieć, że o ile w większości jest zainstalowany sprzęt zgodny z 802.11b, to na półkach sklepowych leży sprzęt zgodny z 802.11g i WPA. W ciągu roku powinien pojawić się sprzęt zgodny z WPA2. Należy też zwrócić uwagę, że nie można zostawiać w urządzeniach konfiguracji domyślnej, ponieważ każdy, kto w okolicy włączy komputer wyposażony w kartę Wi-Fi, automatycznie zaloguje się do naszej sieci jako pełnoprawny użytkownik. Należy o tym pamiętać bezwzględnie. Na stronach Warchalking-gu można znaleźć statystyki, że 75% użytkowników sieci Wi-Fi nie stosuje się do tego zalecenia. Poniżej umieszczono te statystyki oraz dodano autorską mapę centrum Warszawy wraz ze statystykami, która pokazuje, że administratorzy warszawskich sieci Wi-Fi popełniają podobne błędy.

Statystyki (Dane na rok 2002)

- Manhattan
  263 sieci z czego szyfrowanie WEP miało 65 sieci (25%)

- Północna Virginia
  114 sieci z czego szyfrowanie WEP miało 32 sieci (28%)

- Warszawa (mapa opracowana przez autora pracy {rysunek 16}, rok 2004)
  149 sieci – WEP 53 sieci (36%) – Ustawienia domyślne 18 sieci (12%)



Rysunek 16. Mapa urządzeń Wi-Fi w Warszawie (pomiędzy budynkiem Intrako {lewy górny róg} a stacją metra Wilanowska {prawy dolny róg}).

Projekt Programowy

Założenia

W „Analizie problemu" pokazano, że sieci Wi-Fi nie są bezpieczne. Odmienną techniką poprawiającą bezpieczeństwo (w znaczeniu wykrywania zagrożeń) w sieciach są systemy IDS. Systemy te zajmują się analizowaniem ruchu w sieci i alarmowaniem administratora, jeśli zachodzi podejrzenie o naruszenie bezpieczeństwa sieci. Systemy IDS dzielimy na dwa rodzaje. W pierwszym definiujemy standardowe zachowanie sieci. System zgłasza alarm, gdy wykryje odmienny ruch. Niestety takie rozwiązanie naraża administratora na fałszywe alarmy (np. spowodowane zainstalowaniem nowej aplikacji u użytkownika) i osłabia czujność administratora. Drugim rodzajem systemów IDS są systemy ze zdefiniowanymi wzorcami ataków. System taki podgląda zachowanie sieci i zgłasza alarm, kiedy ruch będzie odpowiadał jednemu ze znanych mu wzorców ataku. Ten rodzaj systemów obarczony jest problemem nieczułości na nowe ataki. Jak widać, żadne rozwiązanie nie jest pozbawione wad. Dalsza część tej pracy zajmuje się pewnym rozwinięciem pierwszej idei, a mianowicie analizą ruchu w sieci Wi-Fi na najniższych warstwach. Celem jest sprawdzenie, czy w warstwach tych można znaleźć pewne zachowania sieci działających prawidłowo oraz z naruszonym bezpieczeństwem (np. pojawieniem się stacji atakującej o identycznym adresie MAC jak jeden z legalnych adresów). Aby tego dokonać, została napisana aplikacja WiFi_Analysis (opisana w rozdziale „Opis interfejsu"), która pobiera dane ze skanera sieciowego AiroPeek (opisanego w rozdziale „Narzędzia i sprzęt"), a następnie generuje zestawienia różnych zależności sieciowych (np. zestawienie MAC adresów źródłowych i docelowych). Kolejnym krokiem jest próba znalezienia na takiej analizie grup obiektów, które by wyznaczyły granice bezpiecznej sieci. Do wyznaczenia tych granic został zastosowany algorytm DBScan (opisany w rozdziale „Analiza „Każdy z każdym" oraz DBScan"). Należy zaznaczyć, że są to prace nowatorskie, dlatego jest duża niepewność wyników.

Do tej pory nie spotkałem się z próbą podobnej analizy czasami wręcz niezależnych od siebie parametrów sieciowych. W Internecie można spotkać różnego rodzaju programy realizujące funkcje IDS dla sieci bezprzewodowych. Do najbardziej znanych można zaliczyć wtyczkę do programu „Snort" (aplikacji IDS) zwaną „Snort-wirles" [8]. Jest to rozwinięcie standardowej sondy IDS o sprawdzanie niektórych parametrów związanych bezpośrednio z siecią Wi-Fi, czy nie przekraczają określonych granic. Kolejnym przykładem jest moduł aplikacji „Kismet" do wykrywania intruzów w sieci [7]. Należy jednak zwócić uwagę, że aplikacja ta jest nakierowana na testowanie bezpieczeństwa sieci za pomocą różnych narzędzi służących do włamań sieciowych. Rozwiązania IDS można spotkać także w aplikacjach do zarządzania siecią. Za przykład może posłużyć aplikacja „CiscoWorks" [6] i jej moduł „Wirless LAN Solution Engine". Na rysunku 17 pokazana jest funkcjonalność wykrywania obcych (niebędących na liście uprawnionych) punktów dostepowych przy wykorzystaniu informacji z legalnych urządzeń. Aplikacje te nie wykonują szerokiego badania wielu parametrów sieci

bezprzewodowej w poszukiwaniu ich ukrytych zależności, jakie wykonuje aplikacja WiFi_Analysis, co powoduje, że otrzymujemy nowe, dotąd niebadane dane.

Rysunek 17. CiscoWorks Wireless LAN Solution Engine.

Plan pracy nad projektem.

- Analiza zagadnienia bezpieczeństwa sieci i metod włamań. Opis zawarty jest w rozdziale pt: „Analizie problemu".
- Przygotowanie narzędzi do zbierania informacji o sieci (aplikacja-skaner AiroPeek, z której pobrano dane do dalszej obróbki)
- Przygotowanie narzędzia do badania zależności w sieciach Wi-Fi. Jest to autorska aplikacja WiFi_Analysis.
- Przeprowadzenie nasłuchów sieci, wyszukanie stałych i zmiennych zależności sieci Wi-Fi.
- Interpretacja wyników (zawarta w rozdziale wyniki działania programu).

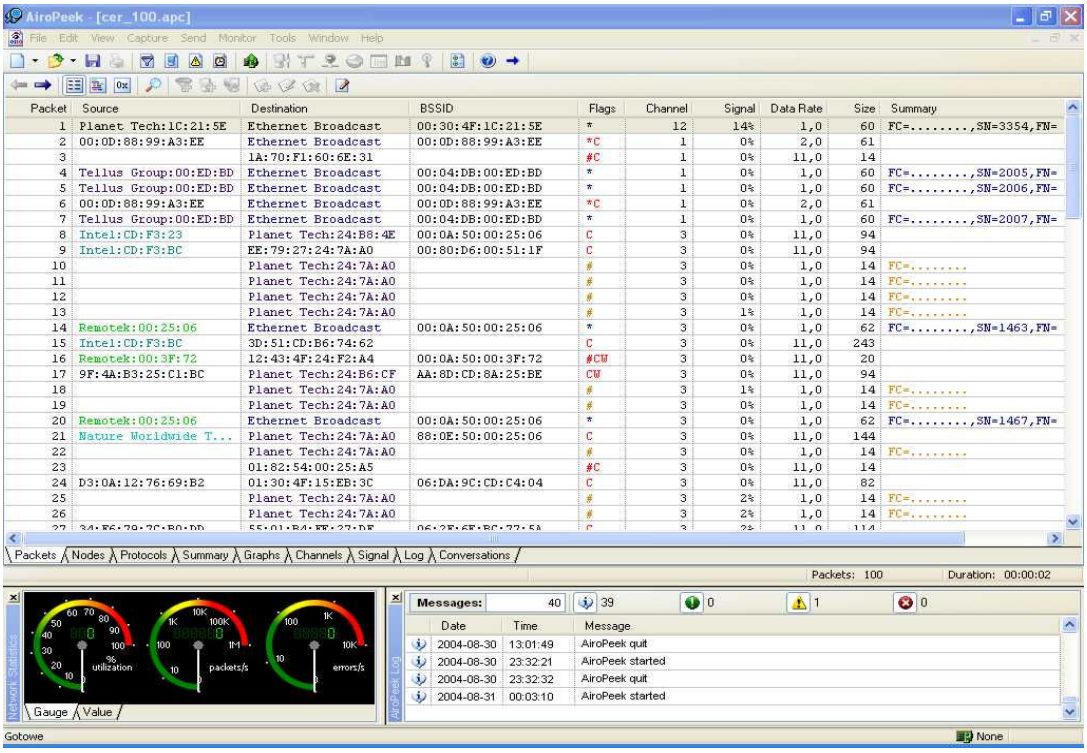
Jak widać podczas prac nie powstała sonda IDS, która by pozwalała na aktywną prace w sieci. Uważam, że aby taka sonda mogła powstać, potrzebne są najpierw prace badawcze, aby sprawdzić, czy proponowane przeze mnie rozwiązania mają sens. Analizując wyniki pracy wydaje się, że dynamika sieci Wi-Fi jest za duża, aby było możliwe znalezienie wzorców pozwalających skutecznie wykrywać ataki.

Narzędzia i sprzęt

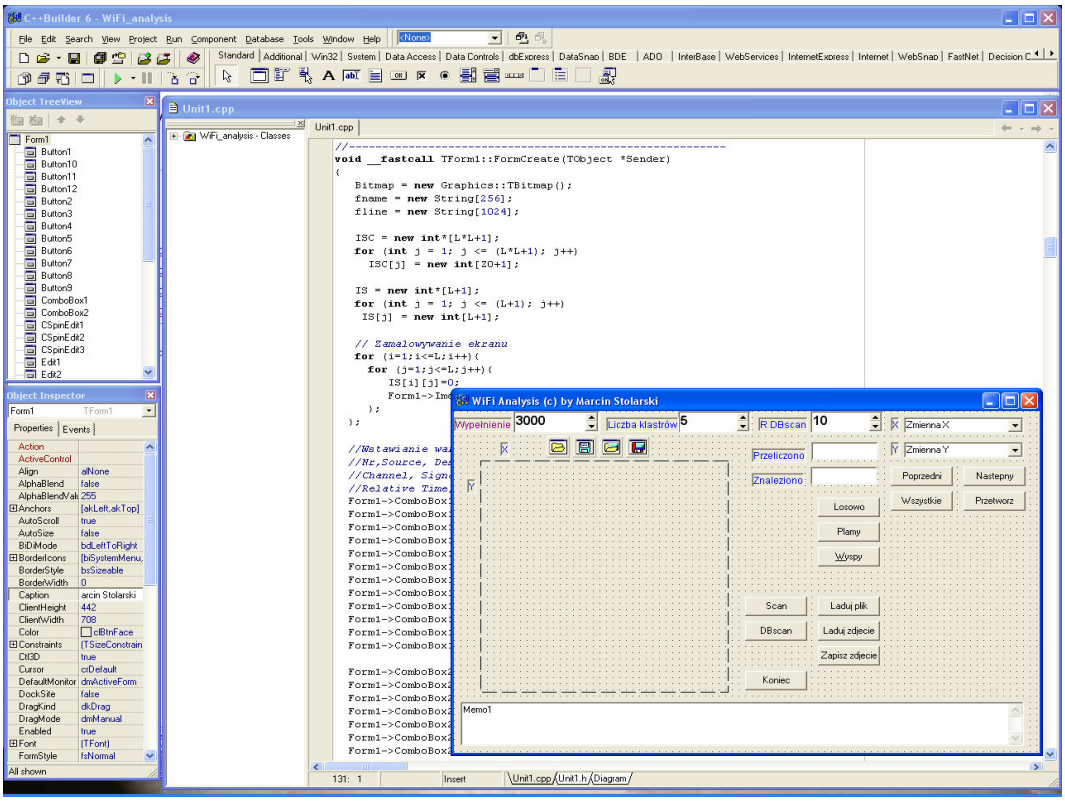
Do przeprowadzenia badań został wykorzystany następujący sprzęt:

- Laptop PIII 1GHz 60GB HD
- Karta Wi-Fi Avaya Gold (Orinoco)
- Panelowa antena kierunkowa 11.5 dBi
- Kable i złączki połączeniowe
- Skaner sieciowy AiroPeek
- C++ Builder 6.0 (do napisania aplikacji)
- Aplikacja WiFi_Analysis
- Serwery obliczeniowe (pięć maszyn klasy P4 2.7GHz)

Przy wyborze komputera kierowałem się pojemnością dysku, ponieważ nasłuch sieci zajmuje sporo miejsca. Karta radiowa, jak wcześniej napisałem, stanowi pewnego rodzaju standard w tego rodzaju pracach. Antena kierunkowa jest potrzebna, aby można było obserwować większy kawałek sieci. Wykorzystuje się też anteny dookolne, jeśli sieć stanowi bliskie otoczenie, bądź robi się nasłuch okolicy z jadącego samochodu (w naszym przypadku nie polecane ze względu na zmiany parametrów sieci związane z naszym ruchem). Skaner sieciowy AiroPeek (Rysunek 18) jest dość uniwersalnym narzędziem do nasłuchiwania sieci. Aplikacja ta, poza samym nasłuchem, interpretuje ramki pokazując siłę sygnału, treść ramek, adresy skąd i dokąd jest kierowana itp. Pozwala tez na eksport danych do plików tekstowych, które następnie są pobierane przez aplikacje WiFi_Analysis.



Rysunek 18. Aplikacja AiroPeek.



Rysunek 19. Aplikacja Borland C++ Builder 6.0.

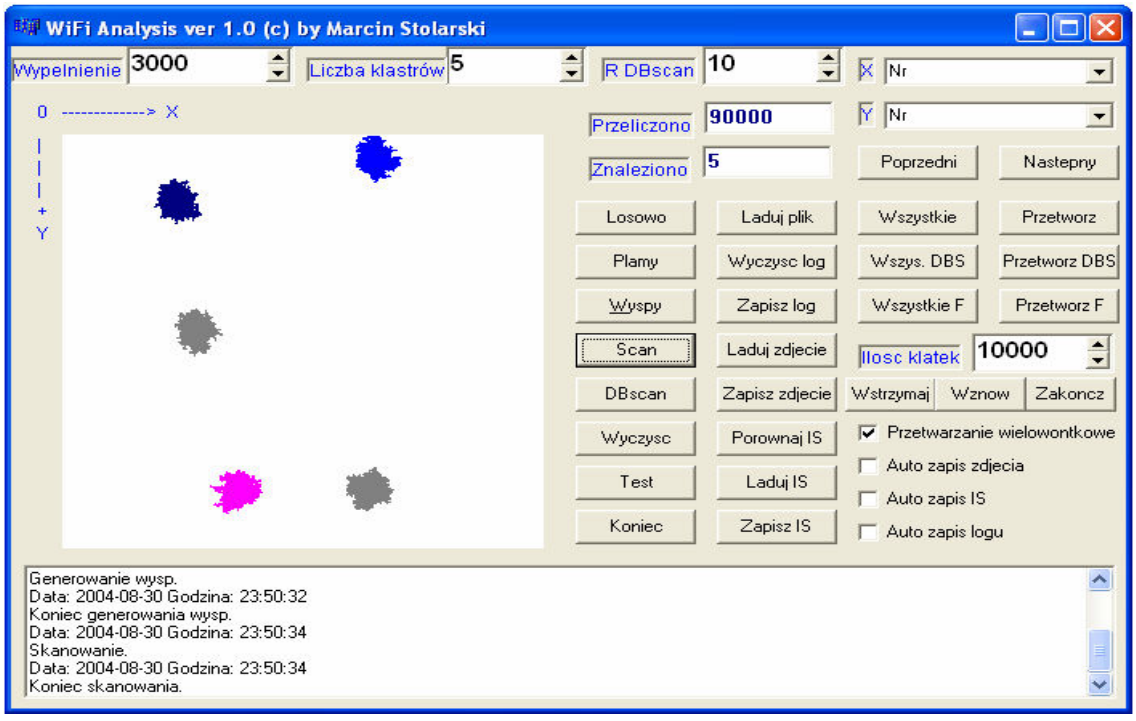
Borland C++ Builder 6.0 (Rysunek 19, [2]) jest przyjemnym środowiskiem programistycznym. Pozwala w łatwy sposób tworzyć kod oraz interfejs aplikacji. Do

wykorzystania są biblioteki MS VC++, autorskie biblioteki firmy Borland, oraz komponenty innych twórców.

Opis implementacji oraz interfejsu użytkownika

Aplikacja WiFi_Analysis (Rysunek 20) została napisana w środowisku Borland C++ Builder 6.0. Ponieważ autor wcześniej nie pisał aplikacji w tym środowisku, jako bazę do aplikacji wybrał przykład z książki [3], który generował obrazy zwane dalej wyspami. Przykład ten wspierał też pracę wielowątkową.

Do przetwarzania danych aplikacja pobiera dane z pliku, bądź generuje własne dane przykładowe. Takie dane przetwarzane są w tablicy IS, która jest matrycą 300x300 liczb całkowitych. Po przetworzeniu dane są wyświetlane w komponencie Imane. Zapis danych możliwy jest poprzez export z komponentu Imane do plików *.jpg, bądź przez eksport Rysunek tablicy IS do plików *.txt. Aplikacja może przetwarzać dane jedno bądź wielowątkowo. W trybie pracy jednowątkowej osiąga się największą prędkość przetwarzania danych, ale podczas przetwarzania aplikacja przestaje reagować na jakiekolwiek bodźce (w aplikacjach BC++B aby możliwe było odświeżanie wyglądu, funkcja odświeżająca musi być uruchomiona jako oddzielny wątek). W trybie wielowątkowym aplikacja cały czas jest aktywna. Można też wstrzymywać obliczenia (aby np. odciążyć maszynę w danej chwili), a następnie wznowić obliczenia. Podczas pracy z aplikacją generowany jest log, który pozwala zaznajomić się z postępem prac (analiza może trwać wiele godzin).



Rysunek 20. Aplikacja WiFi_Analysis.

Ekran aplikacji podzielony jest na cztery zasadnicze części: widok badania, przyciski, pola sterujące oraz log. Poniżej przedstawiam opis poszczególnych pól oraz przycisków.

- Wypełnienie. Ilość punktów do wygenerowania dla funkcji Losowo, Plamy, oraz Wyspy.
- Liczba klastrów . Ilość wysp dla funkcji Wyspy.
- R DBscan. Promień skanowania przez funkcje DBScan.
- Przeliczono (dana wyliczana). Ilość przetworzonych punktów.
- Znaleziono (dana wyliczana). Ilość znalezionych grup przy wykorzystaniu funkcji DBScan.
- X. Rodzaj danych na osi X (przy przetwarzaniu plików).
- Y. Rodzaj danych na osi Y (przy przetwarzaniu plików).
- Ilość klatek. Ilość punktów na pojedyńczą klatkę podczas przetwarzania filmowego.
- Przetwarzanie wielowątkowe. Znacznik pozwalający wybrać tryb pracy aplikacji (wielowątkowo/jednowątkowo). Wileowątkowo mogą być przetwarzane tylko funkcje: DBscan, Przetwórz, Wszystkie, Przetwórz DBS, Wszystkie DBS, Przetwórz F, Wszystkie DBS.
- Auto zapis zdjęcia. Podczas przetwarzania plików generowane są automatycznie pliki z wynikami w postaci obrazów *.jpg.
- Auto zapis IS. Podczas przetwarzania plików generowane są automatycznie pliki z eksportu z tablicy IS w postaci *.txt.
- Auto zapis logu. Podczas przetwarzania plików zapisywany jest automatycznie plik *.log.
- Losowo. Generowane są dane w postaci rozkładu losowego.
- Plamy. Generowane są dane w postaci plam składających się z losowo wybranych punktów.
- Wyspy. Generowane są dane w postaci grup punktów przylegających do siebie.
- Scan. Kolorowanie wygenerowanych bądź wczytanych danych.
- DBscan. Grupowanie za pomocą algorytmu DBScan.
- Wyczyść. Czyszczenie komponentu Image.
- Test. Dodanie do logu zdarzenia test (do sprawdzania aktywności aplikacji).
- Koniec. Wyjście z aplikacji.
- Ładuj plik. Załadowanie pliku *.csv z danymi pochodzącymi ze skanera AiroPeek.
- Wyczyść log. Wyczyszczenie logu przechowywanego w pamięci.
- Zapisz log. Zapisanie logu przechowywanego w pamięci do pliku.
- Ładuj zdjęcie. Załadowanie zdjęcia z pliku do komponentu Image. Dane takie nie mogą być przetwarzane.

- Zapisz zdjęcie. Zapisanie zawartości komponentu Image do pliku *.jpg.
- Porównaj IS. Porównanie danych przechowywanych w tablicy IS z plikiem z danymi IS.
- Ładuj IS. Załadowanie danych z pliku *.txt do tablicy IS.
- Zapisz IS. Zapisanie danych z tablicy IS do pliku *.txt.
- Poprzedni. Wybranie poprzedniej analizy.
- Następny. Wybranie następnej analizy.
- Przetwórz. Wykonanie analizy XY.
- Wszystkie. Wykonanie wszystkich analiz XY.
- Przetwórz DBS. Wykonanie analizy XY, a następnie poddanie wyników funkcji DBScan.
- Wszystkie DBS. Wykonanie funkcji Przetwórz DBS dla wszystkich zastawów XY.
- Przetwórz F. Wykonanie analizy XY filmowo. Do analizy brana jest ilość punktów podana w polu „Ilość klatek". Kolejnymi krokami jest wyczyszczenie tablic z wynikami i przetworzenie następnej porcji punktów.
- Wszystkie F. Wykonanie funkcji Przetwórz F na wszystkich zestawach XY.
- Wstrzymaj. Wstrzymanie wątku przetwarzającego dane.
- Wznów. Wznowienie wątku przetwarzającego dane.
- Zakończ. Zakończenie wątku przetwarzającego dane. Po tej operacji wznowienie nie jest możliwe.

Typowymi działaniami na aplikacji jest testowanie działania algorytmu DBScan na przykładowych danych, przetwarzanie plików oraz porównywanie plików.

Aby przetestować działanie algorytmu DBScan należy wygenerować dane za pomocą funkcji Losowo, Plamy, Wyspy, a następnie obrobić je za pomocą funkcji DBscan (można też załadować dane za pomocą funkcji Laduj IS. Podczas prac można zmieniać wartości pól „Wypełnienie", „Liczba klastrów" oraz „R DBscan".

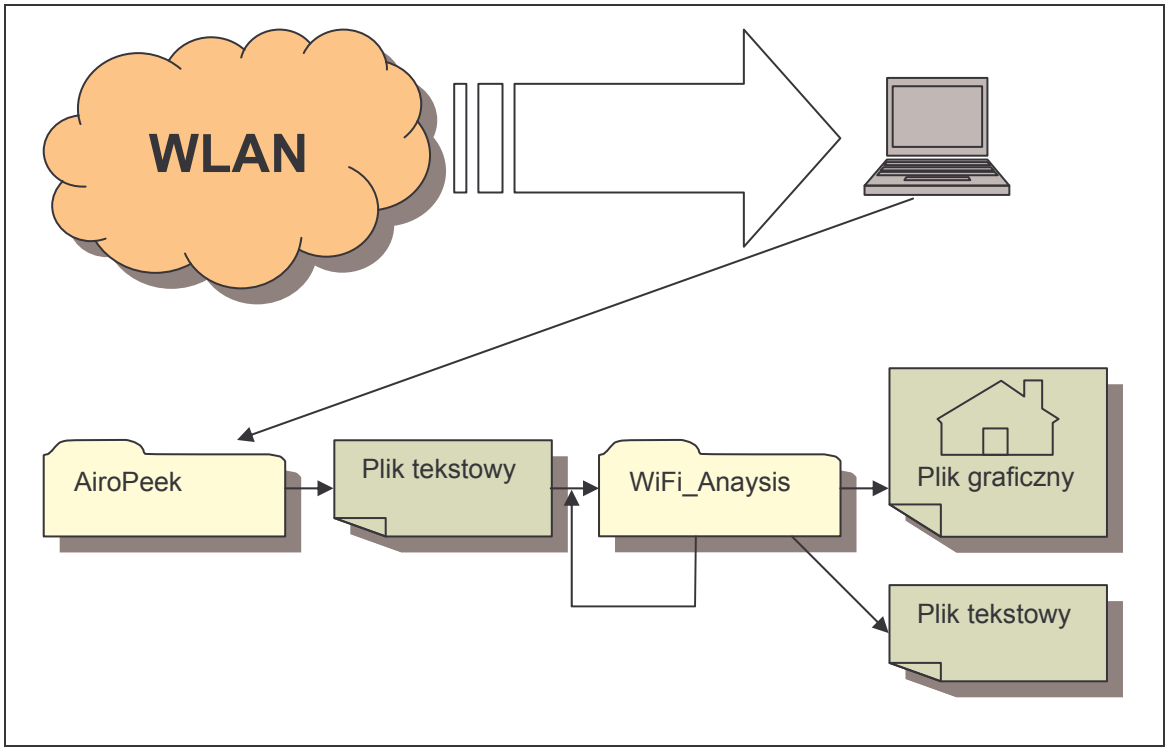
Aby przetworzyć dane w aplikacji należy załadować plik z danymi za pomocą funkcji „Ładuj plik". Podczas tej operacji automatycznie zostaną zaznaczone opcje „Auto zapis zdjęcia", „Auto zapis IS" oraz „Auto zapis logu". Następnie należy wybrać rodzaj danych na odpowiednich współrzędnych XY. Kolejnym krokiem jest naciśnięcie przycisku „Przetwórz", „Przetwórz DBS" bądź „Przetwórz F". Spowoduje to wygenerowanie analizy tj. pojawi się wykres np. MAC adresów nadawcy względem MAC adresów odbiorcy. „Przetwórz DBS" wywoła automatycznie funkcję DBscan. Przy funkcji „Przetwórz F" (filmowo) należy jeszcze ustawić co ile punktów ma być wygenerowany nowy obrazek. Funkcja ta pozwala analizować dynamiczne zmiany tej samej sieci i zapisywać jej stan w określonych odstępach. Szczegóły tego rozwiązania są podane w rozdziale „Analiza porównawcza oraz filmowa". Przyciskami „Następny" lub „Poprzedni" można przemieszczać się po kolejnych zestawach XY bez powtórzeń (jeśli będzie zestaw „Nr-Source", to już nie będzie zestawu „Source-Nr"). Proces ten

można zautomatyzować naciskając odpowiedni przycisk „Wszystkie”. Spowoduje to przetworzenie wszystkich zestawów bez powtórzeń.

Ostatnią operacją, jaką potrafi aplikacja, to porównanie plików IS. Należy załadować (bądź wygenerować) plik IS. Następnie naciskając przycisk „Porównaj IS” należy załadować drugi plik. Punkty z pierwszego pliku IS będą granatowe, z drugiego pliku będą zielone, a część wspólna stanie się czerwona.

Organizacja programu

Przepływ danych podczas analizy sieci pokazany jest na Rysunku 21. Dane z karty sieciowej są przechwytywane przez aplikacje AiroPeek.



Rysunek 21. Przepływ danych podczas analizy sieci Wi-Fi.

Aplikacja ta obsługuje tylko kilka kart i aby współpraca była możliwa, należy zainstalować specjalne sterowniki pozwalające na pracę karty w trybie monitora (dostępne na stronie producenta programu AiroPeek). Skaner sieciowy po nasłuchu umożliwia zapis danych w swoim standardzie lub w postaci plików **.csv* (opis zawartości takiego pliku zawarty jest w tabeli 3). Taki plik można załadować do aplikacji WiFi_Analysis. Następnie po przetworzeniu wyniki są zapisywane w postaci obrazów **.jpg*, bądź w postaci plików **.txt*. Plik zawiera tabele 300 wierszy na 300 kolumn danych pochodzących z tabeli IS. Dane są zapisane w postaci liczb całkowitych rozdzielonych znakami odstępu bądź nowego akapitu.

Kod programu

Kod programu umieszczony jest na dołączonej płycie CD.

Wyniki działania programu

Sposób odczytywania wyników

Program generuje wykresy, które zapisywane są w postaci plików txt oraz jpg. Układ odniesienia jest tak dobrany, że zero znajduje się w lewym górnym rogu. Oś X biegnie od punktu zero w prawą stronę, natomiast oś Y biegnie od punktu zero w dół wykresu. Obrazki mają rozdzielczość 300 na 300 punktów. Podpisy pod rysunkami informują o rodzaju badania (X-Y), pochodzeniu danych, oraz o ewentualnych algorytmach, jakie zostały zastosowane do zobrazowania wyników. Program jako dane wejściowe bierze dane z plików csv, które mają taką strukturę jak pokazana w tabeli poniżej. Dane te traktowane są na trzy sposoby:

- Jako wartości liczbowe zapętlone, tj. jeśli wartość przekroczy 300 to odejmowana jest od niej wielokrotność 300.
- Jako wartości liczbowe skalowane w odpowiedni sposób do wartości [1,300].
- Jako wartości tekstowe, którym przypisywane są potem wartości liczbowe o wartościach [1,300]. Jeśli dany tekst występuje pierwszy raz, to dostaje pierwszą wolną wartość. Jeśli wystąpi po raz kolejny, to program przydziela wcześniej już przyznaną wartość. Pozwala to na skalowanie wartości np. adresów MAC, których pula wynosi 2^14. Należy jednak zwrócić uwagę na to, że jeśli przetwarzamy różne pliki z tej samej sieci, to przyznane wartości najczęściej różnią się miedzy sobą. Kolejna uwaga dotyczy faktu, że pakiety częściej występujące dostaną mniejsze numery, ponieważ istnieje większe prawdopodobieństwo wcześniejszego wystąpienia. Powoduje to, że w pobliżu zera gromadzą się wartości częściej występujące w sieci.

Tabela 3. Opis danych wejściowych do aplikacji WiFi_Analysis.

Nazwa komórki	**Przykładowa wartość**	**Opis**
Nr	1	Numer kolejnego pakietu. Traktowana jest jako wartość liczbowa zapętlona.
Source	Planet Tech:1C:21:5E	Adres MAC nadawcy. Część adresu jest zamieniana na nazwę właściciela puli adresowej. Traktowana jest jako tekst.
Destination	Ethernet Broadcast	Adres MAC odbiorcy. Traktowana jest jako tekst.
BSSID	00:30:4F:1C:21:5E	ID punktu dostępowego. Traktowana jest jako tekst.
Flags	*	Flaga pakietu. Traktowana jest jako tekst.

Channel	12	Nr kanału. Traktowana jest jako wartość liczbowa skalowana zgodnie z wzorem y=x*300/14.
Signal	14%	Siła sygnału. Traktowana jest jako wartość liczbowa skalowana zgodnie z wzorem y=x*300/100.
DataRate	1	Prędkość, z jaką pakiet jest przesyłany. Traktowana jest jako wartość liczbowa skalowana zgodnie z wzorem y=x*10.
Size	60	Wielkość pakietu w bajtach. Traktowana jest jako wartość liczbowa skalowana zgodnie z wzorem y=x*300/1600.
RelariveTime	0	Czas odebrania pakietu od początku skanowania. Traktowana jest jako wartość liczbowa skalowana zgodnie z wzorem y=x
Protocol	802.11 Beacon	Rodzaj pakietu niskiej warstwy. Traktowana jest jako tekst.
Summary	FC=.....,SN=3354,FN=0,BI=100,SSID=wlnet,DS=13	Dodatkowe informacje jak np. Identyfikator sieci. Traktowana jest jako tekst.

Wszystkie wartości mają założone ograniczenie zgodne z zależnością 1<y<300 i jeśli wartości te są przekraczane, to wartość jest skracana do najbliższego ograniczenia.

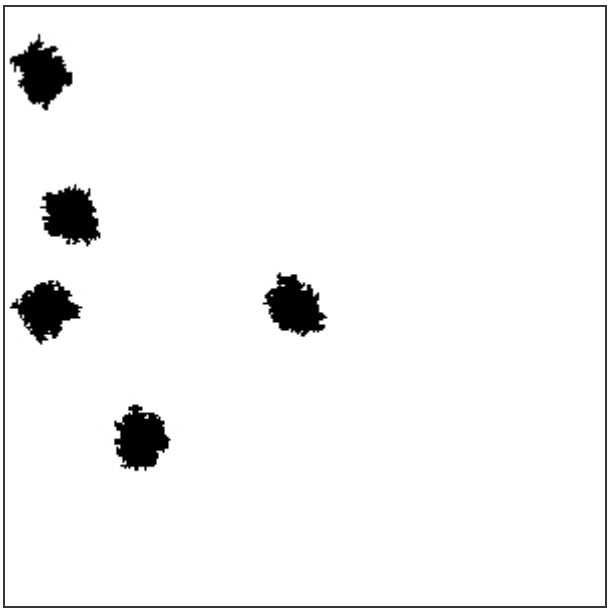
Analiza „każdy z każdym” oraz DBScan

Algorytm DBScan [4] pozwala na grupowanie danych na podzbiory.

Działanie algorytm DBScan:

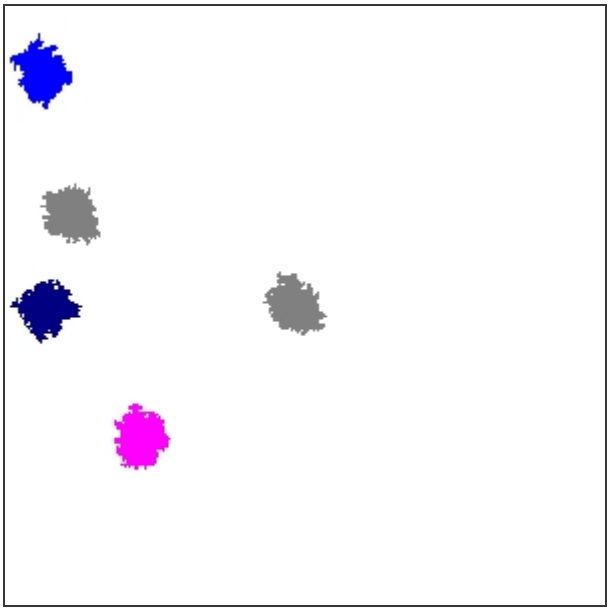
1. Szukaj nieoznaczonego punktu.
2. Jeśli nie znaleziono punktu to idź do punktu 7.
3. Jeśli został znaleziony punkt to go oznacz jako nową grupę.
4. Szukaj w promieniu R od punktu innych nieoznaczonych punktów.
5. Jeśli został znaleziony punkt oznacz go tą samą grupą. Jeśli nie idź do punktu 1.
6. Idź do punktu 4.
7. Koniec.

Algorytm ten dobrze pozwala rozpoznać zbiory jak, na zdjęciu poniżej.



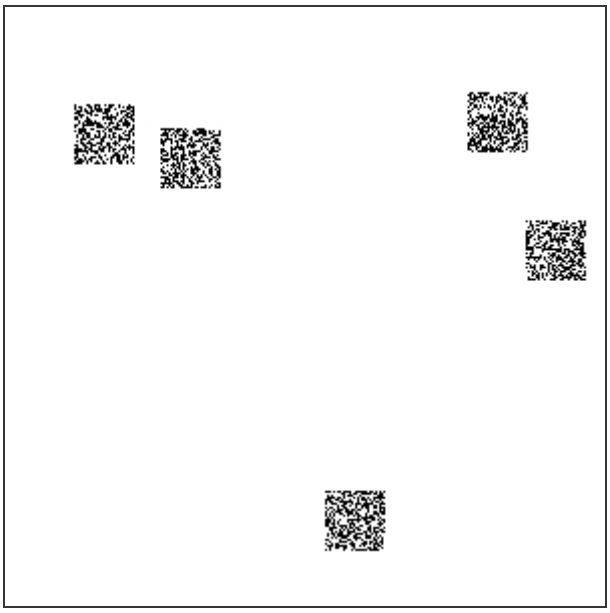
Rysunek 22. Przykładowe zbiory.

W efekcie jego działania otrzymujemy wynik, jak na zdjęciu poniżej.



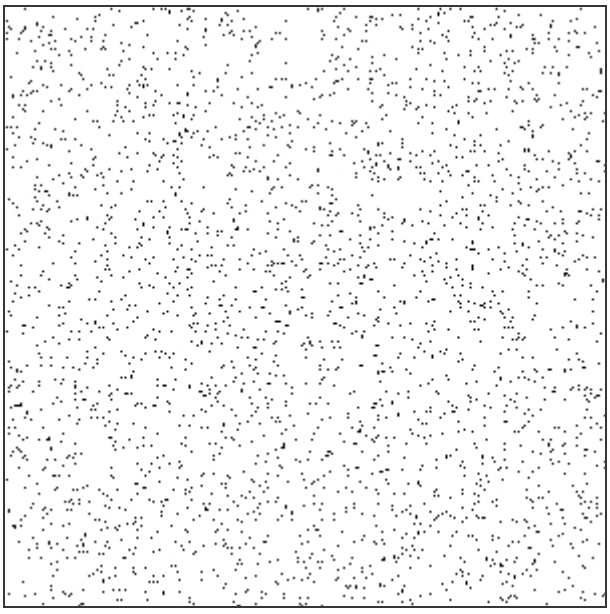
Rysunek 23. Zbiory oznaczone za pomocą algorytmu DBScan.

Zmieniając wartość promienia przeszukującego możliwe jest grupowanie obszarów nieciągłych, jak na zdjęciu poniżej.



Rysunek 24. Przykładowe zbiory nieciągłe.

Niestety algorytm ten zupełnie nie sprawdza się, gdy dane są mocno zaszumione bądź losowe, jak na zdjęciu poniżej.



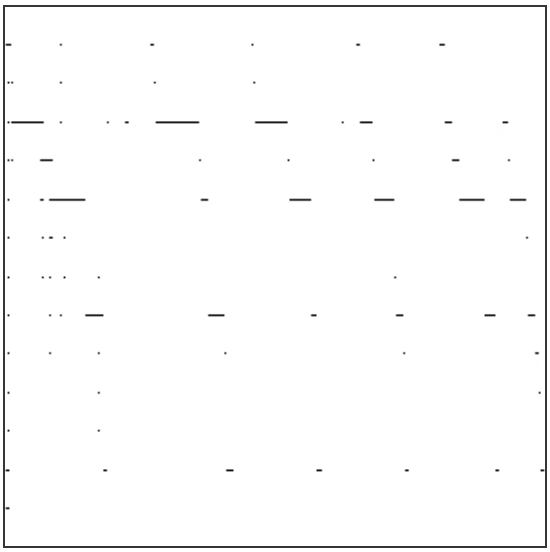
Rysunek 25. Przykładowy rozkład losowy.

Tak więc gdyby w sieci Wi-Fi były wyraźne różne zbiory zachowań pozwalałby on na ich automatyczne grupowanie. Ewentualne szumy można by było usunąć za pomocą filtrów graficznych. Niestety podczas badań okazało się, że zachowania sieci są mocno losowe, a jeśli wyłaniają się jakieś grupy, nie pozwalają się jednoznacznie rozdzielać, ponieważ grupy te mają

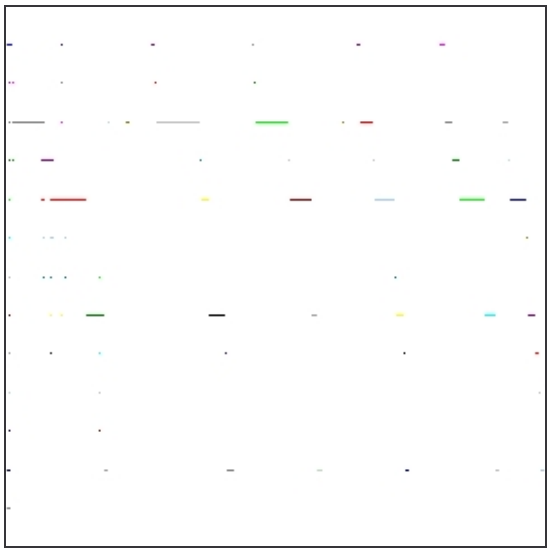
podobne kształty we wszystkich badanych sieciach. Analizy te pozwalają pokazać ciekawe rzeczy, na które zwykle nie zwraca się uwagi.

Na rysunkach o numerach od 26 do 79 są pokazane ciekawsze przykłady analiz sieci, w których można się pokusić o interpretacje wyników. Powstały one na podstawie pliku *cer1.scv*. Pierwszy obrazek pokazuje wynik badania. Kolejny jest przetworzony algorytmem grupowania DBScan. Ponieważ w rozdziale „Analiza porównawcza oraz filmowa" rysunki pozwalają na bardziej dokładne wnioskowanie, poniżej pokazane są jedynie zdjęcia bez komentarza. Można na nich zaobserwować działanie algorytmu DBScan, który niestety nie radzi sobie z prawidłowym grupowaniem ciekawych fragmentów zdjęcia (np. Rysunek 33). Na innych rysunkach (np. rysunek 37) pojawia się wiele grup, tylko dlatego, że jedna z dziedzin (w tym przypadku nr kanału) jest mocno dyskretna (tutaj ma tylko 14 wartości), co powoduje naturalne duże odstępy miedzy danymi, należącymi do tej samej grupy, co powoduje rozbicie grup na mniejsze. W tym wypadku prawidłowym wynikiem było by oznaczenie grupy sygnałów o tej samej wartości na różnych kanałach.

BSSID-Channel

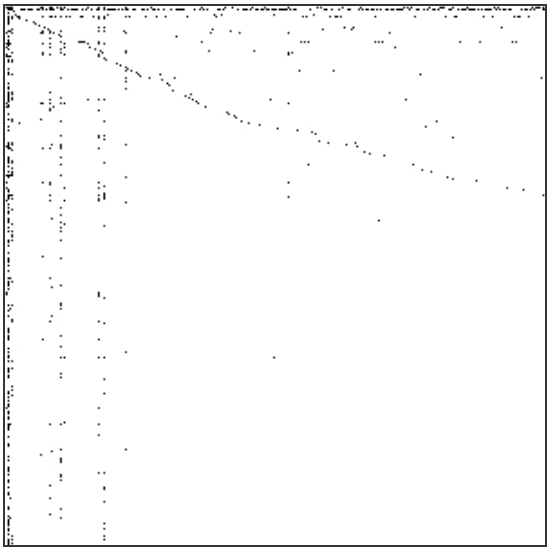


Rysunek 26. BSSID-Channel.

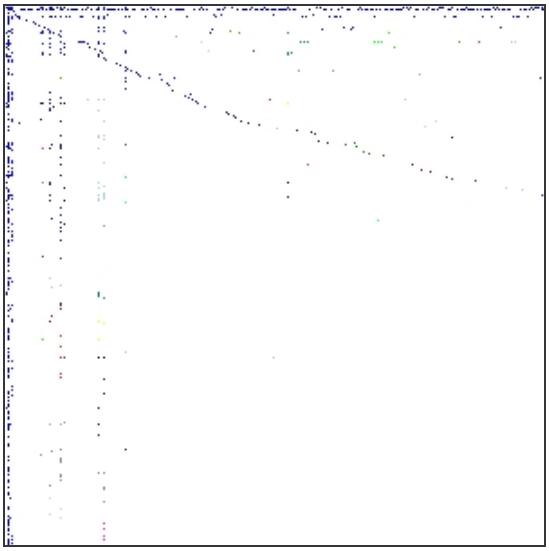


Rysunek 27. BSSID-Channel DBScan.

BSSID-Protocol

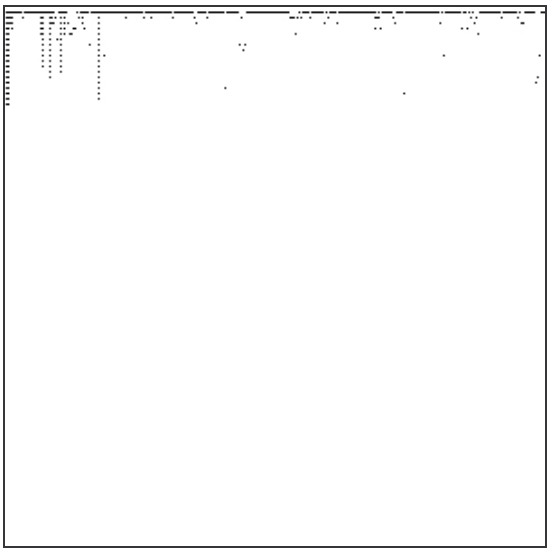


Rysunek 28. BSSID-Protocol.

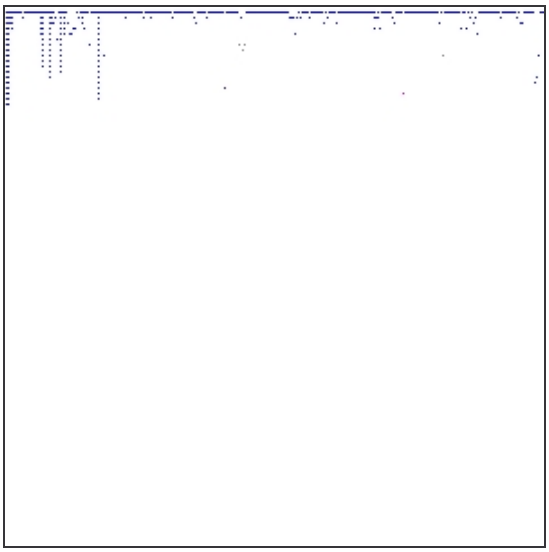


Rysunek 29. BSSID-Protocol DBScan.

BSSID-Signal

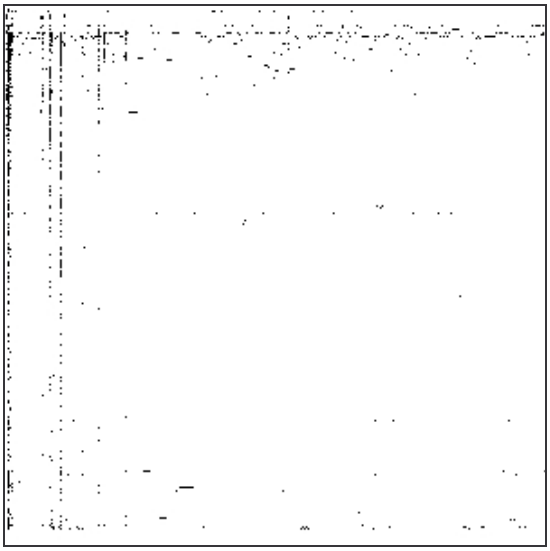


Rysunek 30. BSSID-Signal.

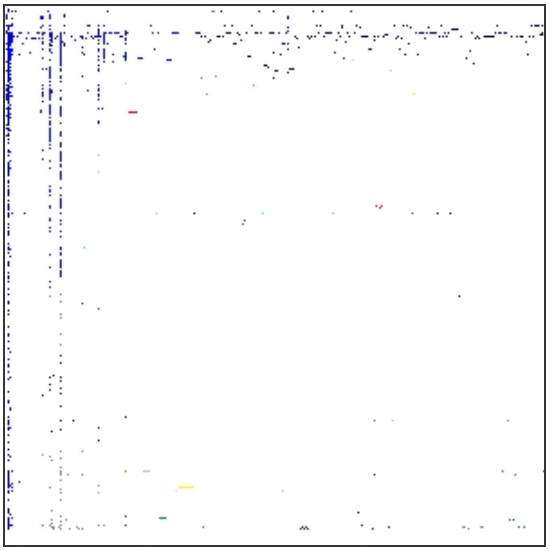


Rysunek 31. BSSID-Sinal DBScan.

BSSID-Size

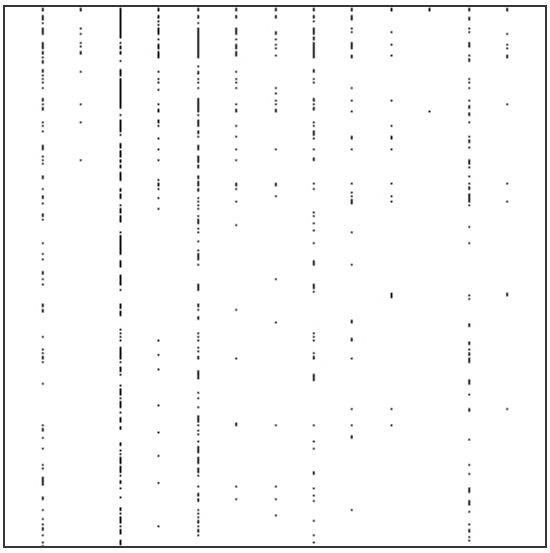


Rysunek 32. BSSID-Size.

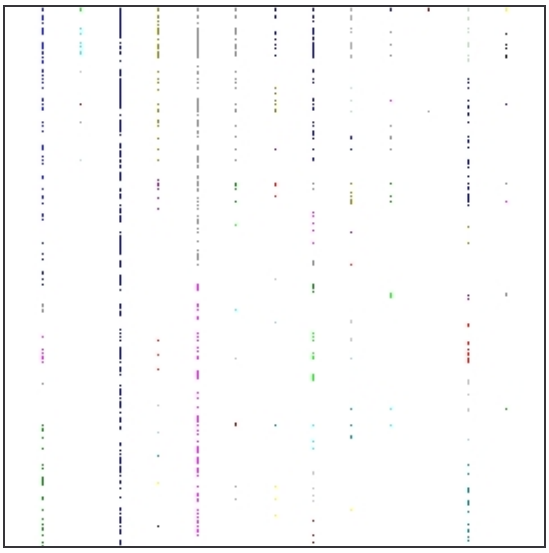


Rysunek 33. BSSID-Size DCScan.

Channel-Protocol

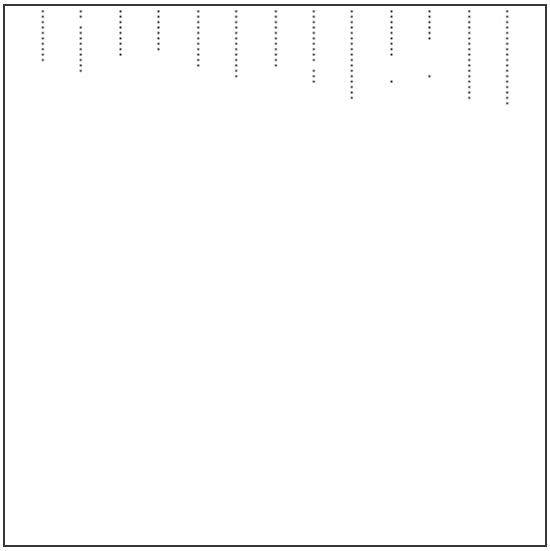


Rysunek 34. Channel-Protocol.

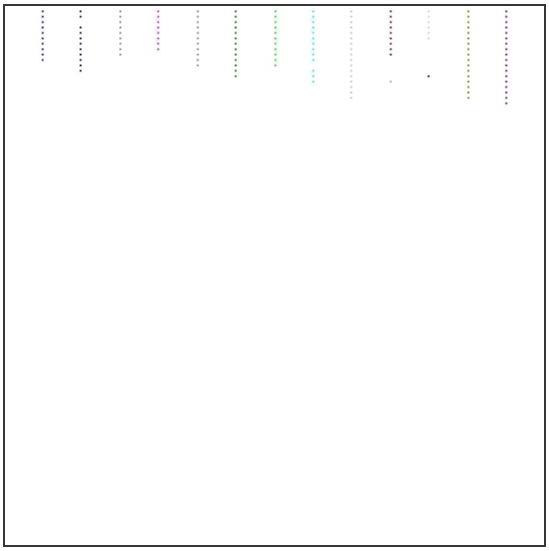


Rysunek 35. Channel-Protocol DBScan.

Channel-Signal

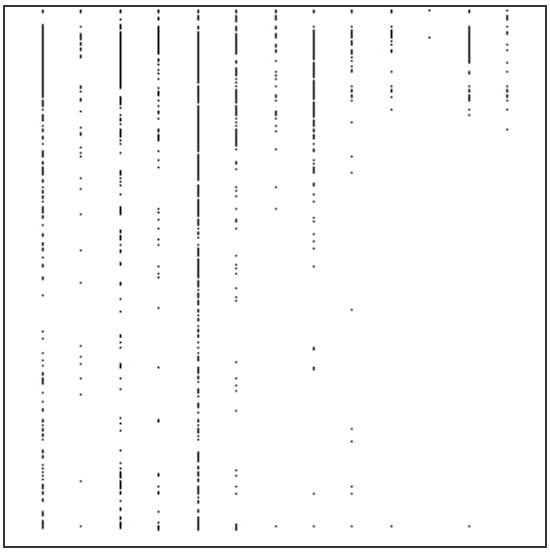


Rysunek 36. Channel-Signal.

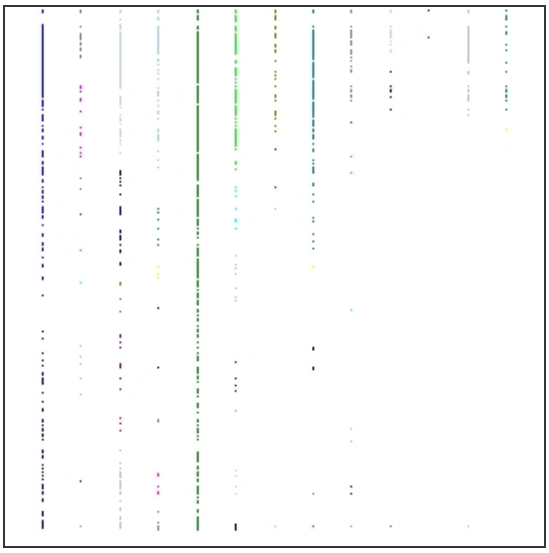


Rysunek 37. Channel-Signal DBScan.

Channel-Size

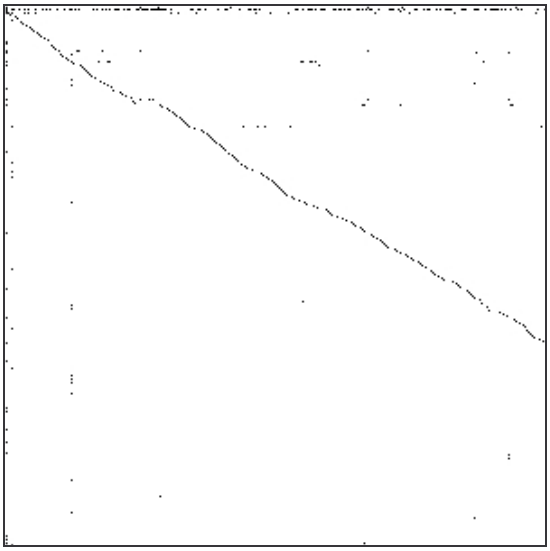


Rysunek 38. Channel-Size.

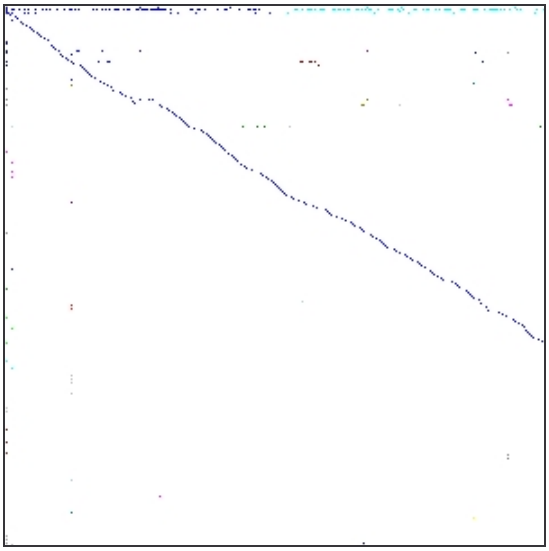


Rysunek 39. Channel-Size DBScan.

Destination-BSSID

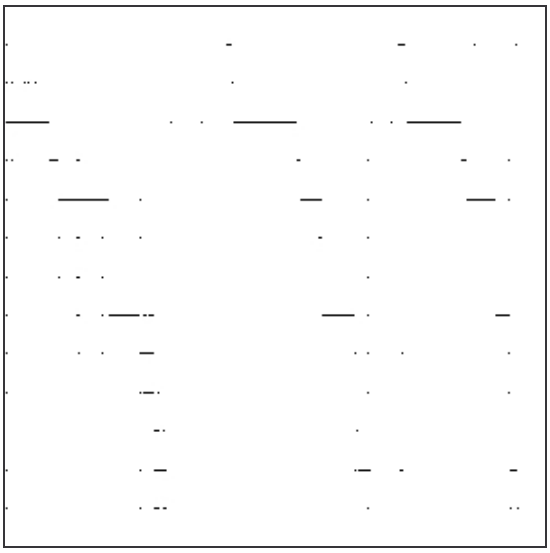


Rysunek 40. Destination-BSSID.

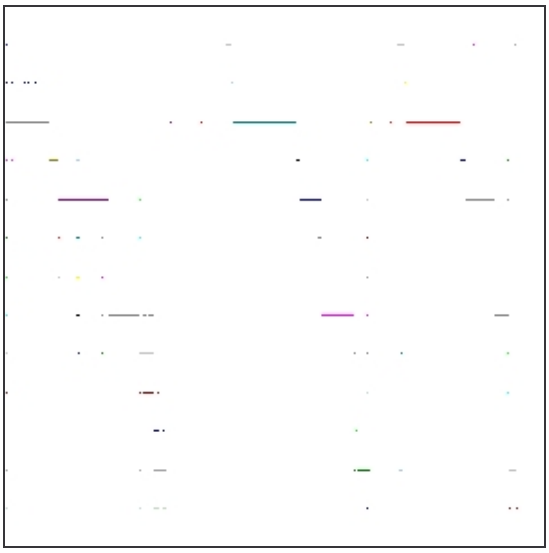


Rysunek 41. Destination-BSSID DBScan.

Destination-Channel

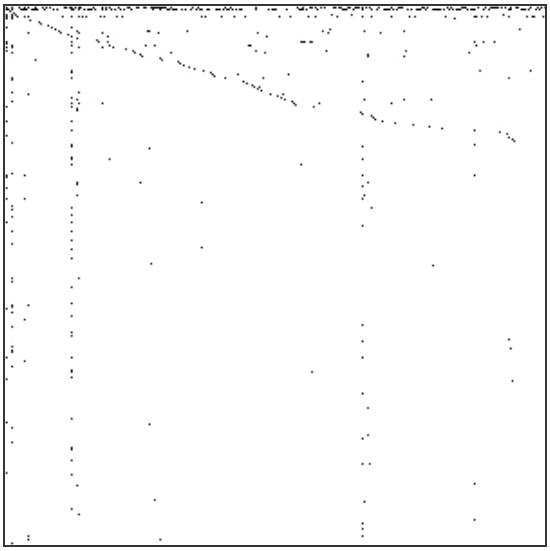


Rysunek 42. Destination-Channel.

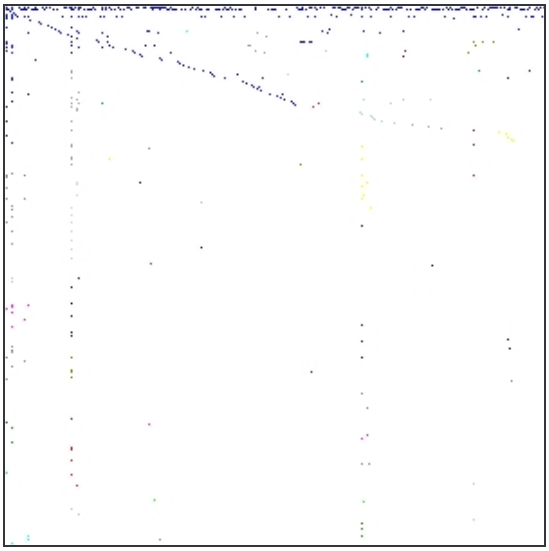


Rysunek 43. Destination-Channel DBScan.

Destination-Protocol

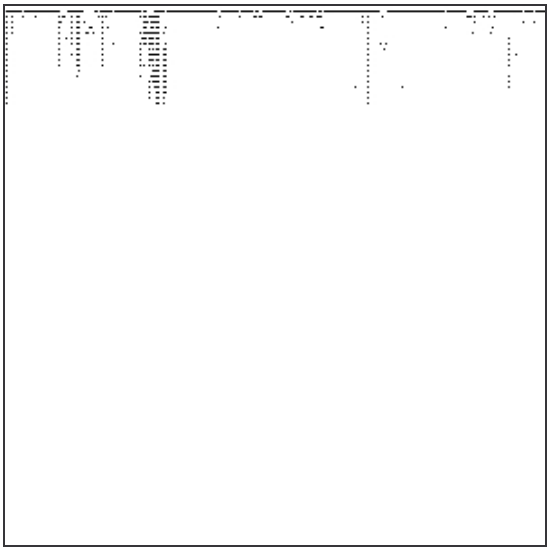


Rysunek 44. Destination-Protocol.

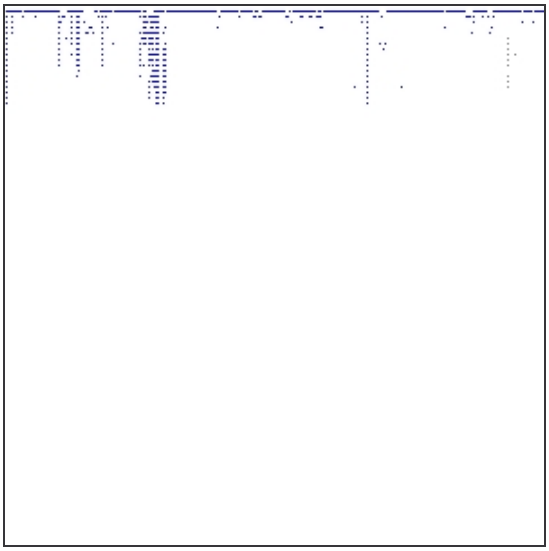


Rysunek 45. Destination-Protocol DBScan.

Destination-Signal

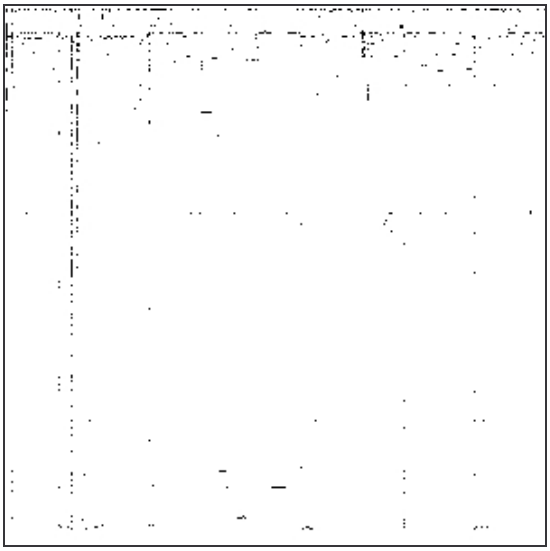


Rysunek 46. Destination-Signal.

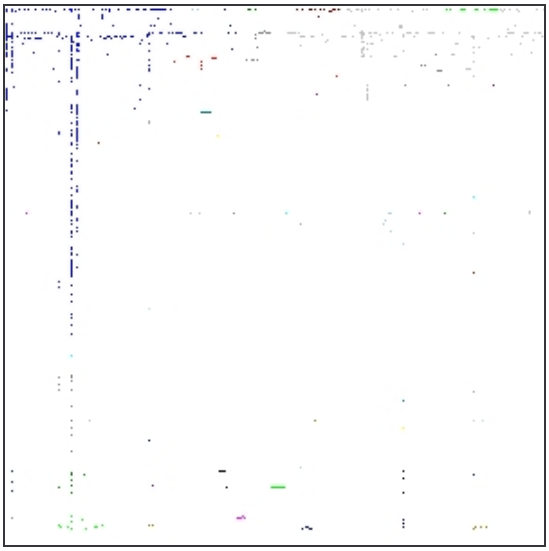


Rysunek 47. Destination-Signal DBScan.

Destination-Size

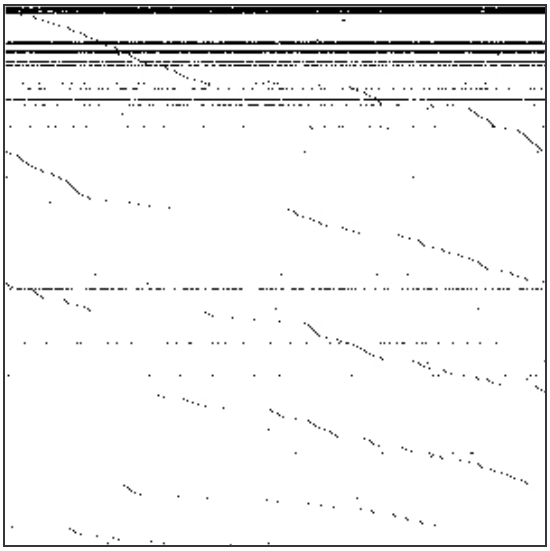


Rysunek 48. Destination-Size.

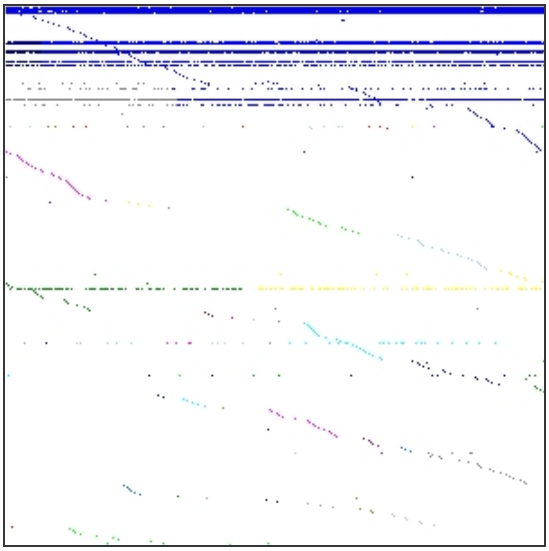


Rysunek 49. Destinon-Size DBScan.

Nr-BSSID

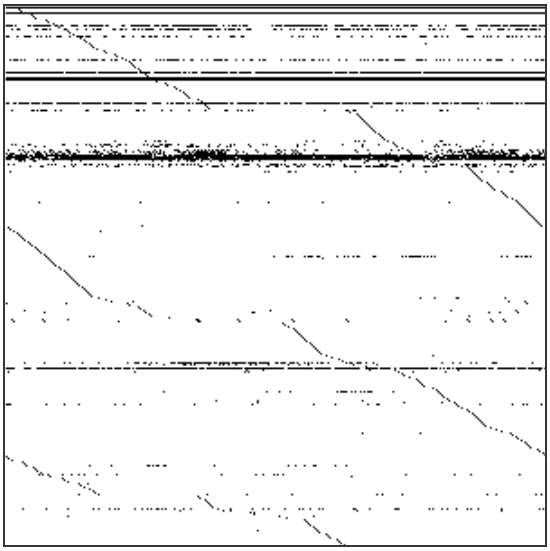


Rysunek 50. Nr-BSSID.

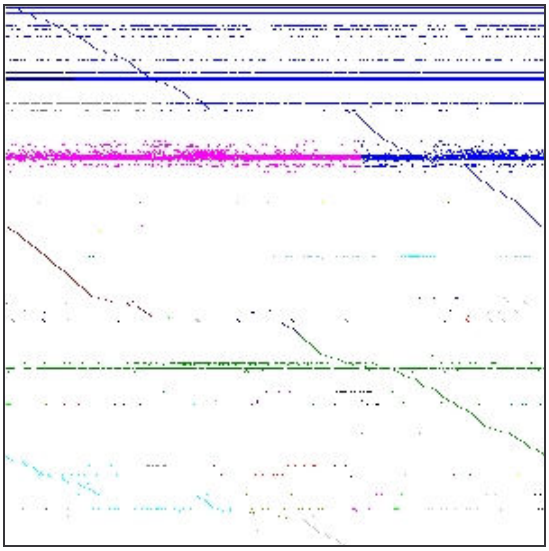


Rysunek 51. Nr-BSSID DBScan.

Nr-Destination

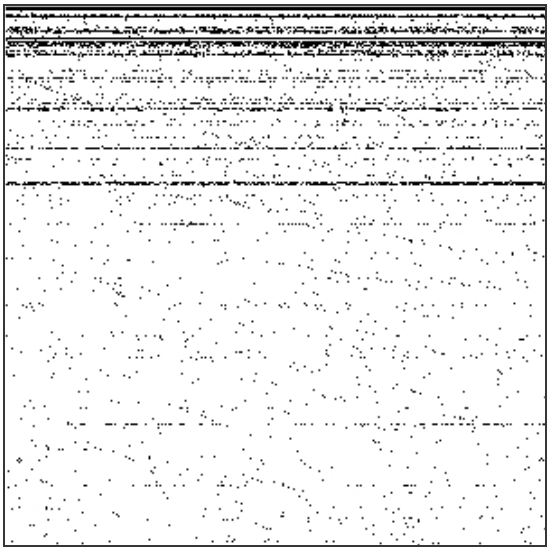


Rysunek 52. Nr-Destination.

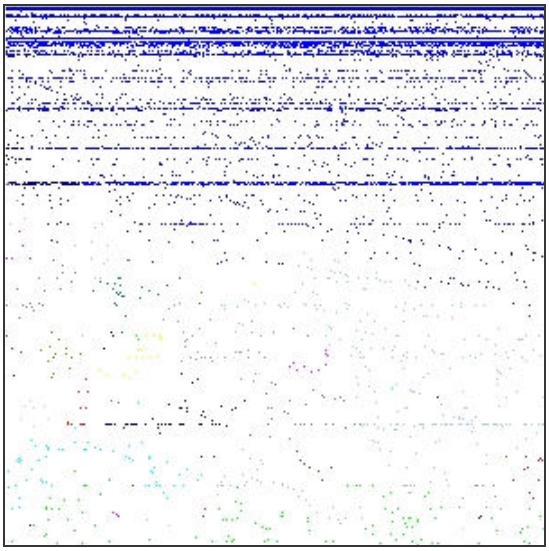


Rysunek 53. Nr-Destination DBScan.

Nr-Protocol

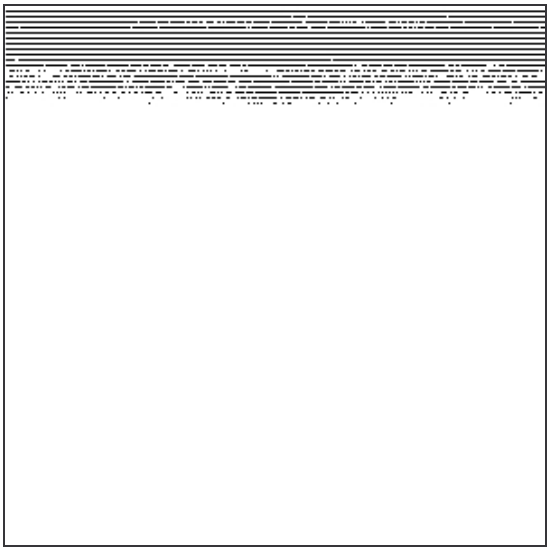


Rysunek 54. Nr-Protocol.

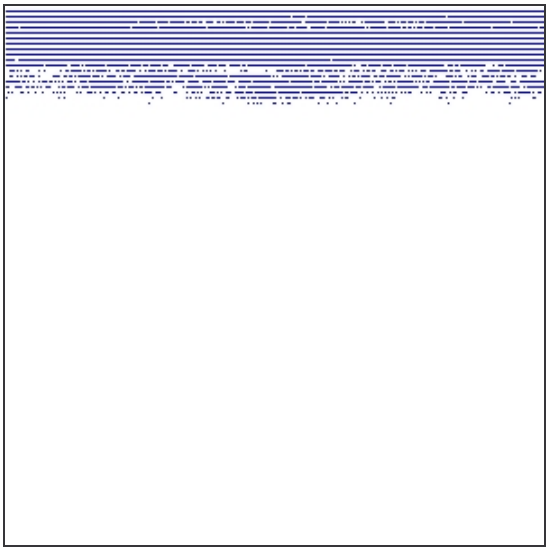


Rysunek 55. Nr-Protocol DBScan.

Nr-Signal

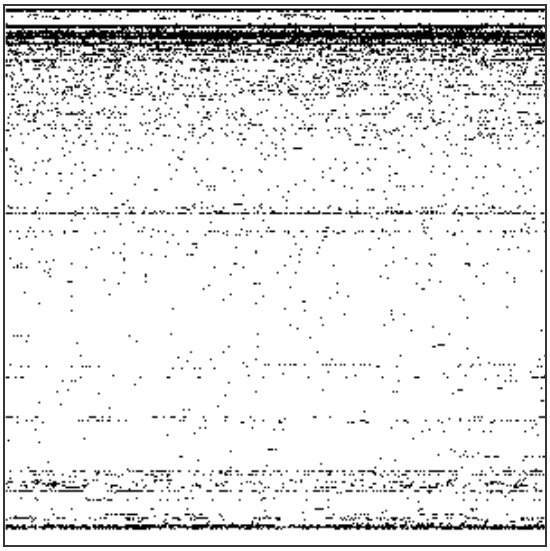


Rysunek 56. Nr-Signal.

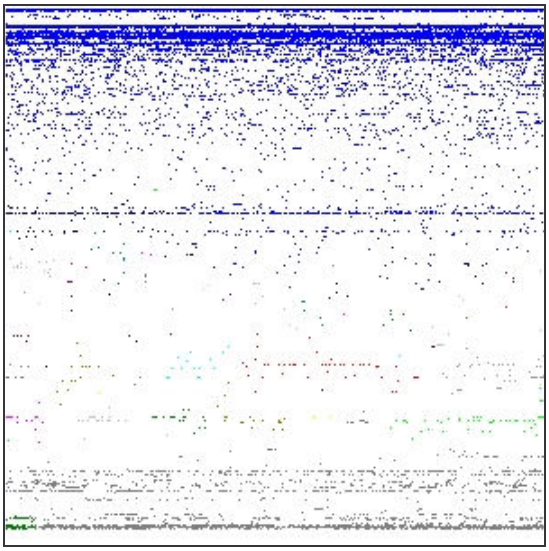


Rysunek 57. Nr-Signal DBScan.

Nr-Size

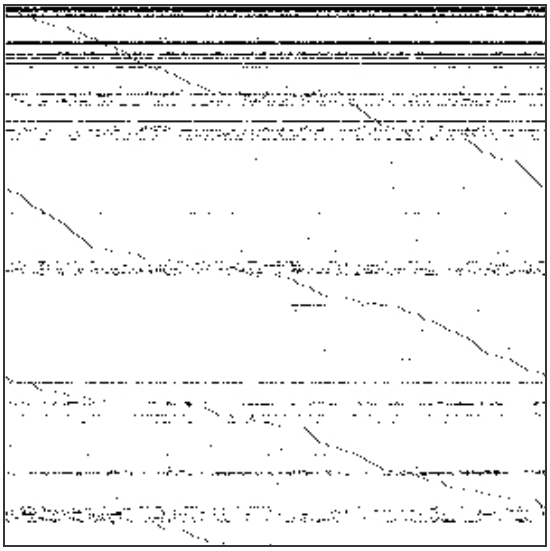


Rysunek 58. Nr-Size.

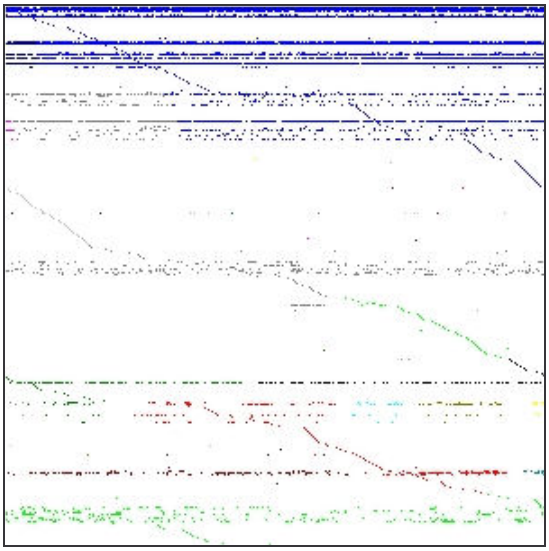


Rysunek 59. Nr-Size DBScan.

Nr-Source

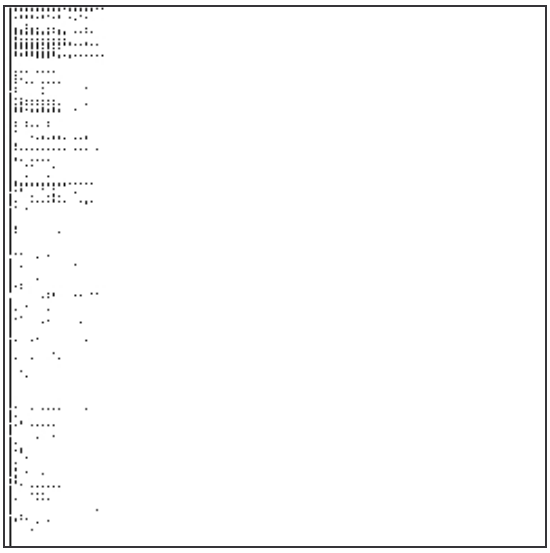


Rysunek 60. Nr-Source.

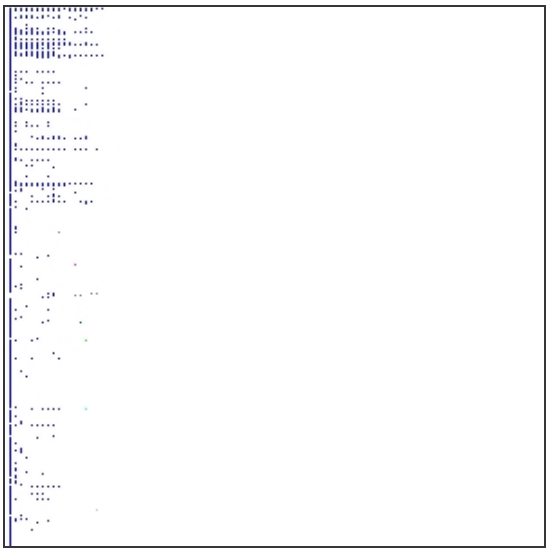


Rysunek 61. Nr-Source DBScan.

Signal-Protocol

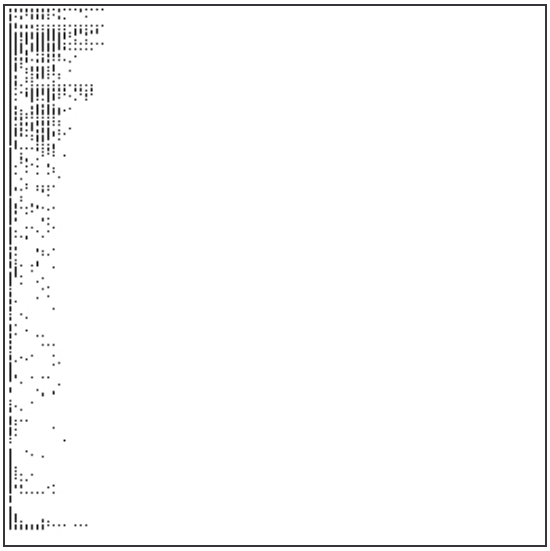


Rysunek 62. Signal-Protocol.



Rysunek 63. Signal Protocol DBScan.

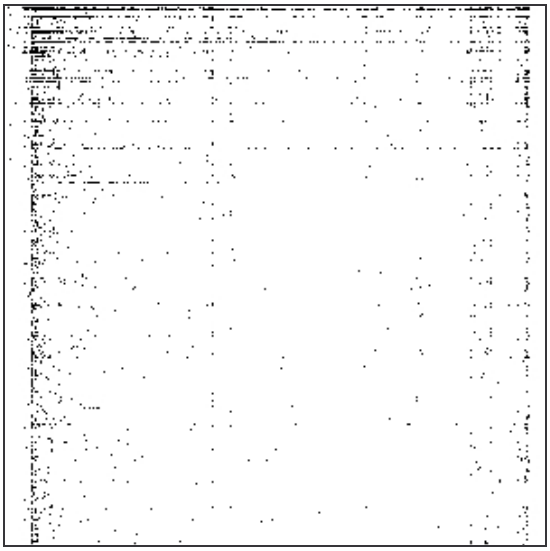
Signal-Size



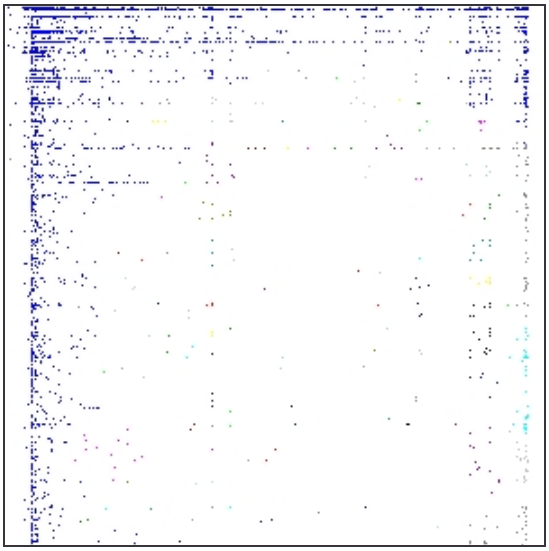
Rysunek 64. Signal-Size.

Rysunek 65. Signal-Size DBScan.

Size-Protocol

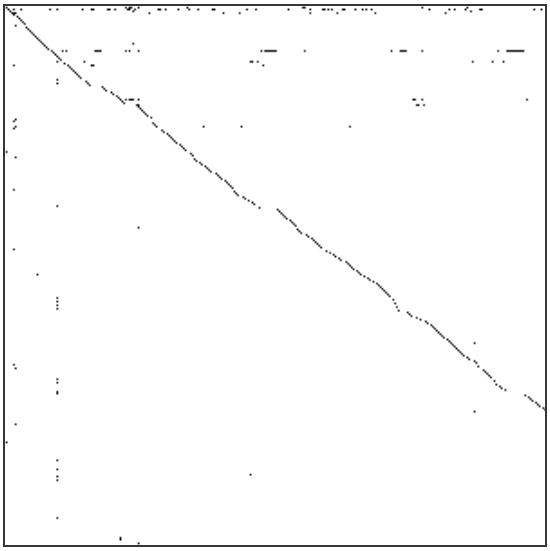


Rysunek 66. Size-Protocol.

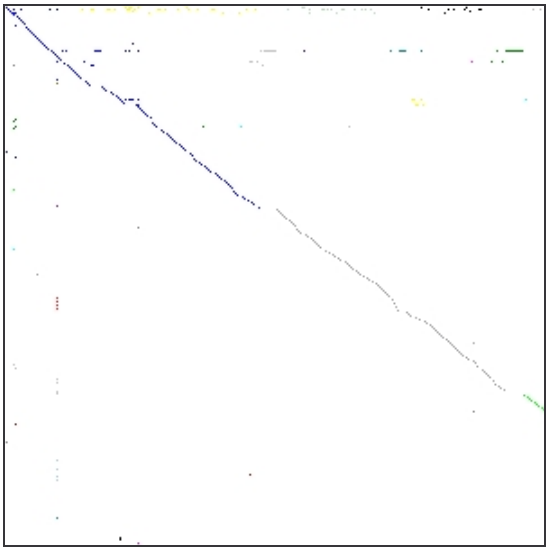


Rysunek 67. Size Protocol DBScan.

Source-BSSID

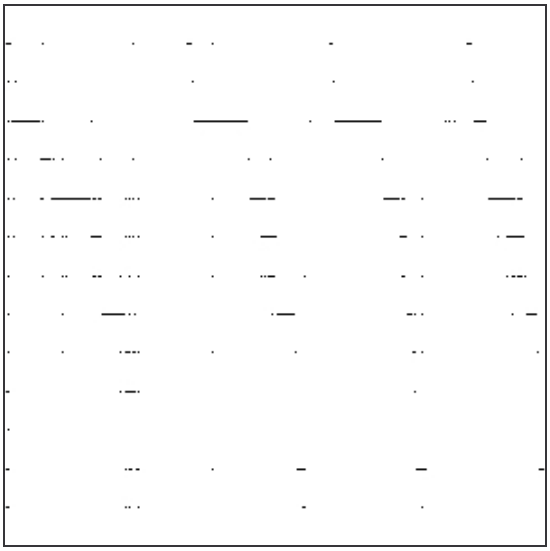


Rysunek 68. Source-BSSID.

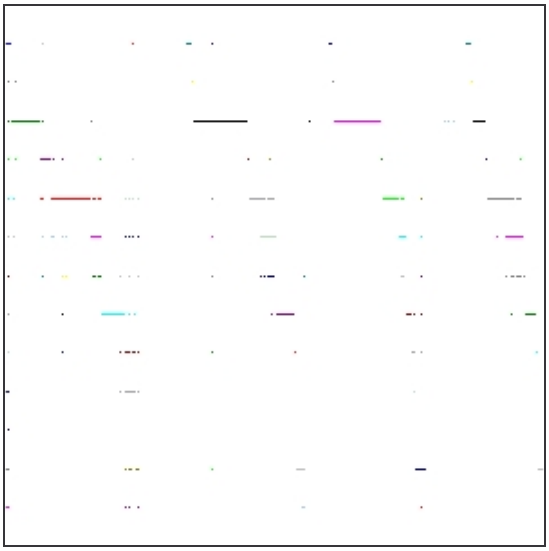


Rysunek 69. Source-BSSID DBScan.

Source-Channel

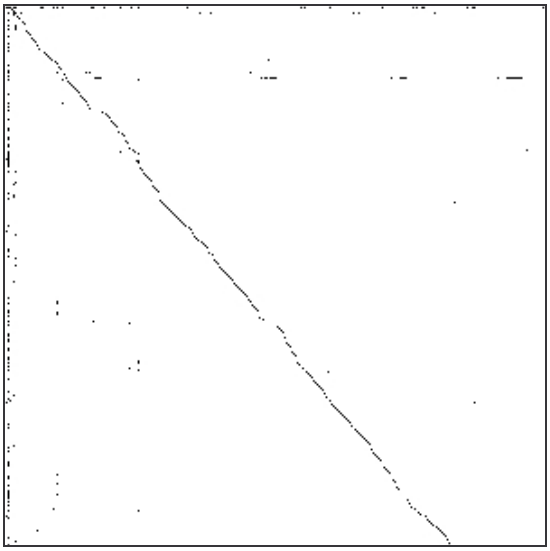


Rysunek 70. Source-Channel.

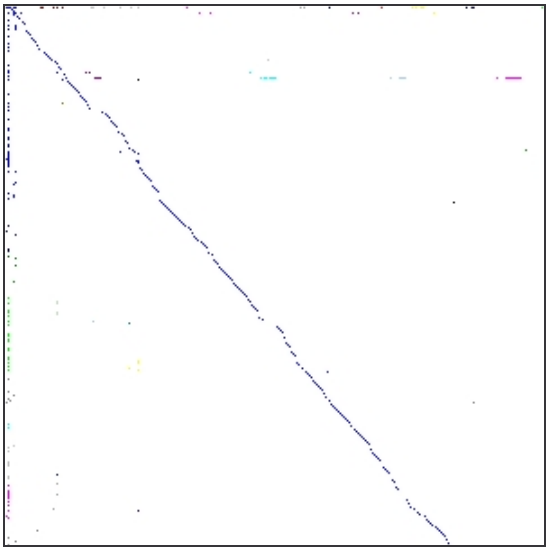


Rysunek 71. Source-Channel DBScan.

Source-Destination

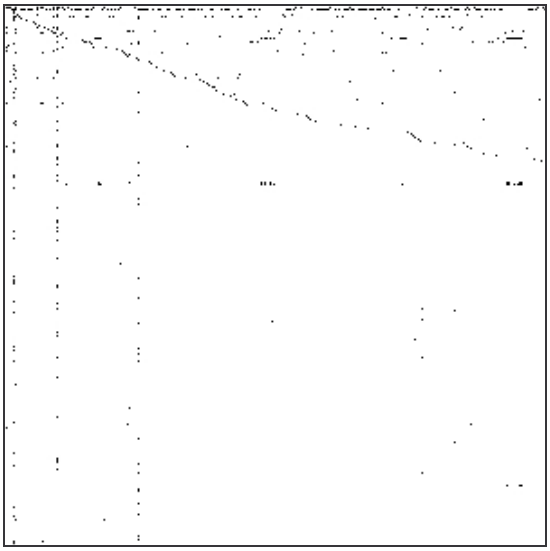


Rysunek 72. Source-Destination.

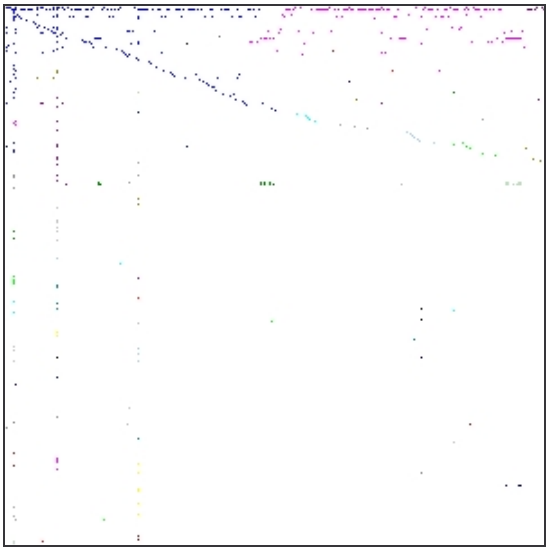


Rysunek 73. Source-Destination DBScan.

Source-Protocol

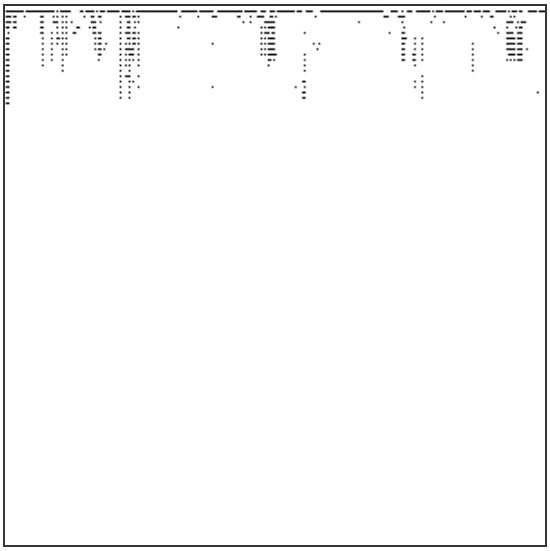


Rysunek 74. Source-Protocol.

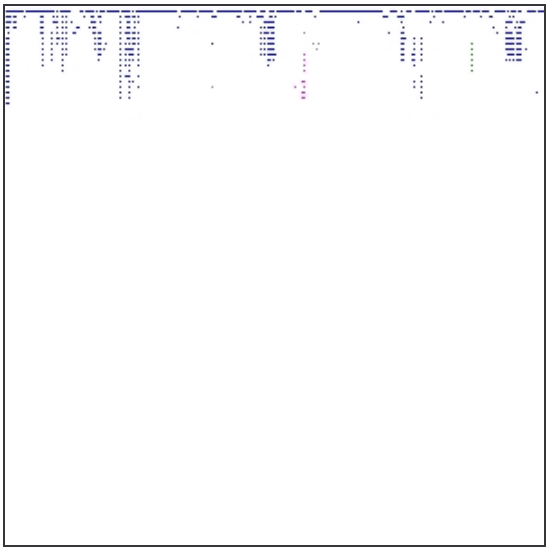


Rysunek 75. Source-Protocol DBScan.

Source-Signal

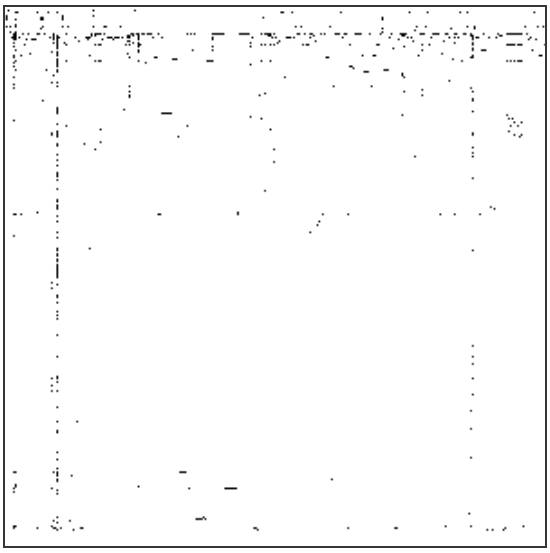


Rysunek 76. Source-Signal.

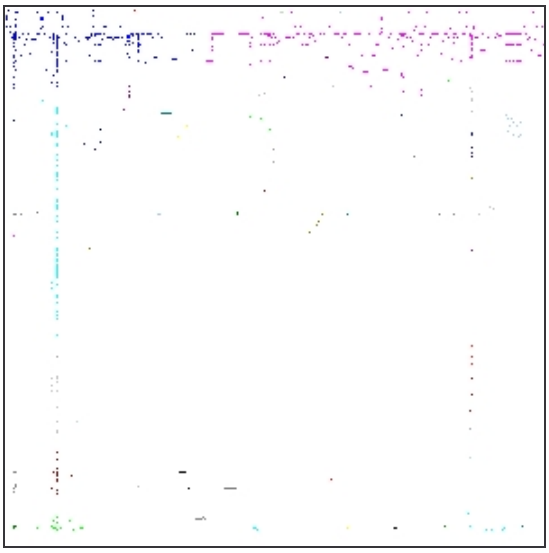


Rysunek 77. Source-Signal DBScan.

Source-Size



Rysunek 78. Source-Size.

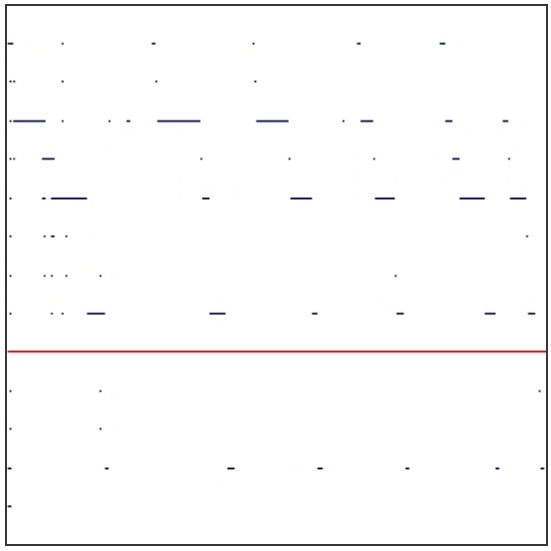


Rysunek 79. Source-Size DBScan.

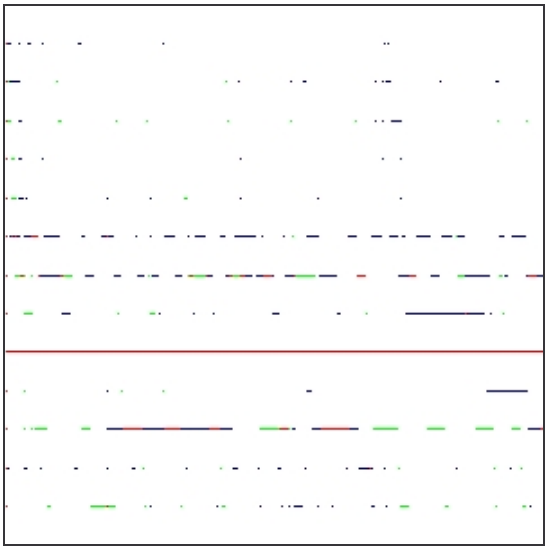
Analiza porównawcza oraz filmowa

Do tej analizy wykorzystano parę grup plików. Pliki od *cer1.csv* do *cer11.csv* pozwoliły na utworzenie analiz „Cer". Jest to zestaw analiz z jednego dnia od godziny 9 rano do godziny 21. Jest to nasza główna badana próba. Pliki *kid2.csv*, *kid2b.csv*, *tarchomin.csv*, *teatr.csv* oraz *zlobek.csv* posłużyły do stworzenia analiz „Inne". Te nasłuchy powstały o różnych porach w różnych miejscach w Warszawie. Jest to nasza próba odniesienia. Plik *cerl1.csv* to nasłuch dziesięciokrotnie większy niż w pliku *cer1.csv* z kolejnego dnia nasłuchu sieci „Cer". Pozwolił on na utworzenie analiz „Cerl film". Jest to rozszerzenie badanej próby pozwalające na analizę zmian w sieci związanych z upływającym czasem. Analizy „All" powstały przez „Porównanie IS" plików *cer*.txt* z plikami *inne*.txt*. W wyniku otrzymujemy nałożenie się ruchu jednej określonej sieci na ruch innych różnych sieci. Z wyżej wymienionych analiz zostały wybrane te, których wyniki wydają się być ciekawe do omówienia. Analizy te zostały pokazane na rysunkach o numerach od 79 do 186 i zostały one pogrupowane na zestawy XY, w ramach których, omówione zostały wyniki.

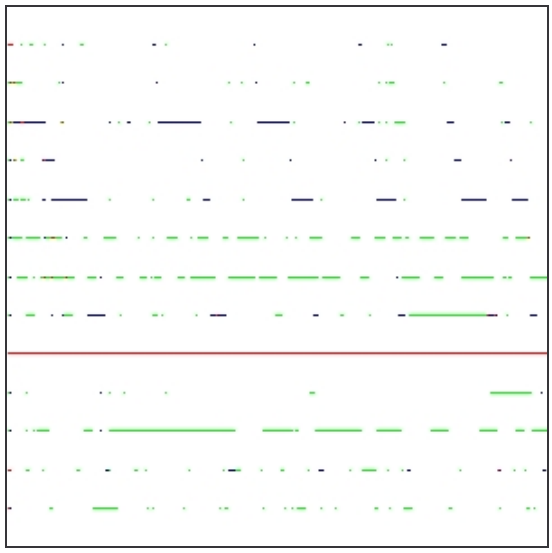
BSSID-Channel



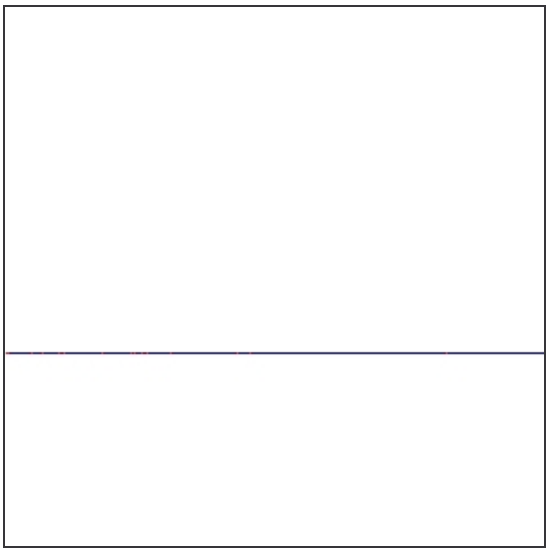
Rysunek 80. Cer BSSID-Channel.



Rysunek 81. Inne BSSID-Channel.



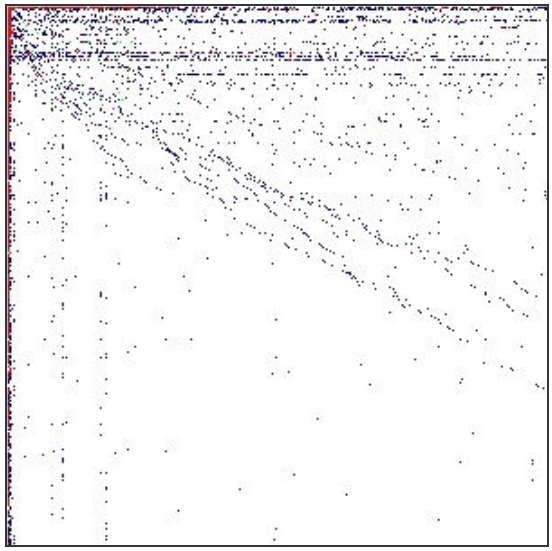
Rysunek 82. All BSSID-Channel.



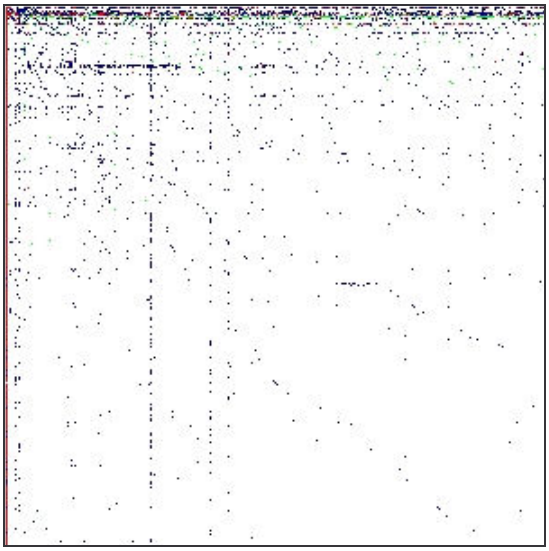
Rysunek 83. Cerl BSSID-Channel film.

Analiza „Cer" pokazuje, że w danym terenie punkty dostępowe korzystają głównie z 3-4 kanałów. Wynika to z tego, że kanały nachodzą na siebie i na danym obszarze można wyselekcjonować, co najwyżej 3-4 kanały nie przeszkadzające sobie nawzajem. Sygnał na innych kanałach pochodzi od sieci bardziej oddalonych. W przypadku analizy filmowej rozkład ten jest niewidoczny z powodu, że nasłuch był realizowany na jednym kanale.

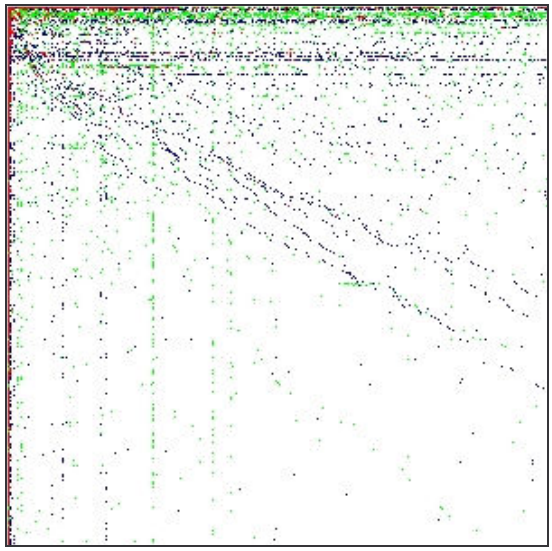
BSSID-Protocol



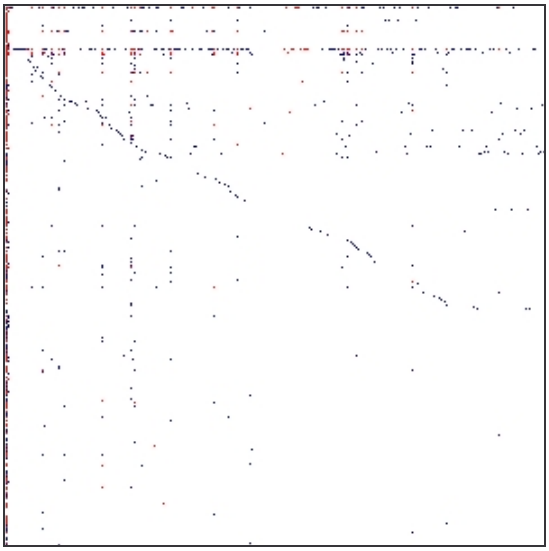
Rysunek 84. Cer BSSID-Protocol.



Rysunek 85. Inne BSSID-Protocol.



Rysunek 86. All BSSID-Protocol.

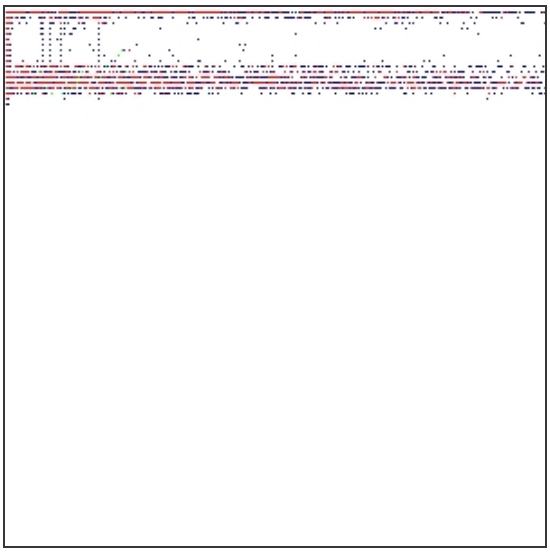


Rysunek 87. Cerl BSSID-Protocol film.

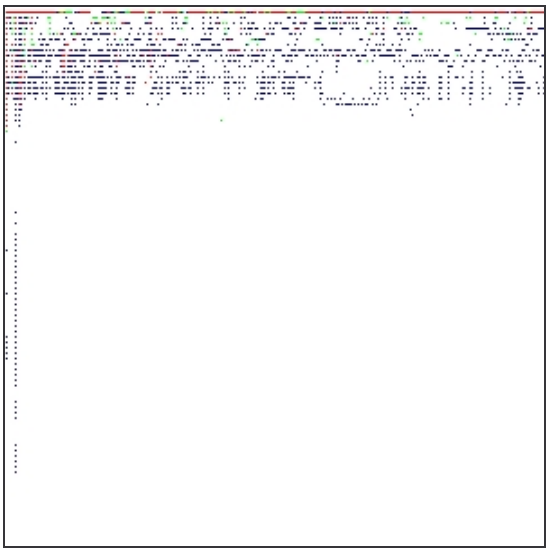
Analizy te pokazują, że urządzenia sieciowe można podzielić na dwie grupy. Urządzenia, które wykorzystują niewielką ilość protokołów (być może są to proste punkty dostępowe oraz karty klienckie) oraz urządzenia wykorzystujące dużą ilość protokołów (być może urządzenia profesjonalne, bądź mosty sieciowe). Linia ukośna jest wynikiem uczenia się analizatora nowych wartości tekstowych (pierwsze wystąpienia obu typu wartości). Z nachylenia tej linii wynika, że więcej jest różnych BSSID niż protokołów. Jeśli przyjrzymy się analizie „All", to widać, że sieci nie mają zbyt wielu wspólnych punktów. Niestety rozrzut parametrów podczas analizy tej samej sieci, widoczny w analizie „Cer", nie pozwala na identyfikacje danej sieci. Z analizy filmowej należałoby wnioskować, że ruch w eterze pochodzi

z urządzeń w ramach kilku aktywnych BSSID, lub z urządzeń w ramach nowych BSSID. Zastanawiający jest duży udział nowych BSSID.

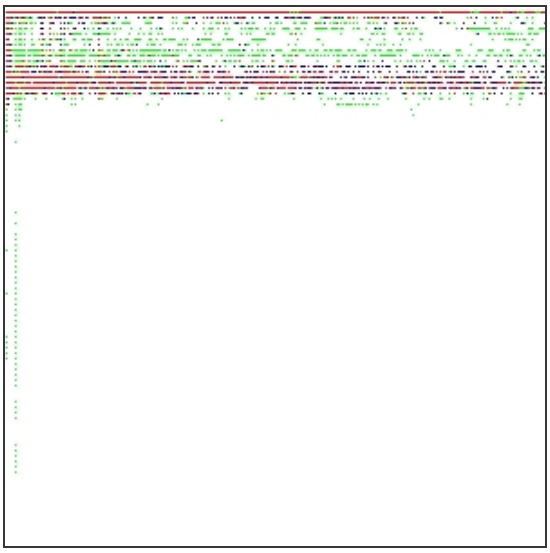
BSSID-Signal



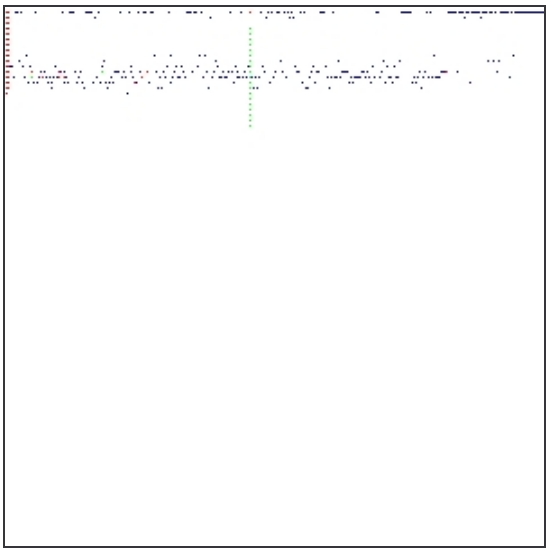
Rysunek 88. Cer BSSID-Signal.



Rysunek 89. Inne BSSID-Signal.



Rysunek 90. All BSSID-Signal.

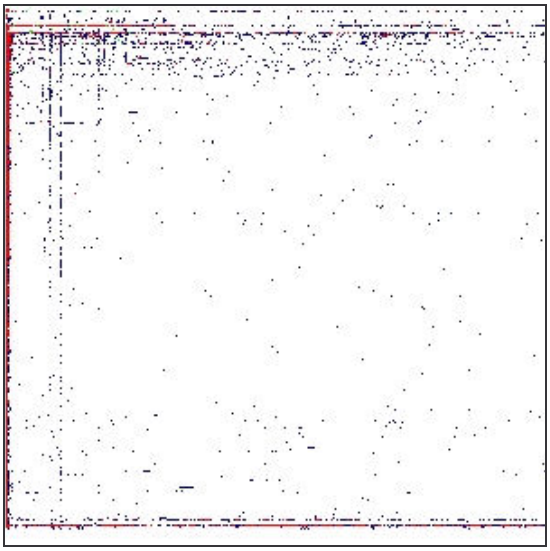


Rysunek 91. Cerl BSSID-Signal film.

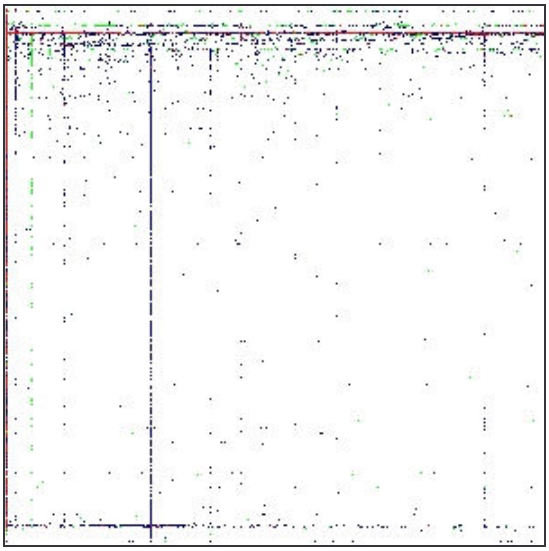
Analiza ta pokazuje, że cześć komunikacji jest dobrze słyszalna, ale nie jest to sygnał jednostajny. Dla BSSID częściej występujących widać cały (od zera do najsilniejszego sygnału) przedział słyszalności. Sugerowałoby to, że słyszymy silny sygnał z punku dostępowego oraz zróżnicowane sygnały od klientów. W innych przypadkach sieć operuje na określonych wartościach sygnału odbieranego. Wskazuje to na punkty sieci wykorzystujące dobre anteny zewnętrzne zapewniające dużą stabilność sygnału. Na analizie „All" widać dość duże

rozbieżności między obserwowanymi sygnałami i można by się pokusić o próbę rozróżnienia sieci, ale taka analiza byłaby trudna do zrobienia ze względu na dużą czułość analizy na warunki pomiaru, oraz konieczność pracy w sieci anten stacjonarnych.

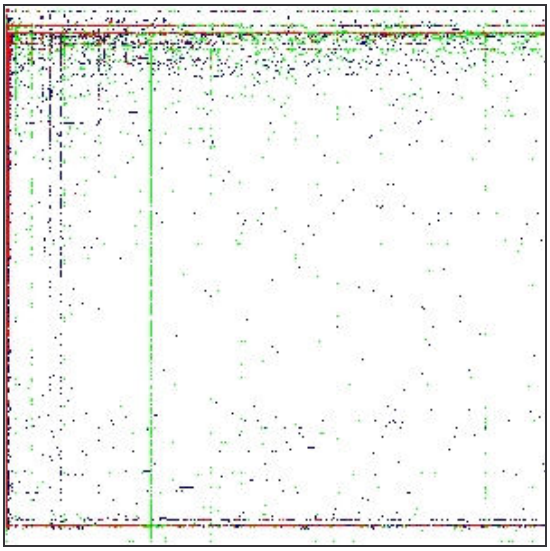
BSSID-Size



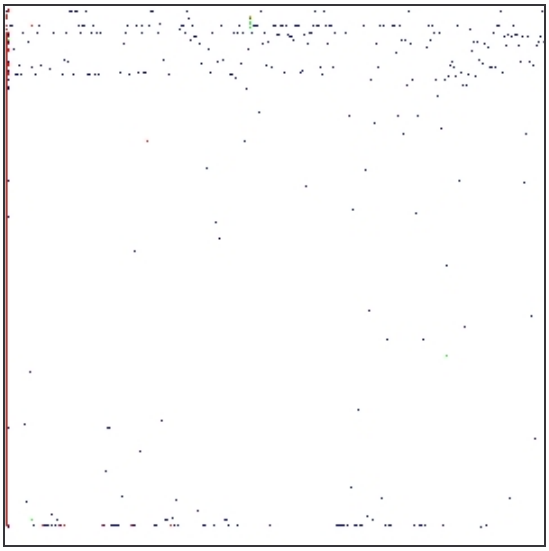
Rysunek 92. Cer BSSID-Size.



Rysunek 93. Inne BSSID-Size.



Rysunek 94. All BSSID-Size.

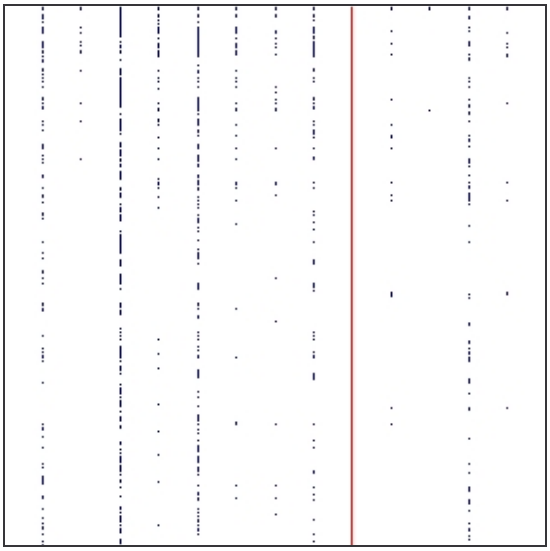


Rysunek 95. Cerl BSSID-Size film.

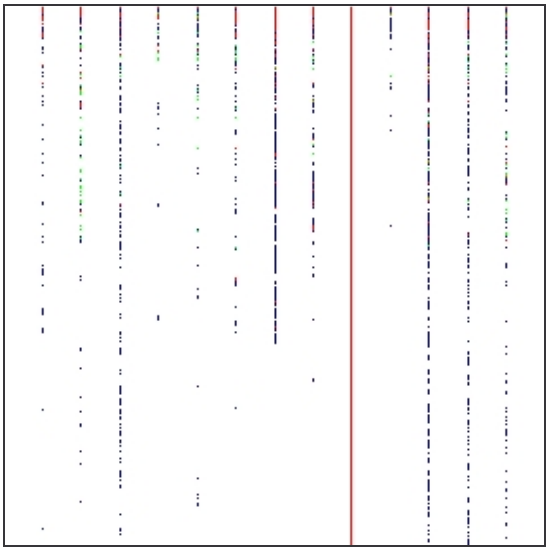
Analiza ta pokazuje, że sieci korzystają z różnej wielkości pakietów, ale z przeważeniem pakietów małych. Spowodowane jest to tym, że w sieci jest dużo pakietów związanych z utrzymaniem sieci, które są możliwie małych wielkości. Widać też, że niektóre wartości są faworyzowane, co oznacza, że cześć popularnego ruchu korzysta z pakietów stałej wielkości. Kolejną rzeczą, jaką widać to fakt, że tak jak w poprzedniej analizie część BSSID

transmituje wyraźnie więcej pakietów, które są mocno zróżnicowane. Na analizie „All" widać różne rozdzielone BSSID różnych sieci, ale należy pamiętać, że taka analiza jest mało powtarzalna ze względu na to, że nr BSSID traktowany jest jako tekst. Uniemożliwia to wykorzystanie do bezpośredniego rozpoznania danej sieci. Z analizy filmowej można wywnioskować, że większość ruchu w danej sieci odbywa się w ramach jednego BSSID. Prawdopodobnie jest to punkt dostępowy wykorzystywany przez mocno obciążoną sieć komputerową (być może jest to sieć osiedlowa z dużą ilością klientów).

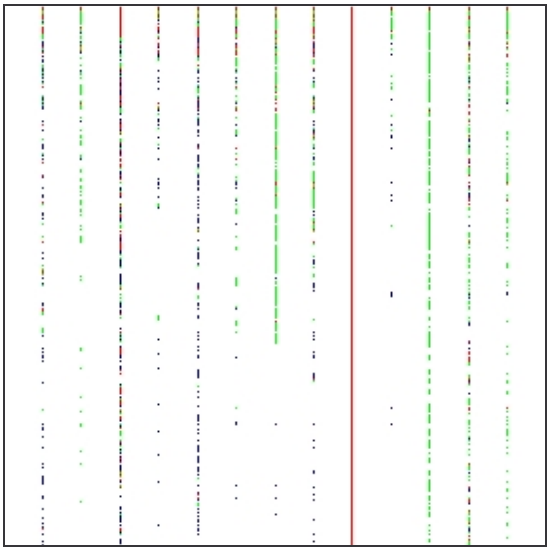
Channel-Protocol



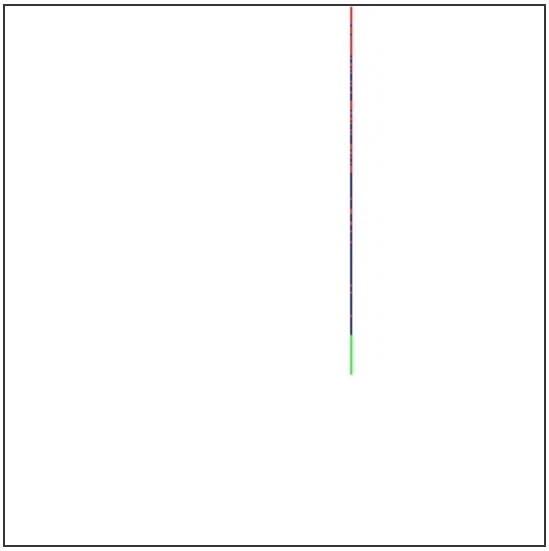
Rysunek 96. Cer Channel-Protocol.



Rysunek 97. Inne Channel-Protocol.



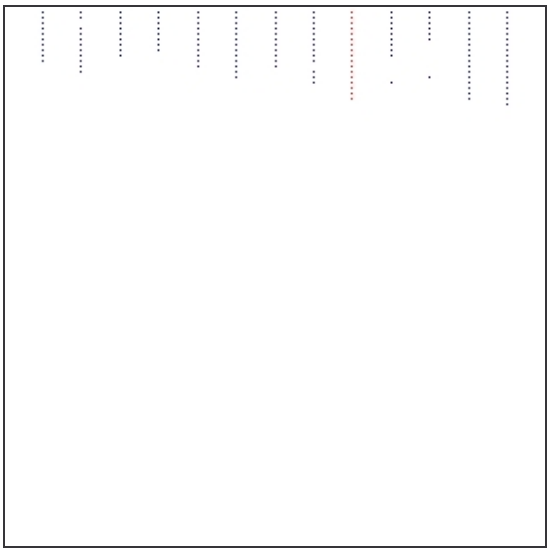
Rysunek 98. All Channel-Protocol.



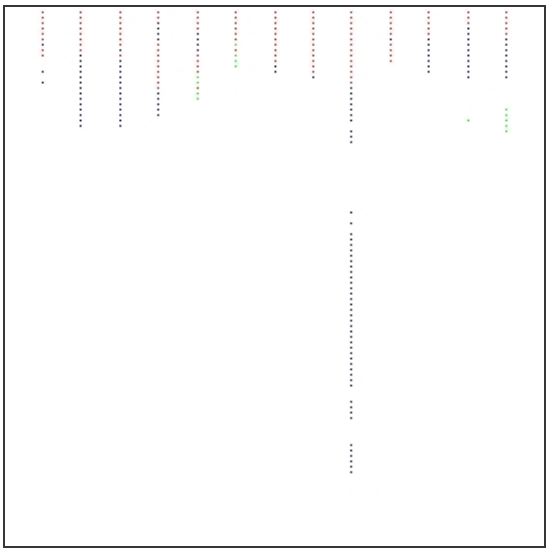
Rysunek 99. Cerl Channel-Protocol film.

Analiza ta pokazuje podobne wyniki do analizy BSSID-Chanel. Dyskretność wartości parametru Channel powoduje ubogość wyników. Po wypełnieniu punktami można wnioskować o obciążeniu kanałów (3-4 najbardziej obciązone). Widać też, że na różnych kanałach rozkład protokołów jest w miarę jednostajny. Drobne rozbieżności mogą być spowodowane różną ilością przesyłanych pakietów na innych kanałach.

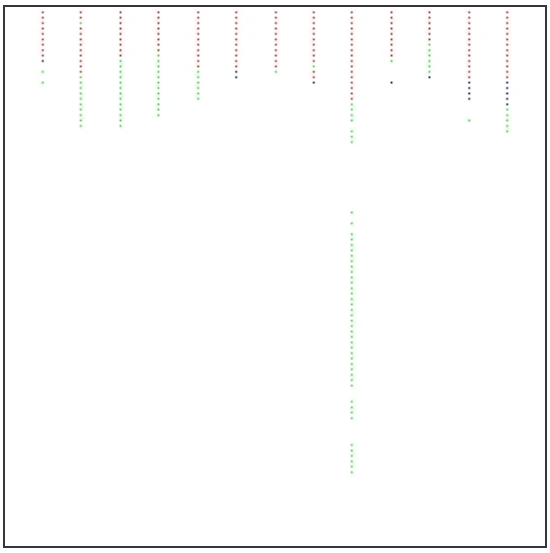
Channel-Signal



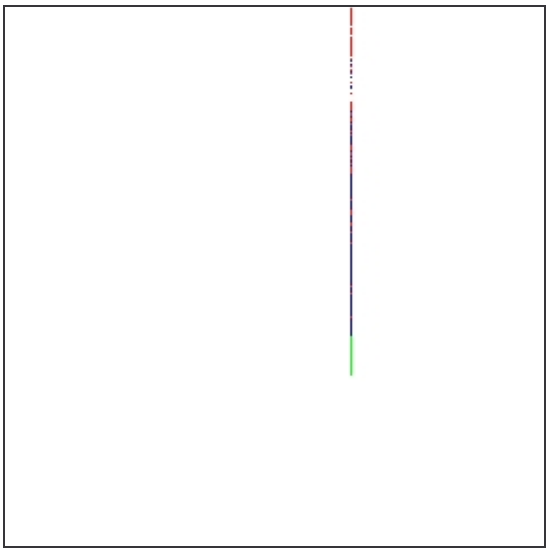
Rysunek 100. Cer Channel-Signal.



Rysunek 101. Inne Channel-Signal.



Rysunek 102. All Channel-Signal.

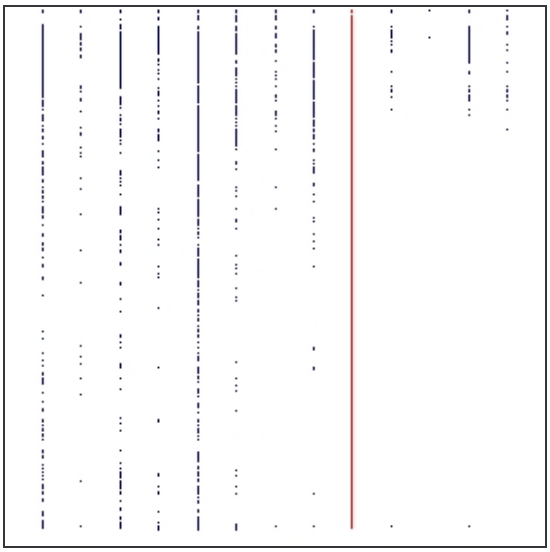


Rysunek 103. Cerl Channel-Signal film.

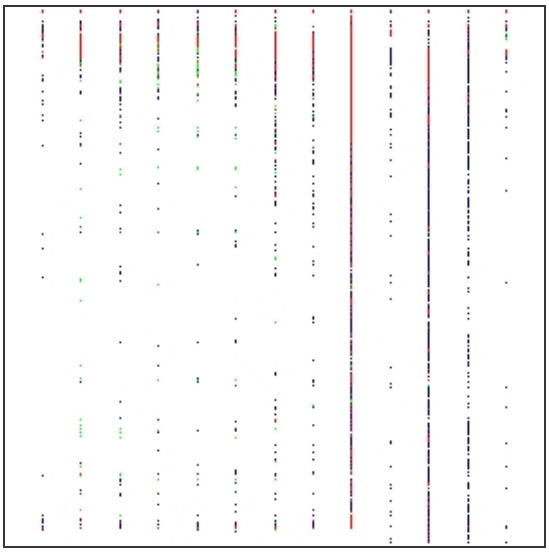
Analiza ta pokazuje różną siłę sygnału na różnych kanałach. Należy pamiętać (jak było to opisane w poprzednich analizach), że w bliskiej odległości (na tym samym obszarze) mogą

jedynie współistnieć sieci na trzech (maksymalnie czterech) rożnych kanałach, ponieważ kanały nachodzą na siebie. Na rysunkach widać cztery maksima, które pochodzą prawdopodobnie od najbliższych sieci względem miejsca skanowania. Jest to typowy obraz jaki można zobaczyć na analizatorze sygnału radiowego. W przypadku analizy filmowej wynik jest zdeformowany z powodu nasłuchu robionym na jednym kanale.

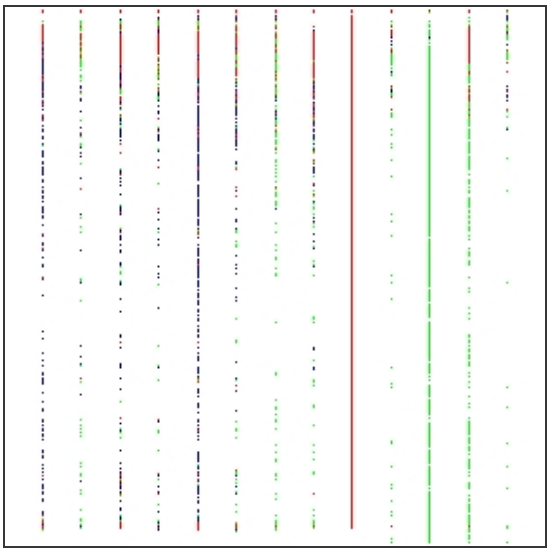
Channel-Size



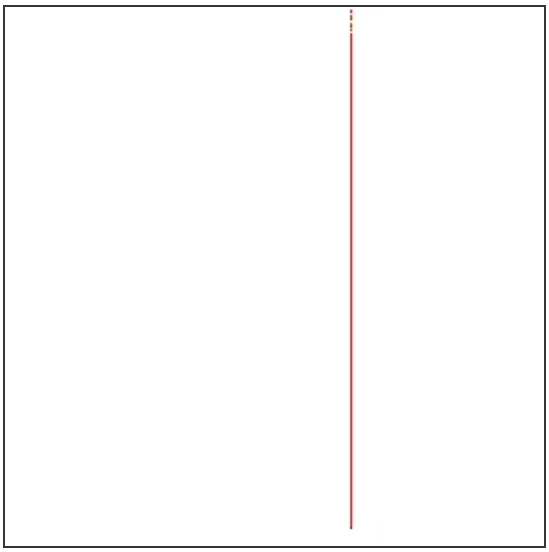
Rysunek 104. Cer Channel-Size.



Rysunek 105. Inne Channel-Size.



Rysunek 106. All Channel-Size.

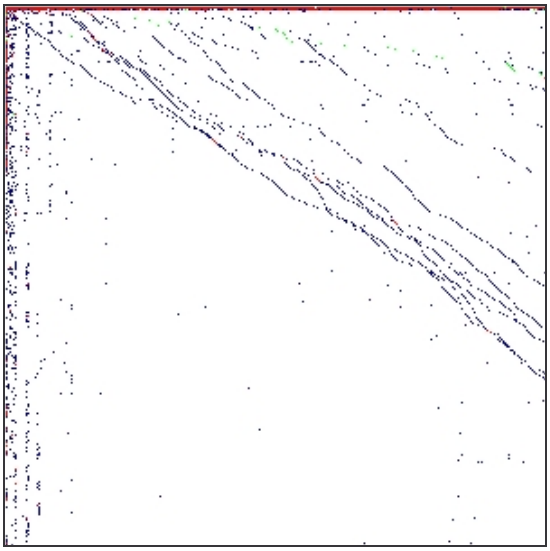


Rysunek 107. Cerl Channel-Size film.

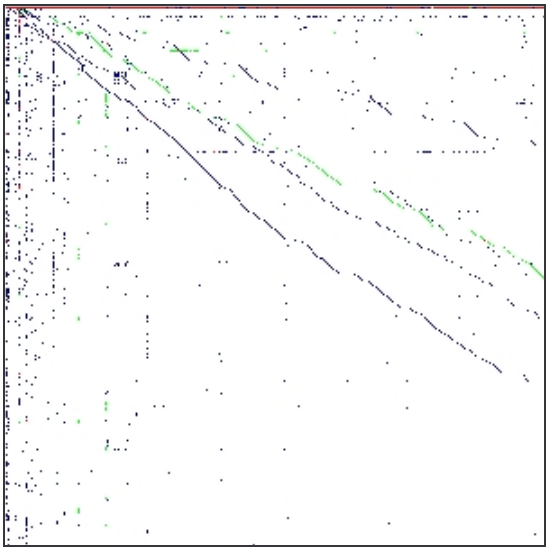
Analiza ta pokazuje dość jednolity rozkład wielkości przekazywanych danych na rożnych kanałach. Ewenualne różnice mogą wynikach z nierównomiernego obciążenia

kanałów. Jak w każdym badaniu z parametrem „Cannal" rozróżniamy maksima oraz degenracje analizy filmowej.

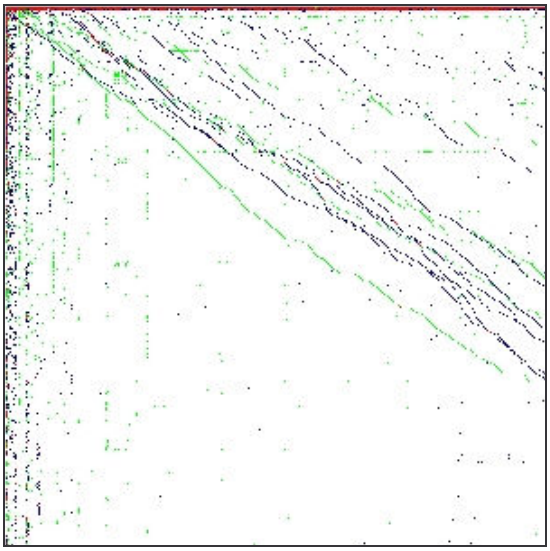
Destination-BSSID



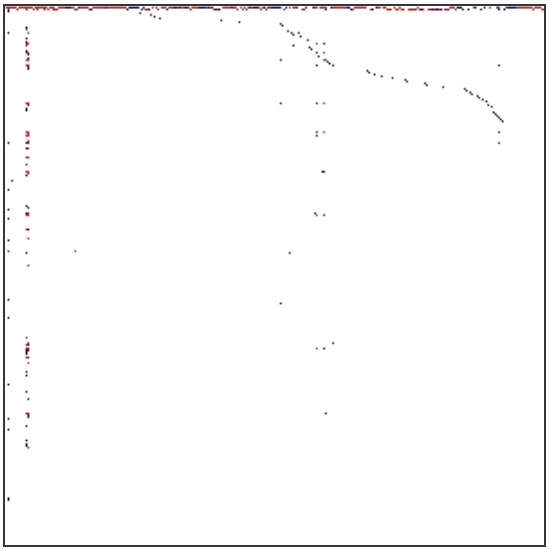
Rysunek 108. Cer Destination-BSSID.



Rysunek 109. Inne Destination-BSSID.



Rysunek 110. All Destination-BSSID.

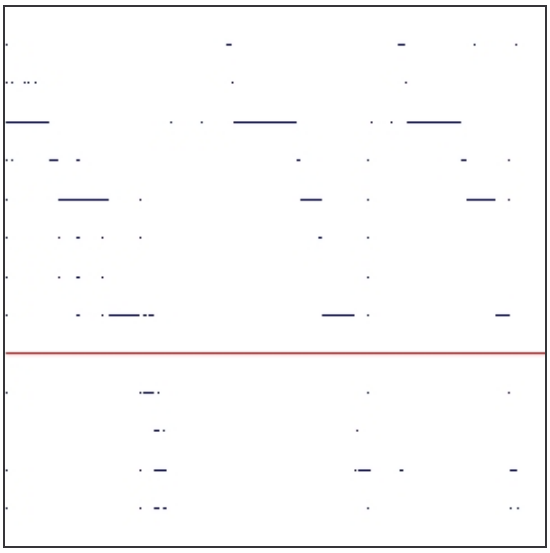


Rysunek 111. Cerl Destination-BSSID film.

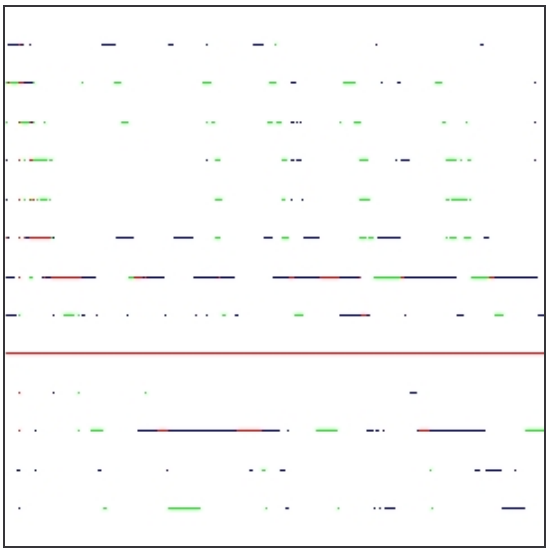
Analiza ta pokazuje trzy grupy ruchu. Pierwszy rodzaj to taki, w którym w ramach jednego BSSID jest wiele adresów docelowych (wiele klientów z jednym punktem dostępowym), Wiele numerów BSSID z jednym adresem docelowym (być może roaming miedzy punktami dostępowymi), oraz jeden adres docelowy na jeden BSSID (jest to efekt

uczenia się analizatora wartości tekstowych). Kąt nachylenia tej linii sugeruje, że obie wartości narastają porównywalnie do siebie. Analiza filmowa pokazuje, że większość ruchu odbywa się w ramach jednego BSSID i prawdopodobnie w ramach tego BSSID jest punkt dostępowy, do którego kierowana jest większość pakietów.

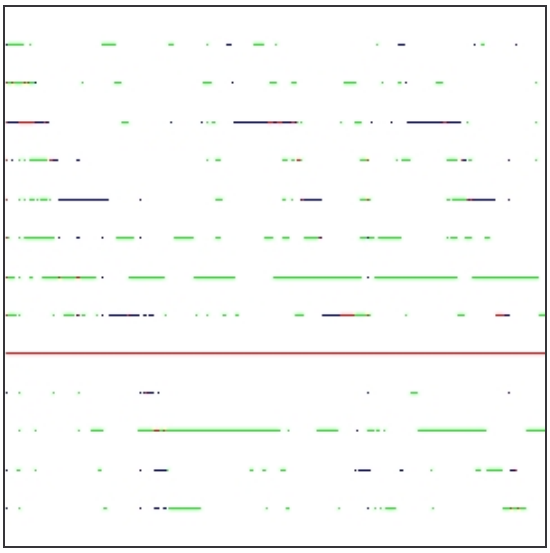
Destination-Channel



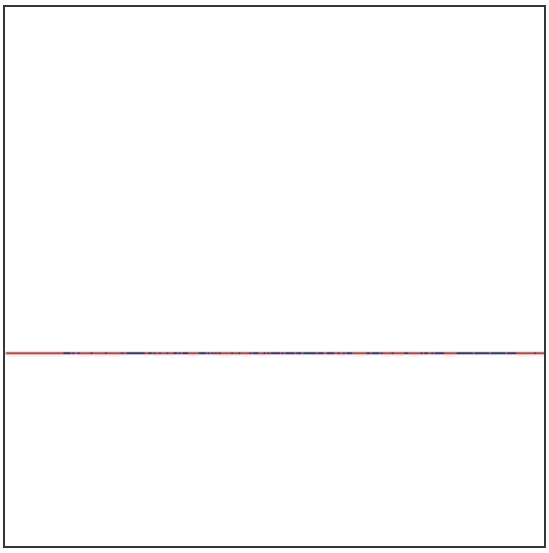
Rysunek 112. Cer Destination-Channel.



Rysunek 113. Inne Destination-Channel.



Rysunek 114. All Destination-Channel.

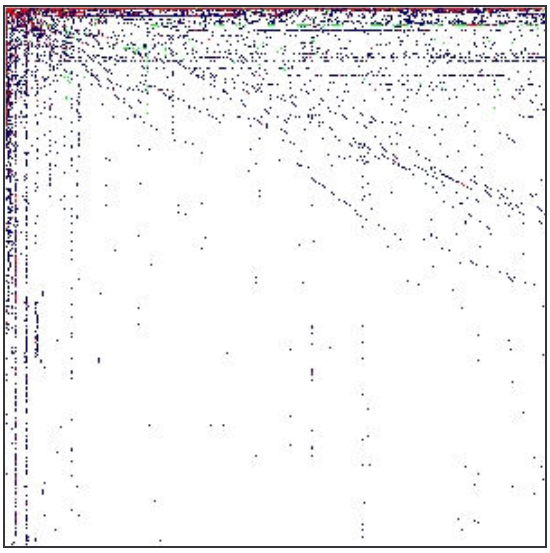


Rysunek 115. Cerl Destination-Channel film.

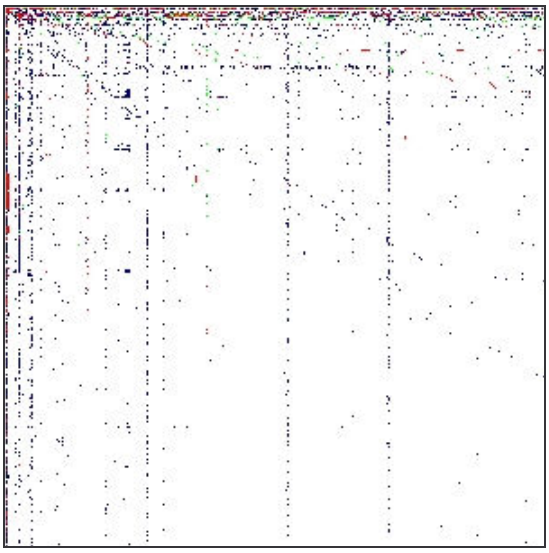
Poza standardowymi wnioskami z analizy, w której jednym z parametrów jest „Channel", na obrazach powyżej można zauważyć ciekawą właściwość. Niektóre urządzenia docelowe korzystają z różnych kanałów (roaming, bądź poszukiwanie sieci do załogowania), ale większość urządzeń docelowych wykorzystuje jeden kanał. Niestety analiza filmowa

pochodzi z nasłuchu jednego kanału, natomiast analizy „Cer", „Inne" oraz „All" zostały utworzone przez nałożenie wielu obrazów. Jeżeli jednak powrócimy do rysunku 41, który nie jest składanym obrazem sieci, przewidywane przeze mnie zachowanie jest tam widoczne.

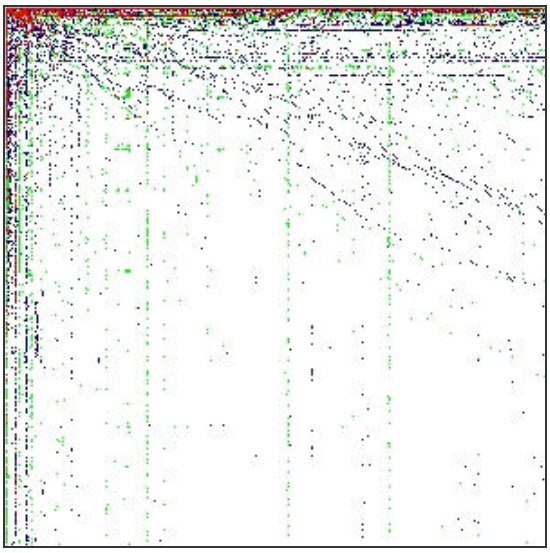
Destination-Protocol



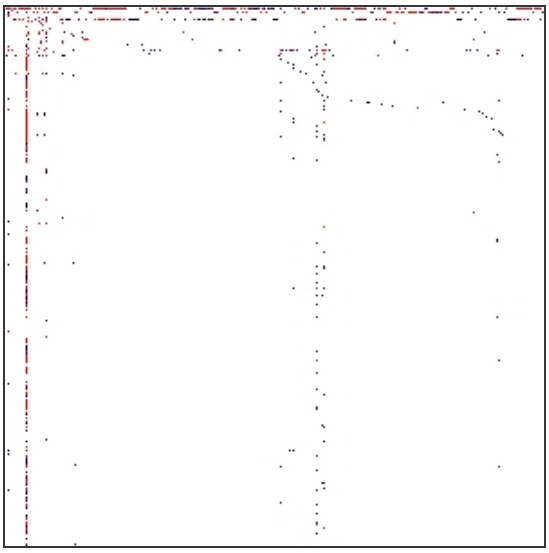
Rysunek 116. Cer Destination-Protocol.



Rysunek 117. Inne Destination-Protocol.



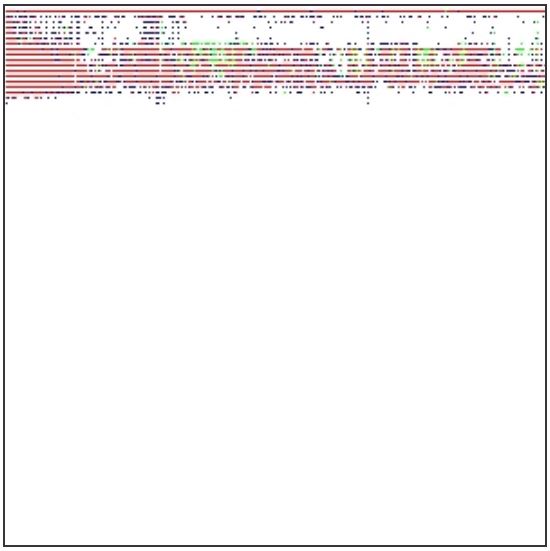
Rysunek 118. All Destination-Protocol.



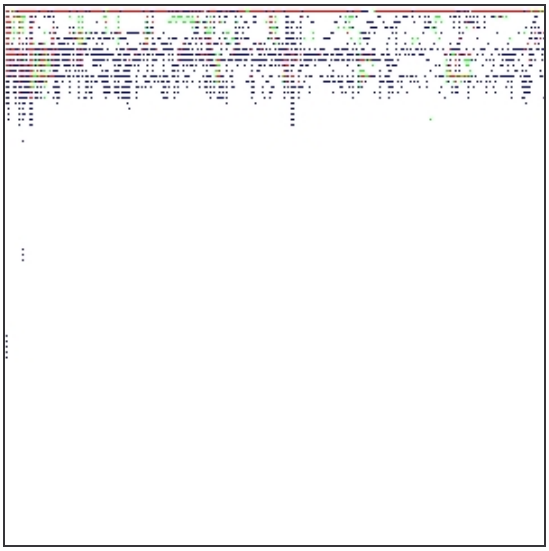
Rysunek 119. Cerl Destination-Protocol film.

Analiza ta pokazuje podobną zależność jak „BSSID-Protocol". Jest to związanie z tym, że większość urządzeń korzysta z jednego punktu dostępowego.

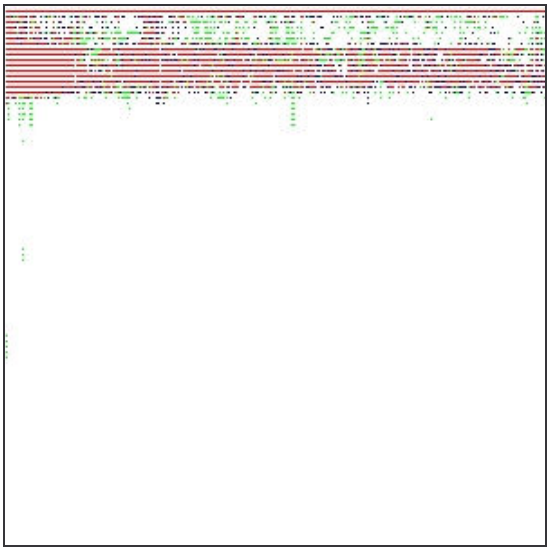
Destination-Signal



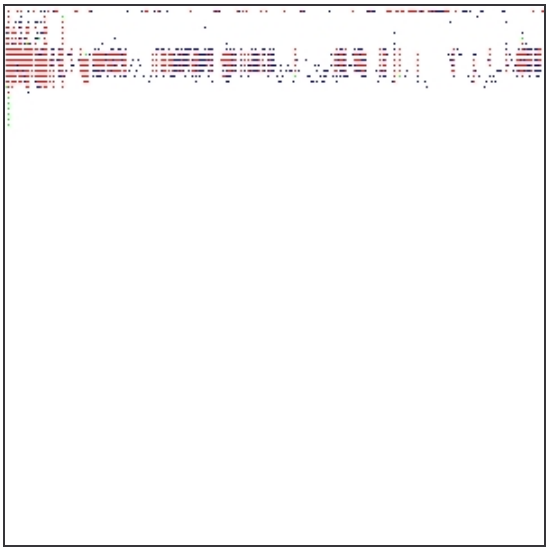
Rysunek 120. Cer Destination-Signal.



Rysunek 121. Inne Destination-Signal.



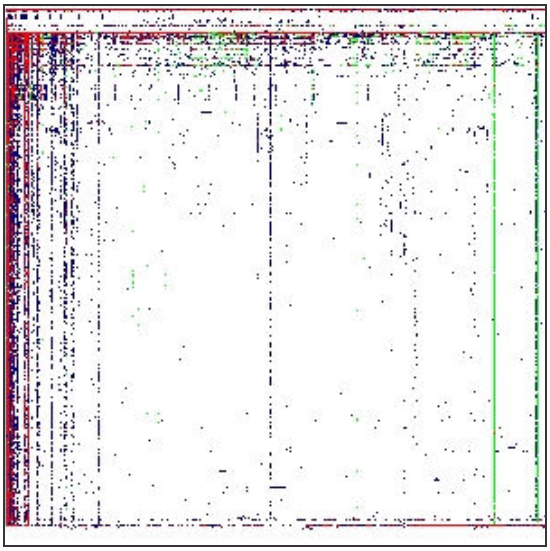
Rysunek 122. All Destination-Signal.



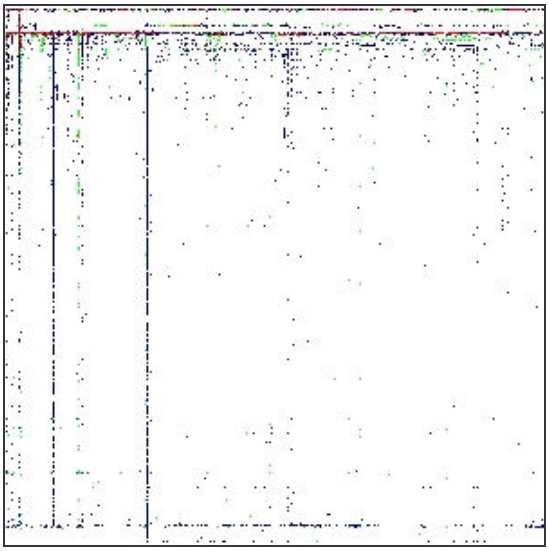
Rysunek 123. Cerl Destination-Signal film.

Analiza ta pokazuje, że różne adresy docelowe są nadawane z różnymi siłami sygnału (od 0 do wartości maksymalnej). Oznacza to, że adresy docelowe są wybierane przez różne urządzenia docelowe. Część adresów docelowych ma ograniczony przedział wartości sygnału, co sugerowałoby, że ruch do tego adresu odbywa się z jednego urządzenia, bądź z grupy urządzeń o podobnym sygnale. Analiza filmowa wskazuje jednak, że taki ruch pasuje bardziej do grupy urządzeń o podobnej odbieranej sile sygnału.

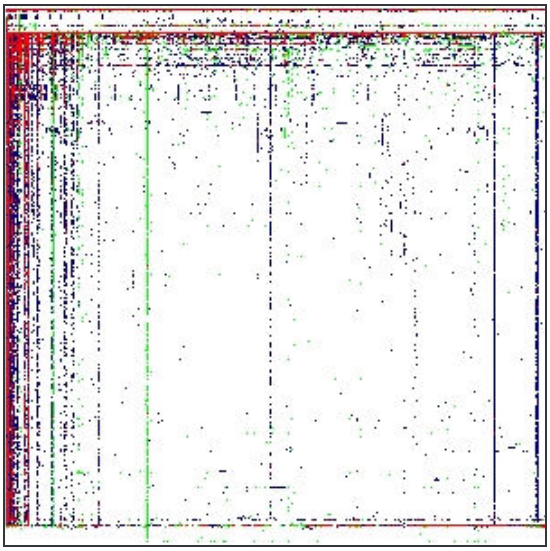
Destination-Size



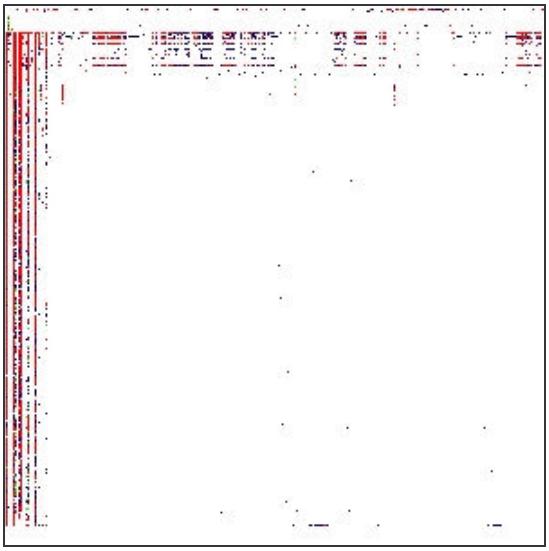
Rysunek 124. Cer Destination-Size.



Rysunek 125. Inne Destination-Size.



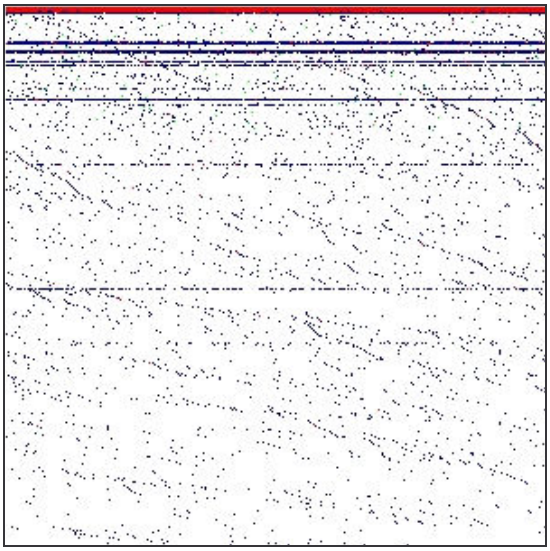
Rysunek 126. All Destination-Size.



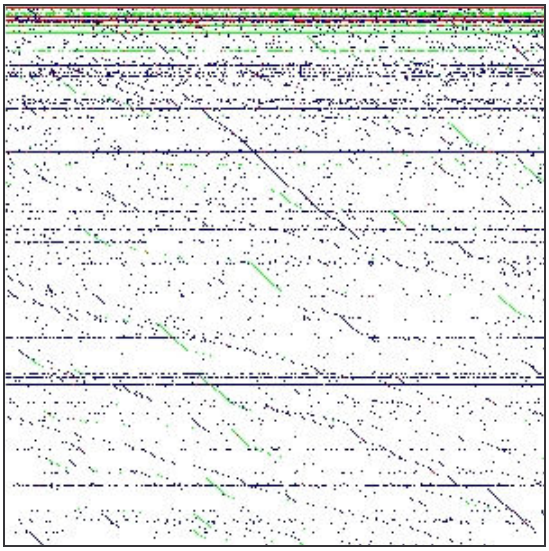
Rysunek 127. Cerl Destination-Size.

Analiza ta jest podobna do analizy „BSSID-size”. Jest to spowodowane tym, że dane urządzenie korzysta z jednego punktu dostępowego.

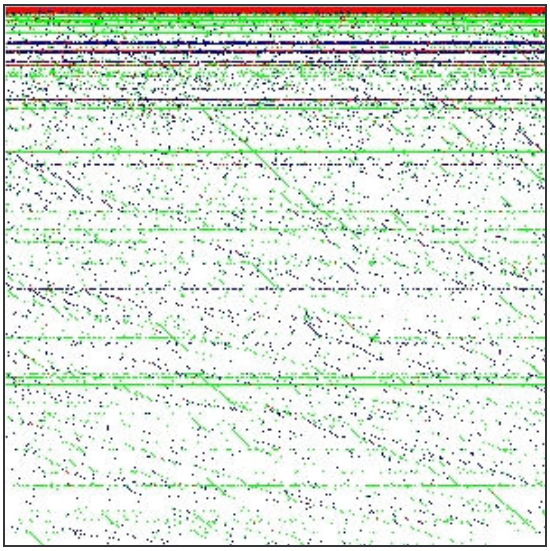
Nr-BSSID



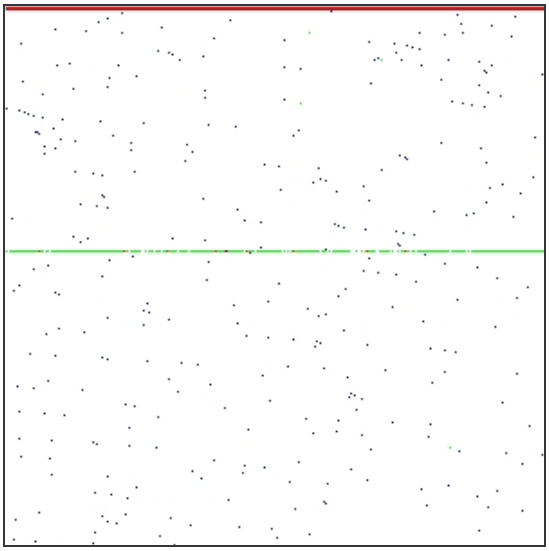
Rysunek 128. Cer Nr-BSSID.



Rysunek 129. Inne Nr-BSSID.



Rysunek 130. All Nr-BSSID.

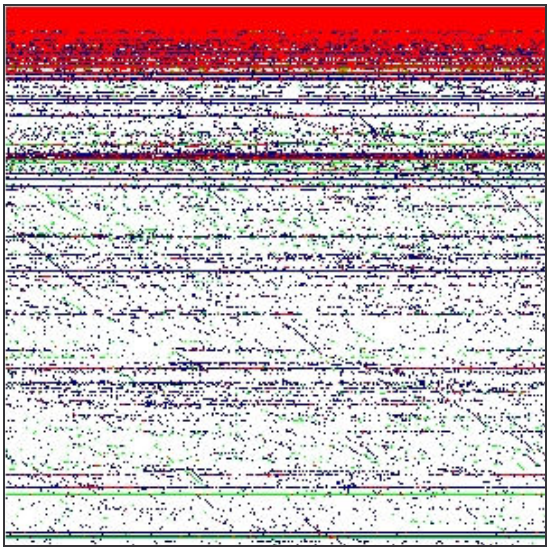


Rysunek 131. Cerl Nr-BSSID film.

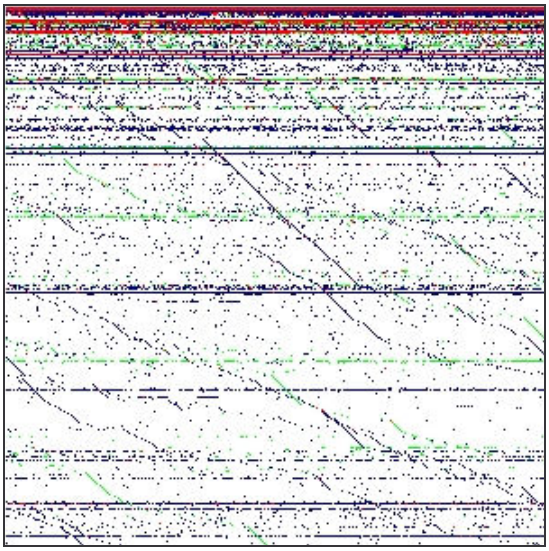
Badania, w których jednym z parametrów jest „Nr" pokazują analizę ilościową drugiego parametru, który jest równomiernie rozłożony względem osi z parametrem „Nr". Po gęstości punktów tworzących linie możemy wnioskować o ilości pakietów powiązanych z drugim paramerem.

Analiza ta pokazuje, że w ramach niektórych BSSID odbywa się zdecydowanie większy ruch (poziome linie). Ale większość urządzeń w ramach jednego numeru BSSID raczej mało przesyła pakietów. Linie pod kątem to efekt uczenia się analizatora nowych wartości tekstowych. Analiza filmowa pokazuje wyraźnie, ze ruch odbywa się głównie za pomocą dwóch BSSID.

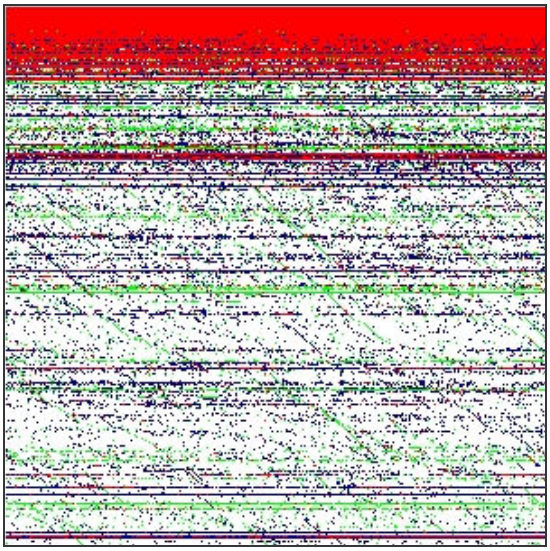
Nr-Destination



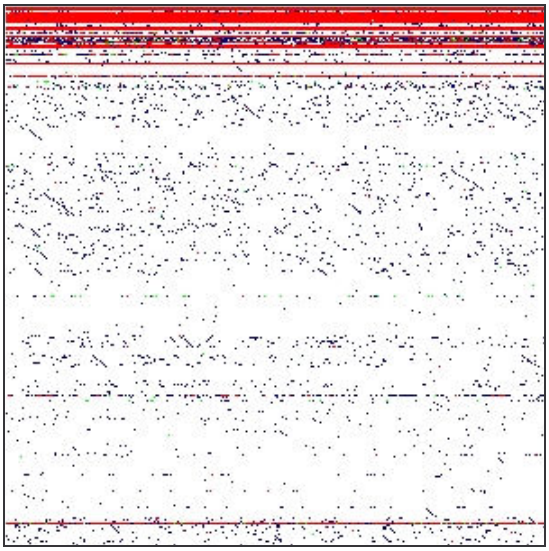
Rysunek 132. Cer Nr-Destination.



Rysunek 133. Inne Nr-Destination.



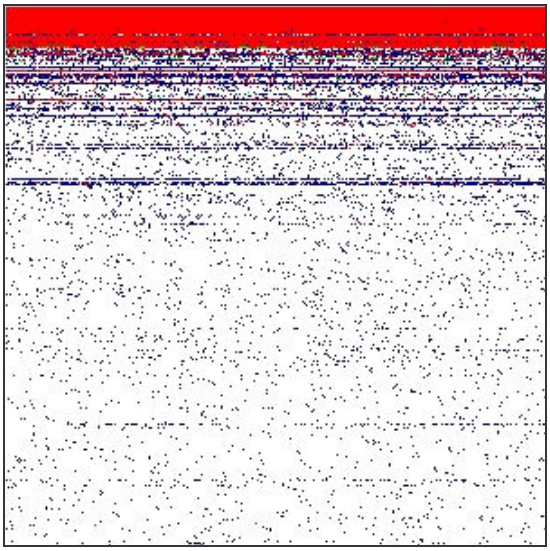
Rysunek 134. All Nr-Destination.



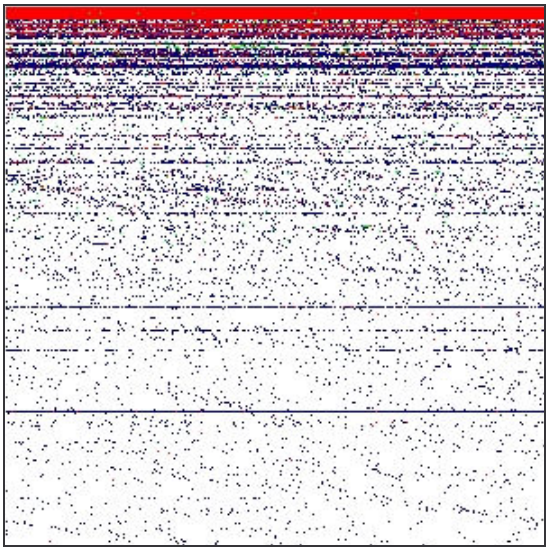
Rysunek 135. Cerl Nr-Destination film.

Analiza ta jest nieco podobna do poprzedniej, ale widać znaczące różnice. Oznacza to, że niektóre adresy docelowe odbierają zdecydowanie więcej pakietów. Wydaje się więc, że w ramach jednego BSSID różne urządzenia powinny generować różny ruch, lub (co jest bardziej prawdopodobne) duży ruch jest kierowany na adresy MAC powiązane z numerem BSSID, w ramach którego odbywa się duży ruch. Widać, że duży ruch jest generowany na wiele adresów docelowych, co sugeruje, że jest to układ „punkt dostępowy" i wielu aktywnych klientów (np. siec osiedlowa).

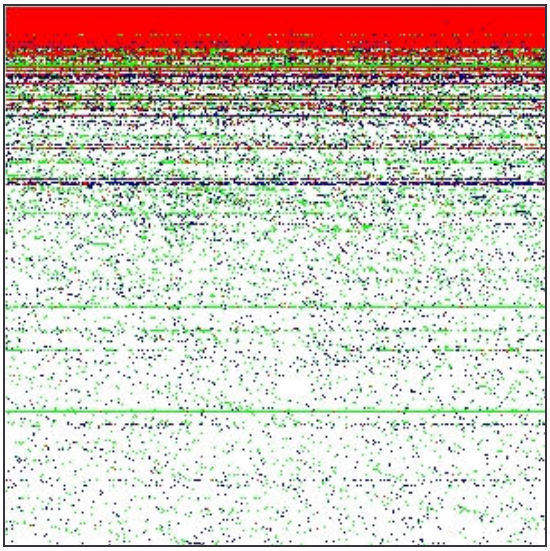
Nr-Protocol



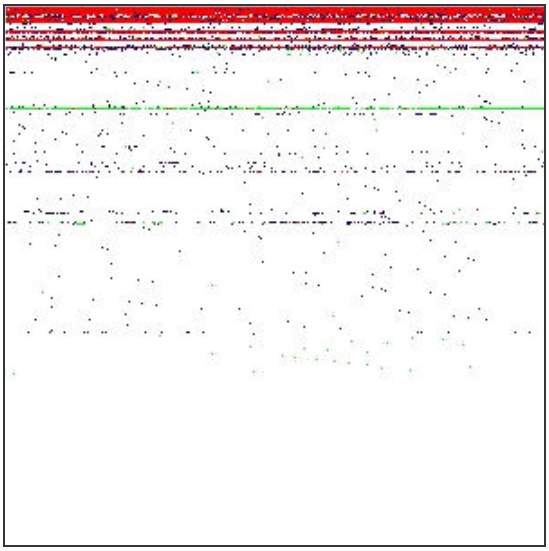
Rysunek 136. Cer Nr-Protocol.



Rysunek 137. Inne Nr-Protocol.



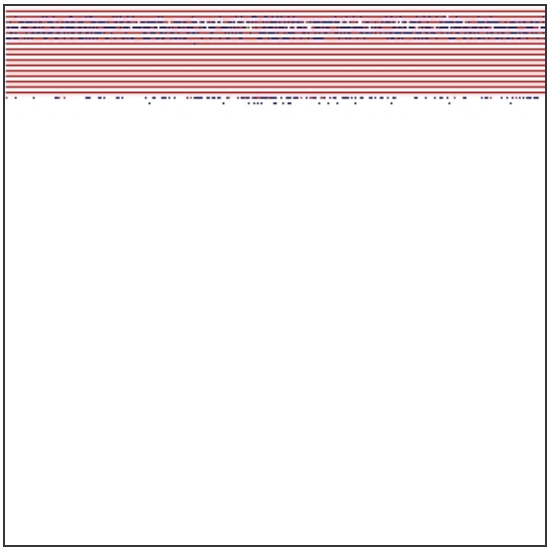
Rysunek 138. All Nr-Protocol.



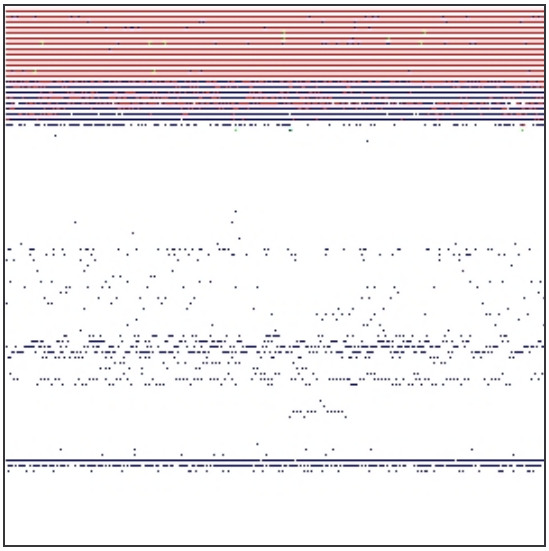
Rysunek 139. Cerl Nr-Protocol film.

Analiza ta pokazuje, że niektóre protokoły w sieci występują częściej. Zapewne najczęściej spotykanymi protokołami będą te związanie z utrzymaniem sieci. Nie widać jakiejś jakościowej różnicy między analizami „Cer" i „Inne", a jedynie analiza „Inne" zawiera mniej punktów, ponieważ została złożona z mniejszej liczby pakietów.

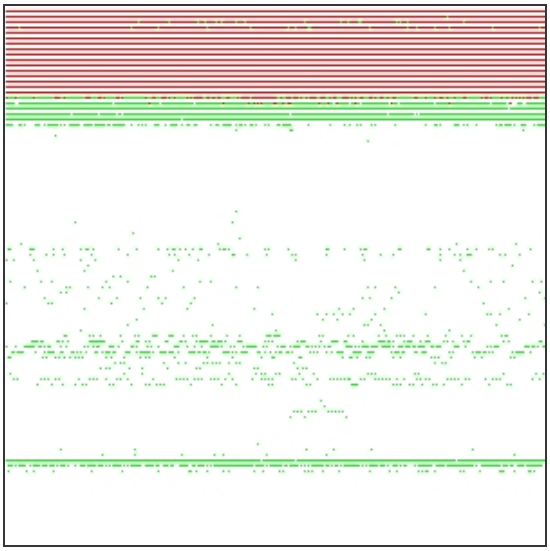
Nr-Signal



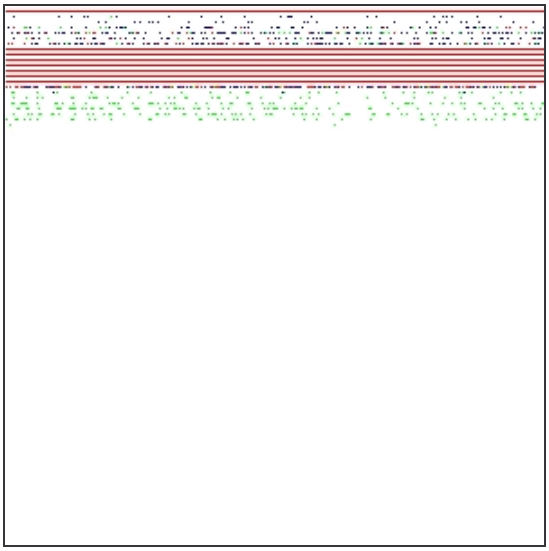
Rysunek 140. Cer Nr-Signal.



Rysunek 141. Inne Nr-Signal.



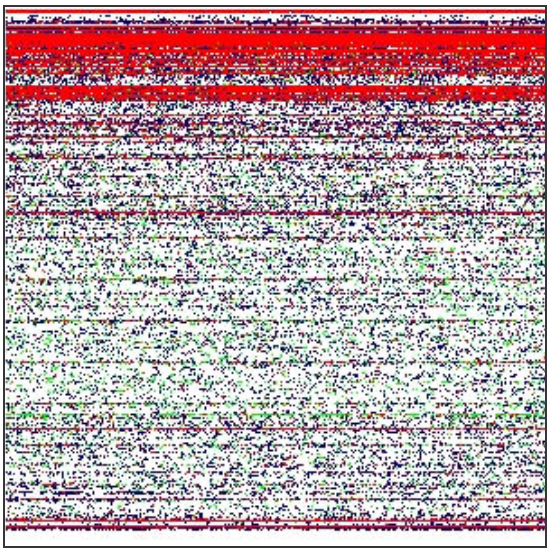
Rysunek 142. All Nr-Signal.



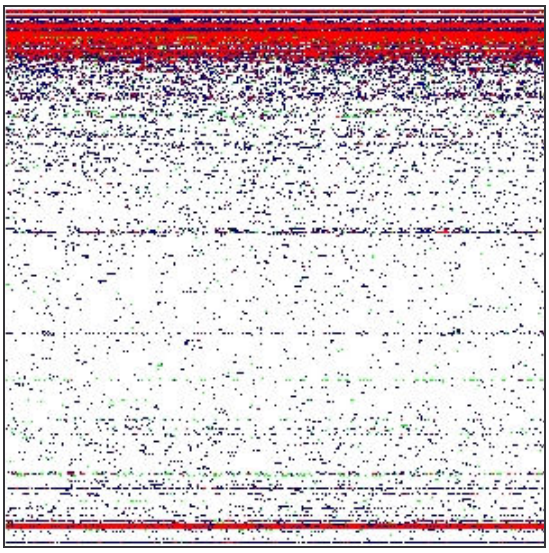
Rysunek 143. Cerl Nr-Signal film.

Analiza „All" pokazuje, że w sieciach można rozpoznać parę maksimów oraz stały duży ruch o niezbyt dużym sygnale. Maksima te pochodzą prawdopodobnie od silnych klientów, ponieważ nie zawierają bardzo dużej liczby pakietów. Większość ruchu odbywa się z niskimi sygnałami i tam zapewne umiejscowione są punkty dostępowe. W analizie filmowej znajduje się maksimum, które jest wyraźnie silniejsze. Prawdopodobnie podczas nasłuchu pakietu *cerl1.csv* sieć była dobrze słyszana, lub było zbyt wcześnie (9 rano dnia roboczego) aby były aktywne słabe stacje klienckie i większość ruchu jest ruchem rozgłoszeniowym punktów dostępowych, których sygnały są dobrze słyszalne i są stabilne w czasie.

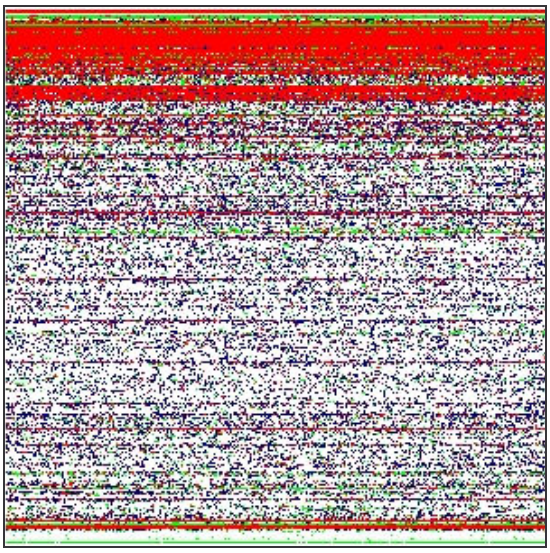
Nr-Size



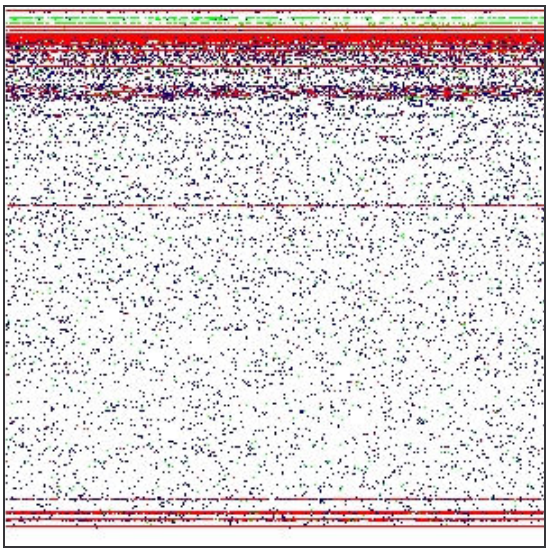
Rysunek 144. Cer Nr-Size.



Rysunek 145. Inne Nr-Size.



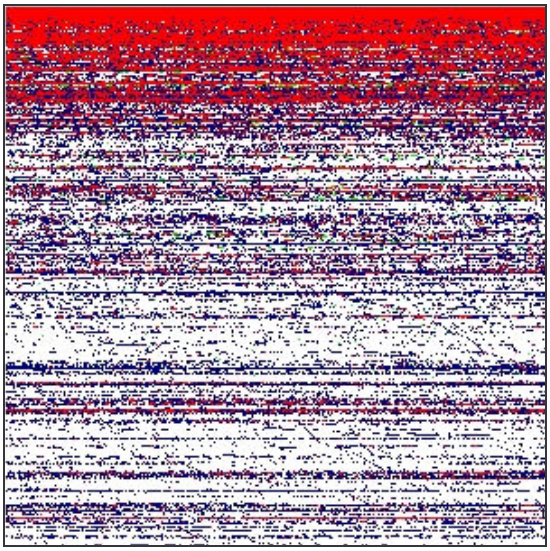
Rysunek 146. All Nr-Size.



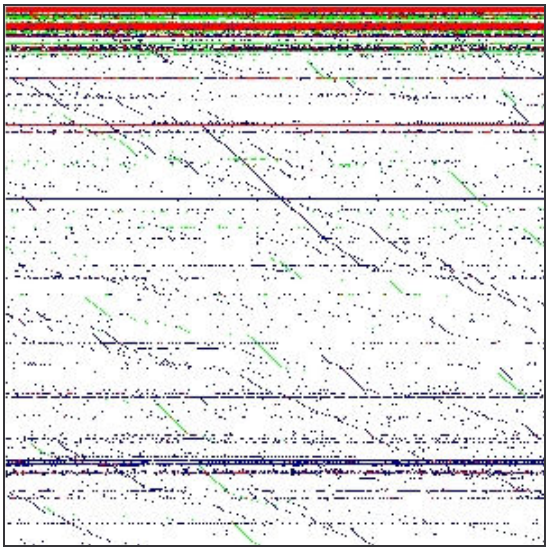
Rysunek 147. Cerl Nr-Size film.

Analiza ta pokazuje, że sieci bezprzewodowe transmitują najczęściej małe pakiety (najczęstsze zapewne są związane z utrzymaniem sieci). Widać też parę większych maksimów większych wartości, które mogą być wynikiem stałych wielkości segmentów pakietów z wyższych warstw.

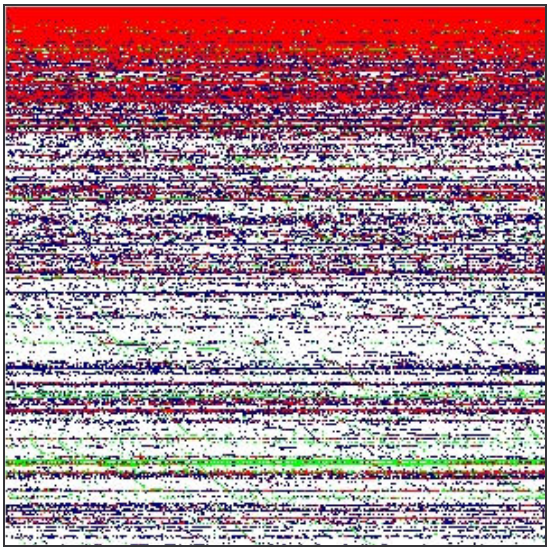
Nr-Source



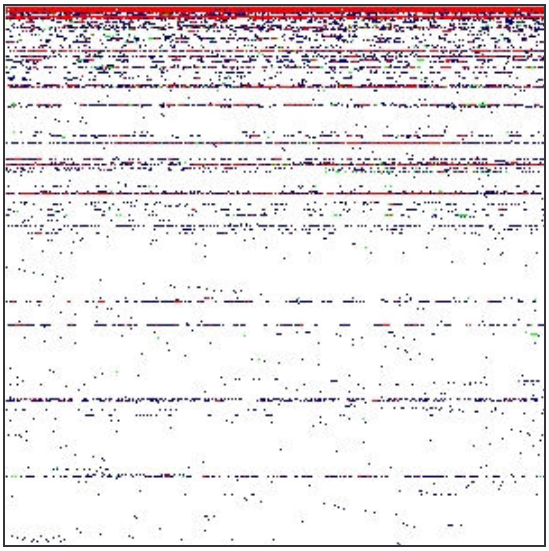
Rysunek 148. Cer Nr-Source.



Rysunek 149. Inne Nr-Source.



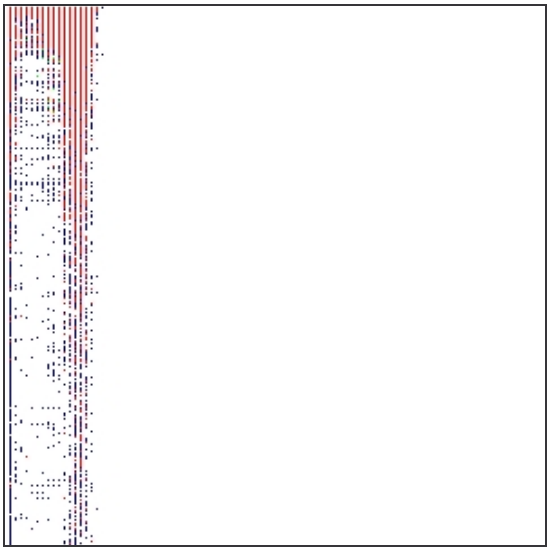
Rysunek 150. All Nr-Source.



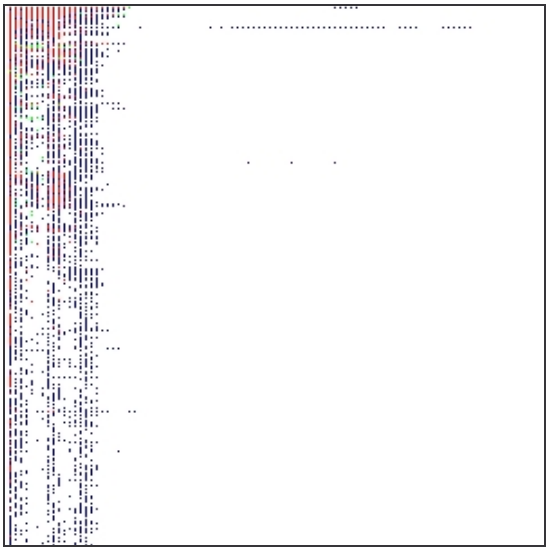
Rysunek 151. Cerl Nr-Source film.

Analiza ta jest nieco różna od analizy „Nr-Destination”, co wydaje się sprzeczne z intuicją. Jest to związane z brakiem symetrii między ruchem wychodzącym a wchodzącym i będzie omówione na analizie „Source-Destination”.

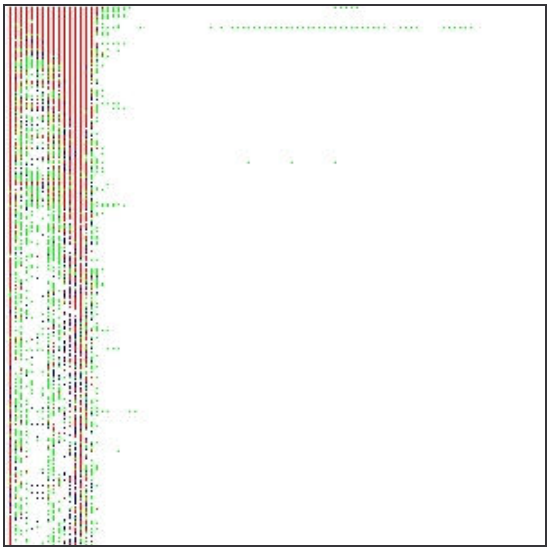
Signal-Protocol



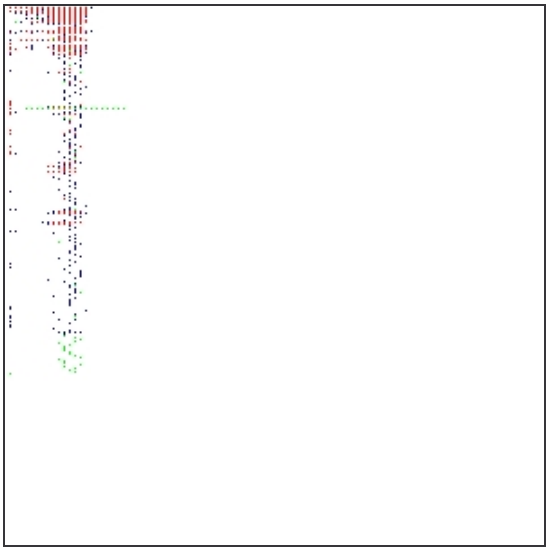
Rysunek 152. Cer Signal-Protocol.



Rysunek 153. Inne Signal-Protocol.



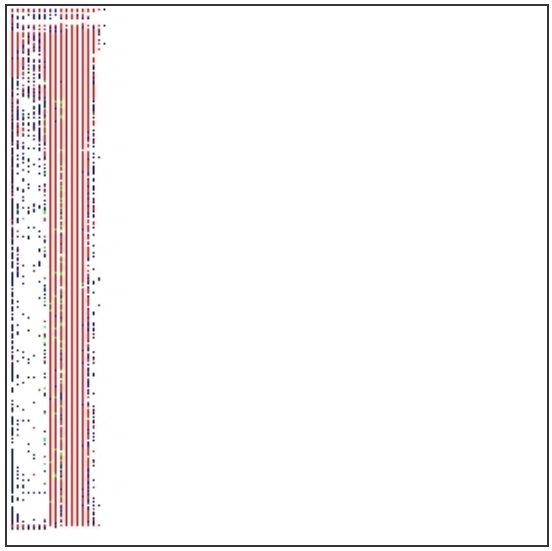
Rysunek 154. All Signal-Protocol.



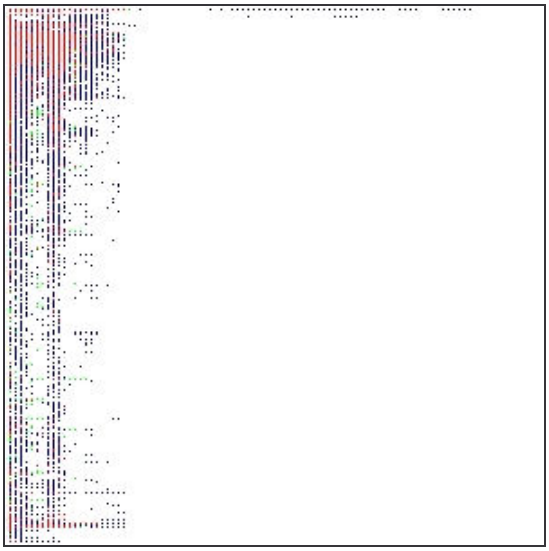
Rysunek 155. Cerl Signal-Protocol film.

Analiza ta pokazuje, że część protokołów jest bardziej popularna. Wypełnienie równomiernie sygnału od 0 do wartości maksymalnej wskazuje na korzystanie z takiego protokołu przez wiele urządzeń. Niejednorodne wypełnienie wskazuje na rzadkie korzystanie z takiego protokołu przez różne urządzenia. Porównując tę analizę z innymi, można wnioskować, że jest to kolejne zobrazowanie, że największy (w znaczeniu ilościowym) ruch w sieci związany jest małymi pakietami utrzymującymi sieć.

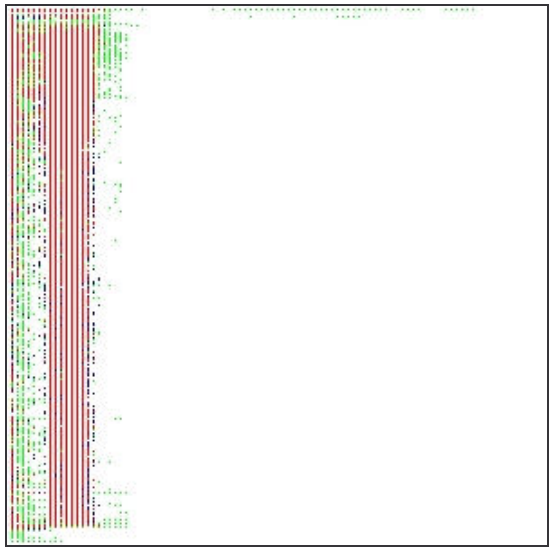
Signal-Size



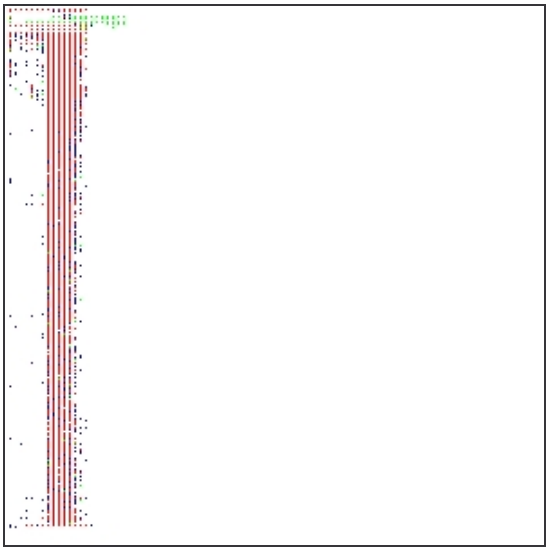
Rysunek 156. Cer Signal-Size.



Rysunek 157. Inne Signal-Size.



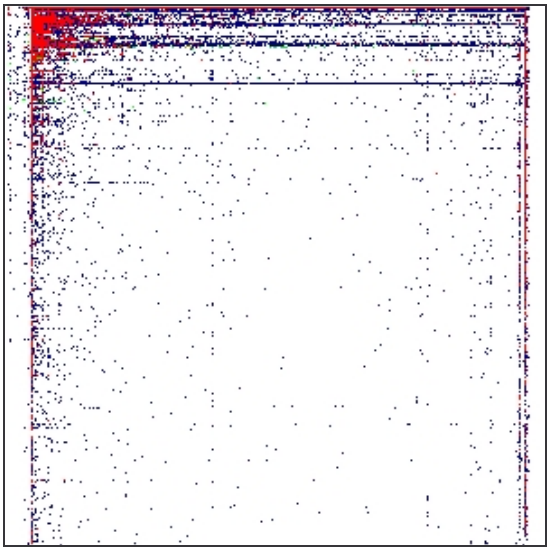
Rysunek 158. All Signal-Size.



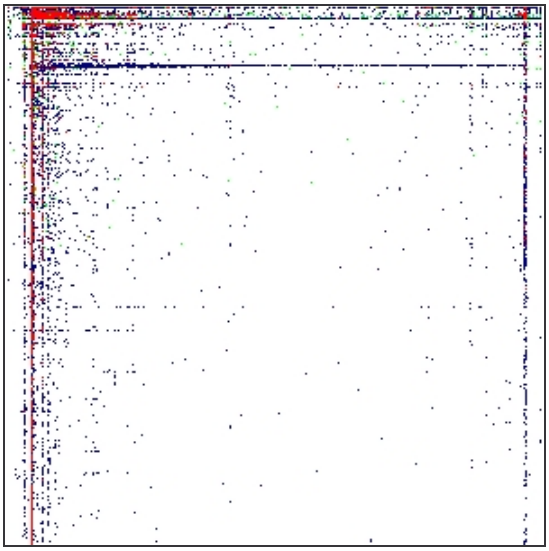
Rysunek 159. Cerl Signal-Size film.

Kolejna analiza wskazująca na częściej występujące pakiety małego rozmiaru. Należy to przypisać utrzymaniu sieci. Można by próbować znaleźć zależność pomiędzy częstością powtórzeń pakietu w zależności od jakości sygnału oraz ilości błędów, ale takie zachowanie jest niezauważalne (częstsze powtarzanie dużych pakietów). Być może reakcja sieci i przełączenie się na mniejszą prędkość niweluje ten efekt. Inną przyczyną takiego stanu rzeczy może być to, że co dla skanera radiowego może się wydawać słabym sygnałem, nie koniecznie jest słabym sygnałem dla korespondenta, ponieważ oba urządzenia mają różne położenie.

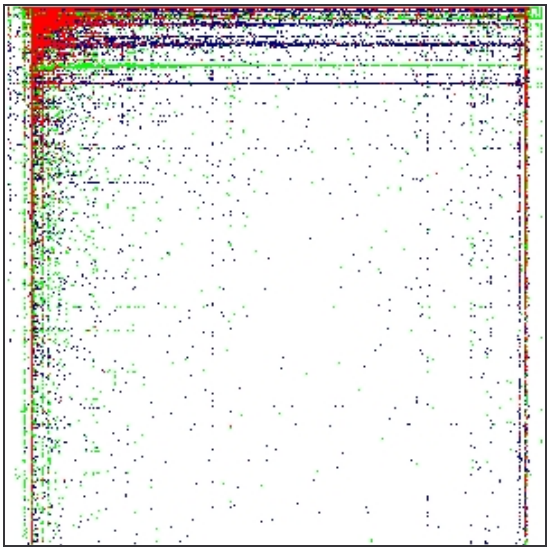
Size-Protocol



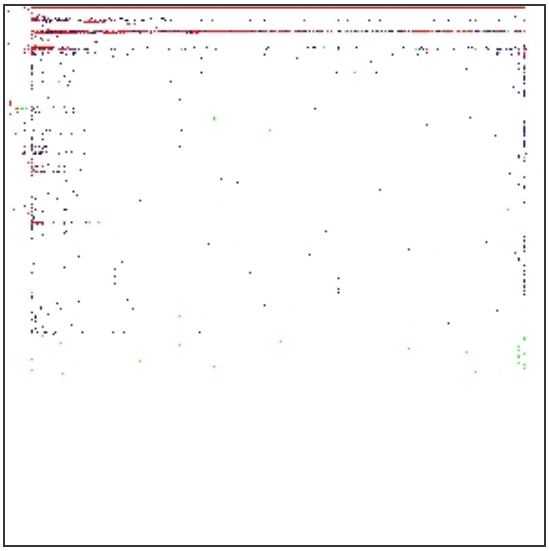
Rysunek 160. Cer Size-Protocol.



Rysunek 161. Inne Size-Protocol.



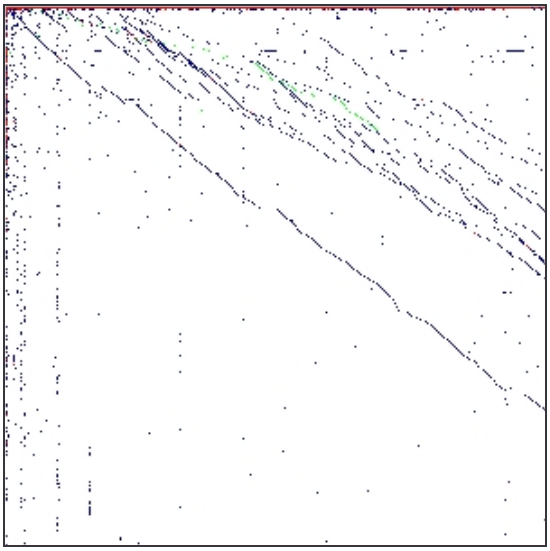
Rysunek 162. All Size-Protocol.



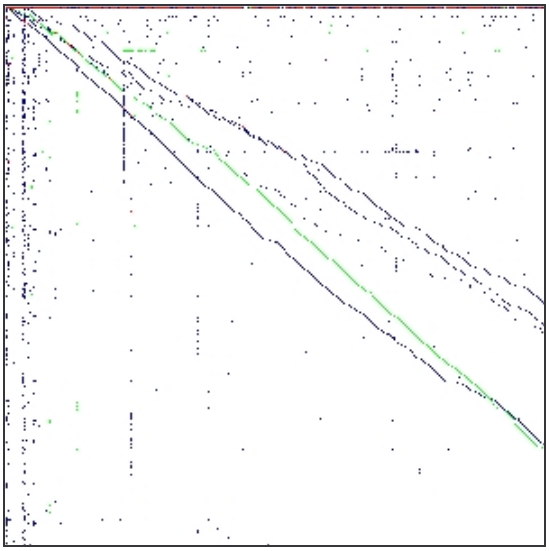
Rysunek 163. Cerl Size-Protocol film.

Analiza „All" pokazuje, że część protokołów korzysta z różnych wielkości pakietu, a część korzysta z wielkości dyskretnych (tylko kilka wartości). Analiza filmowa dostarcza informacji, że kolejna grupa wykorzystuje dokładnie jedną wielkość. Niektóre dyskretne wielkości pakietów są faworyzowane przez większą wielkość pakietów. Prawdopodobnie wielkość tych pakietów jest powiązana z zależnością 2^x, przy czym wielkość pakietu zapisuje się na niewielkiej liczbie bitów.

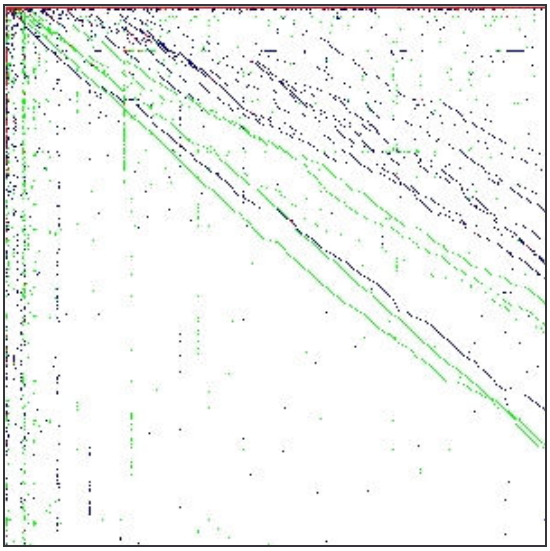
Source-BSSID



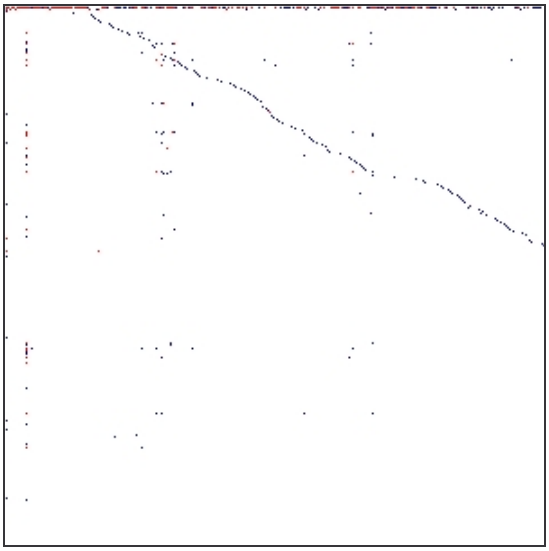
Rysunek 164. Cer Source-BSSID.



Rysunek 165. Inne Source-BSSID.



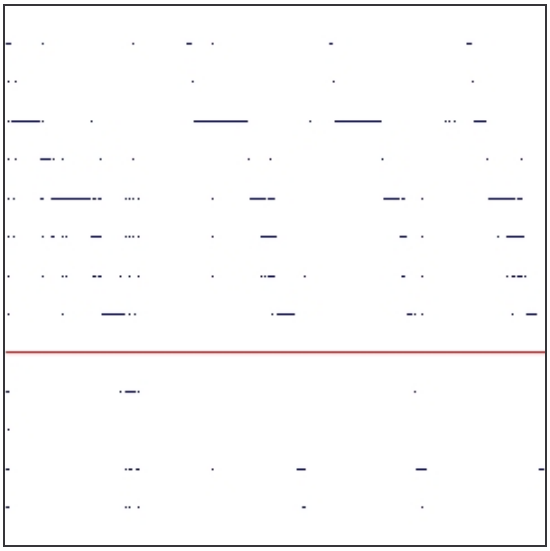
Rysunek 166. All Source-BSSID.



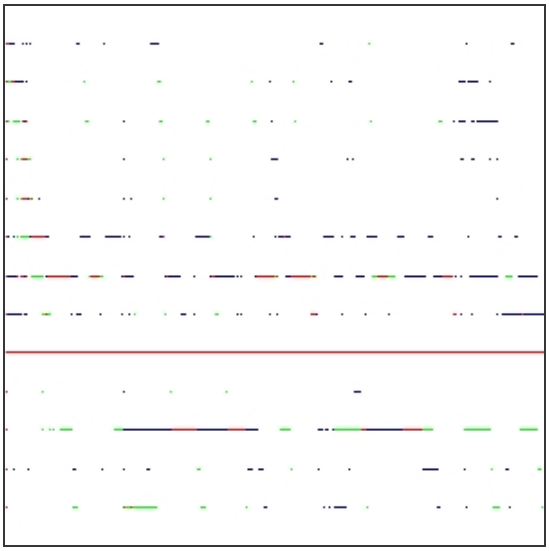
Rysunek 167. Cerl Source-BSSID film.

Analiza ta jest podobna do analizy „Destination-BSSID”. Różnice są omówione podczas analizy „Source-Destination”.

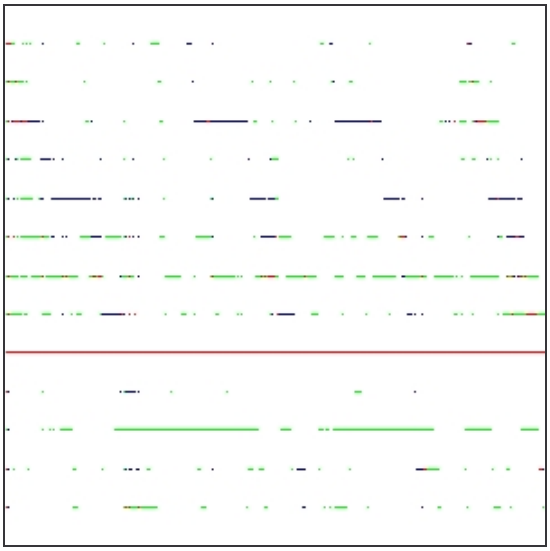
Source-Channel



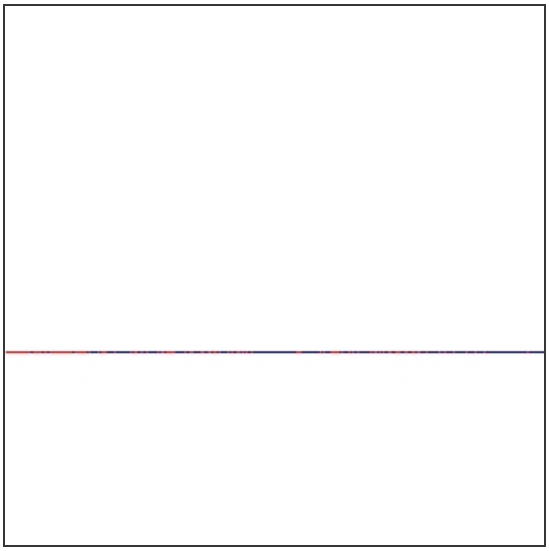
Rysunek 168. Cer Source_Channel.



Rysunek 169. Inne Source_Channel.



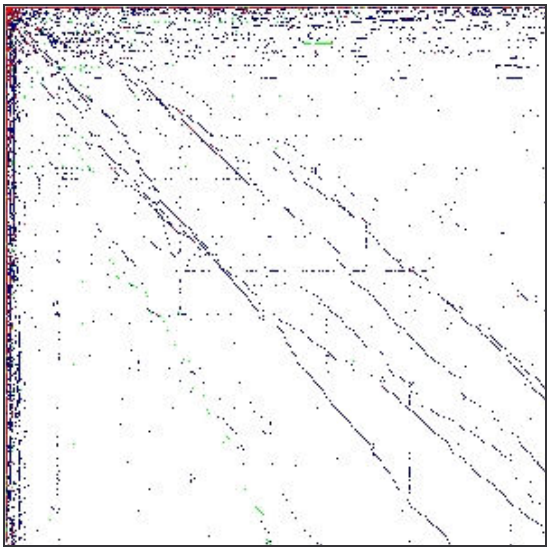
Rysunek 170. All Source_Channel.



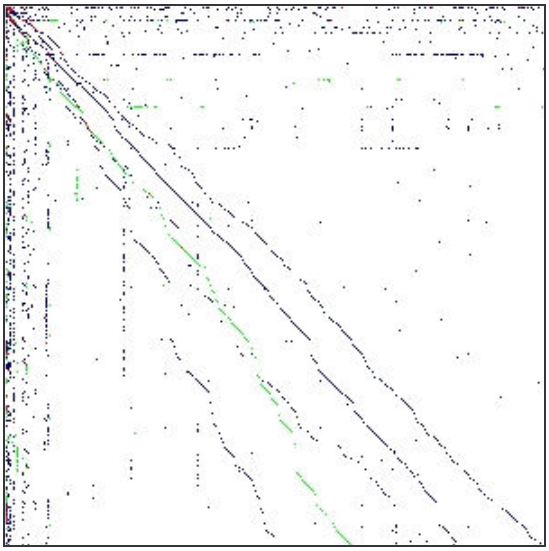
Rysunek 171. Cerl Source_Channel film.

Analiza ta jest podobna do analizy „Destination-Channel". Ponieważ analiza filmowa pochodzi z nasłuchu jednego kanału, nie ma pewności co do tego, czy w analizie „All" ruch jednego MAC adresu na różnych kanałach jest wynikiem zmiany częstotliwości, czy nałożeniem się obrazów z różną kolejnością MAC adresów. Efekt zmiany kanałów przez urządzenia nadawcze można zaobserwować na rysunku 70.

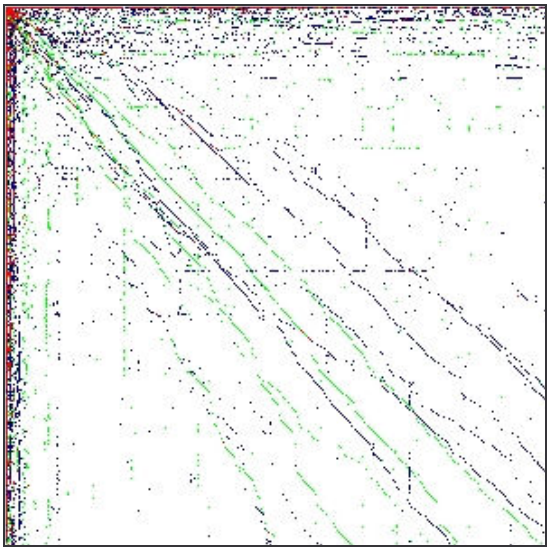
Source-Destination



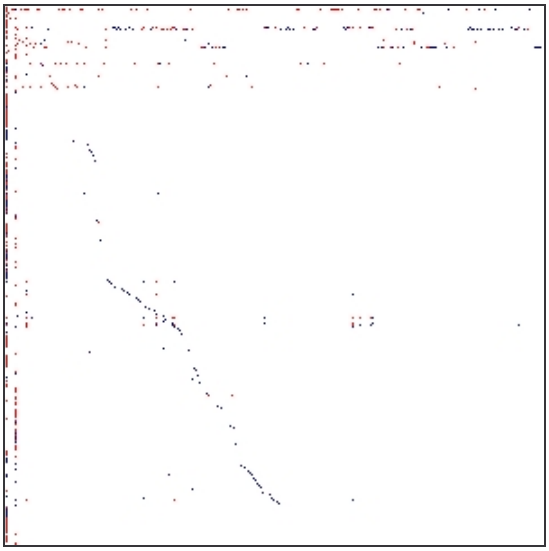
Rysunek 172. Cer Source-Destination.



Rysunek 173. Inne Source-Destination.



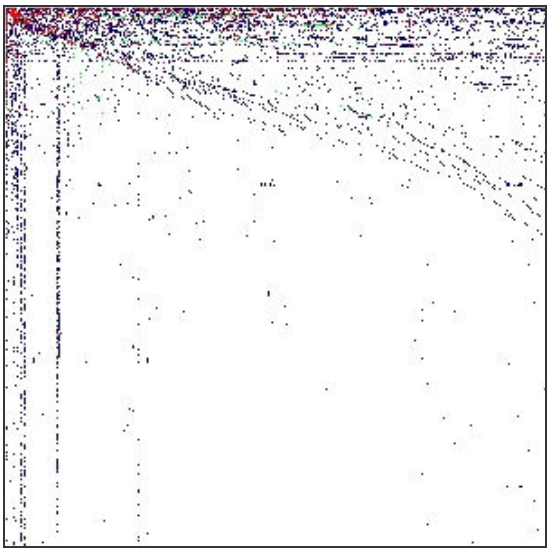
Rysunek 174. All Source-Destination.



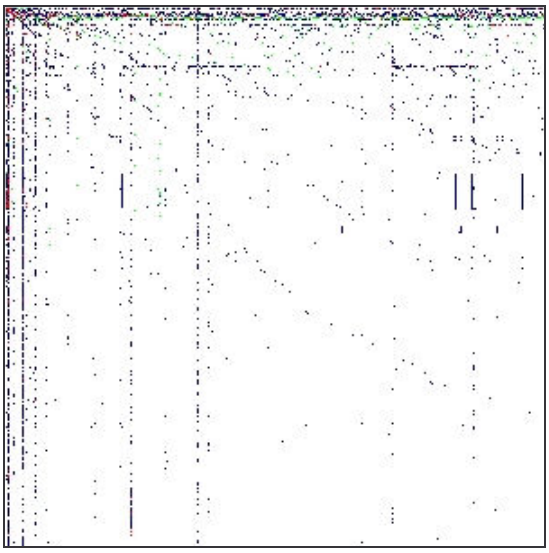
Rysunek 175. Cerl Source-Destination film.

To jedna z ciekawszych analiz. Analiza ta pokazuje, że ruch źródłowo-docelowy nie jest symetryczny. Są maksima zarówno w adresach źródłowych jak i docelowych, ale maksima w adresach źródłowych są wąskie i mocne, a w adresach docelowych szerokie i rozmyte. Wygląda na to, że spora część ruchu w sieci pozostaje bez odpowiedzi. Może być to spowodowanie tym ze sporo ruchu pochodzi ze skanowania sieci (np. celem znalezienia punktu dostepowego) lub innego protokołu, który dużo rozgłasza zapytań, a rzadko otrzymuje odpowiedzi. Drugą przyczyną takiego stanu rzeczy może być ruchu rozgłoszeniowy (broadcast), na który nie oczekuje się odpowiedzi.

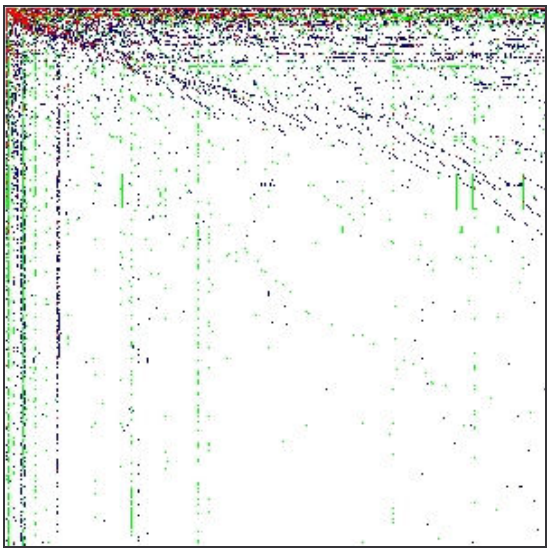
Source-Protocol



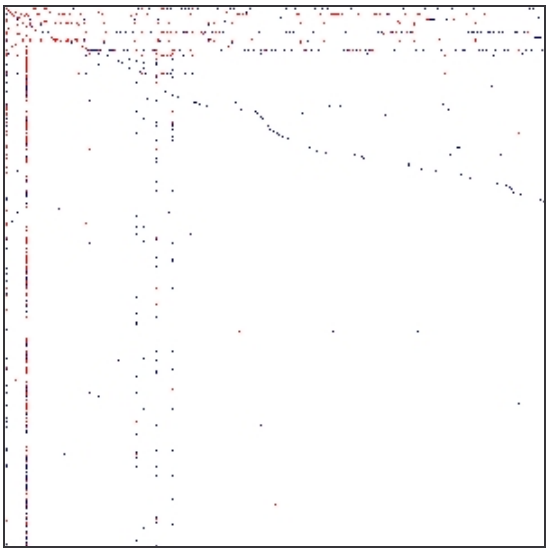
Rysunek 176. Cer Source-Protocol.



Rysunek 177. Inne Source-Protocol.



Rysunek 178. All Source-Protocol.

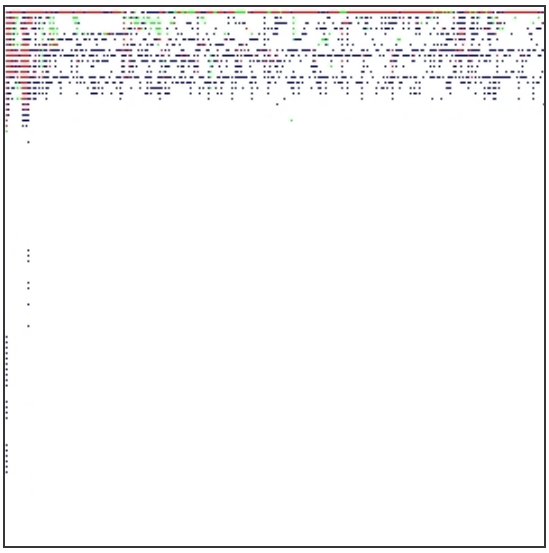


Rysunek 179. Cerl Source-Protocol film.

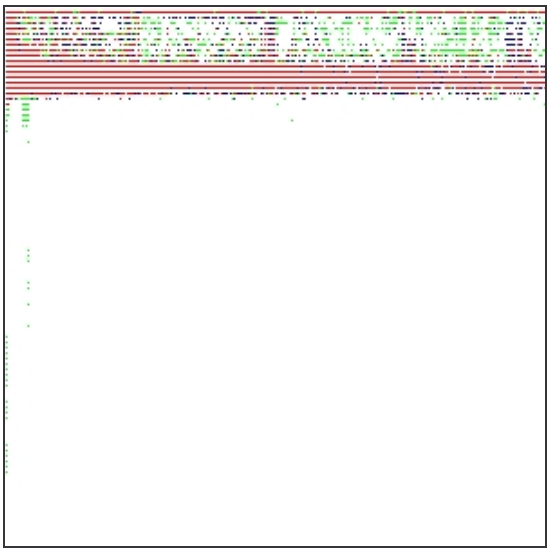
Analiza ta jest podobna do analizy „Destination-Protocol". Widać, że część protokołów jest preferowana (zapewne związane z utrzymaniem sieci), a także że część adresów wykorzystuje dużą ilość protokołów. Analiza filmowa wskazuje, że przyczyna nie leży tylko w tym, że niektóre urządzenia transmitują więcej ruchu, ale wygląda na to, że część urządzeń wykorzystuje tylko niewielką cześć protokołów (widać ze powtarzają się popularne protokoły {czerwone punkty}, ale nie występują punkty z innych protokołów). Być może część urządzeń obsługuje mniejszą ilość protokołów.

Source-Signal

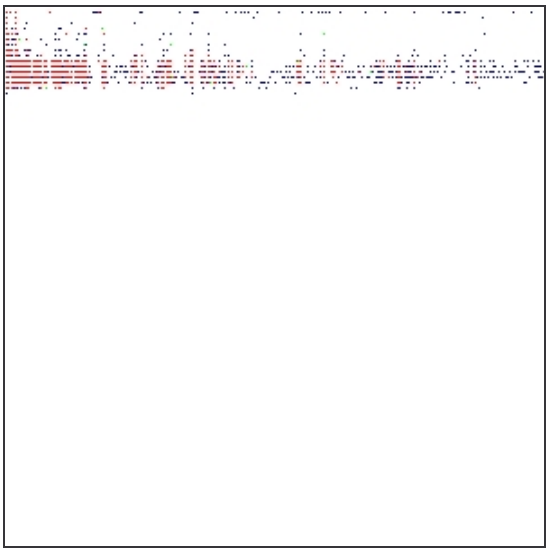
Rysunek 180. Cer Source-Signal.



Rysunek 181. Inne Source-Signal.



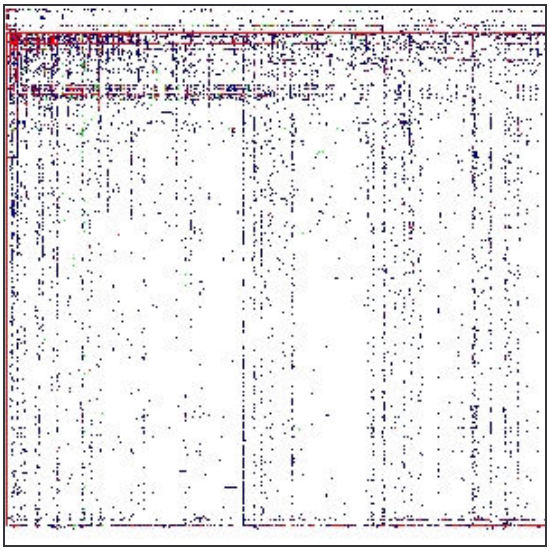
Rysunek 182. All Source-Signal.



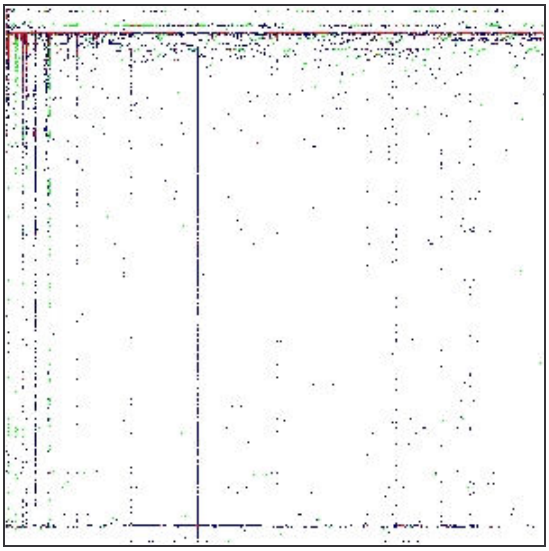
Rysunek 183. Cerl Source-Signal film.

Ta analiza jest podobna do analizy „Destination-Signal". Nie widać łatwo interpretowalnych różnic. Być może jest to spowodowane tym, że brak symetrii „Source-Destination" jest na tyle mały, aby odbił się na różnicy sygnałów.

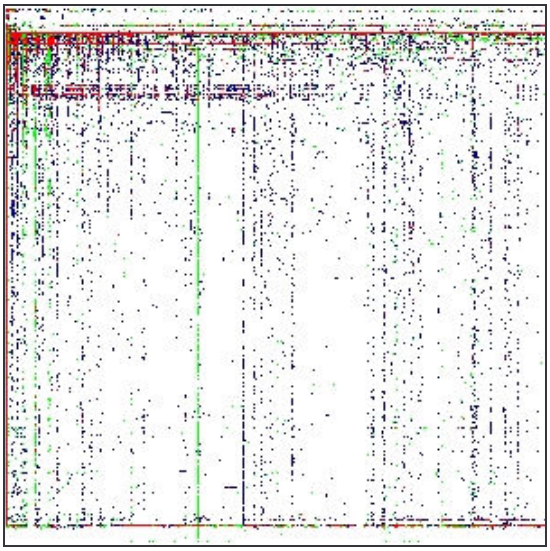
Source-Size



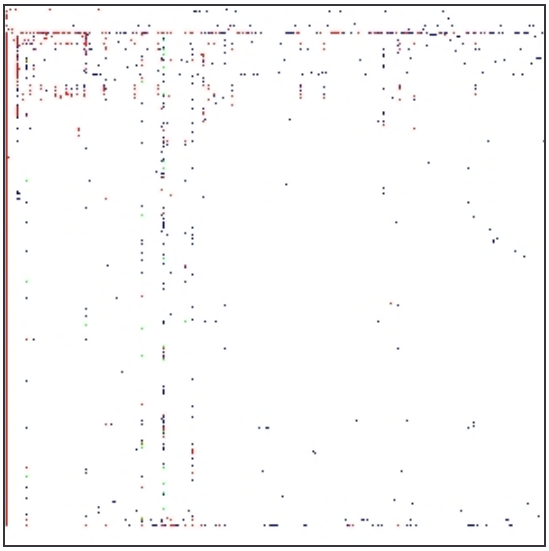
Rysunek 184. Cer Source-Size.



Rysunek 185. Inne Source-Size.



Rysunek 186. Inne Source-Size.



Rysunek 187. Cerl Source-Size film.

Analiza ta jest analizą pokrewną do analizy „Destination-Size". Różnice są jednak duże. Widać, że o ile w analizie „Destination-Size" były wyraźnie widoczne maksima (pionowe kreski), to w analizie „Source-Size" maksima takie uległy rozmyciu. Sporo urządzeń preferuje jedno urządzenie odbiorcze pod względem wysyłania pakietów różnej wielkości. Jest to niejako fragment symetrii z analizy „Source-Destination". Chodzi konkretnie o to, że większość nadawców wybiera określoną grupę odbiorców, przy czym w analizie „Source-Destination" grupa odbiorców jest dość szeroka. Tak wiec na efekt złamanej symetrii źrodłowo-docelowej pływa dodatkowo efekt, który powoduje, że niektóre adresy docelowe są preferowane pod względem zróżnicowania wielkości pakietów.

Podsumowanie

Podsumowując prowadzone badania należy podkreślić, że analiza ruchu sieci bezprzewodowych, na podstawie uzyskanych wyników, nie jest łatwa. Występują różne zjawiska zakłócające pomiary, duże zaszumienie wyników, oraz duże ich podobieństwo w badaniach różnych sieci, przy jednoczesnych dużych różnicach w tej samej sieci (analizy „All" mają zdecydowanie więcej wspólnych punktów niż analizy filmowe) Powoduje to, że wykorzystywanie mechanizmów eksploracji danych do analizy sieci i wykrywania zagrożeń jedynie w oparciu o dane z warstwy MAC, wydaje się mało skuteczne. Na pewno algorytm DBScan nie nadaje się do grupowania obserwowanych tu zachowań. Można by spróbować zbudować sondę IDS analizującą siłę sygnału i adresy MAC, ale takie pomiary obarczone są dużym błędem, spowodowanym faktem, że skanowanie ruchu w sieci jest wykonywane zwykle z innego miejsca niż jest odbierany przez urządzenia docelowe, oraz siła sygnału stabilna jest jedynie w antenach stacjonarnych. Poza tym analizy sygnału wydają się być zbyt rozmyte, by dało się jednoznacznie wykryć takie włamanie (sygnał urządzenia wykorzystywanego we włamaniu musiałby się różnić o około 50% od sygnału właściwego. Kolejnym problemem jest to, że metody te wymagają dużej mocy obliczeniowej, co je niweluje z zastosowań czasu rzeczywistego. Aby możliwe było wykonanie powyższej analizy, autor musiał wykonać wielodniowe nasłuchy sieci Wi-Fi z różnych lokalizacji. Zebrane dane (około 2GB) zostały następnie przetwarzane na pięciu maszynach klasy P4 2.7 GHz przez 7 dni dając w wyniku około 5GB plików. Z otrzymanych plików zostały ręcznie wybrane te, które pozwalały na dalszą analizę (były interpretowalne), a następnie zostały one poddane analizie porównawczej. Po wykonaniu powyższych czynności otrzymane obrazy poddane były interpretacji.

Badania takie pokazują zupełnie innego rodzaju zjawiska w sieci Wi-Fi. Do najbardziej spektakularnych należy niezrównoważenie badania „Source-Destination", które prawdopodobnie jest wynikiem wysyłania dużej ilości pakietów, na które sieć nie otrzymuje odpowiedzi. Efekt asymetrii „Source-Destination" jest dodatkowo wzmacniany przez niesymetryczny rozkład wielkości pakietów, ponieważ o ile nadawcy rozsyłają różne wielkości pakietów podobnie do siebie, to niektórzy odbiorcy są wyraźnie preferowani pod względem różnorodności pakietów. Kolejnymi zaobserwowanymi efektami jest to, że sieci Wi-Fi transmitują głównie małe pakiety (prawdopodobnie związane z utrzymaniem sieci), a jeśli transmitują większe pakiety, to część z nich zmienia wielkość dość płynnie, część dyskretnie, a część wykorzystuje stałą wielkość pakietu. Na uwagę zasługuje fakt, że część urządzeń pracuje z bardzo ograniczoną ilością protokołów (być może urządzenia kilenckie). Podczas obserwacji ruchu względem kanałów można zaobserwować 3-4 maksima, które są związane z ograniczeniem ilości nie zakłócających się kanałów na danym obszarze. Można też zaobserwować to, że większość ruchu w sieci odbywa się w ramach paru numerów BSSID, które zapewne są związane z punktami dostępowymi. Reszta numerów BSSID jest wykorzystywana dość sporadycznie.

Według mojej wiedzy zastosowanie algorytmu DBScan do analizy danych warstwy MAC w sieciach Wi-Fi nie było dotychczas próbowane. Choć uzyskane wyniki pokazują małą

skuteczność wykrywania włamań do sieci, to pozwoliły one zaobserwować niektóre istotne zachowania sieci, które mogą być przydatne do optymalizacji wydajności przyszłych standardów sieciowych. Wydaje się, że praca jest jedynie wstępem do tego typu badań. W kolejnym kroku można spróbować badać korelacje pomiędzy danymi zbieranymi z nagłówków różnych warstw hierarchii protokołów.

Załączniki

Załącznik 1: Płyta CD.

Do pracy załączona jest płyta CD, na której umieszczono:

- Elektroniczną wersję pracy.
- Kod źródłowy oraz pliki wynikowe aplikacji WiFi_Analysis.
- Dane źródłowe pochodzące ze skanowania sieci.
- Wyniki działania aplikacji WiFi_Analysis (badania i porównania).
- Dodatkową dokumentację:
    - opracowanie „Data maining”,
    - opracowanie „Hacking the Invisible Network”,
    - dokumentację kart radiowych firmy Orinoco,
    - normy IEEE 802.11
    - dokumentację Cisco.
- Inne aplikacje wykorzystywane podczas pisania tej pracy:
    - NetStumbler,
    - AiroPeek,
    - Network Associates Sniffer Pro Wireless 802.11,
    - WinRar.

Bibliografia

1. Michael Sutton *Hacking the Invisible Network. Insecurities in 802.11x*, 2002r.
2. Mathew S. Gast *802.11 Sieci bezprzewodowe Przewodnik encyklopedyczny.* HELION 2003r.
3. Andrzej Daniluk *Kompendium programisty C++ BUILDER.* Helion 2003.
4. Opracowanie *Data maining* na podstawie:
   - Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos, Prabhakar Raghavan, *Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications*, Proc. of the ACM SIGMOD Int'l Conference on Management of Data, Seattle, Washington, June 1998.
   - M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, *A density-based algorithm for discovering clusters in large spatial databases with noise*, Proc. of the 2nd Int'l Conference on Knowledge Discovery in Databases and Data Mining, Portland, Oregon, August 1996.
5. Normy:
   - *802.11 IEEE Computer Society, 1999r.*
   - *802.11a IEEE Computer Society, 1999r.*
   - *802.11b IEEE Computer Society, 1999r.*
   - *802.11i IEEE Computer Society, 2002r.*
   - *802.11x IEEE Computer Society, 2001r.*
6. Dokumentacja Cisco:
   - Mariusz Baczyński *Bezpieczeństwo WLAN* 2003
   - Maciej Szeptycki *Wdrażanie zaawansowanych rozwiązań VoIP w środowisku WLAN* 2003
   - Maciej Szeptycki *Projektowanie i wdrażanie bezprzewodowych mostów w technologi 802.11b* 2003
7. Dokumentacja Kismet http://www.kismetwirless.net
8. Dokumentacja Snort-Wirless http://www.snort-wirles.org
9. Dokumentacja internetowa związana z bezpieczeństwem sieciowym:
   - http://hacking.pl
   - http://www.wardriving.com