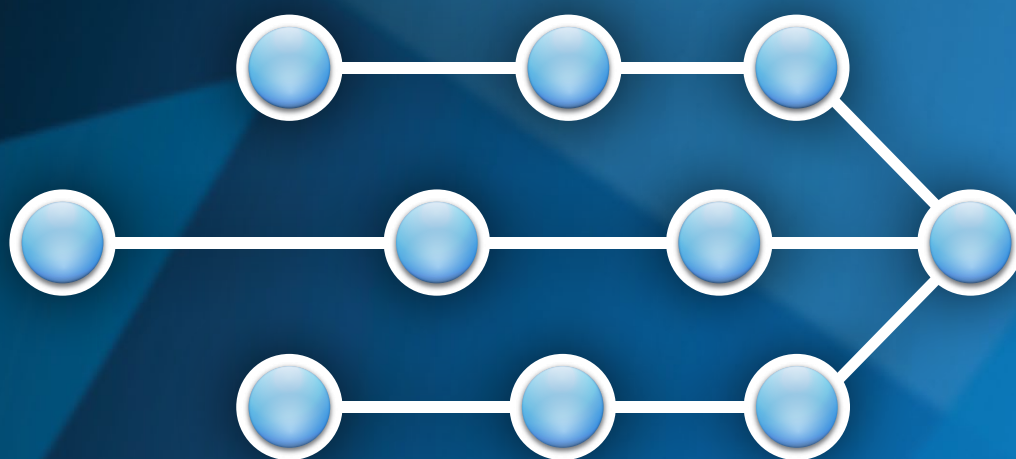


Hackathon Wielkie Wyzwania Programistyczne

MAP THE GAP

Warszawa 16 – 17 lutego 2017 r.

Podsumowanie rezultatów



Patronat honorowy:

Partnerzy:

Patronat medialny:

Podsumowanie rezultatów **hackathonu** Wielkie Wyzwania Programistyczne – Map the Gap

Opis problemu

Zadanie postawione przed uczestnikami pierwszej edycji hackathonu polegało na zaprojektowaniu optymalnego modelu rozwoju sieci światłowodowej na terenie Polski. Obecnie stałe łącze wykorzystuje w Polsce 7,1 mln osób, co przekłada się na 18,5% penetracji rynku. To poniżej średniej europejskiej, która wynosi 32,7%. Polska wypada gorzej także pod względem liczby osób, które korzystają z internetu o bardzo wysokiej prędkości, za jaką uważa się łącza o przepływności co najmniej 100 Mbps. W Europie takie szybkie łącza posiada 15,5% użytkowników internetu stacjonarnego, w Polsce – 13,7%. Co ciekawe, w Unii Europejskiej przoduje pod tym względem Rumunia, gdzie jest aż 57,1% łączy stacjonarnych o przepływności równej 100 Mbps lub większej (<http://www.telko.in/czy-decydenci-wyciagna-i-przedstawia-wnioski-z-pierwszego-konkursu-1-1-popc>). Szczególnie ważnym wyzwaniem jest zapewnienie dostępu do szerokopasmowego internetu w obszarach mniej zurbanizowanych. Rozbudowa sieci światłowodowej pozwalającej na zaoferowanie połączeń internetowych o dużej przepustowości jest jednym z wyzwań, przed którymi stoją operatorzy telekomunikacji oraz instytucje rządowe.

Należy pamiętać, że inwestycje w infrastrukturę teleinformatyczną są trudne i kosztowne, a okres ich zwrotu liczony jest w dekadach (średnio 20 lat). Co więcej, zaangażowany jest nie tylko kapitał firm telekomunikacyjnych, ale także znaczne środki publiczne, na przykład w ramach Programu Operacyjnego Polska Cyfrowa. Zakończony nabór wniosków w POPC zakłada dofinansowanie podłączenia co najmniej 880 tys. lokalizacji z listy białych plam (z uwzględnieniem szkół).

Dotychczasowe metody projektowania rozwoju sieci światłowodowej opierają się na danych statystycznych, takich jak liczba potencjalnych klientów na jednostkowym obszarze. Metody te uwzględniają także analizę statystyczną potencjalnego zainteresowania odbiorców i ich ewentualnych preferencji. Dostępne są również rozwiązania pozwalające na projektowanie i optymalizację układu sieci światłowodowej na niewielkich obszarach – obejmujących jedną lub kilka sąsiednich gmin (np. https://www.softelnet.pl/wp-content/uploads/2017/01/Koncepcja-sieci-FTTH-dla-POPC1_QGIS_GNIFREE_1.1.pdf). Brakuje jednak metod pozwalających na analizę lokalizacji i optymalizację rozwoju sieci światłowodowej w szerszym ujęciu, co umożliwiłoby obniżenie kosztów inwestycji lub podłączenie większej liczby odbiorców.

Z tego powodu warto wykorzystać nowe podejście do problemu i dokonać w skali całego kraju optymalizacji połączeń światłowodowych na podstawie listy adresów oraz innych dostępnych informacji, takich jak lista punktów dostępowych, do których mogą być podłączane rozbudowywane sieci światłowodowe, oraz lista ulic i dróg, wzdłuż których powinny być prowadzone inwestycje światłowodowe. Ze względu na niedostępność listy punktów styku z istniejącą siecią światłowodową jako najbliższe rzeczywistości przybliżenie wykorzystano listę BTS-ów LTE1800 dostępną na stronach Urzędu Komunikacji Elektronicznej (lista z dn. 25.01.2017). Lista ta może być łatwo zastąpiona listą punktów styku z istniejącą siecią światłowodową jednego lub wielu operatorów.

Analiza danych w skali kraju wymaga użycia dużych mocy obliczeniowych obejmujących wiele procesorów, do czego potrzebne jest wykorzystanie odpowiednich algorytmów i aplikacji. Niestety wiedza na temat ich stosowania jest ograniczona i dostępna jedynie w nielicznych instytucjach, takich jak ICM. Uczestnicy hackathonu Wielkie Wyzwania Programistyczne otrzymali dostęp do zasobów obliczeniowych pozwalających na analizę dużych danych, zaś ich efektywne wykorzystanie było jednym z zasadniczych wyzwań, z którymi musieli się zmierzyć.

Wyniki hackathonu

Wszystkie zespoły biorące udział w hackathonie zaprezentowały własne rozwiązania postawionego problemu (podłączenie jak największej liczby punktów adresowych z listy białych plam), które różniły się sposobem podejścia do zagadnienia oraz otrzymanym wynikiem. Zespoły rozwiązywały problem w dwóch wariantach:

1. łącząc punkty za pomocą geometrycznie najkrótszych odcinków,
2. prowadząc odcinki światłowodów wyłącznie wzdłuż dróg (na podstawie OpenStreetMap).

Zespół w składzie: **Aleksandra Kardaś, Rafał Kowalczyk, Piotr Konorski i Piotr Witkiewicz** (studenci inżynierii obliczeniowej ICM UW) przedstawił sposób połączenia wszystkich punktów adresowych z listy białych plam (ponad 2,7 mln lokalizacji) z punktami dostępowymi (BTS) w wariantcie pierwszym (najkrótsza droga) przy wykorzystaniu ok. 650 tys. km światłowodów (mniej niż podany w zadaniu limit 1 mln km światłowodów oszacowany na podstawie środków przeznaczanych w POPC na ten cel i średniego kosztu położenia 1 km sieci światłowodowej). Zespół przedstawił wstępną propozycję poprowadzenia światłowodów wzdłuż dróg zdefiniowanych w OpenStreet Map. Jednym z interesujących aspektów zaproponowanego rozwiązania była analiza gęstości punktów z listy białych plam.

Oprogramowanie zostało wykonane w języku Java z wykorzystaniem biblioteki PCJ. Do obliczeń wykorzystano 60 węzłów komputera Okeanos (2880 rdzeni obliczeniowych). Jednorazowa analiza danych z wykorzystaniem wszystkich zasobów trwała 5–10 minut (przy wykorzystaniu jednego procesora wymagałoby to 250 godzin!).

Zespół w składzie **Michał Żak** i **Kamil Żyła** (studenci 4 roku informatyki MIM UW) wykonał szereg analiz, korzystając z informacji o lokalizacjach znajdujących się najbliżej siebie. Ostatecznie zespół przedstawił koncepcję podłączenia około 90 tys. miejsc z wykorzystaniem 1 mln km światłowodów. Lokalizacje bez internetu zostały podłączone do najbliższego punktu dostępowego (BTS) bezpośrednim połączeniem światłowodowym.

Zespół w składzie **Marianna Zadrożna** (inżynieria obliczeniowa ICM UW) i **Adam Zadrożny** (NCBJ) wykonał analizy na podstawie map rozmieszczenia punktów bez internetu w odrębnych lub sąsiadujących ze sobą gminach. Analizy pozwoliły oszacować długość światłowodu niezbędną do podłączenia punktów na rozpatrywanym terenie. Na przykładowym obszarze 8 gmin można było podłączyć ponad 39 tys. lokalizacji. Oprogramowanie przygotowane przez zespół pozwala na optymalizację połączeń na wybranym obszarze oraz wyznaczenie całkowitej długości połączeń.

Rezultaty hackathonu – wnioski ogólne

Hackaton dowiódł, że analiza dużych danych obejmujących ponad 2,7 mln lokalizacji jest możliwa pod warunkiem wykorzystania w tym celu odpowiednich zasobów obliczeniowych, narzędzi programistycznych oraz kompetencji z zakresu programowania równoległego.

Analiza i optymalizacja dużych sieci połączeń może być przeprowadzona w skali całego kraju z dowolną dokładnością (w tym wypadku do pojedynczego budynku).

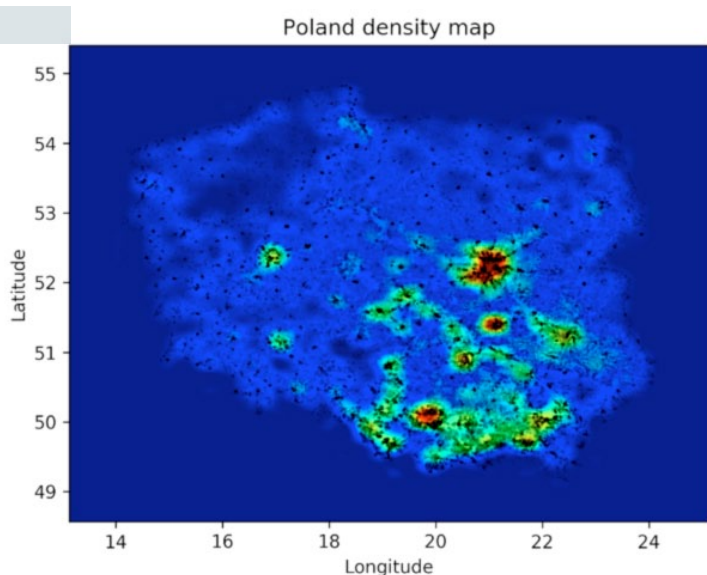
Hackathon ukazał potencjał ukryty w publicznie dostępnych bazach danych pochodzących z różnych źródeł (Ministerstwo Cyfryzacji, UKE, domena publiczna). Niestety, dane są udostępniane w różnych formatach, co utrudnia ich komplementarne wykorzystanie (jednym z wyzwań była konwersja danych do wspólnego formatu). Hackathon pokazał, że choć bariery te można pokonać, to jednak wskazane byłoby ujednoczenie danych i zmniejszenie liczby ich formatów wykorzystywanych przez jednostki administracji publicznej.

Zespoły dysponowały szeroką gamą narzędzi programistycznych. Najbardziej efektywne rozwiązania powstały jednak przy wykorzystaniu oprogramowania rozwijanego w ICM UW (biblioteka PCJ do obliczeń równoległych) oraz z użyciem języka Java, który pozwolił na wykorzystanie dostępnych bibliotek do czytania plików zapisanych w różnych formatach (w tym danych OpenStreet Map w formacie XML). Posłużenie się tradycyjnymi narzędziami, takimi jak biblioteka MPI (w języku C++), okazało się znacznie trudniejsze i spowodowało wiele problemów z napisaniem poprawnego kodu oraz jego uruchomieniem (problemy te są dobrze znane i stanowią motywację do szukania innych rozwiązań w zakresie metod i języków programowania).

Rezultaty hackathonu – przykładowe wyniki

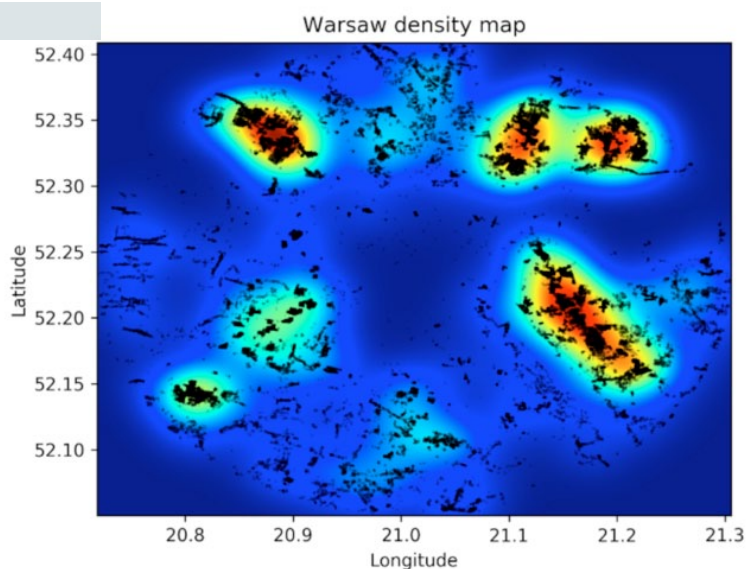
Poniżej przedstawiamy przykładowe wykresy obrazujące wyniki hackathonu. Wykresy zostały przygotowane przez zespół w składzie **Aleksandra Kardaś**, **Rafał Kowalczyk**, **Piotr Konorski** oraz **Piotr Witkiewicz**. Są to wstępne wyniki wymagające dopracowania, niemniej obrazują możliwości oprogramowania przygotowanego przez uczestników hackathonu.

Publikacja wykresów jest możliwa pod warunkiem podania nazwisk autorów oraz informacji, że wyniki zostały uzyskane podczas hackathonu Wielkie Wyzwania Programistyczne – Map the Gap (<http://www.icm.edu.pl>).



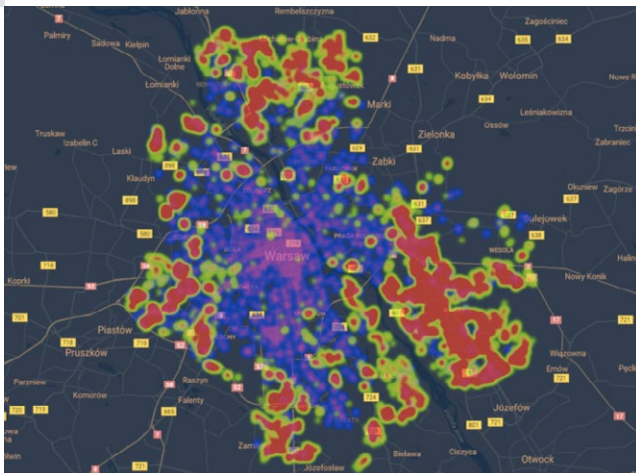
Rysunek 1. Gęstość lokalizacji bez szerokopasmowego dostępu do internetu w Polsce (na podstawie listy białych plam)

Wykres pokazuje obszary o dużym zagęszczeniu punktów (kolor żółty i czerwony) w otoczeniu dużych aglomeracji (np. Warszawa, Kraków, oraz na południu Polski). Na tych obszarach wsparcie publiczne może nie być konieczne ze względu na dużą liczbę potencjalnych klientów.

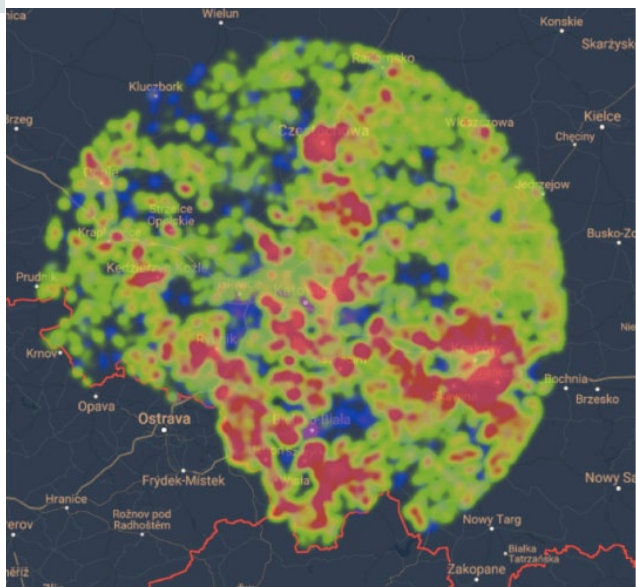


Rysunek 2. Gęstość lokalizacji bez szerokopasmowego dostępu do internetu w Warszawie i okolicy (na podstawie listy białych plam)

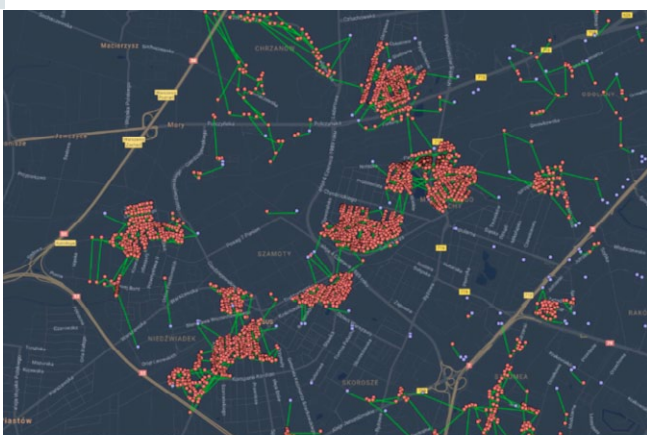
Wyraźnie widoczne obszary gęsto zamieszkałe z ubogą infrastrukturą światłowodową w miejscowościach podwarszawskich (Wawer, Jabłonna).



Rysunek 3. Gęstość lokalizacji bez szerokopasmowego dostępu do internetu w Warszawie i okolicy (na podstawie listy białych plam, kolor czerwony oznacza wysoką gęstość lokalizacji)



Rysunek 4. Gęstość lokalizacji bez szerokopasmowego dostępu do internetu w Krakowie i okolicy (na podstawie listy białych plam, kolor czerwony oznacza wysoką gęstość lokalizacji)



Rysunek 5. Przykładowe połączenia światłowodowe w modelu najkrótszych odległości (zielone linie). Punkty oznaczone kolorem czerwonym oznaczają lokalizacje z listy białych plam, niebieskie punkty oznaczają lokalizacje BTS-ów. Przedstawiony obszar obejmuje okolice węzła drogowego Konotopa w Warszawie.



Rysunek 6. Przykładowe połączenia światłowodowe w modelu, w którym połączenia (zielone linie) prowadzone są wzdłuż dróg*

Punkty oznaczone kolorem czerwonym oznaczają lokalizacje z listy białych plam, niebieskie punkty oznaczają lokalizacje BTS-ów. Przedstawiony obszar obejmuje dzielnicę Włochy (Warszawa).



Rysunek 7. Zaproponowany przebieg światłowodów w modelu najkrótszych połączeń w skali Polski*

Całkowita długość połączeń wymaganych do podłączenia wszystkich punktów z listy białych plam wynosi 650 tys. km.

* Wyniki przedstawione na wykresach 6 i 7 stanowią fragment rozwiązania obejmującego całą Polskę.

Więcej informacji na temat rezultatów hackathonu można uzyskać na stronie <http://wwp.icm.edu.pl>, a także od pomysłodawcy wydarzenia, prof. Piotra Bały (email: bała@icm.edu.pl).

Patronat honorowy nad hackathonem Map the Gap objęły Ministerstwo Cyfryzacji, Ministerstwo Rozwoju oraz Urząd Komunikacji Elektronicznej. Partnerami turnieju były firmy Cisco, Orange oraz Codilime. Patronat medialny objęły serwisy Computerworld.pl oraz Telko.in.