

System Wspomagania Decyzji (SWD)

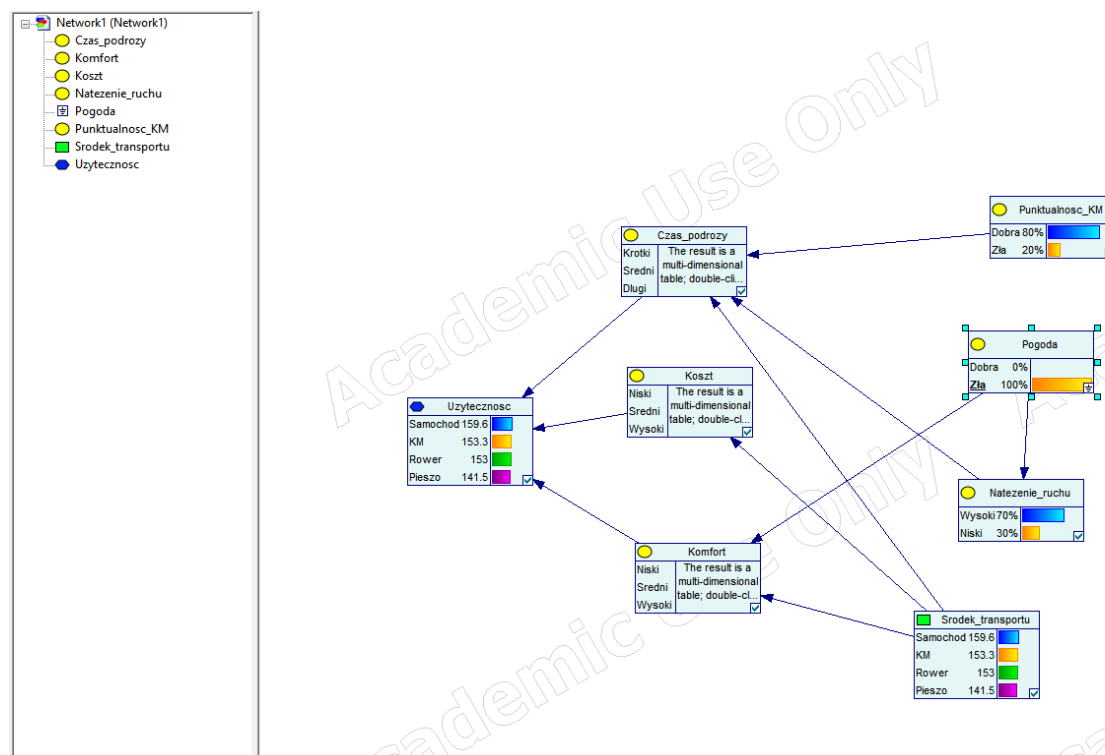
- *Autorzy: Dominik Bojkiewicz, Dominik Jaroszuk, Dominika Kamińska*
- *Data: 21 stycznia 2026*
- *Repozytorium: <https://github.com/Domaka2001/SWD-projekt.git>*

Celem projektu było zaprojektowanie i implementacja Systemu Wspomagania Decyzji (SWD), którego zadaniem jest optymalizacja codziennego wyboru środka transportu. System rozwiązuje klasyczny problem teorii decyzji: wybór jednej z wielu alternatyw w środowisku dynamicznym, charakteryzującym się niepewnością informacji oraz subiektywnymi preferencjami użytkownika.

2.1 Diagramy Wpływu

System został sformalizowany jako **Diagram Wpływu**, będący acyklicznym grafem skierowanym (DAG).

Wizualizacja struktury modelu (GeNIe)



Struktura sieci składa się z 8 węzłów, z których każdy pełni określoną rolę w procesie modelowania decyzji:

A. Węzły losowe (Chance nodes) – modelowanie niepewności:

- **Pogoda (Weather):** Węzeł bazowy (root), określający stany: "Dobra", "Zła". Wpływa bezpośrednio na natężenie ruchu oraz subiektywny komfort podróży.

- **Natężenie ruchu (Traffic):** Węzeł zależny od pogody. Określa prawdopodobieństwo wystąpienia korków ("Niskie", "Wysokie"), co kluczowo wpływa na czas przejazdu samochodem i autobusem.
- **Punktualność KM (PT Punctuality):** Modeluje niezależną zmienną losową dotyczącą sprawności komunikacji miejskiej ("Dobra", "Zła").
- **Czas podróży (Travel Time):** Węzeł kryterialny zależny od decyzji, ruchu oraz punktualności KM.
- **Koszt (Cost):** Węzeł kryterialny bezpośrednio determinowany przez wybrany środek transportu.
- **Komfort (Comfort):** Węzeł kryterialny zależny od środka transportu oraz warunków pogodowych (szczególnie istotne dla roweru i drogi pieszej).

B. Węzeł decyzyjny (Decision node):

- **Środek transportu (Transport Decision):** Centralny punkt modelu, reprezentujący zbiór alternatyw: *Samochód, Komunikacja miejska, Rower, Pieszo*. Wybór w tym węźle modyfikuje rozkłady prawdopodobieństwa w węzłach kryterialnych.

C. Węzeł użyteczności (Utility node):

- **Użyteczność (Utility):** Węzeł końcowy (sink), który nie posiada własnych stanów, lecz przechowuje tabelę wartości skalarnych. Agreguje on wyniki z węzłów *Czas, Koszt i Komfort* w oparciu o subiektywne wagi użytkownika.

2.2 Zasada Maksymalizacji Oczekiwanej Użyteczności (MEU)

Sercem systemu jest silnik wnioskowania realizujący zasadę **MEU (Maximum Expected Utility)**. Proces ten dzieli się na dwa kluczowe etapy teoretyczne:

A. Wielokryterialna Teoria Użyteczności (MAUT): Zastosowano **addytywny model użyteczności**, który zakłada, że całkowita satysfakcja użytkownika jest sumą satysfakcji z poszczególnych atrybutów. Podstawą teoretyczną jest tutaj *założenie o niezależności preferencyjnej* – tzn. ocena czasu podróży jest stała niezależnie od poziomu komfortu. Dzięki temu funkcja użyteczności $U(s, d)$ przyjmuje postać: $U(s, d) = \sum_{i \in \{T, C, K\}} w_i \cdot u_i(s_i)$ gdzie w_i to wagi zdefiniowane przez użytkownika, a u_i to znormalizowane oceny cząstkowe dla czasu (T), kosztu (C) i komfortu (K).

B. Wnioskowanie Bayesowskie i Propagacja Prawdopodobieństw: Dla każdej decyzji d , system musi wyznaczyć rozkład prawdopodobieństwa wyników $P(s|d, e)$. Wykorzystywane jest **wnioskowanie bayesowskie**, które pozwala na aktualizację wiedzy o sieci po wprowadzeniu dowodów e (np. "Pogoda = Zła").

- Dowód propaguje się od węzłów bazowych przez węzły pośrednie (np. wzrost natężenia ruchu).
- Silnik oblicza marginalne prawdopodobieństwa warunkowe dla każdego kryterium ($P(\text{Czas}|\text{Samochód}, \text{Korki})$).

C. Decyzja Optymalna: Ostateczna wartość $EU(d)$ jest sumą ważoną użyteczności wszystkich możliwych stanów kryteriów, pomnożoną przez prawdopodobieństwo ich wystąpienia. System wskazuje decyzję d^* , która spełnia warunek: $d^* = \arg \max_{d \in D} EU(d)$ Takie podejście gwarantuje **racjonalność decyzji** w sensie matematycznym – wybrana opcja jest statystycznie najlepsza przy danej wiedzy o świecie i preferencjach decydenta.

3. Analiza Eksperta: Model Probabilistyczny

Wartości w tablicach prawdopodobieństw warunkowych (CPT) zostały dobrane tak, aby odzwierciedlały logiczne i przyczynowo-skutkowe zależności zachodzące w transporcie miejskim.

3.1 Dynamika ruchu i wpływ czynników zewnętrznych

Model opiera się na założeniu, że niepewność nie jest jednolita – niektóre zdarzenia znacząco podnoszą ryzyko opóźnień.

A. Węzeł: Natężenie ruchu (Traffic) Ekspert założył silną korelację z pogodą. Tablica CPT dla tego węzła wygląda następująco:

- $\$P(\text{Ruch=Wysoki} \mid \text{Pogoda=Dobra}) = 0.2\$$
- $\$P(\text{Ruch=Wysoki} \mid \text{Pogoda=Zła}) = 0.7\$$
- *Uzasadnienie:* W złej pogodzie więcej osób wybiera samochód zamiast roweru czy spaceru, co drastycznie zwiększa prawdopodobieństwo korków.

B. Węzeł: Komfort (Comfort) Wartości modelują subiektywne odczucie wygody, które jest zmienne:

- Dla *Samochodu*: Komfort jest zawsze wysoki (80-100%), niezależnie od pogody.
- Dla *Roweru*: $\$P(\text{Komfort=Wysoki} \mid \text{Pogoda=Dobra}) = 0.6\$$, ale $\$P(\text{Komfort=Niski} \mid \text{Pogoda=Zła}) = 0.8\$$.
- *Uzasadnienie:* Samochód izoluje od otoczenia, podczas gdy rower i przemieszczanie się pieszo są bezpośrednio wystawione na czynniki atmosferyczne.

3.2 Złożoność czasu podróży

Węzeł **Czas podróży** posiada 48 wpisów w CPT, co pozwala na precyzyjne modelowanie ryzyka. Ekspert zróżnicował czas w zależności od interakcji:

- **Samochód:** Czas zależy wyłącznie od natężenia ruchu. Przy wysokim ruchu szansa na "Długi" czas wynosi 40%.
- **Komunikacja Miejska:** Wykorzystuje model podwójnego ryzyka. Jeśli Ruch=Wysoki ORAZ Punktualność=Zła, prawdopodobieństwo "Długiego" czasu wynosi aż 60%. Jeśli jednak oba parametry są optymalne, szansa na "Krótki" czas to tylko 20% (ze względu na sztywne rozkłady jazdy).
- **Rower/Piesz:** Czas jest wysoce stabilny i mało wrażliwy na natężenie ruchu samochodowego (szansa na "Średni" czas to zawsze ok. 60%).

3.3 Stabilność kosztów

Węzeł **Koszt** został zaprojektowany jako niemal deterministyczny:

- Samochód: Zawsze wysokie prawdopodobieństwo stanu "Wysoki" (paliwo, parkowanie).
- Komunikacja: Stan "Średni" (bilet okresowy/jednorazowy).
- Rower/Piesz: 90-100% szansy na stan "Niski". To sprawia, że przy zerowej wadze kosztu, system rzadziej wybierze te opcje, chyba że komfort lub czas będą krytyczne.

4. Analiza Eksperska: Model Preferencji (Użyteczności)

Zastosowano **Wielokryterialną Teorię Użyteczności (MAUT)** w formie addytywnej. Każde kryterium zostało zmapowane na bezwymiarową skalę punktową (0-100).

4.1 Definicja Punktacji Kryteriów

- **Czas podróży:** Krótki (100 pkt), Średni (50 pkt), Długi (0 pkt).
- **Koszt:** Niski (100 pkt), Średni (50 pkt), Wysoki (0 pkt).
- **Komfort:** Wysoki (100 pkt), Średni (50 pkt), Niski (0 pkt).

4.2 Analiza Wagowa i Interakcja z Użytkownikiem

Unikalną cechą systemu jest to, że ostateczna tabela użyteczności jest generowana dynamicznie w oparciu o wagi w_i podane przez użytkownika. Ekspert zdefiniował model jako: $U_{final} = \frac{w_t \cdot U_{time} + w_c \cdot U_{cost} + w_k \cdot U_{comfort}}{\sum w_i}$ Normalizacja do 100% pozwala na porównanie różnych konfiguracji preferencji. Na przykład, użytkownik "ekonomiczny" ustawi wagę kosztu na 5, a czasu na 1, co spowoduje, że system będzie faworyzował rower nawet przy gorszej pogodzie.

5. Architektura Systemu i Implementacja

System został zaprojektowany w architekturze modułowej, co pozwala na niezależne modyfikowanie struktury sieci, parametrów probabilistycznych oraz interfejsu użytkownika.

5.1 Silnik SMILE i integracja z Pythonem

Wykorzystanie profesjonalnej biblioteki C++ (opakowanej w PySMILE) zapewnia wysoką wydajność obliczeniową.

- **Mechanizm wnioskowania:** Silnik wykorzystuje algorytm **Clustering (Lauritzen)**. Jest to metoda dokładnego wnioskowania, która przekształca diagram wpływu w graf złączeniowy (join tree), co pozwala na efektywne obliczanie prawdopodobieństw a posteriori dla wszystkich węzłów przy dowolnych dowodach.
- **Zarządzanie licencją:** Implementacja zapewnia, że moduł `pysmile_license` jest ładowany jako pierwszy, co jest niezbędne do aktywacji natywnego silnika wnioskowania w środowisku wirtualnym Pythona.

5.2 Przepływ Danych (Data Flow)

Proces podejmowania decyzji przebiega przez cztery warstwy:

1. **Warstwa Wejścia (Streamlit):** Użytkownik przesuwając suwaki wag oraz wybiera stany węzłów losowych. Dane te są zbierane do słowników `user_weights` i `evidence`.
2. **Warstwa Logiki (`decision_engine.py`):** Funkcja `solve_decision` inicjuje budowę sieci. Przekazuje wagi do warstwy modelu w celu dynamicznego przeliczenia użyteczności.
3. **Warstwa Modelu (`build_network.py`):** Tworzona jest instancja klasy `pysmile.Network`. Moduły `structure`, `probabilities` i `utilities` konfiguruje odpowiednio łuki, tablice CPT i tabelę użyteczności ważonej.
4. **Warstwa Wyników i Wizualizacji:** Obliczone przez SMILE wartości *Expected Utility* są normalizowane (0-100%) i przesyłane z powrotem do Streamlit, gdzie biblioteka Matplotlib generuje wykres porównawczy.

5.3 Modułowość i Rozszerzalność

Dzięki ścisłemu podziałowi na moduły (np. oddzielenie `variables.py` od `probabilities.py`), system jest łatwy w rozbudowie. Dodanie nowego środka transportu (np. "Hulajnoga elektryczna") wymaga jedynie dodania jednej pozycji w `variables.py` i uzupełnienia odpowiednich wpisów w tablicach CPT i użyteczności, bez konieczności zmiany logiki wnioskowania.

5.4 Interfejs Streamlit (GUI) i obsługa dowodów

Interfejs został zaprojektowany tak, aby oddzielić warstwę preferencji (Sidebar) od warstwy obserwacji (Main Panel). Pozwala to użytkownikowi na szybką analizę typu "co jeśli" (What-if analysis).

Kluczową funkcjonalnością interfejsu jest obsługa dowodów cząstkowych (Partial Evidence Handling): Użytkownik nie musi określać wszystkich warunków zewnętrznych. Dla każdego z nich (Pogoda, Ruch, Punktualność) dostępna jest opcja "**Losowo**".

- **Działanie techniczne:** Gdy użytkownik wybierze tę opcję, system nie ustawia dla danego węzła żadnego dowodu (*evidence*).
- **Skutek decyzyjny:** W takiej sytuacji silnik SMILE automatycznie wykorzystuje rozkłady prawdopodobieństwa *a priori* zdefiniowane w tablicach CPT (np. 70% szans na dobrą pogodę). Dzięki temu system nadal potrafi wyliczyć oczekiwaną użyteczność, uwzględniając statystyczne ryzyko, co pozwala na podjęcie racjonalnej decyzji nawet w warunkach braku kompletnych danych o środowisku.

6. Analiza Wyników i Scenariusze Testowe

System został poddany walidacji w różnorodnych scenariuszach, aby sprawdzić poprawność reakcji na zmianę dowodów oraz wag preferencji.

6.1 Scenariusz: Priorytet Ekonomiczny (Budget Priority)

- **Preferencje:** Koszt = 5.0, Czas = 1.0, Komfort = 0.0.
- **Warunki:** Pogoda = Dobra, Ruch = Średni (Losowo).
- **Wynik:** System zdecydowanie rekomenduje **Pieszko** lub **Rower** (użyteczność > 90%).
- **Analiza:** Nawet jeśli droga zajmie więcej czasu, zerowy koszt tych metod dominuje w końcowej ocenie ze względu na wysoką wagę przypisaną do wydatków.

6.2 Scenariusz: Skrajnie Trudne Warunki (Extreme Weather)

- **Preferencje:** Równe wagi (1.0).
- **Warunki:** Pogoda = Zła, Ruch = Wysoki, Punktualność KM = Zła.
- **Wynik:** Rekomendacja: **Samochód**.
- **Analiza:** Choć czas przejazdu samochodem w korkach jest długi, to drastyczny spadek komfortu na rowerze/pieszko oraz ryzyko spóźnienia w komunikacji miejskiej sprawiają, że samochód (izolacja, komfort) staje się opcją o najwyższej oczekiwanej użyteczności.

6.3 Scenariusz: Brak Obserwacji (Incomplete Information)

- **Preferencje:** Komfort = 5.0, inne = 1.0.
- **Warunki:** Wszystko ustawione na "Losowo".
- **Wynik:** Rekomendacja: **Samochód**.

- **Analiza:** Przy braku dowodów system opiera się na statystyce. Ponieważ samochód oferuje najwyższy komfort w większości stanów (nawet przy złej pogodzie), a użytkownik uznał komfort za kluczowy, system wybiera opcję "najbezpieczniejszą" statystycznie.

6.4 Scenariusz: Szybki Dojazd (Time is Money)

- **Preferencje:** Czas = 5.0, Koszt = 0.5, Komfort = 1.0.
- **Warunki:** Pogoda = Dobra, Ruch = Wysoki.
- **Wynik:** Rekomendacja: **Rower**.
- **Analiza:** W warunkach dużych korków i dobrej pogody, rower staje się najszybszym środkiem transportu. Wysoka waga czasu sprawia, że system ignoruje mniejszy komfort jazdy rowerem na rzecz punktualnego dotarcia do celu.

Wszystkie scenariusze potwierdzają, że model poprawnie integruje wiedzę ekspercką z subiektywnym systemem wartości użytkownika.