



**AGH**

**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE**

**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI,  
INFORMATYKI I INŻYNIERII BIOMEDYCZNEJ**

KATEDRA INFORMATYKI STOSOWANEJ

Praca dyplomowa magisterska

*Portal ankiet porównawczych  
Pairwise comparisons poll portal*

Autor:

*Karol Wójcik*

Kierunek studiów:

*Informatyka*

Opiekun pracy:

*dr Konrad Kułakowski*

Kraków, 2014

*Oświadczam, świadomy(-a) odpowiedzialności karnej za poświadczanie nieprawdy, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.*

*Serdecznie dziękuję ... tu ciąg dalszych podziękowań np. dla promotora, żony, sąsiada itp.*

## Spis treści

<b>1. Wprowadzenie</b>	7
1.1. Cele pracy	7
1.2. Zawartość pracy	7
<b>2. Metoda porównań parami</b>	9
2.1. Macierz porównań parowych	10
2.2. Ranking	11
<b>3. Niespójność w metodzie porównań parami</b>	13
3.1. Metoda wartości własnych - Saaty	14
3.2. Metoda odległościowa - Koczkodaj	15
<b>4. Problem doboru skali w porównaniach parowych</b>	17
4.1. Skala Saaty'ego	18
4.2. Skala Likerta	19
4.3. Inne	20
<b>5. Grupowanie hierarchiczne</b>	21
5.1. Klasyfikator	21
5.2. Zastosowanie hierarchicznego klasyfikatora	22
5.3. Hierarchia w metodzie PC	22
5.3.1. Przykład	24
<b>6. Projekt</b>	29
6.1. Analiza wymagań	29
6.1.1. Wymagania funkcjonalne	29
6.1.2. Wymagania нефункционалне	36
6.2. Architektura systemu	38
6.3. Projekt interfejsu użytkownika	40
6.4. Schemat bazy danych	41
<b>7. Implementacja systemu</b>	43
7.1. Konfiguracja	43

7.2. Interfejs użytkownika .....	44
7.3. Część analityczna .....	50
7.3.1. Niespójność .....	50
7.3.2. Ankiety częściowe .....	50
<b>8. Podsumowanie i wnioski.....</b>	<b>53</b>

# 1. Wprowadzenie

## 1.1. Cele pracy

Celem poniższej pracy jest przedstawienie portalu ankiet porównawczych. Został on zaimplementowany z myślą o rozpowszechnieniu metody porównań parami jako skutecznego instrumentu ankietowego. Powstał na bazie małego narzędzia porównań parowych, wykorzystanego do przeprowadzenia eksperymentu ewaluacji jakości ośrodków naukowych w Polsce za pomocą metody parowej [21]. Współpracowałem z autorami tej pracy zapewniając wykonanie serwisu i obsługę techniczną. Z jednorazowej aplikacji zrodził się pomysł zbudowania portalu ankiet porównawczych. Ma za zadanie świadczyć wysokiej jakości usługi związane z szeroko rozumianą metodą porównań parami. Brak odpowiednich narzędzi był motywacją do stworzenia aplikacji badawczej. Naturalną konsekwencją było stworzenie takich możliwości, aby kolejne badania tworzyć szybciej i łatwiej, a sam proces przeprowadzania ankiety zrobić przyjazny dla respondentów.

Współpracując z naukowcami zajmującymi się w swoich badaniach metodą porównań parami, udało mi się określić zbiór podstawowych wymagań i zależności. Wszystko po to aby nowy system był w możliwie największym stopniu zgodny z zasadami metody jak i oczekiwaniami potencjalnych odbiorców i użytkowników.

## 1.2. Zawartość pracy

Na kolejnych kartach tej pracy zostanie przedstawiona metoda porównań parami od strony matematyczno-formalnej. Zaprezentowane będą także główne problemy i zależności z nią związane. Po obszernym wstępie teoretycznej wiedzy i definicji nastąpi prezentacja projektu systemu. Jego zawartość stanowi opis wymagań funkcjonalnych i нефункциональных wraz z przypadkami użycia. Jako ostatni punkt znajduje się prezentacja implementacji. Zawiera ona informacje o środowisku i opcjach konfiguracji systemu oraz kilka szczególnych punktów zasługujących na poświęcenie im uwagi.



## 2. Metoda porównań parami

W rozdziale tym przedstawiono metodę porównań parami. Jest to prosta metoda zaliczająca się do metod podejmowania decyzji. Służy do subiektywnej oceny wariantów poprzez zestawienie ich ze sobą w parach i wyboru najlepszego z nich.

Początki takich porównań sięgają zamierzchłych czasów, kiedy pojęcie pieniądza ani nawet liczb jeszcze nie istniało. Z łatwością możemy sobie wyobrazić naszych praprzodków, porównujących wagę ptaka trzymanego w jednej ręce z wagą ryby trzymanej w drugiej ręce. Także na porównaniach ludzie opierali barterową wymianę dóbr. Istniała potrzeba określenia jaka ilość rzeczy typu A jest konieczna do otrzymania przedmiotu typu B, np. ile łososi można dostać w zamian za jednego zająca.

W naukach ścisłych metoda porównań parami jest obecna od XVIII wieku, kiedy jej uproszczoną wersję w postaci teorii wyboru społecznego przedstawił Nicolas de Condorcet w 1785 w [4]. Wykorzystana była do stworzenia sprawnego systemu głosowania zwanego kryterium Condorceta. W 2001 roku jego autorstwo zostało przypisane jednak średniowiecznemu filozofowi, Rajmundowi Llull, gdyż zostały wtedy odkryte jego zagubione manuskrypty pochodzące z XIII w. Warto zaznaczyć iż wielu innych uczonych zajmowało się teorią wyboru społecznego i kontynuowało badania w tej dziedzinie. Byli to między innymi Jean-Charles de Borda (1733–1799), Lewis Carroll (1832–1898) oraz Duncan Black (1908–1991). Kolejnym po Condorcecie lecz zdecydowanie bardziej formalnym użyciem metody porównań parowych była praca autorstwa Fechnera z roku 1860 [5]. Wiek XX przyniósł znaczne ożywienie, najpierw za sprawą opracowania L.L. Thurstone’a pt. *A Law of Comparative Judgments* ([18]), które rozszerzyło metodę porównań z użycia sposobu binarnego (większy/mniejszy) na dowolny wybór stosunku. A następnie za sprawą Thomasa Saatyego w [17] który jako pierwszy zaproponował użycie hierarchicznych struktur danych. Stanowiło to istotną zmianę w rozstrzyganiu problemów złożonych z wielu kryteriów.

Dla problemu mającego  $n$  kategorii mamy  $\frac{n(n-1)}{2}$  porównań parowych do wykonania. Na przykład: 10 kategorii generuje 45 par, 30 kryteriów daje 435 porównań parowych, a zaledwie 100 obiektów porównywanych podnosi liczbę par do 4950. Jak widać mamy tu do czynienia ze złożonością  $O(n^2)$ . Znakomita większość ostatnich badań prowadzonych z użyciem porównań parami składała się z nie więcej niż 7 kategorii. Liczba 7 daje 21 par do porównania, co wydaje się być swego rodzaju barierą psychologiczną dla większości ludzi zanim popadną w znużenie lub znudzenie tematem. Dzięki zastosowaniu hierarchii można znacząco obniżyć złożoność obliczeniową, aż do poziomu  $O(n \log_n)$ . Więcej na ten temat zostanie przedstawione w dalszej części pracy.



Metoda porównań parowych może być wykorzystana do różnorodnych pomiarów. Zarówno przedmiotowych, typu odległość, wartość czy też waga, jak i niemierzalnych typu, poziom bezpieczeństwa publicznego czy też stopień zanieczyszczenia. Wartości początkowe dla porównań parami są zazwyczaj określone poprzez pomiary znanych i stałych atrybutów jak długość, cena itd. lub subiektywne preferencje respondentów. (zob. [26]).

Przeprowadzono kilka badań dowodzących przewagę metody porównań parami. W 1998 roku W.W. Koczkodaj w pracy [14] przedstawił wyniki przeprowadzonego eksperymentu, potwierdzające zwiększenie dokładności wyniku przy wykorzystaniu porównań parami. Do tego celu został zaprojektowany eksperyment statystyczny, wykorzystujący metodę Monte Carlo, mający odpowiedzieć na pytanie czy rzeczywiście używanie metody parowej poprawia jakość oszacowania dla czynnika. Eksperyment został zaprojektowany i zaimplementowany w taki sposób aby zminimalizować błędy statystyczne. Uzyskane wyniki po zestawieniu z wynikami otrzymanymi metodą bezpośrednią wykazały znaczny wzrost dokładności.

## 2.1. Macierz porównań parowych

Metoda porównywania parami wykorzystuje do obliczeń tzw. macierz porównań parowych której elementy określają stosunek dwóch odpowiadających sobie obiektów.

Rozważmy  $n$  kategorii będących obiektami porównywanymi. Element  $m_{i,j}$  przedstawia subiektywną preferencję elementu  $a_i$  nad  $a_j$ ,  $i, j = 1, \dots, n$  i wynosi  $m_{i,j} = \frac{a_i}{a_j}$ . Macierz porównań parowych  $A$  ma postać

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{21}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{n1}} & \frac{1}{a_{n2}} & \cdots & 1 \end{bmatrix} \quad (2.1)$$

O macierzy  $A = [a_{ij}] \in R^{n \times n}$  powiemy że jest macierzą porównań parowych (macierz PC) jeśli  $a_{ij} > 0$  dla każdego  $i, j = 1, \dots, n$ . Macierz porównań parowych  $A$  jest nazywana *wzajemną* (ang. *reciprocal*) jeśli  $a_{i,j} = \frac{1}{a_{j,i}}$  dla każdego  $i, j = 1, \dots, n$ . Na tej podstawie możemy stwierdzić iż  $a_{i,i} = 1$  dla każdego  $i = 1, \dots, n$ . Istotnym a także naturalnym jest tutaj założenie niezmienności właściwości. Kiedy jedną z kategorii porównywanych jest np. samochód, to przy każdorazowym porównaniu z inną kategorią mamy na myśli obiekt reprezentujący, a nie konkretny model wybranego producenta. Pomimo intuicyjnej zgodności nie możemy przyjąć że każda macierz spełnia ten warunek. Opisuje to W.W. Koczkodaj w [25]. Obrazowym przykładem może być *ślepy test* win, podczas którego istnieje możliwość iż w naszej ocenie  $i - ty$  trunek jest lepszy niż ten sam  $i - ty$ , podany w innym, nieoznaczonym kieliszku. Dla testującego są to dwie osobne kategorie, kiedy w rzeczywistości był to ten sam napój. Problem ten został poruszony w [23] gdzie autorzy zaprezentowali propozycję zrzucenia domyślnego warunku wza-

jemności z rozważanych macierzy. Na potrzeby pracy przyjąłem że będę pracował jednak na tych, które posiadają tę właściwość.

Macierz PC  $A$  jest nazywana *spójną* (ang. *consistent*) jeśli  $a_{ij} * a_{jk} = a_{ik}$  dla każdego  $i, j, k = 1, \dots, n$ . Zachodzi przy tym zależność iż każda macierz spójna, jest także wzajemna. W drugą stronę implikacja ta z reguły nie jest prawdą. Problem spójności zostanie szerzej opisany w kolejnym rozdziale.

## 2.2. Ranking

Wynikiem którego oczekujemy od badań metodą porównań parowych jest wektor wag badanych obiektów/kategorii. Dzięki niemu wiemy czym kierują się nasi respondenci i co dla nich jest najważniejsze lub ma największą wartość. Do ustalenia priorytetów na podstawie macierzy porównań parowych możemy wykorzystać dwie metody: obliczając wektor własny macierzy bądź średnie geometryczne. Przyjmijmy że mamy daną macierz wzajemną, określoną dla  $n = 5$  kategorii:

$$\begin{bmatrix} 1 & 2 & 3 & 5 & 9 \\ \frac{1}{2} & 1 & 2 & 4 & 9 \\ \frac{1}{3} & \frac{1}{2} & 1 & 2 & 8 \\ \frac{1}{5} & \frac{1}{4} & \frac{1}{2} & 1 & 7 \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{8} & \frac{1}{7} & 1 \end{bmatrix}$$

Znormalizowany wektor wyznaczony w oparciu o wartości własne macierzy wynosi zatem:

$$\begin{bmatrix} 0.426 \\ 0.281 \\ 0.165 \\ 0.101 \\ 0.027 \end{bmatrix}$$

W konsekwencji możemy stwierdzić że wagi kategorii wyliczone za pomocą metody wartości własnych wynoszą:

$$w(x1) = 0.426, w(x2) = 0.281, w(x3) = 0.165, w(x4) = 0.101, w(x5) = 0.027$$

Niemal identyczny wynik otrzymamy obliczając średnią geometryczną każdego z wierszy macierzy oraz następnie dokonując normalizacji wyników.

Skoro istnieją dwie równoważne metody warto zastanowić się czy są pomiędzy nimi różnice i w końcu - którą z nich stosować? Rozważanie należy zacząć od tego, iż obie metody dają identyczny wynik jeśli obliczenia prowadzimy na macierzy spójnej (a co za tym idzie także wzajemnej), gdyż jej wektor własny jest równy średnim geometrycznym. Różnica pojawia się w samym sposobie wykonywania obliczeń. Znacznie prościej i szybciej jesteśmy w stanie obliczyć średnie geometryczne, nawet macierzy większych rozmiarów typu  $n \geq 5$ , niż wyznaczyć jej wartości własne. Jest to główny argument dla którego w pracy będę wykorzystywał metodę średnich geometrycznych.

Mając wektor  $V$  będący rezultatem działania  $(\prod_{j=1}^n a_{ij})^{1/n}$  obliczonego dla każdego wiersza macierzy, należy wyznaczyć sumę jego elementów. Aby otrzymać znormalizowany wektor wag  $V'$  należy podzielić każdy element wektora  $V$  przez sumę obliczoną w poprzednim kroku.

Formalnie:

$$V[*] = [(\prod_{j=1}^n a_{1j})^{1/n}, (\prod_{j=1}^n a_{2j})^{1/n}, (\prod_{j=1}^n a_{3j})^{1/n}, \dots, (\prod_{j=1}^n a_{nj})^{1/n}] \quad (2.2)$$

$$s = \sum_{i=1}^n V(i) \quad (2.3)$$

$$V'[*] = [\frac{v_1}{s}, \frac{v_2}{s}, \frac{v_3}{s}, \dots, \frac{v_n}{s}] \quad (2.4)$$

Na podstawie wektora  $V'$  można stworzyć poprawioną wersję macierzy źródłowej. Nowy element  $a'_{ij}$  jest ilorazem  $v'_i/v'_j$ .

### 3. Niespójność w metodzie porównań parami

Przeciwnicy metody porównywania parami opierają swoje zarzuty głównie na problemie wnioskowania na podstawie niespójnych danych. Mimo iż metoda ma silne fundamenty matematyczne, najsłabszym elementem jest człowiek i jego odpowiedzi na pytania ankiety. Przyczyną jest niespójność danych. Mamy z nią do czynienia w sytuacji kiedy wynik porównania danych dwóch kategorii podany przez respondenta odbiega od wartości która wynika z jego poprzednich porównań których składnikami były te kategorie. Przykładem może być sytuacja gdy trzymając w ręku trzy zapalki mamy za zadanie oszacować ich długość wykorzystując porównania parowe. Porównując pierwszą z drugą, wynikiem jest 1. Pierwszą z trzecią:  $1/2$ . Aby całe porównanie było spójne, wynik zestawienia zapalki drugiej i trzeciej musi zatem wynosić także  $1/2$ .

Przekładając to na język matematyki powiemy że macierz porównań parowych  $A$  (2.1) jest *spójna* (lub także *przechodnia*) jeśli  $a_{ij} * a_{jk} = a_{ik}$  dla każdego  $i, j, k = 1, \dots, n$ . Macierz spójna odpowiada sytuacji wzorcowej gdzie istnieją dokładne wartości  $s_1, \dots, s_n$  dla właściwości. Elementy macierzy  $A$  zdefiniowane są jako ilorazy  $a_{ij} = s_i/s_j$  i tworzą spójną macierz. W praktyce jednak bardzo rzadko zdarza się pracować na idealnych danych, dlatego stosuje się różne metody filtrowania wyników, a nawet ich dostrajania tak aby zapewniały minimum wymaganej dokładności.

Do określania stopnia poprawności danych wprowadzony został tzw. współczynnik niespójności. Analizując jego wartość możemy podejmować decyzję o uwzględnieniu lub odrzuceniu danej macierzy ze zbioru wynikowego. Na przestrzeni lat powstało wiele metod wyznaczania wartości współczynnika. Oprócz prac Saaty'ego i Koczkodaja, opisanych poniżej, także inni uczeni mają swój wkład. W 1993 roku Dodd, Donegan i McMaster przedstawili odwrotny wskaźnik niespójności [6]. Następnie w 1996 roku Monsuur zastosował transformacje maksymalnych wartości własnych [16], a siedem lat później Pelaez i Lamata badali wszystkie trójki elementów macierzy i użyli wyznacznika do określenia niespójności [12], ponadto w 2007 roku Stein oraz Mizzi wprowadzili harmoniczny współczynnik niespójności [20].

Innym typem współczynnika jest odległość od macierzy spójnej. Chu, Kalaba i Spingarn w 1979 roku użyli szacowania błędu metodą najmniejszych kwadratów [1], sześć lat po nich, w 1985 roku Crawford i Williams użyli logarytmicznego szacowania metodą najmniejszych kwadratów [7], a w 2003 roku Aguaron i Moreno-Jimenez zastosowali geometryczny współczynnik niespójności dla logarytmicznego oszacowania błędu metodą najmniejszych kwadratów (metoda średniej geometrycznej wiersza macierzy) [8]. W swojej pracy skupiłem się na publikacjach *A scaling method for priorities in hierarchical structures* [17] oraz *A new definition of consistency for pairwise comparisons* [13] i *Generalization of a new*

*definition of consistency for pairwise comparisons* [27], autorstwa dwóch wspomnianych na początku uczonych i zdecydowałem użyć ich metod radzenia sobie z niespójnością danych.

### 3.1. Metoda wartości własnych - Saaty

Jednym z pierwszych uczonych który podjął się tego tematu jest Thomas Saaty. W 1977 roku zaproponował metodę wyznaczania współczynnika niespójności macierzy porównań parowych [17]. Jego sposób opiera się na analizie wartości własnych macierzy. Wykorzystał w tym celu teorię macierzy rzeczywistych dodatnich Frobeniusa-Perrona [22] mówiącą iż największa wartość własna  $\lambda_{max}$  macierzy  $A$ , jest zawsze  $\lambda_{max} \geq n$ . Zauważył iż w przypadku gdy  $\lambda_{max}$  równa jest  $n$  to wtedy i tylko wtedy macierz  $A$  jest spójna. Indeks średniej niespójności ( $CI_n$ ) macierzy jest określony równaniem

$$CI_n = \frac{\lambda_{max} - n}{n - 1},$$

Wynika z niego iż indeks jest tak naprawdę przeskalowaniem największej wartości własnej.  $CI_n$  jest także zawsze nieujemny, ponieważ  $\lambda_{max} \geq n$ . Sam w sobie nie reprezentuje jednak żadnej wartości dopóki nie zostanie zestawiony z jakimkolwiek punktem odniesienia. Dzięki temu możemy określić wielkość odchylenia od macierzy spójnej. Punktem referencyjnym zaproponowanym przez Saaty'ego jest losowy współczynnik  $RI_n$ .

Przyjmijmy że jest dany zbiór np. 500 losowo wygenerowanych macierzy porównań parowych rozmiaru  $n \times n$ , takich że każdy element  $a_{ij} (i < j)$  został losowo wybrany spośród wartości skali

$$\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \dots, \frac{1}{2}, 1, 2, \dots, 8, 9$$

oraz  $a_{ji} = \frac{1}{a_{ij}}$ . Wtedy współczynnik  $RI_n$  oznacza średnią wartość wyliczonych dla tych losowych macierzy indeksów niespójności. Zależy on od rozmiaru macierzy  $n$  i wybranej skali. Wyliczone wartości współczynnika zostały przedstawione w tabeli 3.1. Współczynnik niespójności  $CR_n$  macierzy  $A$  wynosi

$$CR_n = \frac{CI_n}{RI_n}$$

Dla macierzy spójnej  $\lambda_{max} = n$ , więc  $CI_n = 0$  a także  $CR_n = 0$ . Saaty przyjął że niespójność na poziomie 10% lub mniej może być uznana za akceptowalną, a co za tym idzie macierz może zostać wykorzystana do przeprowadzenia wnioskowania na jej podstawie. Intuicyjne znaczenie reguły dziesięciu procent jest pomijane przez kilku autorów. Reguła została wyjaśniona statystycznie przez Vargas (1982) [19]. W roku 1994, kilkanaście lat po pierwszej publikacji, Saaty poprawił swoją teorię zmieniając maksymalne akceptowalne wartości niespójności na 5% dla macierzy  $3 \times 3$  oraz 8% dla  $4 \times 4$ .

Istnieje również tzw. reguła kciuka (ang. *thumb rule*) która upraszcza proces badania właściwości macierzy do analizy współczynnika  $CI$ . Przyjmuje się że jeżeli  $CI < 0.1$  to macierz jest wystarczająco spójna.

Pomimo iż jedyną szeroko akceptowalną i stosowaną zasadą niespójności dla dowolnego rodzaju macierzy jest praca Saaty'ego, jego definicja niespójności ma kilka słabych punktów. Głównym z nich

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$RI_n$	0	0	0	0.52	0.89	1.11	1.25	1.35	1.4	1.45	1.49	1.51	1.54	1.56	1.57	1.58

Tablica 3.1: Wartości losowego współczynnika niespójności dla różnej ilości kryteriów porównania (n)

wyduje się być zasada 10 procent, określona bardziej na podstawie doświadczeń i obserwacji aniżeli definicji matematycznej. Kolejnym elementem wskazywanym przez krytyków jest traktowanie niespójności w ujęciu globalnym - bez podania nawet przybliżonego jej źródła. Wartości własne są wartościami charakteryzującymi całą macierz, a więc przez ich analizę nie można stwierdzić które elementy macierzy najbardziej wpływają na zwiększenie niespójności.

### 3.2. Metoda odległościowa - Koczkodaj

Druga wybrana przeze mnie metoda została przedstawiona przez Waldemara Koczkodaj (1993) [13] oraz później przez Koczkodaj i Duszak (1994) [27]. Jak sam napisał w [13]: *“Autor tej pracy z całą mocą wierzy iż zdecydowanie za niska popularność metody porównań parami ma swoje źródła w definicji spójności.”*

Każda macierz porównań parowych rozmiaru  $3 \times 3$  ma 3 obiekty porównywane  $a, b, c$ . W celu przedstawienia współczynnika niespójności Koczkodaj rozważmy standardową macierz parową  $3 \times 3$ . Zakładając że macierz jest wzajemna możemy ją zredukować do wektora trzech współrzędnych  $(a, b, c)$ . W przypadku spójności, równość  $b = ac$  jest prawdziwa. Zawsze istnieje możliwość wyznaczenia trzech spójnych, wzajemnych macierzy parowych (reprezentowanych przez trzy wektory) poprzez wyliczenie jednej współrzędnej jako kombinacji pozostałych dwóch współrzędnych. Te trzy wektory są postaci:  $(\frac{b}{c}, b, c)$ ,  $(a, ac, c)$  oraz  $(a, b, \frac{b}{a})$ .

Współczynnik niespójności (nazywany także *miarą spójności CM*) dla macierzy parowej  $3 \times 3$  zdefiniowany jest przez Koczkodaj jako względna odległość od najbliższej, spójnej macierzy parowej  $3 \times 3$  reprezentowanej przez jeden z tych trzech wektorów.

**Def. 1.** *Miara spójności dla standardowej  $3 \times 3$  macierzy porównań parowych wynosi:*

$$CM(a, b, c) = \min \left\{ \frac{1}{a} \left| a - \frac{b}{c} \right|, \frac{1}{b} |b - ac|, \frac{1}{c} \left| c - \frac{b}{a} \right| \right\} \quad (3.1)$$

Współczynnik niespójności macierzy wzajemnej  $A$   $n \times n$  ( $n > 2$ ) wynosi:

$$CM(A) = \max \left\{ \min \left\{ \left| 1 - \frac{b}{ac} \right|, \left| 1 - \frac{ac}{b} \right| \right\} \quad \text{dla każdej triady } (a, b, c) \text{ w } A \right\} \quad (3.2)$$

W przypadku macierzy wyższych rzędów, wskaźnik niespójności elementu macierzy jest równy maksymalnej wartości  $CM$  ze wszystkich możliwych triad zawierających ten element.

Ilość wszystkich możliwych triad dla macierzy  $n \times n$  porównań parowych wynosi (Duszak i Koczkodaj, 1994):

$$\frac{n(n-1)(n-2)}{3!} \quad (3.3)$$

Do głównych zalet tej metody można zaliczyć:

- łatwość interpretacji (względne odchylenie od macierzy spójnej, która może zostać określona przy zachowaniu dwóch ocen niezmiennych),
- tworzy lepsze podstawy dla doboru progu (np. CM dla  $[1, 3, 2]$  lub  $[3, 6, 3]$  lub  $[3, 9, 4]$ , a nawet  $[2, 8, 3]$  wynosi 0.25 ),
- określa położenie spójności (CM jest zawsze połączony z konkretnym elementem macierzy, a nie enigmatycznym globalnym wskaźnikiem jak wartość własna).

W przypadku macierzy  $4 \times 4$  i skali *od 1 do 5* próg akceptowalnego rozwiązania nie powinien przekraczać  $1/3$  [24]. Ze względu jednak na lokalny charakter rozważanej niespójności, podany próg można uznać za prawdziwy i rozważać także dla macierzy innych wymiarów oraz dla innej skali porównawczej. Jak do tej pory nie stwierdzono zależności między rozmiarem, a zakresem skali w obszarze wpływu na wysokość progu akceptowalnego rozwiązania. Potencjalna obecność takiego związku może być przedmiotem kolejnych badań nad metodami porównań parowych.

## 4. Problem doboru skali w porównaniach parowych

Wszystkie teorie i prace traktujące o porównaniach parowych obowiązkowo muszą poruszyć problem skali porównania. Jest to niezwykle istotny element metody porównywania parami. Jak zauważają autorzy w [10] ma wpływ na całość metody w tym także może być przyczyną wpływającej na nią krytyki. Przekonują że właściwa skala pozwoli znacznie poprawić jakość danych początkowych, zmniejszając błędy niespójności i tym samym wytrącając argumenty przeciwników metody.

Najprostsza wersja to skala binarna *wygrana/przegrana* (Condorcet [4]). Odpowiada ona na pytanie: *Który?*, rozumiejąc przez to który z dwóch obiektów ma większą wartość w ramach porównywanej cechy: jest większy/mniejszy, tańszy/droższy, itd. Przykładem zastosowania mogą być wybory. Za każde zwycięstwo, kandydat otrzymuje 1 punkt, a za remis 0.5 punktu. Po przeliczeniu głosów otrzymujemy listę priorytetową kandydatów z których pierwszy jest zwycięzcą. Obecnie tego typu skala nie jest już wykorzystywana, lub ma to miejsce niezwykle rzadko. Zastąpiona została przez wielopoziomowe skale o różnym zakresie. Głównym elementem odróżniającym je od wersji podstawowej jest wartość liczbowa. Dzięki niej możemy odpowiedzieć na jeszcze jedno pytanie: *Ile?*. Jest to niezwykle cenna informacja, dająca realny punkt odniesienia. Pytanie może istnieć w dwóch formach, tj. *Ile razy?* oraz *O ile?*. Najczęściej spotykana i wykorzystywana jest forma pierwsza, ale sam problem decyzyjny mógłby posłużyć za inspirację do kolejnej pracy naukowej z pogranicza psychologii i matematyki.

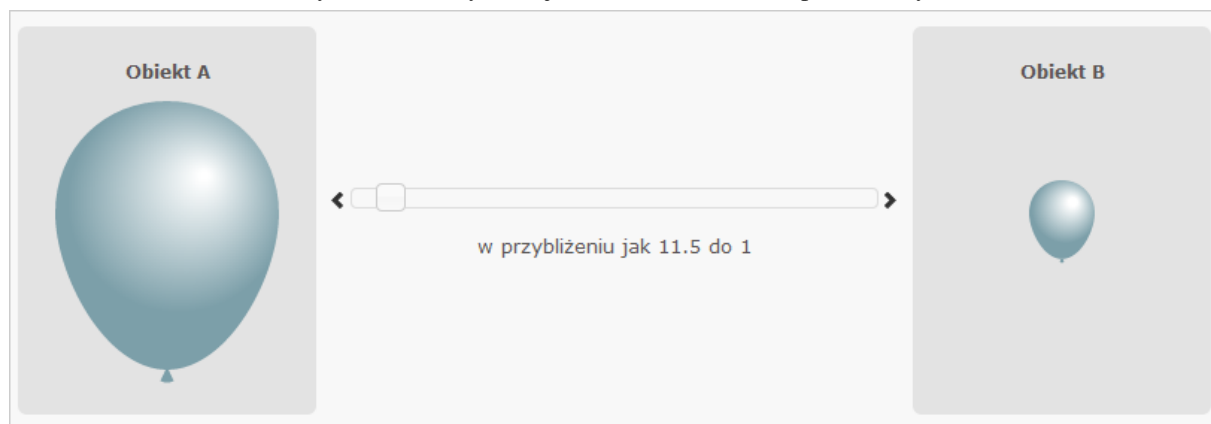
Nie ustają spory wśród naukowców na temat zakresu skali. Wielu autorów badających metodę porównań parowych najczęściej przedstawia swoje propozycje.

Zgodnie z pracą [9] istnieją teoretyczne podstawy by sądzić iż najlepsza skala powinna być mniejsza od stałej Fülöpa:  $\sqrt{\frac{1}{2}(11 + 5\sqrt{5})} \approx 3.330191$ . Odpowiednio dobrana skala wraz z adekwatnie prowadzoną stymulacją wzrokową respondentów pozwala na uzyskanie lepszej jakości wyników. Respondenci zwracają większą uwagę na różnice wielkości obrazków niż wartości liczbowych. W roli narzędzia służącego do ustalenia wartości najlepiej spisuje się suwak, który ma swoją pozycję startową na środku osi po której może się poruszać. Ilustracje 4.1 - 4.3 przedstawiają realizację powyższych wytycznych na potrzeby prototypowego narzędzia [21]. Porównywane ze sobą są dwa dowolne obiekty. Zmiana rozmiaru obiektu - przestawienie suwaka na inną pozycję - jest proporcjonalna w stosunku do zmiany wartości porównania.

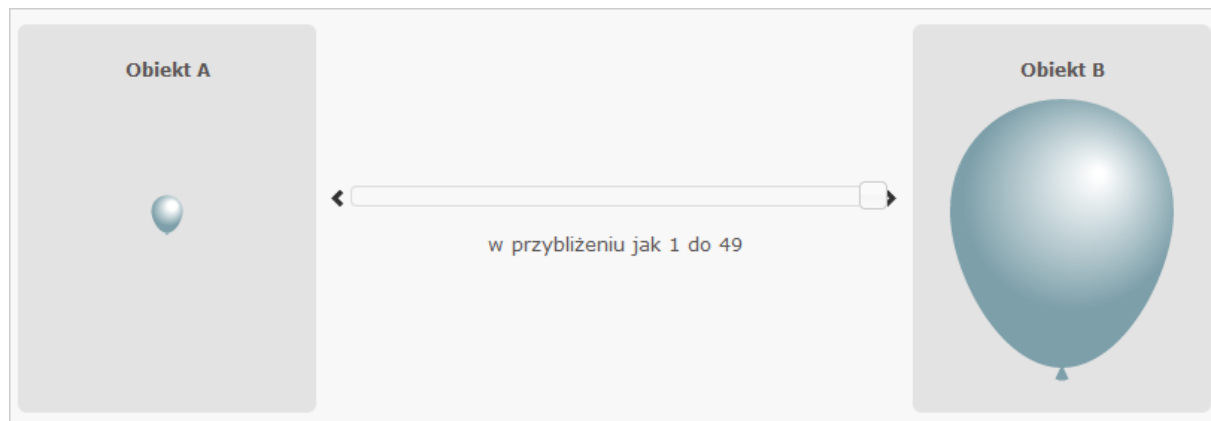




Rysunek 4.1: Stymulacja wzrokowa skali, stan początkowy



Rysunek 4.2: Ponad jedenastokrotna przewaga obiektu A w stosunku do B



Rysunek 4.3: Przytłaczająca różnica wartości na korzyść obiektu B

## 4.1. Skala Saaty'ego

Najszerzej znana i stosowana obecnie w porównaniach parowych, praca Saaty'ego ([17]) opiera się na dziewięciostopniowej skali (tabela 4.1). Daje respondentowi bardzo szeroki wybór. Pomiędzy elementami o mocno zdefiniowanym znaczeniu znajdują się punkty pośrednie, niejako zmiękczające znaczenie. Skala przedstawia stopnie w porządku rosnącym. Gdy dokonujemy porównania skala staje się symetryczna, a więc możemy wybrać wartości odwrotne, ułamkowe, gdy w naszej ocenie obiekt główny jest

*mniejszy/mniej ważny/itd...* od obiektu z nim zestawionego. Uzyskane z porównań wartości są elementami macierzy PC.

Wartość	Definicja	Znaczenie
1	taka sama ważność	Oba obiekty wpływają w tym samym stopniu na porównanie
3	średnio ważniejsze	Doświadczenie i ocena lekko faworyzuje jeden z obiektów
5	ważniejsze	Doświadczenie i ocena zdecydowanie wyróżnia jeden z obiektów
7	dużo ważniejsze	Doświadczenie i ocena bardzo mocno faworyzuje jeden z obiektów. Jego waga jest wykazana w praktyce
9	nieporównywalnie ważniejsze	Dowód za słuszością jednego z obiektów ma największą możliwą zasadność
2,4,6,8	wartości pośrednie	Kiedy trzeba zdobyć się na pewne ustępstwa

Tablica 4.1: Skala ocen wg Saaty'ego

## 4.2. Skala Likerta

To pięciostopniowa skala, którą wykorzystuje się w kwestionariuszach ankiet i wywiadach kwestionariuszowych, dzięki której można uzyskać odpowiedź dotyczącą stopnia akceptacji zjawiska czy też poglądu. Jest intensywnie eksploatowana w wielu metodach badań społecznych. Jej nazwa pochodzi od nazwiska Rensisa Likerta, który wynalazł ją w 1932 r.[15].

Skala składa się z pięciu poziomów odpowiedzi, ułożonych w kolejności od stopnia całkowitej akceptacji do całkowitego odrzucenia. Zadaniem ankietowanego jest określić w jakim stopniu zgadza się z danym twierdzeniem. Przykładowe warianty opisane na skali:

1. - zdecydowanie nie zgadzam się
2. - raczej się nie zgadzam
3. - nie mam zdania
4. - raczej się zgadzam
5. - zdecydowanie się zgadzam

Liczba możliwych do wyboru odpowiedzi powinna być nieparzysta i najczęściej jest to 5. Pozwala to na centralne położenie możliwie najbardziej neutralnego stwierdzenia. Badany wybiera tę możliwość, która najbardziej odpowiada jego odczuciom.

Skala powstała ze względu na potrzebę liczenia wskaźników zbudowanych z sumy (lub średniej) pytań kwestionariuszowych i w związku z tym różni się od innych skal porządkowych, że może być i z reguły jest traktowana jako ilościowa.

W badaniach kwestionariuszowych wyznacza się tzw. szczególne upodobanie. Pozwala to za pomocą w/w skali określić paradygmat (najogólniejszy model) zbioru ankietowanych.

### 4.3. Inne

Autorzy pracy [10] dowodzą iż najlepsza skala to taka, która ma możliwie najmniejszy zakres. Dając respondentowi szeroki przedział wartości, typu 1 – 100 narażamy się na wysokie błędy niespójności. Ponadto można zaobserwować iż w ocenach dominują wartości w okolicach pełnych dziesiątek 20, 30, 40, ..., 70, 80

Są 2 opcje pracy ze skalą i wybieranych wartości. Mogą być dyskretyzowane wartości lub dowolne z ustawioną dokładnością.

Przykłady innych skal można znaleźć w publikacji [11].

Jak widać wybór jest szeroki, każda opcja ma swoje wady i zalety. Dlatego w swojej pracy postanowiłem pozostawić tą decyzję twórcy ankiety. Dzięki temu może skorzystać z dowolnej wybranej lub nawet zastosować własną skalę.

## 5. Grupowanie hierarchiczne

Znakomita większość systemów spotykanych na co dzień posiada strukturę hierarchiczną. Składają się one z małych grup, np. rodzina i rozrastają do różnorodnych dużych organizacji jak społeczność miasta, województwa czy też cały naród. Na czele narodu stoi często rząd, który także ma ściśle określoną strukturę.

Teoria Hierarchii jest formą ogólnej teorii systemów [2]. Wyłoniła się jako osobna część w trakcie intensywnych prac nad złożonością. Mająca swoje korzenie w pracach o tematyce ekonomicznej, chemicznej oraz psychologicznej, teoria hierarchii skupia się na poziomach organizacji i problemach skali. Ma bardzo szerokie zastosowanie i jest powszechnie używana nie tylko przez naukowców.

W teorii hierarchii nie ma rozbudowanych i skomplikowanych reguł, wręcz przeciwnie, stosunkowo mały zbiór zasad pozwala uporządkować nawet skomplikowaną strukturę do postaci wielu poziomów. Z matematycznego punktu widzenia hierarchia jest zbiorem częściowo uporządkowanym. Poziomy wyższe są ponad poziomami niższymi oraz relacja w górę jest niesymetryczna względem relacji w dół. Poziom to nic innego jak zbiór elementów charakteryzowany przez ich właściwości, np. podział ludności ze względu na wiek czy kolor oczu.

Technicznie rzecz biorąc, hierarchia to zbiór zależności pomiędzy poziomami. Dla dwóch następujących po sobie poziomów, poziom wyższy jest nazywany *rodzicem*, a niższy *dzieckiem*. Hierarchia może reprezentować struktury organizacyjne, geograficzne lub dowolnego innego typu. Dlatego ważne jest wspieranie różnych rodzajów hierarchii wraz z odpowiednio dużą elastycznością konfiguracji. Dzięki temu możemy tworzyć struktury modelujące wiele różnych procesów.

### 5.1. Klasyfikator

Klasyfikator hierarchiczny pozwala na przetworzenie danych wejściowych dzięki czemu otrzymujemy odpowiednio pogrupowane wyniki [3]. Proces klasyfikacji rozpoczyna się od najniższego poziomu i do działania wykorzystuje oryginalne dane wejściowe. Następnie ocena poszczególnych jednostek jest sukcesywnie łączona i następuje klasyfikacja na wyższym poziomie. Proces jest kontynuowany do momentu uzyskania jednego, spójnego wyniku. Jest to więc czysty przykład metody bottom-up. Warto odnotować, iż proces klasyfikacji hierarchicznej opiera się nie na możliwościach obliczeniowych wejściowych komponentów, ale na samej strukturze hierarchii. W konsekwencji, dzięki temu hierarchia może być prostą i łatwo rozszerzalną strukturą.

## 5.2. Zastosowanie hierarchicznego klasyfikatora

Na przestrzeni kilku ostatnich dekad zostało stworzonych wiele zastosowań dla grupowania hierarchicznego. Jednym ze znakomitych przykładów jest cała gałąź rozpoznawania obrazów. Rozpoznawanie obrazów jest uznawane za jeden z najbardziej skomplikowanych problemów, który wymaga znacznych mocy obliczeniowych. Dość powiedzieć, że większość zabezpieczeń anty-robotowych znajdujących się na stronach www opiera się właśnie na rozpoznawaniu obrazów, gdyż nikt tak jak człowiek nie potrafi sobie z nimi poradzić. Na pomoc technologii przychodzi przetwarzanie hierarchiczne, które dobrze współpracuje z rozpoznawaniem obrazów. A to dlatego, iż zdjęcia/obrazy mogą być przedstawione jako zbiór komponentów lub obiektów. Te obiekty z kolei mogą być postrzegane jako zbiór mniejszych kształtów. Na nie składają się zbiory linii itd. Widać wyraźnie, iż takie przetwarzanie jest zbieżne z tym jak działa klasyfikacja hierarchiczna.

Podobnie proces może zachodzić w drugą stronę, przetwarzając kształty i linie w obiekty, przy czym najprawdopodobniej będą musiały zostać wykonane pewne dodatkowe kroki pośrednie. W rezultacie, mając wszystkie elementy ułożone hierarchicznie w skierowanym, acyklicznym grafie, zaczynając od pikseli i pnać się w górę możemy przeprowadzić wnioskowanie, co znajduje się na obrazku.

Istnieje wiele innych zastosowań, gdzie z powodzeniem wykorzystamy grupowanie hierarchiczne, jak np. rozpoznawanie tekstu pisanego (OCR), czy też samo-świadomość robotów.

## 5.3. Hierarchia w metodzie PC

Struktury hierarchiczne w metodzie porównań parowych mają swój początek w latach 70. ubiegłego wieku w pracach Thomasa Saaty'ego. Przedstawił on metodę *Analytic Hierarchy Process* (AHP). Jest to wielokryterialna metoda hierarchicznej analizy problemów decyzyjnych. Umożliwia ona przedstawienie skomplikowanego i złożonego problemu decyzyjnego za pomocą prostej struktury hierarchicznej i na jej podstawie utworzenie rankingu dla określonych wariantów. Metoda jest wykorzystywana w wielu dziedzinach, od politologii i socjologii, aż po zarządzanie i transport. Radzi sobie z problemem uszeregowania skończonej ilości kategorii (wariantów) ocenianych pod względem również skończonej liczby kryteriów. Złożoność problemu jest najlepiej widoczna poprzez ilość poziomów w hierarchii po zdekomponowaniu. AHP wykorzystuje algebrę macierzową do poukładania wariantów w sposób optymalny pod względem matematycznym. Jest to sprawdzona metoda, używana do podejmowania decyzji w sprawach o wielomilionowych stawkach.

Najpopularniejszymi obszarami zastosowania metody AHP są:

- klasyfikowanie wymogów i czynników które mają wpływ na procesy powstawania i rozwoju oprogramowania,
- wybieranie najlepszej z możliwych strategii dla poprawy bezpieczeństwa samochodów i innych pojazdów silnikowych,

- szacowanie kosztów i ustalanie opcji dla planowania zapotrzebowania materiałowego (MRP),
- wybieranie najbardziej pożądaných komponentów oprogramowania spośród grona dostawców,
- ocenianie jakości badań rozwojowych lub planów inwestycyjnych.

W obliczeniach, oprócz macierzy wynikowych porównań, wykorzystywane są również właściwe wartości charakterystyczne wariantów jak wartość, cena, ilość, waga, itd. Źródłem błędów metody jest niespójność macierzy parowych, pozostałe etapy przetwarzania oparte są na obliczeniach matematycznych.

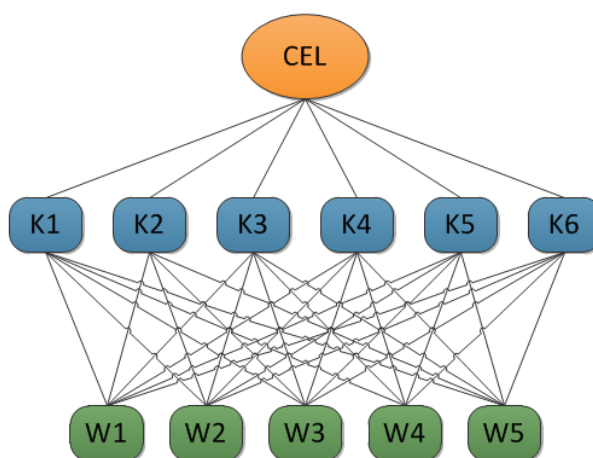
Algorytm metody AHP składa się z czterech etapów:

1. Definicja problemu i określenie celu do którego dążymy.
2. Definicja kategorii i wariantów mających wpływ na cel oraz uporządkowanie ich w strukturze hierarchicznej.
3. Zbieranie danych o preferencjach respondenta.
4. Tworzenie rankingu końcowego.

Niekiedy jako ostatni krok dodaje się tzw. analizę wrażliwości, polegającą na badaniu w jaki sposób poszczególne czynniki wpływają na końcowy rezultat.

Słowo wyjaśnienia dla kroku trzeciego. W tym miejscu wykorzystywane są porównania parowe. Zestawiane są ze sobą warianty na każdym poziomie hierarchii, ale nie pomiędzy poziomami. Porównania wykorzystują skalę Saaty'ego opisaną w rozdziale 4.1. Po dokonaniu porównań ma miejsce wstępne przetwarzanie danych dla każdej macierzy, obliczana jest niespójność (metodą przedstawioną w r. 3.1) oraz tworzony jest ranking.

Rysunek 5.1 przedstawia skondensowaną postać przykładowej struktury dla problemu polegającego na uszeregowaniu pięciu wariantów ocenianych względem sześciu kryteriów.



Rysunek 5.1: Model hierarchicznej struktury procesu decyzyjnego AHP

### 5.3.1. Przykład

Aby lepiej zobrazować proces oceny posłużę się przykładem.

Wyobraźmy sobie iż ekspert X staje przed wyborem nowego telefonu - smartfona dla siebie. Pod uwagę bierze tylko najnowsze i najlepsze konstrukcje, tzw. flagowce. Zakłada iż budżet w wysokości 2000 zł będzie wystarczający. Wstępne rozeznanie ogranicza jego wybór do 4 modeli ( $A, B, C, D$ ). Główne cechy, które mają wpływ na ostateczną decyzję to cena, długość pracy na baterii, rozmiar ekranu, ilość pamięci wewnętrznej oraz jakość wbudowanego aparatu fotograficznego. Korzystając z AHP, jako pierwsze należy określić problem. Oznacza to zdefiniowanie struktury najlepiej odpowiadającej porównaniu smartfonów w ramach wybranych cech. Tabela 5.1 przedstawia zebrane urządzenia i ich cechy które rozważa ekspert X. Są to prawdziwe dane telefonów powstałych w 2014 roku. Cena jest przybliżoną średnią cen z popularnego polskiego serwisu aukcyjnego z dnia 15 marca 2015. Dane o żywotności baterii pochodzą z testów serwisu GsmArena. Dla eksperta X dostępne jest oprogramowanie AHP zajmujące się wszystkimi wyliczeniami. Od niego wymagamy dokonania porównań.

Telefon A - iPhone 6 Plus

Telefon B - Samsung Galaxy S5

Telefon C - Nokia Lumia 930

Telefon D - Sony Xperia Z3

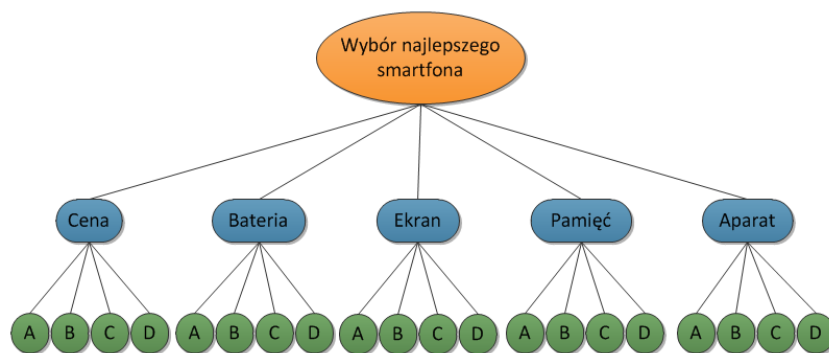
Cecha	A	B	C	D
Cena	3100 zł	1550 zł	1850 zł	1900 zł
Bateria	79h	72h	46h	85h
Ekran	5.5"FHD	5.1 FHD	5"FHD	5.2"FHD
Pamięć	64 GB	16GB	32 GB	16GB
Aparat	8Mpx	16Mpx	20Mpx	20.2 Mpx

Tablica 5.1: Cechy i rozważane alternatywy w ramach wyboru smartfona

Drugim krokiem jest określenie zależności rodzic-dziecko oraz zbudowanie hierarchii kryteriów i alternatyw przedstawionej na rysunku 5.2.

Najwyższy poziom, korzeń, przedstawia cel "wybór najlepszego smartfona". Poziom drugi zawiera cechy produktów w ramach których będą porównywane telefony z poziomu trzeciego.

Dysponując wszystkimi danymi, ekspert przystępuje do trzeciej, najdłuższej części procesu AHP. Musi dokonać szeregu porównań parowych na każdym poziomie oraz w kontekście cechy. Oprogramowanie wykorzystuje skalę Saaty'ego 4.1. Pierwsze porównanie to *cena vs bateria*. Nasz bohater uznaje iż czas pracy na baterii jest nieco ważniejszy od ceny (ocena 3), a cena ma zdecydowanie większy wpływ



Rysunek 5.2: Hierarchia atrybutów i telefonów

na ocenę końcową niż pamięć (ocena 5). Pan X kontynuuje oceny, które zebrane tworzą macierz:

$$\begin{array}{c}
 \begin{array}{ccccc}
 & \text{cena} & \text{bateria} & \text{ekran} & \text{pamiec} & \text{aparat} \\
 \text{cena} & 1 & \frac{1}{3} & 5 & 3 & \frac{1}{2} \\
 \text{bateria} & 3 & 1 & 5 & 4 & 2 \\
 \text{ekran} & \frac{1}{5} & \frac{1}{5} & 1 & \frac{1}{3} & \frac{1}{6} \\
 \text{pamiec} & \frac{1}{3} & \frac{1}{4} & 3 & 1 & 2 \\
 \text{aparat} & 2 & \frac{1}{2} & 6 & \frac{1}{2} & 1
 \end{array}
 \end{array} \quad (5.1)$$

Dla tak powstałej macierzy przeprowadzane są obliczenia. Najpierw, korzystając z metody wartości własnych (3.1) wyznaczana jest niespójność. Rozmiar macierzy wynosi  $n = 5$ . Największa wartość własna  $\lambda_{max} = 5.559$ , a zatem współczynnik  $CI_5 = \frac{5.559-5}{5-1} \approx 0.14$ . Mając powyższe dane, określany jest współczynnik niespójności

$$CR_5 = \frac{0.14}{1.11} \approx 0.126$$

Wyliczona wartość jest nieco powyżej progu akceptacji wyznaczonego przez Saaty'ego, lecz dalej na poziomie tolerancji, a ze względu na większy rozmiar macierzy i niewielkie odchylenie od normy, ekspert X postanawia zaakceptować te dane i kontynuować pracę. Wyznaczony zostaje wektor wag dla cech:

Cecha	Priorytet
Cena	0.194
Bateria	0.420
Ekran	0.045
Pamięć	0.140
Aparat	0.201

Tablica 5.2: Główny wektor wag dla cech



Warto zauważyć że suma priorytetów (wag) sumuje się do wartości 1. Wyniki porównania najważniejszych cech nie pozostawiły wątpliwości - największe znaczenie dla eksperta ma długość pracy na baterii.

Kolejny krok to ocena każdego z wybranych modeli pod względem cech. Kolejność dokonywania porównań nie ma znaczenia, ale pan X zdecydował użyć powyżej zdefiniowanej kolejności. Porównań jest wiele, więc aby nieco ułatwić sobie pracę stworzył tabelkę, przedstawioną na rysunku 5.3, ułatwiającą podejmowanie decyzji.

Długość pracy na baterii (więcej = lepiej)								Ocena pana Jana	
Telefony			Długość pracy		Lepsza bateria	Bateria lepsza jako		Lepszy tel.	Ocena wg skali
#	A	B	A	B		Ilość h	Stosunek		
1	Iphone 6 Plus	Samsung Galaxy S5	79h	72h	A	7h	1.1	A	2
2	Iphone 6 Plus	Nokia Lumia 930	79h	46h	A	33h	1.72	A	7
3	Iphone 6 Plus	Sony Xperia Z3	79h	85h	B	6h	1.08	B	2
4	Samsung Galaxy S5	Nokia Lumia 930	72h	46h	A	26h	1.57	A	6
5	Samsung Galaxy S5	Sony Xperia Z3	72h	85h	B	13h	1.18	B	3
6	Nokia Lumia 930	Sony Xperia Z3	46h	85h	B	39h	1.85	B	8

Rysunek 5.3: Ocena telefonów pod względem długości czasu pracy na baterii

W rezultacie otrzymał kolejny wektor wag:

Alternatywa	Priorytet lokalny	Priorytet globalny
iPhone 6 Plus	0.300	0.124
Samsung Galaxy S5	0.180	0.076
Nokia Lumia 930	0.040	0.018
Sony Xperia Z3	0.480	0.201
<b>SUMA</b>	<b>1.000</b>	<b>0.420</b>

Tablica 5.3: Priorytety telefonów z perspektywy czasu pracy na baterii

Priorytet globalny wylicza się jako iloczyn lokalnego oraz iloczynu priorytetów wszystkich poprzednich poziomów. Jako że rozważany przykład ma 2 poziomy, dla baterii używamy wagę 0.420. Jak widać w wierszu podsumowania, suma priorytetów globalnych jest równa wadze nadrzędnej.

Nasz bohater przeprowadza do końca pozostałe porównania, otrzymując następujące priorytety:

Alternatywa	Priorytet lokalny	Priorytet globalny
iPhone 6 Plus	0.043	0.009
Samsung Galaxy S5	0.142	0.028
Nokia Lumia 930	0.317	0.064
Sony Xperia Z3	0.498	0.100
<b>SUMA</b>	<b>1.000</b>	<b>0.201</b>

Tablica 5.4: Priorytety telefonów z perspektywy aparatu

Alternatywa	Priorytet lokalny	Priorytet globalny
iPhone 6 Plus	0.038	0.007
Samsung Galaxy S5	0.524	0.102
Nokia Lumia 930	0.257	0.050
Sony Xperia Z3	0.181	0.035
<b>SUMA</b>	<b>1.000</b>	<b>0.194</b>

Tablica 5.5: Priorytety telefonów z perspektywy ceny

Alternatywa	Priorytet lokalny	Priorytet globalny
iPhone 6 Plus	0.664	0.093
Samsung Galaxy S5	0.057	0.008
Nokia Lumia 930	0.221	0.031
Sony Xperia Z3	0.057	0.008
<b>SUMA</b>	<b>1.000</b>	<b>0.140</b>

Tablica 5.6: Priorytety telefonów z perspektywy ilości pamięci

Alternatywa	Priorytet lokalny	Priorytet globalny
iPhone 6 Plus	0.050	0.002
Samsung Galaxy S5	0.290	0.013
Nokia Lumia 930	0.490	0.022
Sony Xperia Z3	0.180	0.008
<b>SUMA</b>	<b>1.000</b>	<b>0.045</b>

Tablica 5.7: Priorytety telefonów z perspektywy ekranu

Tym samym etap trzeci został zakończony, a odpowiedzi przeliczone. Oprogramowanie prezentuje końcowe wyniki w formie tabeli 5.4 wag rozważanych alternatyw i ich cech wraz z podsumowaniem. Suma ostateczna wynosi 1.000, a zatem jest równa priorytetowi celu.

Model	Bateria	Aparat	Cena	Pamięć	Ekran	Suma
iPhone 6 Plus	0.124	0.009	0.007	0.093	0.002	0.236
Samsung Galaxy S5	0.076	0.028	0.102	0.008	0.013	0.228
Nokia Lumia 930	0.018	0.064	0.050	0.031	0.022	0.184
Sony Xperia Z3	0.201	0.100	0.035	0.008	0.008	0.353
<b>SUMA</b>	<b>0.420</b>	<b>0.201</b>	<b>0.194</b>	<b>0.140</b>	<b>0.045</b>	<b>1.000</b>

Rysunek 5.4: Globalne priorytety dla eksperta X wybierającego telefon

Sony Xperia Z3, posiadająca globalny priorytet 0.353, jest alternatywą mającą największy udział w celu będącym wyborem najlepszego smartfona dla eksperta X. Na drugim miejscu można umiejscowić

wspólnie iPhone 6 Plus oraz Samsung Galaxy S5. Różnica ich wyników znajduje się dużo poniżej błędu wynikającego z niespójności wszystkich porównań. Ostatni model, telefon firmy Nokia, znalazł się na końcu stawki lecz jego rezultat pokazuje że nie traci kontaktu z czołówką.

## 6. Projekt

Podstawowy pomysł na system był następujący: zrobmy internetową aplikację ankietującą, która do badania opinii będzie wykorzystywać metodę porównywania parami. Zgodnie ze stanem obecnej wiedzy nie ma dostępnego podobnego rozwiązania, dzięki czemu istnieje duża szansa na zaistnienie na rynku. Kolejnym potencjalnym zastosowaniem są badania naukowe nad metodą porównań parowych, dowodzenie jej poprawności i popularyzacji zarówno wśród specjalistów jak i użytkowników prywatnych.

### 6.1. Analiza wymagań

#### 6.1.1. Wymagania funkcjonalne

Głównym zadaniem i celem systemu jest umożliwienie tworzenia ankiet i brania w nich udziału (odpowiadania), a także gromadzenie i przetwarzanie zebranych danych na potrzeby statystyczne. Rysunek 6.1 przedstawia główne przypadki użycia opisujące powyższe funkcjonalności z perspektywy użytkowników systemu.

Na etapie projektowania wyróżniłem trzech podstawowych aktorów, odpowiadają oni rolom użytkowników w aplikacji. Są to:

Administrator	ADMIN
Zarejestrowany użytkownik	USER
Gość	GUEST

Każdy z nich może realizować różne scenariusze. Najważniejsze z nich przedstawiam poniżej.

<b>UC 1</b>	<b>Załącz konto/rejestruj</b>
<i>Cel:</i>	Założenie konta w systemie
<i>Główny Aktor:</i>	GUEST
<i>Scenariusz główny:</i>	

1. Gość chce założyć konto w systemie
2. Gość uzupełnia formularz rejestracji podając dane kontaktowe i dostępowe
3. System weryfikuje poprawność i unikalność danych
4. System wysyła wiadomość email z linkiem aktywacyjnym na podany adres
5. Gość odbiera wiadomość i przechodzi na stronę aktywacji konta
6. System weryfikuje poprawność kodu aktywacyjnego i zatwierdza konto
7. Gość może się zalogować do systemu

---

**UC 2                      Weź udział w ankiecie**

---

*Cel:*                      Udzielenie odpowiedzi i jej zapis w systemie

---

*Poziom:*                Funkcja systemowa

---

*Główny Aktor:*      GUEST, USER

---

*Scenariusz główny:*

1. Użytkownik wybiera interesującą go ankietę z katalogu ankiet: gość widzi tylko ankiety publiczne, zalogowany także te do których wypełnienia został zaproszony
  2. Wypełnia formularz danymi podstawowymi, jeśli jest zalogowany, dane są uzupełnione automatycznie
  3. Wyraża zgodę na przechowywanie danych
  4. Udziela odpowiedzi na wszystkie pytania
  5. Zatwierdza wprowadzone dane
  6. System przetwarza i zapisuje dane
  7. Użytkownik widzi ekran podsumowania zapisanej odpowiedzi
-

*Rozszerzenia i scenariusze alternatywne:*

- 1.a Użytkownik otwiera link który otrzymał w wiadomości e-mail będącej zaproszeniem do wzięcia udziału w ankiecie
- 3.a Użytkownik nie wyraził zgody, więc nie może wziąć udziału w badaniu

**UC 3****Zarządzaj ankietami***Cel:*

Zarządzanie i organizacja własnych ankiet

*Poziom:*

Funkcja systemowa

*Główny Aktor:*

USER

*Warunki początkowe:*

- Użytkownik posiada konto w systemie
- Użytkownik jest zalogowany

*Scenariusz główny:*

1. Użytkownik otwiera domyślny widok panelu sterowania
2. Przegląda ankietę, może podjąć akcje:
  1. Wyszukiwanie
  2. Rozpoczęcie tworzenia nowej ankiety
  3. Edycja
  4. Usuwanie
  5. Zmiana statusu

**UC 4****Dodaj nową ankietę***Cel:*

Dodanie nowej ankiety w stanie roboczym

<i>Poziom:</i>	Funkcja systemowa
<i>Główny Aktor:</i>	USER
<i>Warunki początkowe:</i>	<ul style="list-style-type: none"> <li>– Użytkownik posiada konto w systemie</li> <li>– Użytkownik jest zalogowany</li> </ul>
<i>Scenariusz główny:</i>	
<ol style="list-style-type: none"> <li>1. Użytkownik inicjuje proces definiowania ankiety</li> <li>2. Uzupełnia formularz danych podstawowych i zapisuje</li> <li>3. System weryfikuje poprawność danych i prezentuje główny widok szczegółów ankiety</li> <li>4. Twórca definiuje obowiązkowe elementy: <ul style="list-style-type: none"> <li>– Przedziały znaczeniowe skali UC 5</li> <li>– Porównywane obiekty UC 6</li> </ul> </li> <li>5. Ankieta jest gotowa do aktywacji</li> </ol>	
<i>Rozszerzenia i scenariusze alternatywne:</i>	
<ol style="list-style-type: none"> <li>3.a Użytkownik nie podał wszystkich wymaganych danych, lub zawierają one błędy, system prezentuje informacje o nich i wraca do pkt. 2</li> <li>5.a Opcjonalnie użytkownik dodaje respondentów którzy otrzymają zaproszenia do wzięcia udziału w ankiecie</li> </ol>	
<b>UC 5</b>	<b>Definiuj przedziały znaczeniowe skali porównawczej</b>
<i>Cel:</i>	Określenie podziału zakresu wartości na przedziały wraz z opisami ich znaczenia
<i>Poziom:</i>	Funkcja systemowa

<i>Główny Aktor:</i>	USER
<i>Warunki początkowe:</i>	<ul style="list-style-type: none"> <li>– Użytkownik posiada konto w systemie</li> <li>– Użytkownik jest zalogowany</li> <li>– Istnieje ankieta ze statusem <i>Nowa</i> oraz zdefiniowany jest zakres wartości skali (określona jest minimalna i maksymalna wartość skali)</li> </ul>
<i>Scenariusz główny:</i>	<ol style="list-style-type: none"> <li>1. Użytkownik wybiera ankietę i otwiera widok podstawowy</li> <li>2. Uruchamia definiowanie przedziałów - system prezentuje pasek skali o wartościach z zakresu [min. . . max]</li> <li>3. Użytkownik dodaje nowy marker <ul style="list-style-type: none"> <li>– Ustawia pozycję na pasku</li> <li>– Podaje tekst opisujący marker, może używać symboli zastępczych w miejscach których zostaną wstawione nazwy porównywanych kategorii</li> </ul> </li> <li>4. Użytkownik powtarza poprzedni krok tyle razy ile markerów (przedziałów znaczeniowych) chce utworzyć</li> </ol>
<b>UC 6</b>	<b>Definiuj kategorie</b>
<i>Cel:</i>	Zdefiniowanie kategorii które będą porównywane w ankiecie oraz ich organizacja
<i>Poziom:</i>	Funkcja systemowa
<i>Główny Aktor:</i>	USER



- Warunki początkowe:*
- Użytkownik posiada konto w systemie
  - Użytkownik jest zalogowany
  - Istnieje ankieta ze statusem *Nowa*

*Scenariusz główny:*

1. Użytkownik wybiera ankietę i otwiera widok podstawowy
2. Przechodzi do sekcji definiowania kategorii
3. Naciska przycisk dodawania i z dwóch opcji wybiera *Kategoria*
4. Wypełnia formularz nowej kategorii, obowiązkowo podaje nazwę, zapisuje
5. System weryfikuje poprawność danych i po poprawnym zapisie wraca do widoku kategorii
6. Użytkownik powtarza kroki 3 - 5 w celu zdefiniowania wszystkich potrzebnych obiektów

*Rozszerzenia i scenariusze alternatywne:*

3.a Użytkownik wybrał opcję *Grupa*

1. Podaje nazwę grupy i zapisuje
2. Przenosi wybrane kategorie do grupy aby utworzyć hierarchię
3. Powtarza scenariusz główny (3 - 5) lub alternatywny 3a. do momentu osiągnięcia pożądanej struktury grup i kategorii
4. Zapisuje hierarchię

**UC 7**

**Aktywuj ankietę**

*Cel:* Udostępnienie ankiety respondentom w celu przeprowadzenia badań

*Poziom:* Funkcja systemowa

---

Główny Aktor:           USER

---

- Warunki początkowe:
- Użytkownik posiada konto w systemie
  - Użytkownik jest zalogowany
  - Istnieje ankieta ze statusem *Nowa*
- 

Scenariusz główny:

1. Użytkownik wybiera ankietę i otwiera widok podstawowy
  2. Naciska przycisk aktywacji
  3. System prezentuje widok kompletnej ankiety, w takiej postaci jak ją będą widzieć respondenci
  4. Użytkownik zatwierdza widok
  5. System zapisuje pytania i prezentuje widok wyboru trybu ankiety
  6. Użytkownik wybiera opcję *Standardowy*
  7. System zapisuje tryb, publikuje ankietę i wysyła zaproszenia w formie wiadomości e-mail do zdefiniowanych respondentów
- 

Rozszerzenia i scenariusze alternatywne:

6.a Użytkownik wybrał tryb *Częściowy*

1. Użytkownik podaje jak wiele różnych pod-ankiet  $n$  chce przygotować i zatwierdza
  2. System generuje  $n$  ankiet, z których każda jest unikalną permutacją zbioru wszystkich pytań
  3. Powtarza krok 6.a.1 jeśli wynik losowania nie jest satysfakcjonujący lub zmienił zdanie w stosunku do wartości  $n$
  4. Zatwierdza wybór i rezultat losowania
  5. Kontynuacja scenariusza głównego od punktu 7
-

<b>UC 8</b>	<b>Przeglądaj statystyki</b>
<i>Cel:</i>	Odczytanie wyników przeprowadzonych badań
<i>Poziom:</i>	Funkcja systemowa
<i>Główny Aktor:</i>	USER
<i>Warunki początkowe:</i>	<ul style="list-style-type: none"> <li>– Użytkownik posiada konto w systemie</li> <li>– Użytkownik jest zalogowany</li> <li>– Istnieje ankieta ze statusem <i>W trakcie</i> lub <i>Zakończona</i></li> </ul>
<i>Scenariusz główny:</i>	
<ol style="list-style-type: none"> <li>1. Użytkownik wybiera ankietę i otwiera widok podstawowy</li> <li>2. Przechodzi do statystyk</li> <li>3. System prezentuje widok wyników badania wyliczony na podstawie udzielonych odpowiedzi</li> <li>4. Użytkownik zmienia parametry widoku i zatwierdza</li> <li>5. System prezentuje wyniki badania z uwzględnieniem parametrów wybranych powyżej</li> </ol>	

### 6.1.2. Wymagania niefunkcjonalne

Względem system określiłem następujące wymagania niefunkcjonalne:

1. Zewnętrzny interfejs systemu będzie zbudowany w oparciu o technologię HTML w postaci stron WWW.
2. System będzie dystrybuowany jako oprogramowanie internetowe w modelu SaaS - użytkownik uzyskuje dostęp do całej platformy poprzez założenie konta, bez konieczności instalacji dodatkowego oprogramowania.
3. Wymagana włączona obsługa języka JavaScript w przeglądarce użytkownika, bez tego korzystanie z aplikacji nie będzie możliwe.



Rysunek 6.1: Diagram głównych przypadków użycia dla aplikacji ankietowej

4. Zgodność z przepisami prawa - niedopuszczalne jest aby system w jakikolwiek sposób naruszał prawo, w szczególności prawo autorskie.
5. Stabilność i szybkość działania
  - Łatwa i intuicyjna obsługa.
  - Estetyczny i czytelny wygląd.
  - Sprawne działanie przy 100 użytkownikach pracujących równocześnie - czas odpowiedzi/ładowania strony nie dłuższy niż 5 s.
  - Minimalizacja błędów powstałych przez umyślną bądź nieumyślną działalność użytkownika.

- Zapewnienie poprawnej pracy systemu na dostępnych przeglądarkach WWW.

## 6. Bezpieczeństwo

- Identyfikacja i uwierzytelnienie użytkowników - na podstawie loginu oraz hasła. Hasło przechowywane w bazie danych w postaci sumy kontrolnej z solą.
- Autoryzacja - bazuje na roli użytkownika. Dotyczy dostępu do widoków i funkcjonalności.
- Brak dostępu do edycji oraz podglądu nie-własnych ankiet i statystyk - nie dotyczy Administratora.
- Ochrona przeciw SQL Injection - walidacja i binding danych wejściowych.
- Zapobieganie wprowadzaniu nieprawidłowych danych do systemu.

## 6.2. Architektura systemu

System został stworzony w języku PHP przy pomocy frameworka *Symfony* w wersji 2. Jest to bardzo rozbudowana platforma pozwalająca na łatwą i szybką implementację pożądaných funkcjonalności wykorzystując programowanie obiektowe. Aplikacja docelowa składa się z modułów, tzw. *bundle* połączonych przez framework. Wykorzystuje wzorzec Model - Widok - Kontroler (ang. *MVC: Model, View, Controller*) dzięki któremu główne warstwy aplikacji są od siebie wyraźnie oddzielone, nie pozwalając na mieszanie ich ze sobą. Pozwala to na utrzymanie wysokiej jakości kodu źródłowego, a także ma duży wpływ na organizację pracy. Łatwiej taką aplikację utrzymywać i rozwijać, dodając w przyszłości nowe funkcje.

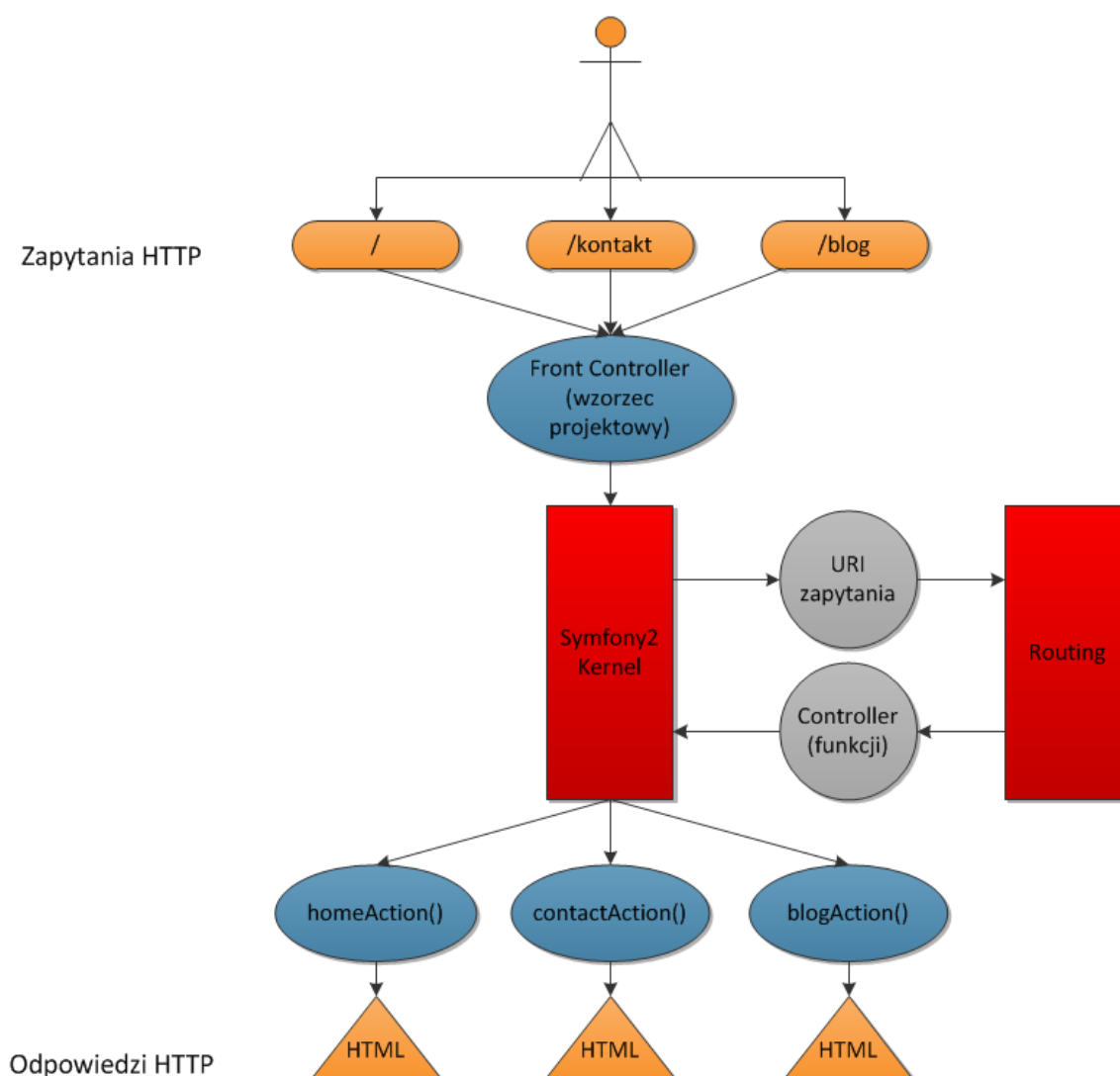
Cała logika biznesowa zaimplementowana jest w *Akcjach* zdefiniowanych w *Kontrolerach*. Połączenie między żądaniami użytkownika, a akcjami jest wynikiem konfiguracji nawigacji (ang. *routing*). Elementy widoczne dla użytkownika, prezentowane w przeglądarce, przygotowywane są przez system szablonów.

Aplikacja zawsze będzie miała za zadanie przetwarzać pewne dane. Aby mogła z nimi pracować potrzebuje *Modelu* - definicji elementów i ich zależności.

Jeśli budowana aplikacja wykorzystuje do działania bazę danych, model jest odwzorowaniem schematu bazy. Mamy wtedy do czynienia z mapowaniem obiektowo-relacyjnym, tzw. *ORM*. *Symfony* używa *Doctrine* jako narzędzie ORM oraz bazodanową warstwę abstrakcji *DBAL*. Dzięki temu nie ma znaczenia silnik bazy danych, który wykorzystujemy w naszej aplikacji. Warstwa *DBAL* potrafi współpracować z większością współczesnych relacyjnych baz danych, np. *SQLite*, *MySQL*, *PostgreSQL*, *Oracle*, *Microsoft SQL*, itd. Istnieje także możliwość stworzenia sterownika dla własnego silnika bazy. Aplikacja stworzona przeze mnie używa bazy danych *MySQL*.

Gdy tworzone rozwiązanie wymaga użycia bazy dokumentowej, można skorzystać z *MongoDB ODM* (*Object Document Mapper*), który w filozofii i sposobie działania jest bardzo podobny do *Doctrine ORM*.

Cykl życia aplikacji internetowej, zgodnie z definicją protokołu HTTP, rozpoczyna się w momencie gdy użytkownik wykonuje zapytanie *request*, następnie żądanie jest przetwarzane i serwer wysyła odpowiedź *response*. Dokładnie ten sam schemat jest realizowany w przypadku Symfony:



Rysunek 6.2: Schemat działania aplikacji we frameworku Symfony

Przychodzące zapytania są interpretowane przez routing i przekazywane do funkcji kontrolerów, które zwracają odpowiedź.

Każda *strona* aplikacji jest zdefiniowana w pliku konfiguracyjnym nawigacji (routingu), który mapuje różne adresy URL z funkcjami PHP. Zadaniem każdej funkcji PHP, nazywanej kontrolerem, jest wykorzystanie informacji z zapytania, wraz z wieloma innymi narzędziami które oferuje Symfony, do stworzenia i odesłania obiektu odpowiedzi.

## 6.3. Projekt interfejsu użytkownika

Podstawowym zadaniem systemu jest gromadzenie i przetwarzanie danych podanych przez użytkownika, w tym celu niezbędne są interfejsy umożliwiające ich wprowadzanie. Podział interfejsów opiera się na opisanych powyżej scenariuszach i przedstawia się następująco:

### 1. Rejestracja

- E-mail
- Login
- Hasło i potwierdzenie hasła

### 2. Logowanie

- Login/E-mail
- Hasło

### 3. Weź udział w ankiecie

- E-mail
- Czy chce otrzymać powiadomienie o wynikach po zakończeniu ankiety
- Akceptacja warunków
- Ustawienie pozycji suwaka dla każdego pytania ankiety

### 4. Dodanie/edycja danych podstawowych ankiety

- Nazwa
- Opis
- Zdjęcie/obrazek - obraz używany w katalogu ankiet
- Typ - określa czy ankieta jest prywatna, czy też publiczna
- Minimum skali - dolna granica przedziału skali
- Maksimum skali - górna granica przedziału skali
- Czy jest wiralem? - respondenci mogą zapraszać kolejnych do wzięcia udziału
- Czy dopuszczalny jest zapis niekompletnej odpowiedzi? - respondent nie udziela odpowiedzi na wszystkie, lecz tylko na część pytań

### 5. Definicja przedziału znaczeniowego

- Pozycja - wartość skali
- Opis - wytłumaczenie znaczenia pozycji

#### 6. Dodanie/edycja kategorii

- Nazwa
- Opis
- Zdjęcie/obraz
- Stała, znana wartość - jeśli znana jest stała wartość którą przyjmuje kategoria w ramach przeprowadzanego badania
- Pozycja - umiejscowienie kategorii w hierarchii

#### 7. Aktywacja ankiety

- Tryb - standardowy lub częściowy
- Ilość pod-ankiet - dla trybu częściowego; ilość mniejszych ankiet które użytkownik chce wygenerować

#### 8. Przeglądanie statystyk

- Poziom niespójności
- Poziom hierarchii - dla ankiet hierarchicznych; wybór przeglądanej gałęzi
- Filtry w oparciu o dane respondentów - wiek, płeć, itd...

### 6.4. Schemat bazy danych

System zbudowany jest na relacyjnej bazie danych. Jej schemat przedstawiony jest na rysunku 6.3.





## 7. Implementacja systemu

Zbudowany portal ankiet porównawczych dzieli się na dwie logiczne części. Pierwszą jest interfejs zarządzania treścią w systemie, do których użytkownicy mają bezpośredni dostęp. Druga z nich to część analityczna, zajmująca się przetwarzaniem danych zgodnie z przyjętymi zasadami przeprowadzania badań metodą porównań parami. Funkcjonalności poświęcone bezpośredniej interakcji użytkownika z systemem zostały opisane w poprzednim rozdziale. W niniejszy skupiam się na przedstawieniu wymaganych elementów wdrożenia systemu, interfejsu użytkownika a także części analitycznej.

### 7.1. Konfiguracja

Aplikacja do działania wymaga odpowiednio skonfigurowanego serwera internetowego z obsługą języka PHP w wersji min. 5.4 oraz kilkoma wymaganymi przez Symfony dodatkami, jak np. biblioteka APC której używanie może znacznie przyspieszyć działanie aplikacji poprzez wykorzystanie mechanizmów cache.

Dzięki zastosowaniu Symfony, konfiguracja samej aplikacji odbywa się poprzez jeden plik tekstowy. W dodatku konstrukcja frameworka wymusza aby parametry aplikacji znajdowały się tylko na serwerze docelowym. Dlatego nie ma niebezpieczeństwa ich ujawnienia w przypadku udostępnienia kodu źródłowego.

Pierwszą wymaganą rzeczą jest baza danych. Jak wspominałem wcześniej system pracuje na bazie relacyjnej. Na etapie konfiguracji możemy zdefiniować który silnik (MySQL, PostgreSQL, itd.) mamy dostępny oraz podajemy dane dostępowe odpowiadające naszej instancji bazodanowej.

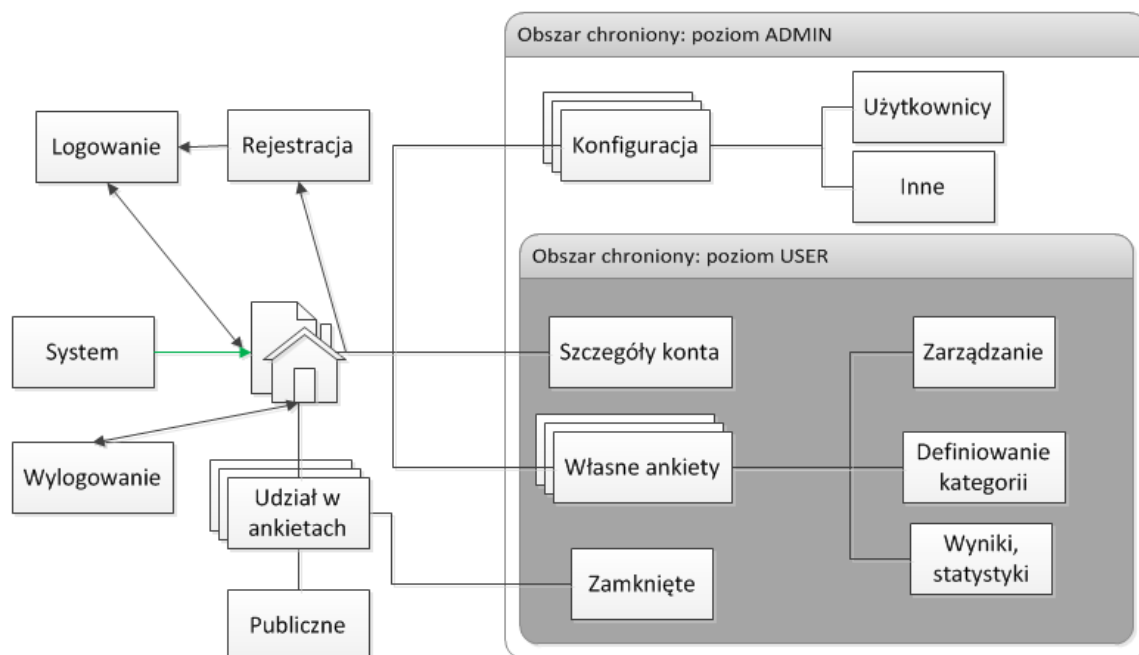
System wysyła pocztę elektroniczną. Ma to miejsce np. podczas rejestracji użytkownika czy też zapraszania do wzięcia udziału w ankiecie, a także przy powiadamianiu o wynikach. W związku z tym niezbędne jest podanie konfiguracji konta pocztowego i np. serwera SMTP za którego pośrednictwem będzie można wiadomości wysyłać. W najprostszej opcji można używać konta zarejestrowanego w domenie *gmail.com*. Należy jednak wziąć pod uwagę obowiązujące limity wysyłanych wiadomości. W przypadku gdy system będzie bardzo popularny i zapotrzebowanie na e-maile wzrośnie, można wykupić usługę konta mailowego u wielu różnych dostawców, w tym np. Amazon SES *Amazon Simple Email Service*, lub skonfigurować własny serwer pocztowy. W obu przypadkach należy dbać o to aby nie wysyłać treści nieodpowiednich, a także weryfikować adresy odbiorców i wprowadzić politykę obsługi tzw. *bounce email*, czyli wiadomości które nie mogły zostać dostarczone do adresata z wielu różnych powodów

i 'odbiły' się od serwera obsługującego domenę adresata. Najczęściej zdarza się tak z powodu błędnego adresu. Odbicia są poważnym problemem, gdyż mogą spowodować iż nasze wiadomości zaczną być traktowane jak spam i blokowane. Tym samym uniemożliwiając pełne funkcjonowanie. Aplikacja stara się minimalizować ten problem wymagając potwierdzenia adresu mailowego podczas rejestracji. Nie ma natomiast możliwości aby potwierdzić adresy do których wysyłane są zaproszenia. Tutaj pozostaje tylko obsługa 'odbić'.

Użytkownicy mają prawo do przesyłania plików do aplikacji. Na potrzeby identyfikacji ankiety lub kategorii twórca może wykorzystać zdjęcia i obrazy, które zapisze w systemie. Taki scenariusz tworzy potrzebę udostępnienia przestrzeni dyskowej o nie do końca zdefiniowanym rozmiarze. Ze względu na to iż serwery internetowe hostujące aplikacje mają najczęściej niewielką przestrzeń dyskową, lub podnosi ona znacznie koszty utrzymania, zdecydowałem się na integrację aplikacji z usługą przechowywania plików Amazon S3 *Amazon Simple Storage Service*. Dzięki temu serwer aplikacji nie przechowuje plików. Są one zapisywane na dyskach usługi S3, a dostępnym miejsce można w łatwy i szybki sposób zarządzać wedle potrzeb. Usługa ta jest płatna, a do założenia konta wymagane jest podanie karty kredytowej. W pierwszym roku można używać jej za darmo w ramach programu promocyjnego *Free Tier*. Prezentowana aplikacja do współpracy z usługami Amazon potrzebuje kluczy: publicznego i prywatnego, nazwę regionu w którym usługa jest zarejestrowana oraz tzw. 'koszyka' *bucket* który jest pojemnikiem zawierającym przechowywane dane.

## 7.2. Interfejs użytkownika

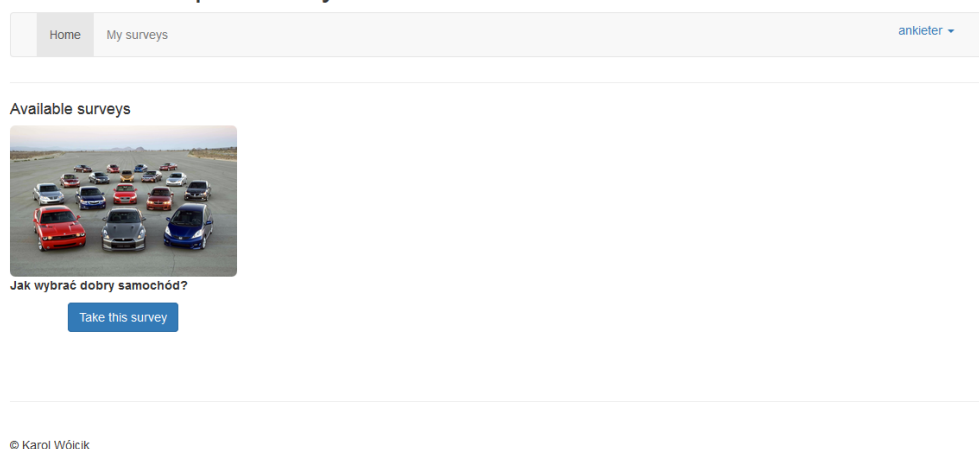
W obecnym podrozdziale prezentuję główne koncepty i najważniejsze widoki z działającej aplikacji, więcej szczegółów można znaleźć na załączonym wirtualnym obrazie systemu. Oprócz całego systemu, zawiera także kilka video-tutoriali będących swoistą instrukcją i prezentacją możliwości systemu.



Rysunek 7.1: Schemat interfejsu użytkownika

Interfejs użytkownika został podzielony na trzy główne obszary, odpowiadające poziomom dostępu użytkowników 7.1.. Można wyróżnić część dostępną dla wszystkich: stronę główną, ekran logowania, formularz rejestracji, a także publiczny katalog ankiet i możliwość wzięcia udziału w ankiecie jako niezarejestrowany użytkownik systemu.

## Pairwise comparison system



Rysunek 7.2: Strona główna aplikacji - katalog ankiet w których użytkownik może wziąć udział

Pairwise comparison system

Home My surveys [ankieter](#)

**Jak wybrać dobry samochód?**  
Wybór odpowiedniego samochodu to trudna decyzja, na którą wpływa wiele czynników. Ankieta ma na celu przyporządkowanie im stopni ważności. Dzięki temu każdy będzie mógł łatwiej znaleźć wymarzone auto.

Step 2/27

Rodzaj paliwa

Your opinion in words:  
Rodzaj nadwozia is just as important as Rodzaj paliwa

approximately 1 : 5.7

Change the aspect ratio by moving the slider toward more important of the criteria

Prev Next

© Karol Wójcik

Rysunek 7.3: Każde pytanie ankiety prezentuje spójny widok, rozmiar balonów jest zależny od pozycji suwaka

Następnie każda zalogowana osoba ma dostęp do edycji danych swojego konta, a także widzi własny katalog, wraz z ankietami o ograniczonym dostępie do których został zaproszony, a nie wziął jeszcze w nich udziału.

Najważniejszą z funkcjonalności oferowanych posiadaczom kont w systemie jest możliwość własnego ankietowania. Uzyskują oni dostęp do części tworzenia ankiety, definiowania obiektów porównań i hierarchii oraz zapraszania respondentów. Mogą zarządzać treściami które stworzyli, a na końcu przeglądać wyniki, filtrować je i wyciągać wnioski.

Pairwise comparison system

Home My surveys [ankieter](#)

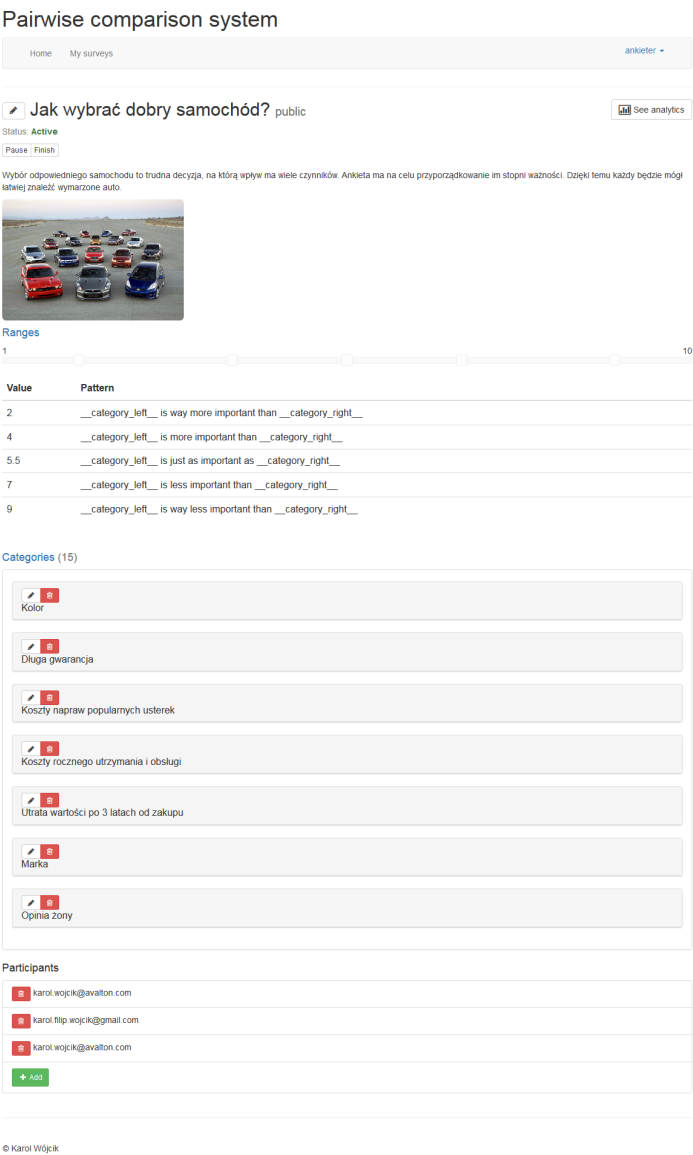
Surveys table

[+ New survey](#)

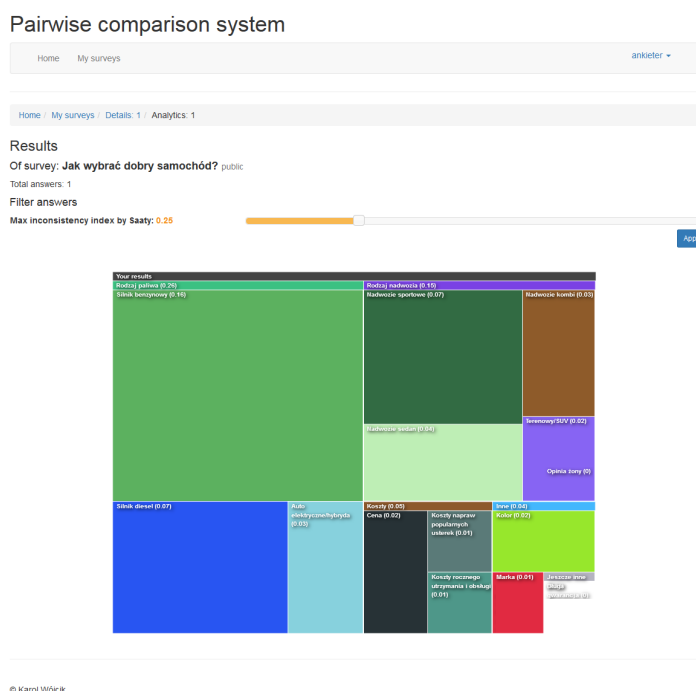
Id	Name	Type	Viral	Incompleteness	Scale Min	Scale Max	Status	Date Created	Actions
1	Jak wybrać dobry samochód?	public	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1	10	Active	2015-01-13 22:55:34	<a href="#">+</a> <a href="#">-</a>
2	Fixed-step scale	public	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1	10	New	2015-05-03 20:40:33	<a href="#">+</a> <a href="#">-</a>

© Karol Wójcik

Rysunek 7.4: Lista ankiet zalogowanego użytkownika







Najwyższy poziom dostępu zajmuje Administrator, który oprócz wymienionych wcześniej elementów, ma także prawo zarządzać systemem, w szczególności użytkownikami. Schemat interfejsu prezentuje rysunek 7.1

Graficzny interfejs użytkownika został zbudowany w oparciu o framework *Bootstrap*. Jest to darmowy i niezwykle popularny zestaw szablonów HTML i stylów CSS, które sprawiają że projektowanie i budowa interfejsów staje się prostsza i szybsza niż ma to miejsce w przypadku tworzenia od podstaw. Proces wytwarzania interfejsu w Bootstrapie w większości przypadków sprowadza się do układania elementów zdefiniowanych we frameworku. Dostępna jest szczegółowa dokumentacja, a wokół projektu istnieje duża społeczność internetowa. Wszystko to sprawia że praca z Bootstrapem jest przyjemna i efektywna.

Jednak nawet najpiękniejszy szablon jest bezużyteczny gdy nie widać na nim prawdziwej zawartości. Z pomocą przychodzą silniki szablonów, zajmujące się 'ożywianiem' statycznych elementów i wypełnianiem ich pożądaną zawartością. W aplikacjach budowanych w oparciu o Symfony, rolę takiego silnika pełni Twig. Oprócz doskonałej współpracy z frameworkiem Symfony, Twig jest bardzo szybki oraz bezpieczny. Ponadto jego możliwości można dowolnie rozszerzać poprzez tworzenie własnych elementów. W modelu MVC, Twig odpowiada za Widok *View*, zapewniając odpowiednią separację kodu źródłowego aplikacji od szablonów i widoków, przez co kod jest znacznie wyższej jakości.



## 7.3. Część analityczna

Część analityczna to umownie nazwana grupa funkcjonalności niezbędnych do prawidłowej pracy systemu, ale nie widocznych bezpośrednio dla użytkownika. W jej skład wchodzi przede wszystkim elementy przetwarzania wyników porównań przesyłanych przez respondentów. Poddawane są wstępnej obróbce już w czasie zapisu do bazy danych, np. wyliczane są współczynniki niespójności wg Saatyego i Koczkodaja.

### 7.3.1. Niespójność

Obliczanie niespójności jest kosztownym procesem. Ciężko wyobrazić sobie możliwość filtrowania wyników względem niespójności w przypadku gdy byłaby ona obliczana każdorazowo dla każdej odpowiedzi. Czas oczekiwania na wynik byłby bardzo duży. Jako że niespójność jest własnością niezmienną (dla niezmienniej macierzy), może być wyznaczona na początku i później używana według potrzeby.

System oferuje dwie metryki niespójności, analizę wektora własnego (metoda Saatyego) oraz metodę odległościową (wg. W. Koczkodaja). Procedura wyznaczania wektora własnego jest operacją pracochłonną dla komputerów, a także dla programistów ponieważ nie istnieją żadne biblioteki dla języka PHP realizujące tę funkcjonalność. Także i ja stanąłem przed wyborem sposobu obliczania wektora własnego macierzy kwadratowej, dowolnego rozmiaru.

Zdecydowałem się na integrację z popularnym serwisem naukowo-obliczeniowym *WolframAlpha*. Korzystając z publicznego API mam dostęp do większości funkcji, w tym także matematycznych metod obliczeniowych wśród których znajduje się także sposób na obliczanie wektora wartości własnych. W przeciwieństwie do samodzielnej implementacji formuły matematycznej, integracja zapewnia znacznie większą jakość obliczonych wartości przy zachowaniu szybkości działania całej aplikacji. WolframAlpha oferuje swoje API w dwóch wersjach: darmowej i komercyjnej. Darmowa nie może być wykorzystywana do celów komercyjnych, a także ma górne ograniczenie ilości zapytań wynoszące 2000/miesiąc. Jest to więc idealna propozycja dla rozwoju takiego systemu który buduję. W przyszłości, gdy zaistnieje taka potrzeba, można bez problemu pozbyć się ograniczeń przechodząc na licencję komercyjną. W związku z tym iż WolframAlpha jest używany do zapewnienia podstawowej funkcjonalności systemu, wymagana jest konfiguracja dostępu do zdalnych zasobów. Składa się na nią założenie konta developerskiego i stworzenie wirtualnej aplikacji, której identyfikator jest niezbędny do podpisania zapytań i musi być wysyłany każdorazowo przy połączeniu z API.

### 7.3.2. Ankiety częściowe

Chciałbym rozwinąć myśl przedstawioną w przypadku użycia 6.7. Zaimplementowałem funkcję ankiet w trybie częściowym. Ten tryb polega na tym iż respondent udziela odpowiedzi tylko na wybrane pytania ankiety, ale nie na wszystkie. Ma to zastosowanie w przypadku gdy obiektów porównań jest wiele (duża ilość pytań), a nie chcemy tworzyć hierarchii lecz porównywać wszystkie kategorie na tym samym

poziomie. Częściowa odpowiedź może wydawać się niewystarczająca gdyż nie mamy całej macierzy PC, lecz istnieją metody na odpowiednią interpretację lub też zbudowanie pełnej macierzy. Najbardziej naturalnym rozwiązaniem jest wyznaczanie brakujących elementów korzystając z równania spójności macierzy przedstawionego w 2.1. W rezultacie otrzymujemy możliwie najbardziej spójną odpowiedź nie tracąc subiektywizmu respondenta. Istnieją jednak pewne ograniczenia w stosunku do minimalnej ilości pytań. Aby odpowiedź ankietowanego mogła zostać uznana za wartościową i użyta w procesie wnioskowania, musi on odpowiedzieć na przynajmniej  $n$  pytań, gdzie  $n$  to ilość kategorii. Tylko taka macierz częściowa będzie mogła zostać przetworzona.

Zbudowany system oferuje możliwość wyboru trybu podczas aktywacji ankiety, tj. w momencie gdy wersja robocza jest ukończona i twórca chce rozpocząć ankietowanie. Tryb częściowy jest inicjalizowany dwoma parametrami. Są to: ilość różnych pod-ankiet wygenerowanych z pełnej ankiety, oraz minimalna ilość pytań w pod-ankiecie. Domyślnie drugi parametr przyjmuje wartość odpowiadającą ilości kategorii. Po zatwierdzeniu parametrów, system losowo generuje zbiory pytań tworząc pod-ankiety. Są one następnie prezentowane operatorowi. Jeśli ten uzna że spełniają jego oczekiwania to zatwierdza i ankietę zostaje opublikowana, w przeciwnym przypadku może powtórzyć proces generacji i w rezultacie otrzymać inne permutacje pytań.



## 8. Podsumowanie i wnioski

Cel pracy jakim było stworzenie i wdrożenie aplikacji internetowej ankiet porównawczych został zrealizowany. Implementacja oparta została na założeniach zdefiniowanych podczas projektowania systemu. System w obecnej formie jest przystosowany do obsługi ankiet parowych tworzonych przez każdego kto tylko posiada konto użytkownika w systemie.

Obecnie nie istnieją żadne nowoczesne narzędzia oferujące kompleksowe usługi z kategorii ankiet parowych. Metoda porównań parami stale się rozwija, obejmując swoim zasięgiem coraz to nowsze obszary nauk, jak np. socjologię. Dlatego potrzebuje dobrych narzędzi które pozwolą w dalszym ciągu się jej rozwijać i popularyzować spośród wszystkich metod podejmowania decyzji.

Moją odpowiedzią na to zapotrzebowanie jest zbudowany system. Może być wykorzystywany zarówno na potrzeby akademickie czy wspierając badania naukowe jak i prywatne podejmowanie decyzji w istotnych dla twórcy ankiety sprawach. Aplikacja jest napisana w popularnym języku PHP i wykorzystuje sprawdzony i polecany framework Symfony 2. Dzięki temu mogła zostać sprawnie zbudowana i dostosowana do potrzeb. Spośród wielu zalet, najważniejszymi które wymienię, są pełna modułowość pozwalająca na łatwą zmianę lub dodanie funkcjonalności poprzez gotowe komponenty, wykorzystanie ORM/ODM do komunikacji z bazą danych czy też silnik szablonów Twig z możliwością łatwego wprowadzenia różnych wersji językowych aplikacji. Z punktu widzenia programisty, przejrzystość kodu źródłowego i możliwość łatwego zrozumienia logiki biznesowej są ogromną wartością projektu, mierzalną np. za pomocą ilości wymaganego czasu do poznania implementacji. W przypadku gdy nad rozwojem aplikacji zaczynać będą pracę nowe osoby.

Plany na dalsze usprawnianie i budowanie systemu są rozległe. Dotyczą obszarów zarówno merytorycznych jak i samego funkcjonowania aplikacji. Najważniejszymi zmianami jest dodanie funkcjonalności z kategorii metody porównań parowych. Wprowadzenie możliwości udzielania niepełnych odpowiedzi i stosowania dla nich heurystyk, dodanie innych metod obliczania niespójności i ich dowolne wybieranie dla tworzonych ankiet, lub nawet definiowanie własnych funkcji obliczeniowych. Oprócz tego chciałbym opracować wspólnie z grafikiem nowy interfejs, który byłby wygodniejszy dla użytkowników, a także zintegrować usługi mediów społecznościowych jak Facebook, Twitter, Wykop itd. Pomoże to w popularyzowaniu portalu.

Gdy wszystkie powyższe elementy będą zrealizowane, chciałbym wprowadzić model subskrypcji na dostęp do serwisu. Podstawowe elementy pozostaną dostępne bezpłatnie dla ograniczonej ilości ankiet, a rozszerzenia i większa ilość prowadzonych badań będzie wymagała wykupienia usługi.



## Bibliografia

- [1] K. Spingarn A. T. W. Chu, R. E. Kalaba. A comparison of two methods for determining the weights of belonging to fuzzy sets. *Journal of Optimization Theory and Applications*, 27:531–538, 1979.
- [2] Valerie Ahl and T. F. H. Allen. Hierarchy theory: A vision, vocabulary, and epistemology. *Columbia University Press*, 1996.
- [3] Alex A. Freitas Carlos N. Silla Jr. A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery*, 22:31–72, 2011.
- [4] N. de Condorcet. *Essay on the Application of Analysis to the Probability of Majority Decisions*. l’Imprimerie Royale.
- [5] G.T. Fechner. *Elemente der Psychophysik*. Breitkopf und Härtel.
- [6] T.B.M. McMaster F.J. Dodd, H.A. Donegan. A statistical approach to consistency in ahp. *Mathematical and Computer Modelling*, 18:19–22, 1993.
- [7] C. Williams G. Crawford. A note on the analysis of subjective judgment matrices. *Journal of Mathematical Psychology*, 29:387–405, 1985.
- [8] J. M. Moreno-Jiménez J. Aguarón. The geometric consistency index: Approximated thresholds. *European Journal of Operational Research*, 147:137–145, 2003.
- [9] S. J. Szarek J. Fülöp, W. W. Koczkodaj. On some convexity properties of the least squares method for pairwise comparisons matrices without the reciprocity condition. *Journal of Global Optimization*, 54:689–706, 2012.
- [10] S. J. Szarek J. Fülöp, W.W. Koczkodaj. A different perspective on a scale for pairwise comparisons. *Transactions on Computational Collective Intelligence I*, 6220:71–84, 2010.
- [11] Aleš Krešta Jiří Franek. Judgment scales and consistency measure in ahp. *Procedia Economics and Finance*, 12:164–173, 2014.
- [12] M.T. Lamata J.I. Peláez. A new measure of consistency for positive reciprocal matrices.

- [13] W.W. Koczkodaj. A new definition of consistency for pairwise comparisons. *Mathematical and Computer Modelling: An International Journal*, 18:79–84, 1993.
- [14] W.W. Koczkodaj. Testing the accuracy enhancement of pairwise comparisons by a monte carlo experiment. *Journal of Statistical Planning and Inference*, 69:21–32, 1998.
- [15] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22:1–55, 1932.
- [16] H. Monsuur. An intrinsic consistency threshold for reciprocal matrices. *European Journal of Operational Research*, 96:387–391, 1996.
- [17] Thomas L Saaty. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3):234–281, 1977.
- [18] L. L. Thurstone. A law of comparative judgment. *Psychological Review*, 34(4):273–286, 1927.
- [19] L.G. Vargas T.L. Saaty. *The Logic of Priorities*. Kluwer-Nijhoff.
- [20] P. J. Mizzi W. E. Stein. The harmonic consistency index for the analytic hierarchy process. *European Journal of Operational Research*, 177:488–497, 2007.
- [21] A. Ligeża W. W. Koczkodaj, K. Kułakowski. On the quality evaluation of scientific entities in poland supported by consistency-driven pairwise comparisons method. *Scientometrics*, 99:911–926, 2014.
- [22] Wikipedia. Perron–frobenius theorem — wikipedia, the free encyclopedia, 2015. [Online; Dostęp 2-Lipiec-2015].
- [23] M. Orłowski W.W. Koczkodaj. Computing a consistent approximation to a generalized pairwise comparisons matrix. *Computers & Mathematics with Applications*, 37(3):79–85, 1999.
- [24] M.W. Herman W.W. Koczkodaj and M. Orłowski. Using consistency-driven pairwise comparisons in knowledge-based systems. *Proceedings of the sixth international conference on Information and knowledge management*, pages 91–96, 1997.
- [25] S. Szarek W.W. Koczkodaj. On distance-based inconsistency reduction algorithms for pairwise comparisons. *Logic Journal of the IGPL*, 18(6):859–869, 2010.
- [26] R. Janicki Y. Zhai. On consistency in pairwise comparisons based numerical and non-numerical ranking. *Proceedings of the International Conference on Foundations of Computer Science*, page 183–186, 2010.
- [27] W.W. Koczkodaj Z. Duszak. Generalization of a new definition of consistency for pairwise comparisons. *Information Processing Letters*, 52:273–276, 1994.