

Symulator procesu badania charakterystyk jakościowych diagnoz medycznych generowanych przez komputerowy system wspomagania decyzji medycznych

A. AMELJAŃCZYK, P. DŁUGOSZ
aameljanczyk@wat.edu.pl

Instytut Systemów Informatycznych
Wydział Cybernetyki WAT
ul. S. Kaliskiego 2, 00-908 Warszawa

W pracy przedstawiono narzędzie symulacyjne zaprojektowane w celu badania wybranych charakterystyk jakościowych diagnoz medycznych generowanych przez komputerowy system wspomagania decyzji medycznych. System ten wykorzystuje metodę wnioskowania diagnostycznego bazującą na określeniu podobieństwa stanu zdrowia pacjenta do wzorców jednostek chorobowych zawartych w repozytorium. Program umożliwia symulacyjne badanie charakterystyk jakościowych procesu diagnozowania, a szczególnie tzw. wiarygodności diagnostycznej wygenerowanych diagnoz medycznych przy różnych scenariuszach badawczych. Program komputerowy powstał w oparciu o platformę .NET Framework oraz biblioteki uzupełniające.

Słowa kluczowe: symulacja komputerowa, symptomy chorobowe, czynniki ryzyka, wzorzec medyczny, wiarygodność diagnozy, wyrazistość diagnozy, niejednoznaczność diagnozy, platforma .NET, Windows Forms.

1. Wprowadzenie

Wykorzystanie techniki komputerowej oraz wyspecjalizowanego oprogramowania w służbie medycznej stało się w dzisiejszych czasach bardzo powszechne. Gwałtowny rozwój technologii spowodował wzrost roli systemów informatycznych w różnego rodzaju zastosowaniach medycznych, jak np.:

- zarządzanie informacją medyczną
- monitorowanie i alarmowanie
- wspomaganie procesów diagnostycznych
- wspomaganie komunikacji i pracy grupowej na odległość
- kompleksowe wsparcie medyczne.

Prezentowana praca dotyczy obszaru wspomagania procesów diagnozowania medycznego [2, 3, 6].

W pracy przedstawiony został komputerowy system wspomagania decyzji medycznych (KSWDM) oparty na wnioskowaniu diagnostycznym na podstawie rozpoznanych symptomów chorobowych i czynnikach ryzyka. Każdy z takich systemów wspomagania implementuje pewne metody wnioskowania opracowane przez zespół ekspertów. Słuszność zastosowania poszczególnych metod oraz ich efekty można badać, korzystając z odpowiednich narzędzi informatycznych zwanych

symulatorami. Głównym celem tej pracy jest przedstawienie narzędzia do przeprowadzania symulacji w oparciu o zaproponowaną metodę wyznaczania wstępnej diagnozy medycznej. Prezentowany symulator umożliwia wszechstronne badanie jakości generowanych przez KSWDM decyzji diagnostycznych. Tym samym może być dodatkowym narzędziem ułatwiającym lekarzowi podjęcie decyzji o praktycznym wykorzystaniu zaproponowanej diagnozy wstępnej w dalszym procesie diagnozowania.

2. Skrócony opis metody diagnostycznej [3, 4]

Repozytorium wykorzystywane w procesie ustalania wstępnej diagnozy medycznej zawiera ustalony zbiór wzorców jednostek chorobowych. Definiowanie wzorców chorobowych odbywa się z udziałem diagnostów (ekspertów medycznych) poprzez określenie zbioru towarzyszących danej chorobie symptomów chorobowych oraz zbioru czynników ryzyka.

Założmy, że w wyniku wstępnego etapu rozpoznania stwierdzono u pacjenta $x \in X$ występowanie zbioru $S_o(x) \subset S$ symptomów

chorobowych oraz zbiór czynników ryzyka $R_o(x) \subset R$

$$\begin{aligned} S_o(x) &= \{s \in S \mid w(x,s) > 0\}, \\ R_o(x) &= \{r \in R \mid w(x,r) > 0\}, \end{aligned} \quad (2.1)$$

przy czym $w(x,s)$ – stopień „nasilenia się występowania” symptomu $s \in S$ (ustalony przez lekarza w skali $[0, 1]$ w czasie pierwszej wizyty) oraz podobnie $w(x,r)$ – stopień nasilenia się występowania czynnika ryzyka nr r u badanego pacjenta (również w skali $[0, 1]$).

Zbiór $M_o(S)$ chorób sugerowanych zestawem występujących symptomów określimy następująco:

$$M_o(S) = \{m \in M \mid S_o(x) \cap S^m \neq \emptyset\} \quad (2.2)$$

Podobnie zbiór $M_o(R)$ chorób związanych z występującymi czynnikami ryzyka określimy następująco:

$$M_o(R) = \{m \in M \mid R_o(x) \cap R^m \neq \emptyset\}$$

Kolejnym krokiem będzie ustalenie łącznego zbioru możliwych chorób wstępnego rozpoznania. Wstępnym oszacowaniem może być zbiór

$$M_o = M_o(S) \cup M_o(R) \quad (2.3)$$

lub bardziej radykalnie:

$$M_o = M_o(S) \cap M_o(R) \quad (2.4)$$

Takie podejście w ustalaniu wstępnego rozpoznania jest jednak ryzykowne ze względu na możliwość występowania czynników ryzyka lub symptomów jednocześnie kilku chorób oraz trudność precyzyjnego ich określenia.

Dysponując danymi pacjenta $x \in X$ odnośnie do występowania symptomów chorobowych oraz czynników ryzyka w postaci liczb $w(x,s)$, $s \in S_o(x)$ oraz $w(x,r)$, $r \in R_o(x)$, możemy określić „odległość stanu zdrowia pacjenta” od wzorców potencjalnych chorób zawartych w zbiorach $M_o(S)$ i $M_o(R)$. Możemy tego dokonać w następujący sposób.

Model aktualnego stanu zdrowia pacjenta $x \in X$, zdefiniowany na podstawie występujących symptomów chorobowych i czynników ryzyka, przyjmie postać pary:

$$f(x) = (f_s(x), f_r(x)), \quad x \in X \quad (2.5)$$

gdzie:

$$\begin{aligned} f_s(x) &= (w(x,s); s \in S_o(x)), \\ f_r(x) &= (w(x,r); r \in R_o(x)). \end{aligned}$$

Symbolami $s^*(m)$ i $r^*(m)$ oznaczymy wzorce choroby nr m odpowiednio w aspekcie symptomów i czynników ryzyka [5, 6].

Symbolem $d_1\left(f_s(x), s^*(m)\right)$, $m \in M$ oznaczymy

odległość stanu zdrowia pacjenta x (wynikającego z występujących symptomów) od wzorca choroby $m \in M$ zdefiniowanego na podstawie symptomów chorobowych i analogicznie oznaczymy symbolem

$d_2\left(f_r(x), r^*(m)\right)$, $m \in M$ odległość stanu

zdrowia pacjenta x (wynikającego z występujących czynników ryzyka) od wzorca choroby $m \in M$, zdefiniowanego na podstawie czynników ryzyka.

Zbiór $M(S_o(x))$ „najbardziej prawdopodobnych” chorób ze względu na symptomy chorobowe wyznaczmy następująco:

$$\begin{aligned} M(S_o(x)) &= \left\{ m \in M_o \mid d_1\left(f_s(x), s^*\left(m^*\right)\right) = \right. \\ &= \left. \min_{m \in M_o} d_1\left(f_s(x), s^*(m)\right) \right\} \end{aligned} \quad (2.6)$$

Z kolei zbiór $M(R_o(x))$, „najbardziej prawdopodobnych” chorób ze względu na występujące czynniki ryzyka wyznaczmy następująco:

$$\begin{aligned} M(R_o(x)) &= \left\{ m \in M_o \mid d_2\left(f_r(x), r^*\left(m^*\right)\right) = \right. \\ &= \left. \min_{m \in M_o} d_2\left(f_r(x), r^*(m)\right) \right\} \end{aligned} \quad (2.7)$$

Część wspólna tych zbiorów $M^* = M(S_o(x)) \cap M(R_o(x))$ najczęściej jest jednak zbiorem pustym [3].

Interesującą propozycję w zakresie wyznaczenia zbioru chorób najbardziej prawdopodobnych jednocześnie z punktu widzenia zbioru występujących symptomów oraz czynników ryzyka oferuje teoria optymalizacji wielokryterialnej oraz teorii podobieństwa [3, 4]. Ustalając odpowiedni model „preferencji diagnostycznych” \bar{R} , możemy zadanie takie zdefiniować w postaci

$$(M_o, d(m), \bar{R}), \quad (2.8)$$

gdzie:

$$d(m) = (d_1(m), d_2(m)), \quad (2.9)$$

zaś:

$$\begin{aligned} d_1(m) &= d_1\left(f_s(x), s^*(m)\right), \\ d_2(m) &= d_2\left(f_r(x), r^*(m)\right), \end{aligned} \quad (2.10)$$

\bar{R} – model preferencji diagnostycznych [1, 2].
W praktyce najczęściej brane są pod uwagę trzy następujące warianty preferencji:

- 1) symptomy chorobowe i czynniki ryzyka są tak samo ważne,
- 2) ważniejsze są symptomy choroby,
- 3) ważniejsze są czynniki ryzyka.

Z punktu widzenia bezpieczeństwa diagnostycznego przypadek pierwszy jest do przyjęcia bez zastrzeżeń, gdyż w efekcie gwarantuje „szerszy” (bezpieczniejszy) zbiór wyników wstępnej diagnozy obejmujący dwa pozostałe przypadki [3, 4]. Obrazem [3] zbioru chorób \mathbf{M}_o w kontekście odległości od stanu zdrowia pacjenta jest zbiór Y .

$$Y = d(\mathbf{M}_o) = \{d(m) \in R^2 \mid m \in \mathbf{M}_o\} \quad (2.11)$$

Rozwiązaniem zadania (2.8) będzie zatem tzw. zbiór Pareto [2, 3], czyli zbiór tych chorób ze zbioru wstępnego oszacowania \mathbf{M}_o , od których nie ma „bardziej prawdopodobnych”. Zbiór ten oznaczymy symbolem:

$$\mathbf{M}_N^{\bar{R}} = \left\{ m \in \mathbf{M}_o \mid \text{nie istnieje } m \in \mathbf{M}_o - \left\{ \overset{o}{m} \right\}, \text{ że} \right. \\ \left. d(m) \leq d\left(\overset{o}{m}\right) \right\}. \quad (2.12)$$

Zbiór $\mathbf{M}_N^{\bar{R}}$ jest przeciwobrazem [3] zbioru Pareto $Y_N^{\bar{R}}$.

$$\mathbf{M}_N^{\bar{R}} = d^{-1}(Y_N^{\bar{R}}) = \left\{ m \in \mathbf{M}_o \mid d\left(\overset{o}{m}\right) \in Y_N^{\bar{R}} \right\} \quad (2.13)$$

Ostatecznym czynnikiem rozstrzygającym może być w tej sytuacji tzw. „rozwiązanie kompromisowe” [1], które prowadzi najczęściej do rozwiązania jednoznacznego.

Obliczając odległość obrazów tych chorób od obrazu y^* „utopijnej” (wirtualnej), „najbardziej prawdopodobnej” (ze względu na stwierdzone symptomy i czynniki ryzyka) choroby, możemy utworzyć ranking chorób do dalszych działań diagnostycznych.

Współrzędne choroby utopijnej $y^* = \left(y_1^*, y_2^* \right)$ wyznaczmy następująco:

$$y_1^* = \min_{m \in \mathbf{M}_o} d_1(m), \quad y_2^* = \min_{m \in \mathbf{M}_o} d_2(m) \quad (2.14)$$

Etap wstępnego diagnozowania [3, 4] jest szczególnie ważny w procedurze rozpoznawania chorób. Trafne „wytypowanie” podzbioru chorób, będących potencjalnym zagrożeniem, ma bowiem decydujący wpływ na

strategię dalszych badań specjalistycznych. Badania specjalistyczne, ich rodzaj i kolejność przeprowadzenia decydują w głównym stopniu o kosztach i czasie badań, a w konsekwencji o koszcie, czasie i skuteczności leczenia pacjenta. Głównym celem opracowanej metody symulacyjnej oraz jej implementacji komputerowej było opracowanie skutecznego narzędzia umożliwiającego wszechstronne badanie jakości uzyskiwanych diagnoz w komputerowym systemie wspomagania medycznego. Szczególnie istotnym wskaźnikiem, badanym w procesie symulacyjnym, jest „wiarygodność diagnostyczna” opisana wzorem:

$$w(x, \mathbf{M}) = 1 - \|y\|_p, \quad p \geq 1, \quad (2.15)$$

gdzie:

$y = (y_1^*, y_2^*)$ – tzw. punkt utopijny – obraz choroby najbardziej prawdopodobnej.

Stopień informacyjności wyników badań (w aspekcie symptomów i czynników ryzyka) dla poszczególnych chorób można określić następującymi liczbami [10,12]:

$$i_S^m(x) = \sum_{s_k^m \in S_o^m(x)} w(x, s_k^m) \alpha(s_k^m), \quad (2.16)$$

$$i_R^m(x) = \sum_{r_l^m \in R_o^m(x)} w(x, r_l^m) \beta(r_l^m), \quad (2.17)$$

$\alpha(s_k^m)$ oraz $\beta(r_l^m)$ są liczbami z przedziału $[0, 1]$ wyznaczanymi przez ekspertów i oznaczają odpowiednio „stopień ważności” poszczególnych parametrów z obszaru symptomów chorobowych i czynników ryzyka w diagnozowaniu jednostki chorobowej nr $m \in \mathbf{M}$.

Możemy zatem napisać, iż stopień wiarygodności diagnozy jest pewną funkcją zawartości repozytorium \mathbf{M} (dokładnie liczności tego zbioru) oraz informacyjności $i_S^m(x)$ oraz $i_R^m(x)$

$$w(x, \mathbf{M}) = f(|\mathbf{M}|, i_S^m(x), i_R^m(x), m \in \mathbf{M}_o(x))$$

Istnieje wiele innych właściwości omawianych mechanizmów wnioskowania diagnostycznego, mających duże znaczenie informacyjne dla lekarza. Wynikają one głównie z analizy własności zbioru Pareto otrzymanych wyników diagnostycznych. Są to między innymi:

- wskaźnik niejednoznaczności diagnozy $n(x)$ pacjenta $x \in X$

$$n(x) = \frac{\max_{y \in Y_N^R} (F^{-1}(y))}{|F^{-1}(Y_N^R)|} \quad (2.18)$$

- wyrazistość diagnozy $s(x)$

$$s(x) = 1 - \frac{|Y_N^R|}{|M|}, \quad (2.19)$$

gdzie:

Y_N^R – zbiór obrazów chorób, od których nie ma bardziej prawdopodobnych,

$F^{-1}(y)$ – klasa równoważności obiektów (jednostek chorobowych) ze względu na ustalone wartości odległości od wzorca w kontekście symptomów chorobowych oraz czynników ryzyka,

$F^{-1}(Y_N^R)$ – przeciwobraz zbioru obrazów chorób, od których nie ma bardziej prawdopodobnych.

Wszystkie te charakterystyki łącznie ze znajomością zawartości zbioru diagnoz, od których nie ma bardziej prawdopodobnych, oraz jego rankingu tworzą zestaw informacji wspomagających proces diagnozowania wstępnego.

Prezentowany w pracy symulator umożliwia graficzne przedstawienie zbioru obrazów podobieństwa stanu zdrowia pacjenta do poszczególnych wzorców chorób zawartych w repozytorium. Umożliwia wyznaczenie i graficzną prezentację zbioru obrazów diagnoz, od których nie ma bardziej prawdopodobnych (front Pareto), oraz tzw. obraz „podobieństwa idealnego” (punkt idealny, utopijny [1]). Przedstawiony w pracy symulator umożliwia wyznaczanie i badanie wymienionych wyżej charakterystyk jakościowych otrzymywanych diagnoz przy różnych scenariuszach symulacyjnych i zmiennych stopniach informacyjności wprowadzanych wyników badań lekarskich.

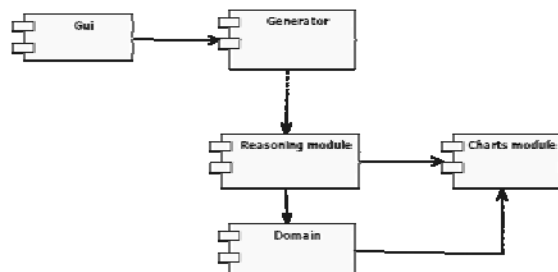
3. Technologia

Aplikacja symulatora została napisana w oparciu o technologie Microsoft .NET Framework. W procesie implementacji został zastosowany język C#. Środowisko projektowe, którego użyto do wykonania oprogramowania, to Microsoft Visual Studio 2010. Repozytorium symulatora jest generowane w sposób dynamiczny (przy określeniu odpowiednich parametrów podanych przez użytkownika). Istnieje także możliwość wykorzystania bazy danych MSSQL w celu gromadzenia danych dotyczących chorób i wykorzystywania ich w procesie symulacyjnym.

Przedstawione oprogramowanie ma postać aplikacji okienkowej wykorzystującej graficzny interfejs Windows Forms. W celu skorzystania z aplikacji wymaga się uprzedniego zainstalowania .Net Framework w wersji 3 lub wyższej. Wyniki poszczególnych symulacji obrazowane są dzięki wykorzystaniu dodatkowej biblioteki Microsoft Charts, pozwalającej na dynamiczne generowanie różnych typów wykresów.

4. Opis symulatora

Uproszczona architektura aplikacji została przedstawiona na rysunku 1. Można tu wyszczególnić 5 kluczowych modułów (komponentów) zaproponowanej architektury aplikacji:



Rys. 1. Architektura aplikacji

gdzie:

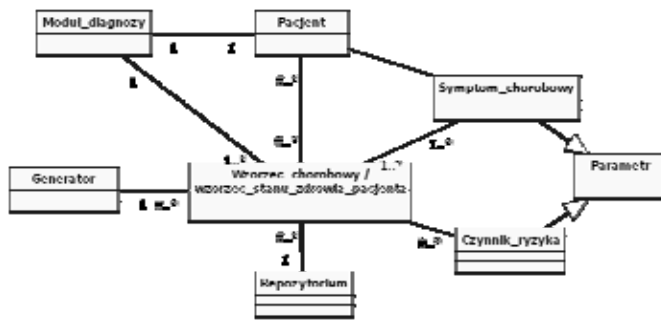
GUI – graficzny interfejs użytkownika (Windows Forms).

Generator – komponent odpowiedzialny za wygenerowanie danych testowych symulacji. W zależności od wybranego scenariusza generowane są odpowiednie dla niego dane.

Reasoning module – komponent odpowiedzialny za wykonywanie obliczeń „odległościowych”, bazując na przedstawionym uprzednio algorytmie.

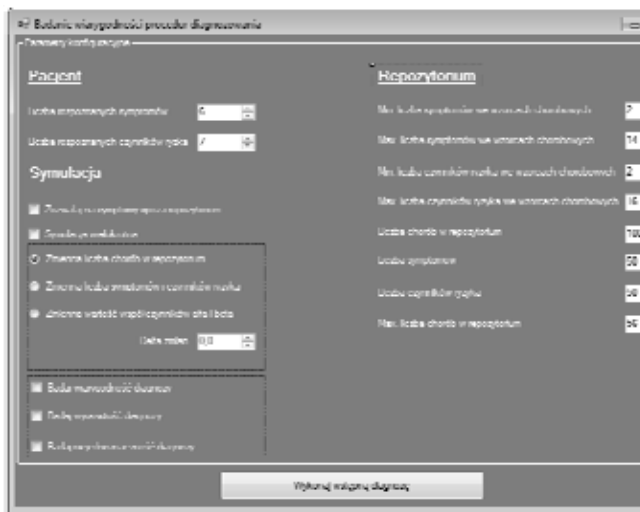
Domain – komponent odpowiedzialny za wykonywanie obliczeń rankingowych.

Charts module – komponent odpowiedzialny za obrazowanie wyników symulacji na dwuwymiarowych wykresach. Uproszczona struktura symulatora została przedstawiona na rysunku 2.



Rys. 2. Uproszczony diagram klas dla aplikacji symulatora

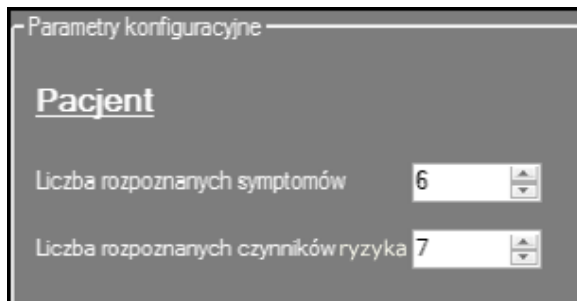
liczby ich odpowiedników w repozytorium. Panel pacjenta odgrywa bardzo ważną rolę w generowaniu i realizacji wielu scenariuszy symulacyjnych. Pozwala na badanie wrażliwości generowanych przez system diagnoz w zależności od zmian wartości parametrów zdrowotnych pacjenta, a w szczególności od zmian stopnia intensywności występowania symptomów chorobowych i czynników ryzyka oraz zmian liczby symptomów rozpoznanych u pacjenta. Inną bardzo ważną funkcjonalnością jest możliwość badania charakterystyk jakościowych generowanych diagnoz w sytuacji występowania jednocześnie wielu symptomów sugerujących choroby współistniejące.



Rys. 3. Interfejs użytkownika

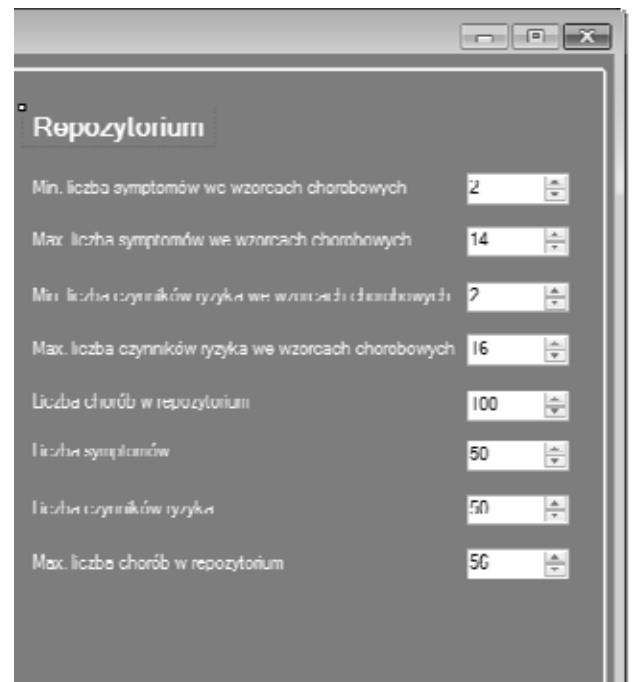
Główne okno aplikacji zostało przedstawione na rysunku 3. Wyszczególniono w nim następujące obszary:

- Panel pacjenta
- Panel repozytorium
- Panel konfiguracji scenariusza symulacji.



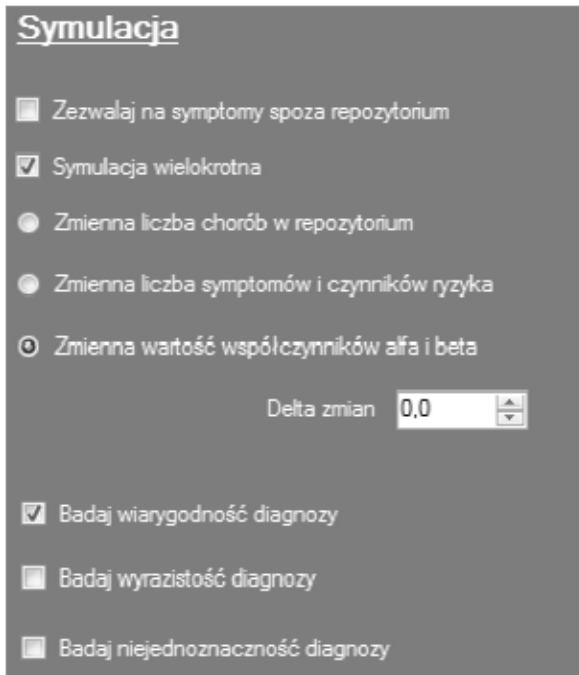
Rys. 4. Panel pacjenta

W panelu pacjenta, przedstawionym na rysunku 4, określa się parametry dotyczące rozpoznanych symptomów chorobowych oraz stwierdzonych czynników ryzyka. Oczywiście liczba każdego z nich nie może być większa od



Rys. 5. Panel repozytorium

Panel repozytorium, zaprezentowany na rysunku 5, zawiera szereg edytowalnych kontrolki, dzięki którym można zdefiniować parametry symulacji związane z repozytorium (liczba chorób w repozytorium, maksymalna liczba symptomów chorobowych i czynników ryzyka etc.). Dostępność poszczególnych kontrolki edycyjnych zależy od wybranego scenariusza symulacji.



Rys. 6. Panel konfiguracji symulacji

Panel symulacyjny, pokazany na rysunku 6, pozwala na wybór jednej z 4 rodzajów symulacji. Aktualnie dostępne są następujące scenariusze symulacji:

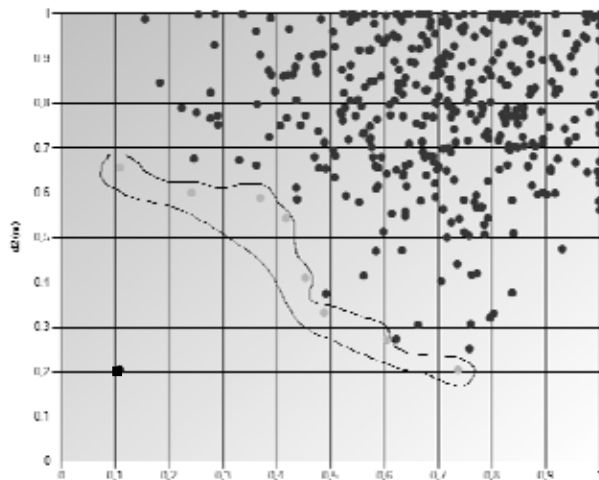
- symulacja 1: przebiega przy stałych niezmieniających się parametrach (odpowiednie obliczenia wykonywane są dla wprowadzonych parametrów jeden raz)
- symulacja 2: przebiega przy zmiennej liczbie chorób w repozytorium (liczba chorób określana jest wartością parametru)
- symulacja 3: przebiega przy zmiennej liczbie symptomów chorobowych oraz czynników ryzyka rozpoznanych u pacjenta oraz stałej liczbie wzorców chorobowych zawartych w repozytorium. Na początku symulacji liczność zbioru zdiagnozowanych symptomów chorobowych jak i czynników ryzyka wynosi 1 i jest zwiększana w trakcie trwania symulacji, aż do osiągnięcia liczności poszczególnych zbiorów wyznaczonej przez użytkownika
- symulacja 4: przebiega przy zmieniających się wartościach współczynników nasilenia symptomów chorobowych i czynnikach ryzyka.

W ramach każdej z wyżej wymienionych symulacji możliwe jest badanie następujących wskaźników:

- wiarygodności diagnozy
- wyrazistości diagnozy
- niejednoznaczności diagnozy.

Po zakończeniu symulacji generowane są wykresy przedstawiające uzyskane wyniki.

Przykładowy wykres został przedstawiony na rysunku 7.



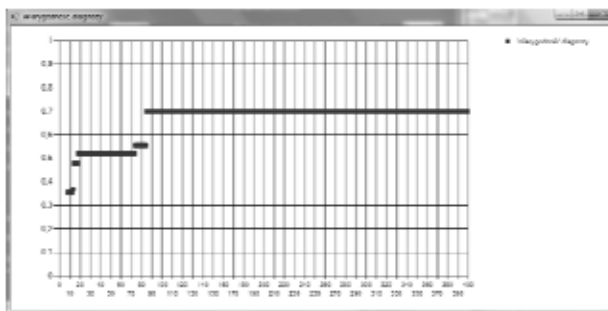
Rys. 7. Wyznaczenie zbioru M_N^R chorób, od których nie ma „bardziej prawdopodobnych”

„Choroby, od których nie ma bardziej prawdopodobnych” (tzw. front Pareto), w myśl zaimplementowanej metody diagnozowania, oznaczane są przez okrągłe punkty o jasnym odcieniu. Pozostałe zaznaczono przez okrągłe punkty o ciemniejszym odcieniu. Punkt oznaczony prostokątem (punkt utopijny) to punkt wyznaczony po uwzględnieniu:

- minimalnej odległości stanu zdrowia pacjenta od wzorca chorobowego w ujęciu symptomów chorobowych w całym repozytorium
- minimalnej odległości stanu zdrowia pacjenta od wzorca chorobowego w ujęciu czynników ryzyka w całym repozytorium.

Jest to obraz „idealnego podobieństwa” stanu zdrowia pacjenta do „idealnie pasującej” choroby „utopijnej”, której niestety nie ma w repozytorium. Jasne punkty na wykresie wyznaczają front Pareto. Użytkownik ma możliwość zdefiniowania, na jakich etapach symulacji generowane są wykresy z wynikami (ile ma ich być). Istnieje także możliwość zapisu otrzymanych wyników do pliku tekstowego. Numer (nazwę) jednostki chorobowej reprezentowanej na wykresie przez określony punkt można uzyskać, klikając kursorem. Szczególnie istotnym wskaźnikiem, który jest badany w procesie symulacyjnym, jest „wiarygodność diagnostyczna” opisana wzorem (2.15). W rozpatrywanym przypadku wartość parametru p ustalono na 2 (odległość geometryczna) [1]. Jak widać ze wzoru, wartość wskaźnika „wiarygodności diagnostycznej” jest ściśle zależna od położenia punktu utopijnego.

Podobnie jak w przypadku badania możliwości wystąpienia danej choroby u pacjenta, tak i tutaj otrzymujemy wykres ilustrujący zmianę omawianego wskaźnika. Może być on badany w każdym z wybranych scenariuszów symulacji. Przykład wykresu wartości wskaźnika wiarygodności diagnostycznej został przedstawiony na rysunku 8. Może być on badany w każdym z wybranych scenariuszów symulacji.



Rys. 8. Wykres przedstawiający wartość wskaźnika wiarygodności w czasie trwania symulacji

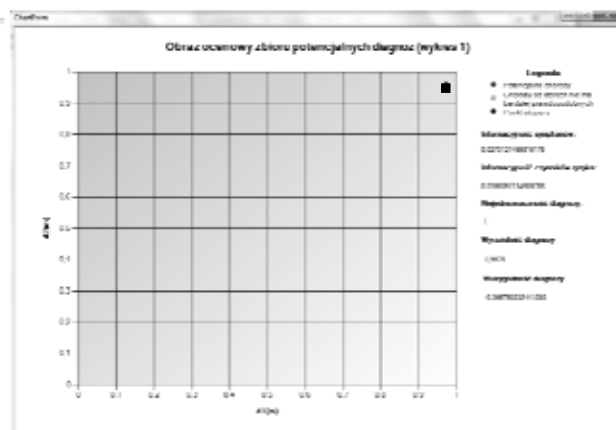
5. Przykładowe przebiegi symulacji

Prezentowany symulator umożliwia też badanie wielu innych charakterystyk jakościowych otrzymywanych diagnoz, w tym również wrażliwości generowanych diagnoz na zmianę stopnia informacyjności wyników badań lekarskich. Pozwala wyznaczanie progów stopnia informacyjności wyników i tym samym kalibrację modelu procedury wspomagania. Poniżej zostanie zaprezentowany przykładowy przebieg symulacji dla scenariusza ze zmienną liczbą chorób w repozytorium. Liczba symptomów chorobowych oraz czynników ryzyka rozpoznana u pacjenta pozostaje bez zmian. Poza obserwacją wykresów przedstawiających zmiany w zbiorze Pareto (dla chorób, od których nie ma bardziej prawdopodobnych), zależy nam także na wygenerowaniu wykresu wartości „wskaźnika wiarygodności”. Po zaznaczeniu odpowiednich pól w aplikacji, zdefiniujemy licznosc zbiorów dla symptomów chorobowych oraz czynników ryzyka rozpoznanych u pacjenta. W omawianym przykładzie wartości parametrów dobrano następująco:

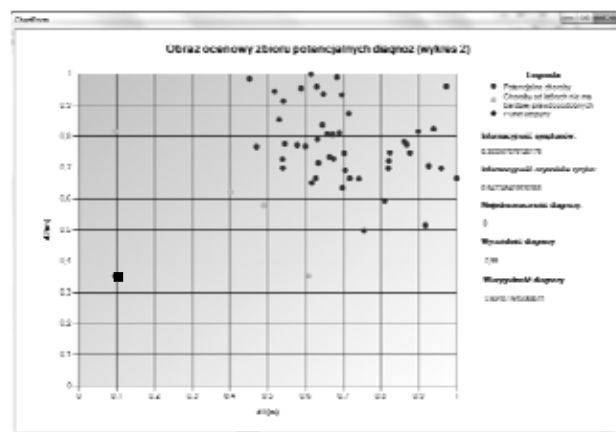
Parametr	Wartość
Liczba rozpoznanych symptomów	32
Liczba rozpoznanych czynników ryzyka	4
Minimalna liczba symptomów opisująca wzorzec chorobowy	2
Maksymalna liczba symptomów opisująca wzorzec chorobowy	8

Minimalna liczba czynników ryzyka opisująca wzorzec chorobowy	2
Maksymalna liczba czynników ryzyka opisująca wzorzec chorobowy	6
Liczba symptomów w repozytorium	110
Liczba czynników ryzyka w repozytorium	70
Maksymalna liczba chorób w repozytorium	400

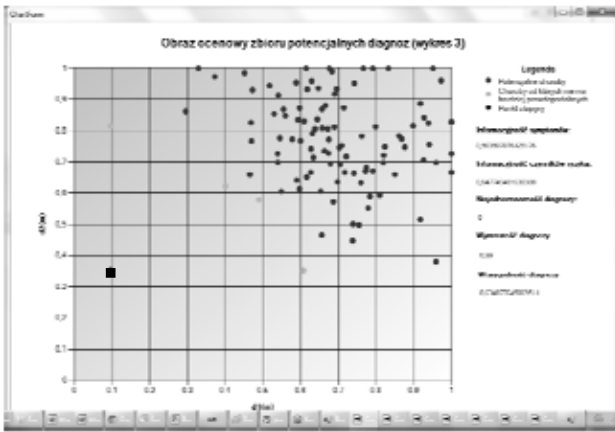
W trakcie trwania symulacji aplikacja generowała wykres przedstawiający odległość stanu zdrowia pacjenta od poszczególnych chorób w repozytorium. Wykresy tworzone co 50 kolejnych wzorców chorobowych dodanych do repozytorium. W pracy przedstawiono wybiórczo jedynie ich część z początku, środka i końca symulacji.



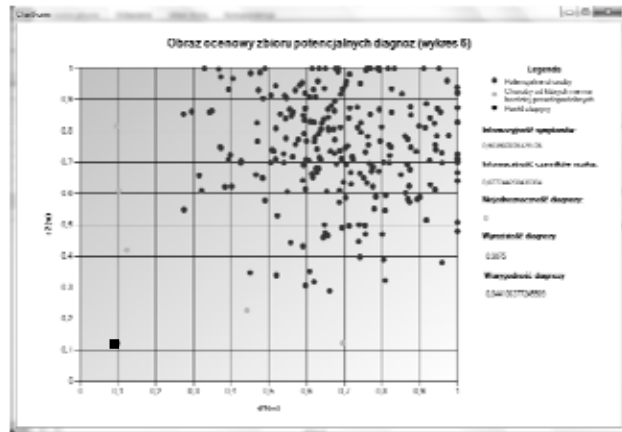
Rys. 9. Początek symulacji – 1 choroba w repozytorium



Rys. 10. Repozytorium – 50 chorób



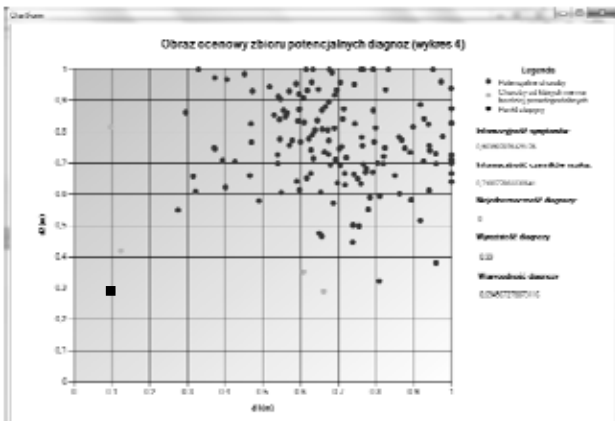
Rys. 11. Repozytorium – 100 chorób



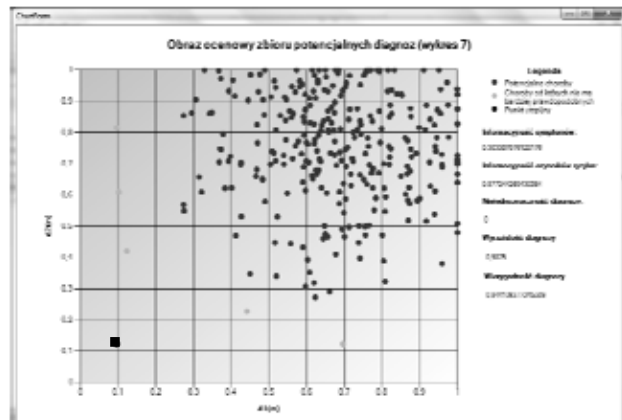
Rys. 14. Repozytorium – 250 chorób

Powyższe wykresy przedstawiają wygenerowane przestrzenie podobieństwa, które są obrazem aktualnej zawartości repozytorium w zestawieniu z rozpoznanym przez lekarza stanem zdrowia pacjenta (zbiór rozpoznanych symptomów chorobowych i czynników ryzyka wraz z ich stopniami nasilenia). Ważnym spostrzeżeniem jest fakt, że przeciwobrazy niektórych punktów, wyznaczane w przestrzeni chorób, mogą być zbiorami o liczności większej niż 1.

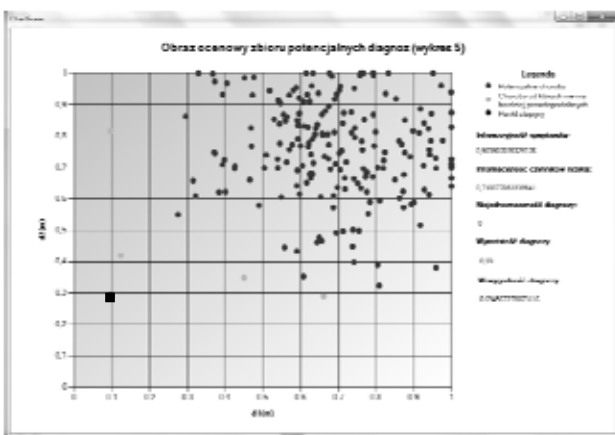
Rysunki o numerach od 9 do 17 przedstawiają kolejne obrazy podobieństwa stanu zdrowia pacjenta, określone przez lekarza na podstawie stwierdzonych symptomów chorobowych i czynników ryzyka do wzorców chorobowych, „rozszerzającego się” w sensie liczby zawartych wzorców jednostek chorobowych repozytorium. Interesująca jest ewolucja tzw. frontu Pareto przedstawiającego obrazy jednostek chorobowych, od których nie ma w repozytorium bardziej „prawdopodobnych”. Rysunek 18 przedstawia zależność wiarygodności diagnoz od zawartości repozytorium.



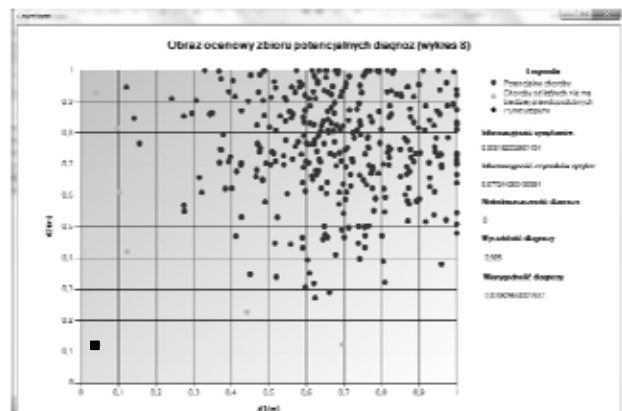
Rys. 12. Repozytorium – 150 chorób



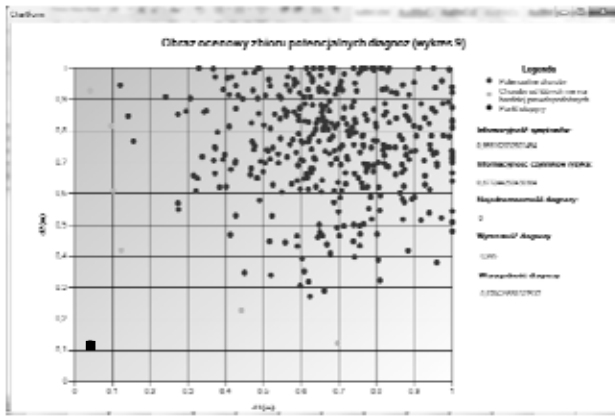
Rys. 15. Repozytorium – 300 chorób



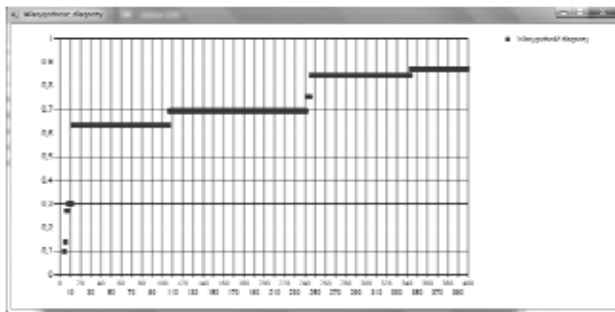
Rys. 13. Repozytorium – 200 chorób



Rys. 16. Repozytorium – 350 chorób



Rys. 17. Repozytorium – 400 chorób



Rys. 18. Wskaźnik wiarygodności diagnozy w zależności od liczby wzorców chorobowych w repozytorium

6. Podsumowanie

Etap wstępnego diagnozowania jest szczególnie ważny w procedurze rozpoznawania chorób. Wykorzystanie specjalistycznych narzędzi bazujących na wiedzy medycznej wspomaga proces ustalania wstępnej diagnozy. Zawężenie procesu diagnozowania do pewnego podzbioru chorób ma decydujący wpływ na strategię dalszych badań specjalistycznych oraz ścieżkę leczenia pacjenta. Oprogramowanie wspierające proces ustalania wstępnej diagnozy medycznej przyczynia się do minimalizacji kosztów i czasu badań, a także do skuteczności leczenia pacjenta.

Przedstawiony w pracy symulator pozwala na badanie podstawowych charakterystyk jakościowych diagnoz medycznych tworzonych przez komputerowy system wspomaganie decyzji medycznych. Aktualnie symulator pozwala na badanie takich charakterystyk jak wiarygodność diagnozy, wyrazistość diagnozy oraz niejednoznaczność diagnozy. Planuje się rozwijanie narzędzia poprzez dodanie nowych scenariuszy symulacji oraz umożliwienie badania innych charakterystyk jakościowych procesu medycznego diagnozowania. Szczególnie istotna wydaje się możliwość symulacyjnego badania wrażliwości generowanych diagnoz na

zmiany informacyjności wyników badań lekarskich rozumianej w sensie bardzo szerokim [3, 4]. Kolejnym bardzo interesującym i ważnym z praktycznego punktu widzenia obszarem aplikacji symulatora może być badanie warunków występowania tzw. chorób współistniejących. Tego typu badania umożliwiają wyznaczenie tzw. progów wrażliwości i kalibrację wszystkich parametrów przyjętego modelu wnioskowania diagnostycznego.

7. Bibliografia

- [1] A. Ameljańczyk, *Optymalizacja wielokryterialna w problemach sterowania i zarządzania*, Ossolineum, Warszawa, 1984.
- [2] A. Ameljańczyk, „Matematyczny model przestrzeni życia w komputerowym systemie wspomaganie decyzji medycznych”, *Materiały I Krajowej Konferencji pt. „Systemy komputerowe i teleinformatyczne w służbie zdrowia”*, Warszawa, 2009.
- [3] A. Ameljańczyk, „Properties of the Algorithm for Determining an Initial Diagnosis Based on Two-Criteria Smilarity Model”, *Biuletyn Instytutu Systemów Informatycznych*, Nr 8, str. 9–16 (2011).
- [4] A. Ameljańczyk, „Wielokryterialne mechanizmy wspomaganie podejmowania decyzji medycznych w modelu repozytorium w oparciu o wzorce”, *Biuletyn Instytutu Systemów Informatycznych*, Nr 5, str. 1–6 (2010).
- [5] R.S. Porter, J.L. Kaplan, B.P. Homeier, *The Merc Manual. Objawy kliniczne. Praktyczny przewodnik diagnostyki i terapii*, wyd. I polskie, 2010.
- [6] J. Makal, *Medical Diagnosis as the Result of Multiparameter Indirect Measurement*.
- [7] A. Troelsen, *Język C# 2008 I platform .NET 3.5*, PWN, 2009.
- [8] K. Matulewski, *C# 3.0 I NET 3.5. Technologia LINQ*, Helion 2009.
- [9] J. Sharp, *Microsoft Visual C# 2010. Krok po kroku*, APN PROMISE, 2010.
- [10] K. Krupa, *Modelowanie, symulacja i prognozowanie*, WNT, 2008.

The simulator of the examination process of qualitative characteristics of medical diagnoses generated by the computer medical decisions support system

A. AMELJAŃCZYK, P. DŁUGOSZ

The article presents a simulation tool designed to examine selected qualitative characteristics of medical diagnoses, generated by the computer system supporting medical decisions. The system uses diagnostic conclusion method based on defining a similarity between the patient's condition and diseases' patterns contained in repository. The presented application allows a simulative examination of qualitative indexes of the diagnostic process, particularly: a reliability index of generated medical diagnosis with different test scenarios. The computer program was built using .NET Framework technology and supplement libraries.

Keywords: computer simulation, disease symptoms, risk factors, medical patterns, the credibility of the diagnosis, the expressiveness of the diagnosis, the in equivalence of the diagnosis, platform .NET, Windows Forms.