

UŻYCIE SIECI PROBABILISTYCZNEJ ORAZ METODY NAJBLIŻSZEGO SĄSIADA DO IDENTYFIKACJI OBCYCH RADIOSTACJI OKRĘTOWYCH

Praca przedstawia zastosowanie probabilistycznej sieci neuronowej oraz metody najbliższego sąsiada do identyfikacji obcych radiostacji okrętowych. Badania zaprezentowane w pracy zostały przeprowadzone z użyciem danych rzeczywistych.

USING A PROBABILISTIC NEURAL NETWORK AND THE NEAREST NEIGHBOUR METHOD TO IDENTIFY SHIP RADIOSTATIONS

The paper presents application of a probabilistic neural network and the nearest neighbour method to identify nonself ship radiostations. All the experiments reported in the paper were performed on real data.

1. WSTĘP

W pracy zaprezentowano oraz porównano dwa różne rozwiązania problemu identyfikacji swój-obcy: metodę najbliższego sąsiada (NN – ang. Nearest Neighbour) oraz probabilistyczną sieć neuronową (PNN – ang. Probabilistic Neural Network). Oba rozwiązania są klasycznymi pamięciowymi metodami identyfikacji obiektów. Konstrukcja systemu identyfikacji w przypadku obu metod polega na zapamiętaniu w pamięci pewnej liczby obrazów uczących reprezentujących wszystkie identyfikowane obiekty. Zadanie identyfikacji swój-obcy zakłada jednak brak obrazów uczących dla obiektów uznanych za obce. W tym przypadku system identyfikacji musi bazować wyłącznie na danych o obiektach własnych. Aby sprawdzić jak obie metody radzą sobie w sytuacji braku informacji o jednej z identyfikowanych klas przeprowadzono eksperymenty w trakcie których zadaniem obu metod była identyfikacja radiostacji okrętowych.

2. METODY ZASTOSOWANE PODCZAS BADAŃ

2.1 Metoda najbliższego sąsiada

Najprostszym sposobem identyfikacji obiektu jest porównanie jego sygnatury z sygnaturami obiektów umieszczonymi w systemie. Można tego dokonać poprzez wyszukanie w bazie danych systemu sygnatury najbardziej podobnej do sygnatury obiektu podlegającego identyfikacji. Tożsamość tej sygnatury wskazuje na klasę identyfikowanego

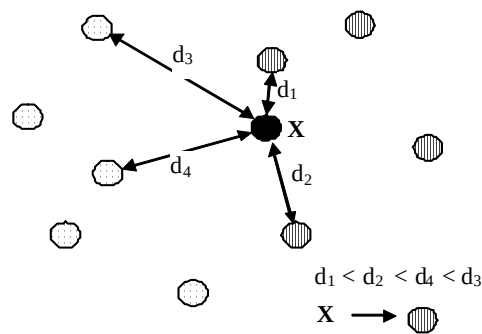
¹ Akademia Marynarki Wojennej, Wydział Nawigacji i Uzbrojenia Okrętowego; 81-103 Gdynia; ul. Śmidowicza 69, Tel: + 48 58 6262703, E-mail: t.praczyk@amw.gdynia.pl

obiektu. Takie rozwiązanie nazywa się metodą najbliższego sąsiada (NN). W metodzie NN do określenia podobieństwa najczęściej stosuje się następujące metryki [3]:

$$\rho(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_{k=1}^n (x_1^k - x_2^k)^2} \quad (1)$$

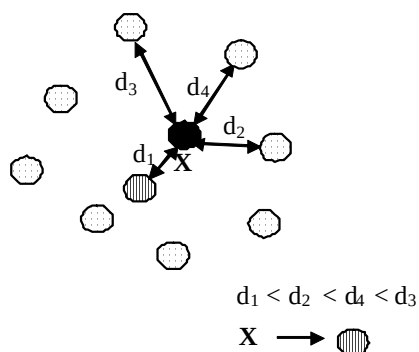
$$\rho(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_{k=1}^n |x_1^k - x_2^k|} \quad (2)$$

gdzie $\mathbf{x}_1, \mathbf{x}_2$ są wektorami: $\mathbf{x}_1 = (x_1^1, x_1^2, x_1^3, \dots, x_1^n)$, $\mathbf{x}_2 = (x_2^1, x_2^2, x_2^3, \dots, x_2^n)$.

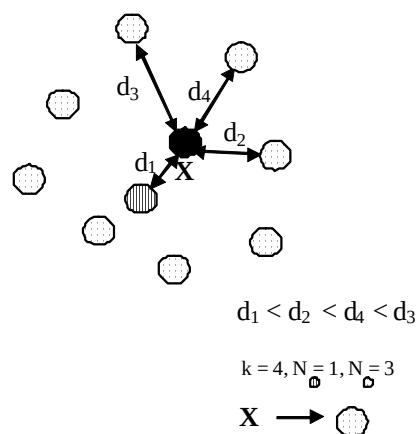


Rys.1. Zasada działania metody NN

Metoda NN posiada jedną podstawową wadę. Jest nią mała odporność na występowanie zakłóceń w bazie danych systemu identyfikacji. Jedna błędna sygnatura umieszczona w systemie potrafi wypaczyć uzyskiwane wyniki. Sytuację taką prezentuje Rys. 2. Błędna sygnatura oznaczona paskami pionowymi umieszczona w samym środku obszaru należącego do kategorii obiektów „kropkowanych” wpływa na niewłaściwą identyfikację wszystkich obiektów o sygnaturach występujących w najbliższym otoczeniu błędnej sygnatury. Zjawisko to możemy ograniczyć poprzez zastosowanie zmodyfikowanej metody najbliższego sąsiada zwanej również kNN. Polega ona na uporządkowaniu sygnatur uczących (sygnatur z bazy danych systemu) według rosnących wartości odległości od sygnatury reprezentującej niezidentyfikowany obiekt. Na początku ciągu znajdują się sygnatury najbliższe, na końcu zaś najdalsze. Z tak skonstruowanego ciągu tworzymy podciąg segregując pierwszych k sygnatur. Niezidentyfikowany obiekt zaliczamy do kategorii, która jest najliczniej reprezentowana w stworzonym podciągu. Powyższa modyfikacja powoduje, że pojedyncza błędna sygnatura ucząca nie jest w stanie w sposób decydujący wpłynąć na ostateczny wynik identyfikacji obiektu.

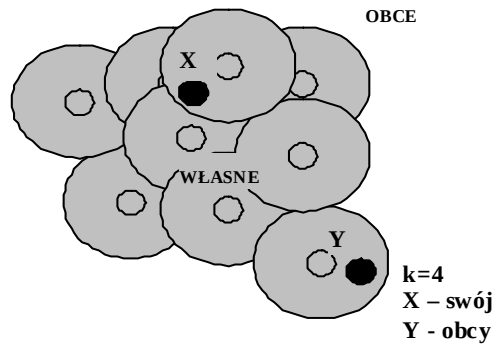


Rys.2. Wpływ błędnej sygnatury na niewłaściwą identyfikację obiektu metodą NN



Rys.3. Zasada działania metody kNN

W eksperymentach przedstawionych w dalszej części pracy zastosowano zmodyfikowany wariant metody przedstawionej powyżej. Do identyfikacji obiektów wykorzystywane były wyłącznie sygnatury obiektów własnych (baza danych nie zawierała sygnatur obiektów obcych). Aby niezidentyfikowany obiekt został uznany za własny jego sygnatura musiała być podobna do co najmniej k sygnatur zapamiętanych w systemie. Zastosowany wariant metody kNN zakładał, że sygnatury są podobne do siebie, jeśli odległość między nimi jest mniejsza lub równa pewnej zakładanej wartości (δ).



Rys.4. Zasada działania wariantu metody kNN zastosowanego w trakcie eksperymentów

2.2 Probabilistyczna sieć neuronowa

Klasyfikacja obiektów przy pomocy sieci probabilistycznej (PNN – ang. Probabilistic Neural Network) odbywa się za pomocą następującej funkcji decyzyjnej [2,3]:

$$g(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \exp \left[- \left(\frac{d(\mathbf{x}, \mathbf{x}_i)}{p\sigma} \right)^2 \right] \quad (3)$$

gdzie

d jest odległością euklidesową pomiędzy wektorami $\mathbf{x}$ oraz $\mathbf{x}_i$ reprezentującymi odpowiednio obiekt identyfikowany oraz i -ty przykład uczący,
 n jest liczbą wektorów uczących danej kategorii,
 p jest rozmiarem wektora $\mathbf{x}$ oraz wszystkich wektorów uczących $\mathbf{x}_i$,
 σ jest parametrem kształtu.

Ostateczna decyzja o przyporządkowaniu obiektu reprezentowanego przez wektor $\mathbf{x}$ do kategorii j -tej opiera się na następującej relacji:

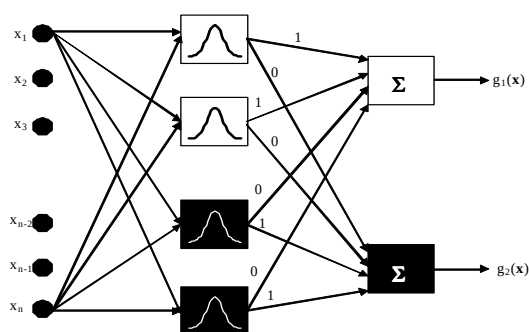
$$h_j c_j g_j(\mathbf{x}) > h_k c_k g_k(\mathbf{x}) \quad \forall_{k \neq j} \quad (4)$$

gdzie

h_i oznacza prawdopodobieństwo a priori przynależności obiektu do i -tej klasy,
 c_i koszt błędnej identyfikacji obiektu należącego do i -tej kategorii.

Sieć probabilistyczna składa się, co najmniej z trzech warstw: wejściowej, radialnej oraz wyjściowej. Zadaniem pierwszej warstwy jest doprowadzenie sygnału wejściowego do sieci. Druga warstwa składa się z neuronów radialnych. Każdy z nich reprezentuje funkcję radialną z centrum nad swoim przypadkiem uczącym. Warstwa ta zawiera tyle neuronów ile wzorców znajduje się w ciągu uczącym – jeden neuron odpowiada pojedynczemu przypadkowi uczącemu (np. sygnaturze okrętu). Warstwa wyjściowa dokonuje zsumowania

sygnałów pochodzących od neuronów z warstwy wcześniejszej. Każdej klasie odpowiada jeden neuron tej warstwy. Sygnały generowane przez neurony radialne klasy reprezentowanej przez neuron wyjściowy podawane są na jego wejście z wagą 1. Połączenie z innymi neuronami radialnymi charakteryzowane jest wagą 0*. Dodatkowymi elementami struktury sieci PNN mogą być kolejne warstwy związane z parametrem h oraz c – wzór (4). Wygląd trzech pierwszych zasadniczych części sieci PNN prezentuje poniższy rysunek.



Rys.5. Sieć probabilistyczna służąca do identyfikacji dwóch klas obiektów

Sieć PNN posiada wszystkie zalety i wady związane z wykorzystaniem danych uczących podczas wnioskowania. Przede wszystkim proces uczenia polega na wkomponowaniu wzorców uczących w drugą radialną warstwę sieci i jest to czynność prawie natychmiastowa. Dodatkowy czas trzeba poświęcić jedynie na eksperymentalne dostrójenie parametru kształtu σ przy pomocy specjalnie stworzonego ciągu walidacyjnego.

Wielce użyteczną własnością sieci PNN jest jej zdolność do poszerzania swojej wiedzy. Pojawienie się nowego wzorca uczącego, który może wносить dodatkowe, ważne informacje wzbogacające dotychczasowy stan wiedzy, nie powoduje konieczności przeprowadzenia procesu uczenia od samego początku. Możliwe jest bezpośrednie wprowadzenie go do sieci w postaci kolejnego neuronu radialnego. Jest to bardzo istotna cecha sieci PNN która umożliwia jej ciągłą pracę, bez przerw przeznaczonych na długotrwały proces uczenia. Wykorzystując sieć PNN jesteśmy w stanie dokonywać identyfikacji obiektów zawsze w oparciu o dane aktualne.

Wykorzystanie wzorców uczących rodzi niestety również swoje negatywne skutki. Potrzeba analizy dużego zbioru danych uczących podczas wnioskowania wywołuje efekt długiej odpowiedzi na zadany sygnał wejściowy. Chcąc nauczyć sieć każdego szczegółu związanego z rozpoznawanymi obiektami musimy dodawać do ciągu uczącego coraz to nowe wzorce a to z kolei wydłuża czas odpowiedzi. Z tego względu w przypadku obszernych zbiorów uczących konieczne jest zastosowanie metod przyspieszających działanie sieci [1]. Mogą one polegać np. na zastosowaniu wzorców średnich przy

* Można oczywiście wyobrazić sobie również alternatywną strukturę sieci PNN, w której neurony radialne danej klasy połączone są tylko z neuronami wyjściowymi tej klasy. Nie występują połączenia do neuronów wyjściowych pozostałych klas.

jednoczesnej modyfikacji funkcji decyzyjnej (3) uwzględniającej zastąpienie wzorców uczących ich średnimi reprezentantami.

Drugą ważną zaletą sieci PNN jest większa niż we wcześniej zaprezentowanych metodach (NN i kNN) bazujących na odległości odporność na występowanie zakłóceń w ciągu uczącym. Jest tak ze względu na wykorzystywaną przez sieć PNN postać funkcji decyzyjnej preferującej pod względem wielkości generowanej odpowiedzi obszary o gęstym upakowaniu wzorców uczących. Pojawienie się w ciągu uczącym pojedynczego błędnego obrazu nie jest w stanie w sposób decydujący wpłynąć na poprawność udzielonej odpowiedzi.

W trakcie eksperymentów, których wyniki zaprezentowane są w niniejszej pracy wykorzystano sieć PNN składającą się z trzech podstawowych warstw. Pierwsza warstwa była odpowiedzialna za dostarczenie wejściowych sygnałów radiowych do sieci. Druga warstwa zawierała neurony radialne z zapamiętanymi sygnałami pochodzącymi od okrętów własnych. Ostatnia warstwa była zbudowana z pojedynczego neuronu, którego sygnał wyjściowy był interpretowany jako prawdopodobieństwo tego, że sygnał radiowy podany na wejście sieci jest sygnałem pochodzącym od jednostki własnej. W przypadku, kiedy sygnał wyjściowy sieci był większy lub równy 0.5 sygnał radiowy był uznawany jako własny. W przeciwnym przypadku sygnał wejściowy był identyfikowany jako obcy.

3. EKSPERYMENTY

3.1 Sygnały radiowe i sygnatury okrętowe

Podczas całych eksperymentów okręty reprezentowane były przez zakodowane sygnały radiowe emitowane przez radiostacje okrętowe. Częstotliwość sygnałów wykorzystywanych w trakcie badań była z zakresu 1.6 – 30 MHz. Aby sygnały mogły stanowić sygnatury okrętowe poddano je wstępnej obróbce. Najpierw z każdego z nich wyznaczono transformatę FFT. W kolejnym kroku określono numer próbki centralnej S_c [1]:

$$S_c = \frac{f_i}{R} \quad (5)$$

$$R = \frac{f_s}{N_s} \quad (6)$$

gdzie

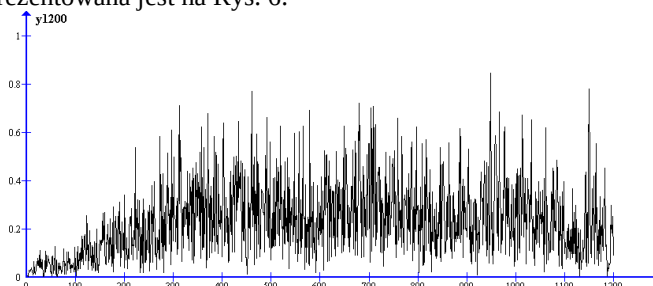
f_i jest częstotliwością pośrednią;

f_s jest częstotliwością próbkowania (4000 Hz – 200 kHz);

N_s jest liczbą próbek w widmie każdego sygnału radiowego (32768).

Próbka S_c wskazywała część widma każdego sygnału zawierającą najwięcej informacji. Ostatnią czynnością przygotowującą sygnały do badań było wyodrębnienie z widm zawierających 32768 próbek wektorów o długości 1200 elementów. Wektory te zbudowano z 600 próbek występujących na lewo od próbki S_c oraz 600 próbek występujących na prawo

od wspomnianej próbki centralnej. Przykładowa sygnatura okrętowa wykorzystana w badaniach zaprezentowana jest na Rys. 6.



Rys.6. Przykładowe widmo sygnału radiowego (sygnatura okrętowa)

W trakcie eksperymentów dysponowano w sumie 2610 sygnałami radiowymi pochodzącymi od sześciu różnych radiostacji okrętowych. Na potrzeby badań cały dostępny zbiór sygnałów podzielony został na trzy podzbiory. Pierwszy z nich (Zbiór nr 1) zawierał 919 sygnałów pochodzących od 3 różnych okrętów uznanych jako własne. Zbiór ten został wykorzystany do budowy każdego z testowanych rozwiązań. Kolejny zbiór (Zbiór nr 2) zawierał 900 sygnałów pochodzących od tych samych trzech okrętów własnych. Zbiór ten został wykorzystany do testowania każdego z rozwiązań. Ostatni zbiór (Zbiór nr 3) zawierał 791 sygnałów pochodzących od trzech okrętów uznanych jako obce. Głównym przeznaczeniem Zbioru nr 3 podobnie jak Zbioru nr 2 były testy.

3.2 Wyniki eksperymentów

W trakcie eksperymentów przetestowano następujące rozwiązania:

- PNN – sieć PNN,
- 1NN – kNN z $k=1$,
- 2NN – kNN z $k=2$,
- 4NN – kNN z $k=4$.

Wszystkie powyższe rozwiązania zostały przetestowane dla różnych wartości parametrów. Optymalne wartości parametrów podano w Tab. 1.

Tab. 1. Wyniki eksperymentów (% błędów typu false positive – okręty własne o sygnaturach ze zbioru nr 2 uznane jako obce; % błędów typu false negative – okręty obce o sygnaturach ze zbioru nr 3 uznane jako własne; % wszystkich błędów – błędne identyfikacje sygnałów ze zbiorów nr 2 i 3)

	% błędów typu false positive	% błędów typu false negative	% wszystkich błędów	parametry
PNN	2.5%	39.4%	19.8%	$\sigma = 6.7$
2NN	2.7%	40.8%	20.5%	$\delta = 4800$
4NN	2.2%	42%	20.8%	$\delta = 4900$
1NN	3.2%	41.4%	21.1%	$\delta = 4730$

Badania pokazały że najskuteczniejszą metodą identyfikacji swój-obcy jest sieć PNN. Uzyskała ona wynik 19.8% błędnych identyfikacji. Poszczególne warianty metody kNN okazały się tylko nieznacznie gorsze. Najgorszy wynik uzyskany został przy pomocy metody kNN i wynosi on 21.1% błędnych identyfikacji. Wynik ten jest jednak tylko o 1.3% gorszy od wyniku sieci PNN. Na uwagę zasługują liczne błędy identyfikacji typu „false negative”. Wynikają one ze zbyt dużego zasięgu działania sieci PNN lub przykładów uczących zapamiętanych w pamięci dla różnych wariantów metody kNN. Zmniejszenie zasięgu oddziaływania dla każdej z metod prowadziło jednak do zwiększenia całkowitej liczby błędów.

3. WNIOSKI

Badania zaprezentowane w pracy pokazały, że zarówno metoda kNN jak i sieć PNN są w miarę skutecznym narzędziem identyfikacji typu swój-obcy. Wyniki uzyskane przez obie metody są porównywalne. Wydaje się że osiągi obu metod mogą jednak zostać jeszcze poprawione. Elementem który wymaga dalszej pracy jest redukcja błędów typu „false negative”. Poprawa osiągnięć w tym zakresie może przyczynić się do znacznej poprawy wyników obu metod.

4. BIBLIOGRAFIA

- [1] Praczyk T.: *Probabilistic neural network application to warship radio station identification*, Computational Methods in Science and Technology, CMST 13(1) pp 53-58, 2007
- [2] Specht D.F.: *Probabilistic neural networks*, Neural Networks Volume 3, Issue 1, pp. 109-118, 1990
- [3] Stateczny A., Praczyk T.: *Sztuczne sieci neuronowe w rozpoznawaniu obiektów morskich*, Gdańskie Towarzystwo Naukowe. 2002