

doi:10.15199/48.2016.11.47

## Nieinwazyjna identyfikacja odbiorników energii elektrycznej z wykorzystaniem metod sztucznej inteligencji

**Abstract.** The paper presents the electrical appliances identification method using artificial intelligence methods, based on the measurements of currents and voltages in the medium frequency band. To identify the change of the particular device state, the random forest and *k* Nearest Neighbours were proposed. The former is characterized by the high classification accuracy and the ability to work in the uncertainty conditions. The latter is based on the distance calculation, missing the machine learning stage. Experiments confirmed the usefulness of both methods to identify the selected appliances. (**Non-invasive identification of electrical appliances using artificial intelligence methods**).

**Streszczenie.** W artykule przedstawiono zastosowanie metod sztucznej inteligencji do identyfikacji odbiorników energii elektrycznej na podstawie pomiarów prądu i napięcia w paśmie średnich częstotliwości. W celu klasyfikacji konkretnego urządzenia wybrano las losowy oraz metodę *k* Najbliższych sąsiadów. Pierwszy algorytm charakteryzuje się wysoką skutecznością oraz zdolnością do pracy w warunkach niepewności pomiarowej, drugi zaś to podejście oparte na obliczaniu odległości, nie wymagające procesu uczenia maszynowego. Badania wykazały wysoką skuteczność klasyfikacji wybranych odbiorników zapewnianej przez obie metody.

**Keywords:** appliance identification, random forest, *k* Nearest Neighbours.

**Słowa kluczowe:** identyfikacja odbiorników energii elektrycznej, las losowy, *k* Najbliższych Sąsiadów.

### Wprowadzenie

Nieinwazyjna identyfikacja odbiorników energii elektrycznej to aktualne i ważne zadanie dla zaawansowanych systemów pomiarowych. Jest ono trudne ze względu na dużą liczbę różnorodnych urządzeń działających w sieci energetycznej, a także charakter ich pracy. Istotą systemu identyfikacyjnego jest poprawne określenie zbioru odbiorników działających zarówno w konkretnym momencie jak i wybranym przedziale czasu, na podstawie cech sygnałów prądu i napięcia mierzonych w pobliżu miernika energii elektrycznej. Do tego celu wykorzystywano dotychczas mniej lub bardziej skomplikowane metody, również należące do dziedziny sztucznej inteligencji. Największą popularnością cieszą się obecnie sztuczne sieci neuronowe [1], ukryte łańcuchy Markowa [2], czy drzewa decyzyjne [3]. Są to podejścia łatwe w implementacji, a przy tym wykazujące wysoką dokładność identyfikacji. Ze względu na dużą różnorodność konfiguracji, w których testowane są metody, trudno określić przewagę jednego konkretnego podejścia nad innymi. Dążenie do maksymalizacji jakości klasyfikacyjnej (czyli przypadków poprawnie zidentyfikowanych odbiorników) uzasadnia implementację nowych algorytmów.

W artykule przedstawiono implementację dwóch nowoczesnych algorytmów sztucznej inteligencji w zadaniu identyfikacji zmiany stanu odbiorników energii elektrycznej. Pomimo dowiedzionej skuteczności w innych dziedzinach, nie zostały one skutecznie przetestowane w przedstawionym celu. Pierwszy algorytm to las losowy (ang. Random Forest), będący rozwinięciem drzewa decyzyjnego. Charakteryzuje się on wysoką skutecznością w zadaniach klasyfikacyjnych, która jednak zależy m.in. od liczby wygenerowanych drzew. Druga metoda to *k* Najbliższych Sąsiadów (ang. *k* Nearest Neighbours), będąca rozwinięciem metody słownikowej, podejmującej decyzję na podstawie najmniejszej odległości pomiędzy przykładem badanym a zapamiętanym w bazie danych (słowniku). Oba podejścia są istotnie różne, z tego powodu ich konfrontacja jest pożądana i może pokazać wady i zalety obu algorytmów.

Ponieważ skuteczność klasyfikatora silnie zależy od danych dostarczonych mu do analizy, jednym z przedstawionych etapów realizacji systemu identyfikacji jest analiza cech sygnałów, na podstawie których podejmowana

jest decyzja wskazująca na odbiornik, którego stan uległ zmianie. Omówiono cechy sygnałów prądowo-napięciowych wykorzystane do uczenia klasyfikatora oraz późniejszej identyfikacji odbiorników. Następnie przedstawiono szczegóły implementacji obu algorytmów. Wyniki eksperymentów udowadniają ich skuteczność w opisanym zadaniu.

### Zasada działania systemu identyfikacji odbiorników energii elektrycznej

System nieinwazyjnej identyfikacji odbiorników energii elektrycznej składa się z dwóch elementów. Pierwszy to specjalizowane urządzenie pomiarowe, instalowane w pobliżu miernika znajdującego się w węzle zbiorczym, zdolne do analizy zagregowanych sygnałów prądu i napięcia w sieci. Drugi to moduł programowy, ekstrahujący wybrane cechy z mierzonych sygnałów oraz podejmujący decyzję na temat aktualnej konfiguracji działających odbiorników. System przystosowany jest do analizy urządzeń dwu- oraz wielostanowych. Do pierwszych zalicza się takie, które można jedynie włączyć lub wyłączyć (np. żarówka, czajniki itp.). Druga grupa obejmuje urządzenia realizujące kilka programów pracy (np. pralki, czy zmywarki). Są one opisywane za pomocą skończonych maszyn stanów (ang. finite state machines), reprezentujących szereg trybów, w których odbiornik może się znajdować. Ponieważ te ostatnie są najtrudniejsze do analizy i wymagają osobnego podejścia, w przedstawionym projekcie skupiono się wyłącznie na pierwszej grupie odbiorników.

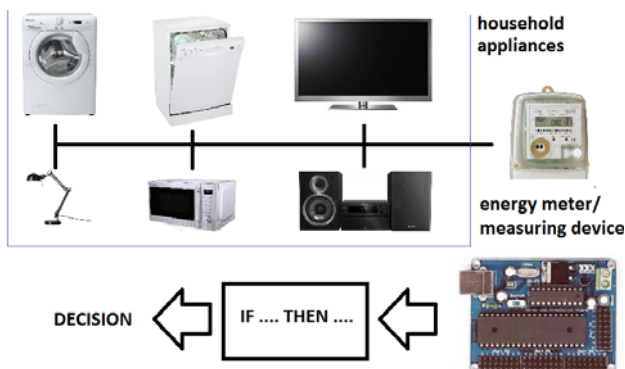
System działa na bieżąco (on-line), najpierw wykrywając zmianę w konfiguracji odbiorników działających w mieszkaniu na podstawie poziomu sygnału prądowego. Przyjęto, że do takiego zdarzenia dochodzi w momencie, gdy różnica pomiędzy poprzednim, a aktualnym poziomem prądu (nazywana dalej „skokiem”) wynosi co najmniej 100mA. Wartość ta została dobrana eksperymentalnie jako kompromis pomiędzy zdolnością do wykrywania odbiorników energooszczędnych (takich jak żarówka LED) oraz pojawianiem się fałszywych alarmów (tzn. wykrywaniem zmian poziomu prądu, które nie są związane z włączeniem lub wyłączeniem żadnego odbiornika). Wyszukiwanie zdarzeń odbywa się dwukrotnie w ciągu sekundy, dzięki czemu do minimum sprowadzono

prawdopodobieństwo przeoczenia zdarzenia zmiany stanu odbiornika. Gdy zmiana zostaje wykryta, dochodzi do identyfikacji urządzenia, które ją spowodowało na podstawie zbioru cech  $s$  zmierzonych sygnałów.

Wykorzystanie metody sztucznej inteligencji wymaga pozyskania danych w trybie off-line, na podstawie których będzie ona ekstrahować wiedzę na temat związku pomiędzy zestawem cech a identyfikatorem odbiornika. W tym celu konieczne są badania laboratoryjne, w wyniku których dla każdego odbiornika z osobna pozyskiwane są wektory cech  $s = \{s_1, \dots, s_m\}$ . Procedura polega na włączeniu jednego urządzenia oraz dokonywania pomiarów w pewnych odstępach czasu. Dzięki temu uzyskuje się szereg wektorów opisujących (za pomocą identyfikatora całkowitoliczbowego  $c$ ) wybrany odbiornik, które mogą następnie zostać wykorzystane podczas uczenia klasyfikatora. Powtórzenie operacji dla wszystkich urządzeń pracujących w mieszkaniu powoduje powstanie zbioru przykładów  $L$ . Tak zdefiniowany zbiór umożliwia przeprowadzenie uczenia z nadzorem, podczas którego algorytm rozróżnia przykłady należące do różnych kategorii na podstawie ich cech.

Głównym problemem zastosowania klasyfikatora do identyfikacji w trybie on-line jest fakt, że został on nauczony na danych dla pojedynczych odbiorników, zaś w rzeczywistej lokalizacji występuje wiele urządzeń jednocześnie, dynamicznie zmieniając swój stan. Z tego powodu analizie za pomocą metody inteligentnej poddawane są nie wektory symptomów obliczone bezpośrednio na podstawie zmierzonych sygnałów, lecz ich różnice w stosunku do poprzedniej iteracji:

$$(1) \quad s = \Delta s_{i,i-1} = s_i - s_{i-1}$$



Rys. 1. Struktura systemu nieinwazyjnego monitorowania energii elektrycznej

System monitorowania energii elektrycznej (Rys. 1) wykorzystuje specjalizowany sprzęt akwizycji danych, którego zadaniem jest próbkowanie przebiegów prądu i napięcia w sieci energetycznej zasilającej badaną lokację. Spośród dostępnych częstotliwości próbkowania (od których zależą następnie pozyskiwane cechy sygnałów) w opisywanym projekcie wykorzystano pasmo średnich częstotliwości (rzędu pojedynczych kHz), dzięki czemu możliwa jest ekstrakcja szeregu cech opisujących zagregowane przebiegi w dziedzinie czasu i częstotliwości. Zadaniem systemu monitorującego jest analiza inkrementacyjna konfiguracji odbiorników pracujących w konkretnym momencie. Oznacza to wykrywanie pojedynczych urządzeń zmieniających swój stan (tzn. włączanych, wyłączanych lub przechodzących w inny tryb pracy). Podstawą decyzji o wykryciu takiego zdarzenia jest analiza wartości chwilowych prądu, w nim bowiem manifestuje się zmiana konfiguracji odbiorników. Zakładając, że każde zdarzenie zostanie poprawnie

wykryte, powinno stać się możliwe pełne scharakteryzowanie sekwencji pracy odbiorników. Po określeniu, że doszło do zdarzenia, w stanie ustalonym uruchamiany jest moduł identyfikacyjny. Wykorzystano tu klasyfikatory lasu losowego oraz kNN, które pracują zamiennie. Operując na zbiorze cech pozyskanych z indywidualnej analizy każdego odbiornika (przy wszystkich pozostałych urządzeniach wyłączonych) podejmują one decyzję odnośnie stanu konfiguracji urządzeń.

### Zestaw wykorzystanych cech sygnałów prądowo-napięciowych

Zaprojektowany system dokonuje pomiarów sygnałów prądowo-napięciowych w paśmie średnich częstotliwości dzięki próbkowaniu z częstotliwością 2 kHz. Dzięki temu możliwe jest wychwycenie większości istotnych cech zarówno w dziedzinie czasu, jak i częstotliwości. Ich dobór zależy od projektanta systemu, rzutują one również na zdolności do rozróżniania odbiorników (szczególnie w sytuacji, gdy charakteryzują je podobne wartości odpowiednich cech). Dobór cech powinien być jednak dopasowany do możliwości klasyfikatora. W przypadku lasu losowego, który tworzony jest podczas procesu uczenia, cechy najistotniejsze z punktu widzenia możliwości rozróżnienia są wybierane automatycznie do struktury każdego drzewa. Dzięki temu uzyskiwany jest podzbiór cech istotnych dla klasyfikatora, które powinny być analizowane podczas działania. Z kolei klasyfikator kNN bierze pod uwagę wszystkie cechy wchodzące w skład analizowanego wektora. Tym samym istotne jest dobranie najważniejszych cech w sensie obliczanej odległości między przykładami. Ponieważ metoda kNN nie podlega uczeniu, konieczne jest zastosowanie zewnętrznej techniki selekcji najważniejszych cech.

Z wymienionych powodów dobrano szeroki zbiór cech do przetwarzanego zbioru. Ponieważ wszystkie one są pozyskiwane z tych samych przebiegów, wzbogacenie wektora  $s$  o kolejne wartości jest związane tylko ze zwiększeniem nakładu obliczeniowego, nie wymaga zaś żadnych modyfikacji w części sprzętowej systemu. Na potrzeby przedstawionych badań wybrano następujące, charakteryzujące się addytywnością (tzn. spełniające równanie (1)) cechy: wartość średnia prądu, składowa stała prądu, moc średnia, pierwsze 16 harmonicznych mocy czynnej, pierwsze 16 harmonicznych mocy biernej, pierwsze 16 harmonicznych składowych rzeczywistych i urojonych widma prądu, pierwsze 16 harmonicznych konduktancji, pierwsze 16 harmonicznych susceptancji. Wszystkie parametry obliczane były na podstawie wektora 1024 próbek pozyskanych z przebiegów prądu i napięcia po wykryciu skoku w przebiegu prądowym. Zbiór uczący  $L$  został stworzony na podstawie stu wektorów dla każdego odbiornika, wliczając w to wektory opisujące brak zmiany w stanie odbiorników (wykrywanie takiego stanu jest również istotne ze względu na możliwość błędnego wykrycia zdarzenia zmiany stanu odbiornika).

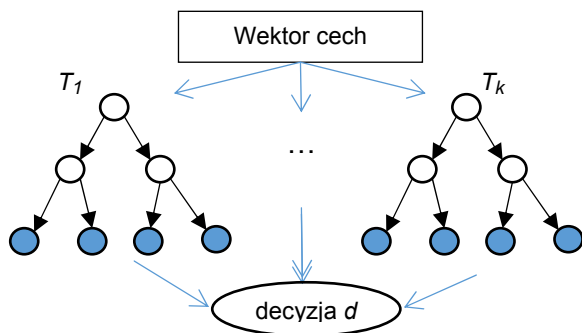
### Charakterystyka zastosowanych metod sztucznej inteligencji

W tym punkcie przedstawiono dwie metody sztucznej inteligencji zaimplementowane w celu rozwiązania opisanego wcześniej problemu. Zwrócono uwagę na cechy różniące obie metody, a także parametry, od których zależy ich skuteczność.

#### Las losowy

Jest to popularna obecnie metoda klasyfikacyjna ze względu na dowiedzioną wysoką skuteczność [4]. Las składa się ze zbioru drzew decyzyjnych (Rys. 2)

konstruowanych na podstawie zbioru uczącego  $L$ . Każde drzewo składa się z węzłów połączonych krawędziami, prowadzącymi do węzłów na niższym poziomie. Drzewo rozpoczyna się od korzenia, kończy zaś na liściach, które zawierają identyfikatory odbiorników energii elektrycznej. Pozostałe korzenie zawierają testy, tzn. wybrane wartości konkretnej cechy, z którymi porównywana jest cecha z aktualnie przetwarzanego wektora. Klasyfikacja za pomocą drzewa polega na wykonaniu szeregu porównań pomiędzy wartościami cech z wektora a wartościami znajdującymi się w poszczególnych węzłach. Tym samym wektor przechodzi ścieżkę od korzenia do jednego z liści, który stanowi decyzję klasyfikacyjną drzewa. Istotą lasu losowego jest wprowadzenie losowości do procesu konstrukcji każdego drzewa. Kandydaci do testów w węzle są szeregowani względem jakości (np. według kryterium entropijnego [5]), następnie dochodzi do losowego wyboru jednego ze zbioru najlepszych o określonej przez projektanta liczności. W ten sposób dla zbioru danych  $L$  można stworzyć dużą liczbę unikatowych drzew, które podczas podejmowania decyzji równoległe przetwarzają wektor cech. Ostateczna decyzja  $d$  zostaje podjęta na podstawie standardowego głosowania, w którym każde drzewo ma jeden głos o równej wadze.



Rys. 2. Struktura lasu losowego

#### $k$ Najbliższych Sąsiadów

Algorytm ten jest rozwinięciem metody słownikowej, w której decyzja na temat kategorii przetwarzanego wektora liczb rzeczywistych podejmowana jest na podstawie wektora ze zbioru uczącego (tu nazywanego słownikiem) znajdującego się najbliżej niego (zgodnie z wybraną metryką) [6]. W przypadku kNN wykorzystuje się  $k$  najbliższych sąsiadów, zaś decyzja podejmowana jest na podstawie głosowania [7]. W zastosowanej metodzie wykorzystano metrykę Euklidesa do obliczania odległości oraz standardową metodę głosowania, w której każdy najbliższy sąsiad ma taki sam głos. Tym samym decyzja  $d$  podejmowana jest zwykłą większością głosów, gdzie  $c_i$  to kategorie wektorów biorących udział w głosowaniu,  $a$  oznacza zbiór najbliższych sąsiadów  $x$ , analizowany jest wektor  $y$ , zaś  $o$  oznacza odległość pomiędzy wektorami ze zbioru  $L$  oraz przypadkiem badanym (musi ona być mniejsza od wartości progowej  $\theta$ ):

$$(2) \quad d = \max\{c_i(x), x \in a : o(x, y) < \theta\}$$

Parametry algorytmu określają również liczbę najbliższych sąsiadów  $k$ , która ma istotne znaczenie podczas podejmowania decyzji. Ponieważ opisywany klasyfikator nie jest poprzedzony fazą uczenia, jego wykorzystanie jest znacząco prostsze, niż lasu losowego. Z drugiej strony uniemożliwia to selekcję przez klasyfikator najbardziej istotnych cech z oryginalnego zbioru  $L$ . Z tego powodu istotne mogą być dodatkowe metody selekcji cech, które pozwalają na wyodrębnienie podzbioru najbardziej istotnego według konkretnego kryterium. W

przedstawionych badaniach wykorzystano miarę informacyjności, czyli kryterium określające, jaki jest wkład każdej cechy w możliwość rozróżnienia pomiędzy konkretnymi odbiornikami. Współczynnik informacyjności  $q$  obliczany jest na zbiorze  $L$  składającym się z  $n$  wektorów dla  $i$ -tej cechy jako:

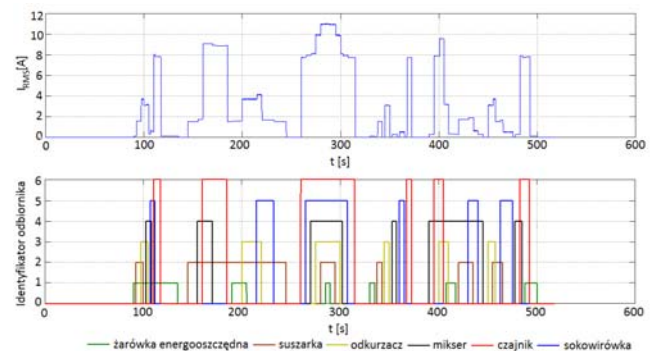
(3)

$$q(i) = \frac{\text{abs}(\bar{s}_i)}{\sqrt{\frac{\sum_{j=1}^C \frac{1}{|c(x_i)=j|} \cdot \sum_{j=1}^C b_j \cdot (|c(x_i)=j|-1)}{n-2}}}$$

gdzie  $\bar{s}_i$  jest średnią wartością  $i$ -tej cechy,  $C$  to liczba wszystkich identyfikatorów odbiorników energii elektrycznej,  $b_j$  to waga poszczególnych kategorii (tutaj przyjęto wszystkie wagi równe), zaś operator  $\|$  oznacza liczbę wektorów opisujących  $j$ -te urządzenie. Obliczone współczynniki  $q$  są sortowane malejąco, a grupa o największych wartościach jest wybierana do analizy. Zastosowanie metody pozwala znacząco zmniejszyć liczbę analizowanych cech przy zachowaniu dokładności identyfikacji. W przedstawionych badaniach klasyfikacji dokonywano na wektorze cech oryginalnym oraz zredukowanym zgodnie z równaniem (3). Zmniejszenie liczby cech pozwala zmniejszyć koszt obliczeniowy systemu, co w przypadku implementacji w systemie mikroprocesorowym może mieć znaczenie.

#### Wyniki eksperymentów

W tym punkcie przedstawiono wyniki eksperymentów polegających na identyfikacji odbiorników energii elektrycznej za pomocą obu klasyfikatorów. Pracowały one na danych pomiarowych zarejestrowanych w laboratorium. W ciągu pół godziny sześć odbiorników włączano i wyłączało w określonej kolejności (Rys. 3). Analizowane urządzenia obejmowały (w nawiasach podano ich identyfikatory): żarówkę energooszczędną (1), suszarkę (2), odkurzacz (3), mikser (4), sokowirówkę (5) oraz czajnik (6). Jednocześnie mogły pracować wszystkie odbiorniki, zaś pomiędzy kolejnymi zmianami stanów dowolnych urządzeń upływało przynajmniej pół sekundy (aby algorytm wykrywający zdarzenia miał szansę na identyfikację wszystkich). Metoda detekcji zmiany stanu wykryła 128 zdarzeń, po których uruchamiany był inteligentny algorytm klasyfikacyjny. Dzięki temu możliwe było porównanie skuteczności obu metod dla tych samych danych.



Rys. 3. Sekwencja pracy badanych odbiorników energii elektrycznej

Przykłady wyników klasyfikacji wybranych zdarzeń dla lasu losowego przedstawiono w Tab. 1 (przebieg identyfikacji dla metody kNN jest analogiczny). W poszczególnych kolumnach znajdują się numer zdarzenia, informacja  $c$  na temat identyfikatora w rzeczywistości odpowiedzialnego za zmianę konfiguracji oraz decyzja  $d$  podjęta przez las losowy.

Tabela 1 Wyniki działania lasu losowego

| No. | c | d |
|-----|---|---|
| 1   | 1 | 1 |
| 8   | 0 | 0 |
| 12  | 5 | 5 |
| 16  | 6 | 6 |
| 31  | 4 | 4 |
| 51  | 3 | 2 |
| 58  | 0 | 1 |

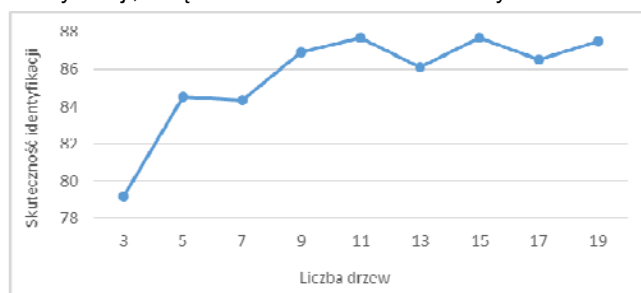
Podczas przetwarzania danych pomiarowych przez oba klasyfikatory zaobserwowano cztery zjawiska:

- Poprawne wykrycie zdarzenia zmiany stanu i identyfikacja odbiornika (zdarzenia 1, 12, 16).
- Pominięcie zdarzenia zmiany stanu odbiornika energooszczędnego, jednak poprawne wykrycie następującego po nim zdarzenia (zdarzenie 51).
- Falszywy alarm, poprawnie zidentyfikowany (pomimo uaktywnienia modułu identyfikacyjnego wskazuje on na brak zmiany stanu) – zdarzenie 8.
- Poprawne wykrycie zmiany stanu odbiornika i jego niepoprawna identyfikacja (zdarzenie 58).

W niektórych przypadkach nie powiodła się detekcja zmiany stanu czajnika, co jest symbolizowane np. przez wartość „11” (powstaje ona poprzez złożenie zdarzeń związanych z odbiornikami nr 5 i 6). Ponadto problem stanowić może identyfikacja odbiornika energooszczędnego, pobierającego stosunkowo mało prądu. Jest ono wówczas przeoczone na tle innych, bardziej energochłonnych urządzeń. Klasyfikatory dobrze sobie radzą z fałszywymi alarmami, tj. sytuacjami, w których wykryto zdarzenie, pomimo iż do takowego nie doszło.

Porównanie obu klasyfikatorów wskazuje na przewagę lasu losowego nad kNN (89.28% skuteczności przeciw 85.71%). Zaznaczyć należy, że obie metody popełniają błędy dla innych przykładów, co oznacza, że zastosowanie kombinacji klasyfikatorów może zwiększyć ich skuteczność w porównaniu z pojedynczymi algorytmami.

Z punktu widzenia klasyfikatora istotna jest optymalizacja jego parametrów, umożliwiająca maksymalizację skuteczności identyfikacji odbiorników. W przypadku lasu losowego najważniejsza jest liczba wygenerowanych drzew (Rys. 4). W opisywanym przypadku najlepszy wynik uzyskano dla 11 drzew. Większa ich liczba nie powoduje poprawy jakości identyfikacji, zwiększa za to koszt obliczeniowy.



Rys. 4. Skuteczność lasu losowego w zależności od liczby drzew.

W przypadku metody kNN istotne parametry to zbiór cech wykorzystanych do obliczenia odległości oraz liczba sąsiadów. W tym pierwszym przypadku przeprowadzono eksperymenty zarówno z oryginalnym zbiorem danych (obejmującym wszystkie cechy), jak i zredukowanym w wyniku zastosowania miary informacyjności (3). Po wyodrębnieniu trzech cech wyraźnie istotniejszych od pozostałych okazało się, że jakość identyfikacji pozostaje niezmienną. Jest tak dlatego, że ich wartości są znacząco większe od wartości pozostałych cech, zatem wkład

pozostałych do wartości odległości jest pomijalnie mały. Z kolei optymalna liczba sąsiadów została ustalona na poziomie 3 – dalsze ich zwiększanie również nie poprawiło jakości identyfikacji.

### Podsumowanie

Zastosowanie metod sztucznej inteligencji pozwoliło zautomatyzować proces identyfikacji odbiorników energii elektrycznej na podstawie cech ich sygnałów prądowo-napięciowych. Porównanie dwóch klasyfikatorów znacząco różniących się możliwościami pokazało, że w przypadku poprawnie wykrytych zdarzeń zmian konfiguracji odbiorników sam proces identyfikacji jest stosunkowo prosty, w niewielu bowiem przypadkach nie udało się zidentyfikować właściwego urządzenia.

Metoda kNN, pomimo iż znacznie prostsza od powszechnie stosowanych klasyfikatorów, odpowiednio sparometryzowana charakteryzuje się podobną skutecznością jak las losowy. Ze względu na brak modułu uczącego zalecane jest zastosowanie osobnej metody selekcji cech, która wybierze najważniejsze do obliczenia odległości. Z kolei porównanie lasu losowego oraz innych klasyfikatorów wypada zdecydowanie na korzyść tego pierwszego, za cenę większego stopnia skomplikowania algorytmu.

Główne problemy klasyfikatorów obejmowały wykrycie odbiornika energooszczędnego na tle innych urządzeń (jego cechy nikną wtedy wśród znacznie silniejszych składowych) oraz fałszywe alarmy. W tym drugim przypadku losowa wartość skoku powoduje niepotrzebne uruchomienie modułu decyzyjnego, który popełnia błąd klasyfikacyjny.

Kolejne badania obejmować powinny analizę urządzeń pracujących w wielu trybach (a przez to mogących znajdować się w wielu stanach).

*Praca wykonana w ramach projektu badawczego NCBiR PBS2/A4/8/2013.*

**Autorzy:** dr hab. inż. Piotr Bilski, Politechnika Warszawska, Instytut Radioelektroniki i Techniki Multimedialnych, ul. Nowowiejska 15/19, 00-650 Warszawa, E-mail: [pbilski@ire.pw.edu.pl](mailto:pbilski@ire.pw.edu.pl).

### LITERATURA

- [1] H.-T. Yang, H.-H. Chang, C.-L. Lin, "Design a Neural Network for Features Selection in Non-intrusive Monitoring of Industrial Electrical Loads", Proc. 11th International Conference on Computer Supported Cooperative Work in Design, (Melbourne, Australia, 26-28 April 2007), pp. 1022-1027.
- [2] J. Z. Kolter and T. Jaakkola "Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation", Proceedings of the 15th International Conference on Artificial Intelligence and Statistics AISTATS, (La Palma, Canary Islands, 2012).
- [3] S. Alshareef W.G. Morsi, "Application of wavelet-based ensemble tree classifier for non-intrusive load monitoring", Proc. IEEE Electrical Power and Energy Conference EPEC, (London, 26-28 Oct. 2015), pp. 397-401.
- [4] J. Ham, Y. Chen, M. M. Crawford, J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data", IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 3, pp. 492-501, March 2005.
- [5] H. Jin, X.-N. Wang, F. Gao, J. Li, J.-M. Wei, "Learning decision trees using confusion entropy", Proc. International Conference on Machine Learning and Cybernetics, (Tianjin, China, 14-17 July 2013), pp. 560-564.
- [6] L. Kuncheva, "Combining Pattern Classifiers. Methods and Algorithms", Wiley and Sons, New Jersey, 2004.
- [7] M.-L. Zhang, Z.-H. Zhou, "A k-nearest neighbor based algorithm for multi-label classification", IEEE International Conference on Granular Computing, Vol. 2, pp. 718 - 721, 25-27 July 2005.