

**Uzasadnienie wniosku o Nagrodę Prezesa Rady Ministrów za wyróżniającą się  
rozprawę doktorską dr. inż. Piotra Maciąga  
pt. „*Methods of sequential patterns discovery, detection of anomalies and prediction from spatio-temporal data with particular use of evolving spiking neural networks*”**

**Opis osiągnięcia kandydata do nagrody**

Dr inż. Piotr Maciąg realizował swoją rozprawę doktorską w Instytucie Informatyki Politechniki Warszawskiej pod kierunkiem prof. dr hab. inż. Marzeny Kryszkiewicz oraz pod kierunkiem dr. inż. Roberta Bembenika, jako promotora pomocniczego. W okresie studiów doktoranckich odbył 3 zagraniczne staże naukowe: 10-miesięczny na *Auckland University of Technology* (Nowa Zelandia), jednomiesięczny na *University of the Basque Country* (Hiszpania) oraz dwumiesięczny na *Carnegie Mellon University* (USA). Swoją rozprawę doktorską obronił 7 grudnia 2021. W styczniu 2022 r. Rada Naukowa Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej nadała dr. inż. Piotrowi Maciągowi stopień doktora oraz wyróżniła jego rozprawę doktorską. Od roku 2019 jest pracownikiem Zakładu Systemów Informacyjnych Instytutu Informatyki PW, gdzie początkowo był zatrudniony na stanowisku asystenta, a od 2022 roku jest zatrudniony na stanowisku adiunkta.

**Publikacje wchodzące w skład rozprawy doktorskiej**

Na rozprawę doktorską dr. inż. Piotra Maciąga składa się 8 publikacji [P1-P8], z których trzy są wyłącznie jego autorstwa, a w pozostałych pięciu współautorskich publikacjach jest pierwszym autorem:

- P1. **Piotr S. Maciąg**, Nikola Kasabov, Marzena Kryszkiewicz and Robert Bembenik, "Air pollution prediction with clustering-based ensemble of evolving spiking neural networks and a case study for London area", *Environmental Modelling & Software*, vol. 118, pp. 262-280, 2019, DOI: 10.1016/j.envsoft.2019.04.012.
- P2. **Piotr S. Maciąg**, Marzena Kryszkiewicz and Robert Bembenik, "Online Evolving Spiking Neural Networks for Incremental Air Pollution Prediction", in *Proceedings of the 2020 International Joint Conference on Neural Networks: IJCNN 2020*, pp. 1-8, DOI: 10.1109/IJCNN48605.2020.9206775.
- P3. **Piotr S. Maciąg**, Marzena Kryszkiewicz, Robert Bembenik, Jesus L. Lobo and Javier Del Ser, "Unsupervised Anomaly Detection in Stream Data with Online Evolving Spiking Neural Networks", *Neural Networks*, vol. 139, pp. 118-139. 2021, DOI: 10.1016/j.neunet.2021.02.017.
- P4. **Piotr S. Maciąg**, "A Survey on Data Mining Methods for Clustering Complex Spatiotemporal Data", in *Proceedings of 13th International Conference Beyond Databases, Architectures and Structures. Towards Efficient Solutions for Data Analysis and Knowledge Representation: BDAS 2017*, vol. 716, pp. 115-126, Springer International Publishing, DOI: 10.1007/978-3-319-58274-0\_10
- P5. **Piotr S. Maciąg**, "Efficient Discovery of Sequential Patterns from Event-Based Spatio-Temporal Data by Applying Microclustering Approach", in *Intelligent Methods and Big Data in Industrial Applications*, vol. 40, pp. 183-199, 2019, Springer International Publishing, DOI: 10.1007/978-3-319-77604-0\_14
- P6. **Piotr S. Maciąg**, "Efficient Discovery of Top-K Sequential Patterns in Event-Based

Spatio-Temporal Data”, in *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems: FedCSIS 2018*, vol. 15, pp. 47-56, [ieeexplore.ieee.org/document/8511229](http://ieeexplore.ieee.org/document/8511229)

P7. **Piotr S. Maciąg** and Robert Bembenik, ”A Novel Breadth-first Strategy Algorithm for Discovering Sequential Patterns from Spatio-temporal Data”, in *Proceedings of the 8<sup>th</sup> International Conference on Pattern Recognition Applications and Methods: ICPRAM 2019*, pp. 459–466, DOI: 10.5220/0007355804590466

P8. **Piotr S. Maciąg**, Marzena Kryszkiewicz and Robert Bembenik, ”Discovery of closed spatio-temporal sequential patterns from event data”, in *Proceedings of the 23<sup>rd</sup> International Conference Knowledge-Based and Intelligent Information & Engineering Systems: KES 2019*, pp. 707-716, DOI: 10.1016/j.procs.2019.09.226

Dwie spośród tych prac ([P1] i [P3]) są opublikowane w prestiżowych czasopiśmie: *Neural Networks* oraz *Environmental Modelling & Software* (łącznie: za 340 punktów), cztery w materiałach z konferencji międzynarodowych wykazywanych w WOS (łącznie: za 250 punktów, w tym 1 publikacja konferencyjna za 140 punktów), a pozostałe dwie w materiałach z innych konferencji międzynarodowych (za 25 punktów). Praca [P2] została opublikowana w materiałach konferencyjnych prestiżowego kongresu *IEEE World Congress on Computational Intelligence*. Łącznie liczba punktów za publikacje stanowiące rozprawę wynosi 615, a sumaryczny IF za nie wynosi 10.34.

### Znaczenie rozprawy

Rozprawa doktorska dr. inż. Piotra Maciąga wnosi istotny wkład naukowy w dziedzinę eksploracji danych w zakresie:

1. nowych metod uczenia i wykorzystania **impulsowych sieci neuronowych** (ang. Spiking Neural Networks) [P1-P3],
2. nowych metod odkrywania wiedzy z danych, a w szczególności nowych metod odkrywania **przestrzenno-czasowych wzorców sekwencyjnych** [P4-P8].

Obydwa te nurty prac badawczych, zrealizowanych w ramach doktoratu przez dr. inż. Piotra Maciąga, należą do intensywnie zgłębianych przez międzynarodowe środowisko naukowo-badawcze.

W ramach pierwszego nurtu badawczego, dr inż. Piotr Maciąg opracował szereg nowych **metod** i **algorytmów** uczenia impulsowych sieci neuronowych (a w szczególności istotnego rozszerzenia takich sieci zwanego *ewoluującymi impulsowymi sieciami neuronowymi*). Wkład naukowy Autora rozprawy w tym nurcie obejmuje:

- Opracowanie nowej metody *CEeSNN (Clustering-based Ensemble of evolving Spiking Neural Networks)* predykcji wartości szeregów czasowych za pomocą zespołu ewoluujących impulsowych sieci neuronowych.
- Opracowanie nowego algorytmu *OeSNN-UAD (Online evolving Spiking Neural Networks for Unsupervised Anomaly Detection)*, umożliwiającego skuteczną i wydajną nienadzorowaną detekcję anomalii w strumieniach danych z wykorzystaniem ewoluujących impulsowych sieci neuronowych.
- Opracowanie nowego algorytmu *OeSNN-IP (Online evolving Spiking Neural Networks for Incremental Prediction)* do predykcji wartości strumieni danych z wykorzystaniem ewoluujących impulsowych sieci neuronowych.

W ramach drugiego z nurtów badawczych dr inż. Piotr Maciąg opracował szereg nowych, wydajnych algorytmów odkrywania przestrzenno-czasowych wzorców sekwencyjnych, a także struktur danych wykorzystywanych przez te algorytmy. Zaproponowane algorytmy stanowią

rozszerzenie lub znaczące usprawnienie algorytmu STS-Miner<sup>1</sup> autorstwa Y. Huanga, L. Zhanga i P. Zhanga. Do osiągnięć badawczych uzyskanych w drugim nurcie przez dr. inż. Piotra Maciąga należy opracowanie:

- Literaturowego przeglądu typów danych przestrzenno-czasowych oraz metod grupowania takich danych.
- Algorytmu *Micro-ST-Miner* efektywnie odkrywającego przestrzenno-czasowe wzorce sekwencyjne poprzez zastosowanie *mikrogrupowania danych*.
- Algorytmu *STBFM* wykorzystującego, w odróżnieniu od algorytmu STS-Miner, strategię generowania *wszerz* wzorców oraz, opracowaną przez dr. inż. Piotra Maciąga, strukturę *drzewa prefikсового SP-Tree*. W eksperymentach przeprowadzonych przez Autora rozprawy, opracowany algorytm *STBFM* okazał się znacząco bardziej wydajny od algorytmu STS-Miner.
- Algorytmów *Top-K-STMiner* oraz *Top-K-STBFM* odkrywających zbiór K najbardziej znaczących wzorców. *Top-K-STMiner* odkrywa najbardziej znaczące wzorce ze względu na wartość indeksu sekwencji, natomiast *Top-K-STBFM* odkrywa najbardziej znaczące wzorce ze względu na wartość indeksu uczestnictwa.
- Algorytmu *CST-SPMiner* oraz *drzewa prefikсового CSP-Tree*, które umożliwiają wydajne odkrywanie zamkniętych przestrzenno-czasowych wzorców sekwencyjnych, stanowiących zwięzłą reprezentację wszystkich przestrzenno-czasowych wzorców sekwencyjnych.

### **Innowacyjny charakter otrzymanych wyników**

Do innowacyjnego i oryginalnego charakteru otrzymanych wyników należy opracowanie przez Autora rozprawy nowych, nietrywialnych metod uczenia impulsowych sieci neuronowych, wydajnych czasowo i pamięciowo. Opracowane metody znajdują zastosowanie w predykcji wartości w strumieniach danych oraz nienadzorowanej detekcji anomalii. Do istotnego wkładu naukowego Autora rozprawy w zakresie *właściwości ewoluujących impulsowych sieci neuronowych* należy również:

- Wprowadzenie nowych rodzajów neuronów wyjściowych o wartościach rzeczywistych ustalanych adaptacyjnie na podstawie okna strumienia danych, a wykorzystywanych do predykcji (sieć *OeSNN-IP*) bądź wykrywania anomalii (sieć *OeSNN-UAD*).
- Sformułowanie ograniczenia górnego na odległość euklidesową pomiędzy wektorem wag synaps każdego z aktualnych neuronów wyjściowych sieci *OeSNN-IP* a wektorem wag nowego kandydata na neuron wyjściowy sieci *OeSNN-IP*. Znalezione ograniczenie znacząco upraszcza dobór wartości progowej determinującej sposób wykorzystania neuronu kandydującego do aktualizacji sieci.

Do najbardziej innowacyjnych osiągnięć dr. inż. Piotra Maciąga w nurcie poświęconym odkrywaniu przestrzenno-czasowych wzorców sekwencyjnych należy zaliczyć:

- Opracowanie algorytmu *Micro-ST-Miner* do wydajnego odkrywania takich wzorców przy użyciu jedynie zaproponowanego przez Autora rozprawy *indeksu mikrogrup* (którego sposób tworzenia także został zaproponowany przez Autora).
- Opracowanie *drzewa prefikсового SP-Tree* i jego wariantu *CSP-Tree* do przechowywania już odkrytych przestrzenno-czasowych wzorców sekwencyjnych, umożliwiających sprawne tworzenie nowych kandydatów na kolejne takie wzorce. Zaproponowane drzewo *SP-Tree* znalazło zastosowanie przy odkrywaniu wszystkich znaczących przestrzenno-czasowych wzorców sekwencyjnych (algorytm *STBFM*) oraz K najbardziej znaczących przestrzenno-czasowych

---

<sup>1</sup> Huang, Y., Zhang, L., Zhang, P.: A framework for mining sequential patterns from spatio-temporal event data sets. IEEE Trans. Knowl. Data Eng. 20(4), 433–448 (2008), DOI: 10.1109/TKDE.2007.190712

wzorców sekwencyjnych (algorytm *TOP-K-STBFM*). Natomiast drzewo *CSP-Tree* zostało zaproponowane i użyte do odkrywania wszystkich znaczących zamkniętych przestrzenno-czasowych wzorców sekwencyjnych.

### **Możliwości zastosowania uzyskanych wyników**

Niezwykle wartościowym aspektem wyników uzyskanych przez dr. inż. Piotra Maciąga i zaprezentowanych w jego rozprawie doktorskiej są ich zastosowania praktyczne. Opracowane rozwiązania zostały wykorzystane do:

- Przewidywania wartości zanieczyszczenia powietrza dla rzeczywistych danych dotyczących Londynu i Warszawy (również w przypadku danych strumieniowych), takich jak dane o niezwykle szkodliwych dla zdrowia i życia ludzi pyłach zawieszonych PM<sub>2.5</sub>, PM<sub>10</sub> oraz ozonie.
- Nienadzorowanej detekcji anomalii w strumieniach danych za pomocą ewoluujących impulsowych sieci neuronowych. Rozwiązanie to zostało przez Autora rozprawy zweryfikowane na ponad czterystu rzeczywistych zbiorach danych (m.in. z takich źródeł jak czujniki ruchu na autostradach, liczba kliknięć w reklamy w serwisie Twitter, wykorzystanie procesorów komputera).
- Odkrywania przestrzenno-czasowych wzorców sekwencyjnych w danych o przestępczości w Bostonie.

Kody źródłowe rozwiązań opracowanych przez dr. inż. Piotra Maciąga zostały przez niego udostępnione publicznie w repozytorium GitHub (<https://github.com/piotrMaciag32?tab=repositories>) i są wykorzystywane przez badaczy z różnych ośrodków naukowych i krajów (w tym m.in. Hiszpanii, Niemiec, Iranu, Malezji). W szczególności w publikacji D. Bäßlera, T. Kortusa, i G. Gühringa z Hochschule Esslingen, zatytułowanej *Unsupervised anomaly detection in multivariate time series with online evolving spiking neural networks*<sup>2</sup>, zaproponowano metodę OeSNN-D, którą autorzy przedstawiają jako bezpośrednie rozszerzenie opracowanej przez dr. inż. Piotra Maciąga metody *OeSNN-UAD* detekcji anomalii w strumieniach danych.

### **Prace po uzyskaniu stopnia doktora tematycznie powiązane z rozprawą**

Warto odnotować, że po uzyskaniu stopnia doktora, dr inż. Piotr Maciąg kontynuował badania w nurtach powiązanych z tematyką swojej rozprawy doktorskiej, obejmujących zarówno zagadnienia dotyczące ewoluujących impulsowych sieci neuronowych, jak i przestrzenno-czasowych wzorców sekwencyjnych. Nowe wyniki zaprezentował w pracach:

- **Piotr S. Maciąg**, Wojciech Sitek, Łukasz Skonieczny and Henryk Rybiński, „A Comparative Study of Short Text Classification with Spiking Neural Networks”, *Proceedings of the 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS)*, DOI: 10.15439/2022F184.
- **Piotr S. Maciąg**, Robert Bembenik, Artur Dubrawski, “Discovery of Constricted Spatio-Temporal Sequential Patterns” (zgłoszone do *Journal of Big Data*; w recenzji).

---

<sup>2</sup> Bäßler, D., Kortus, T. & Gühring, G. Unsupervised anomaly detection in multivariate time series with online evolving spiking neural networks. *Mach Learn* 111, 1377–1408 (2022). <https://doi.org/10.1007/s10994-022-06129-4>