Modelling and Optimization of Toll Stations on a Highway by Using Nonstationary Poisson Process

Kamil Książek
Institute of Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland
KamilKsiazek95@gmail.com

***Abstract*—The Poisson process is used to simulate streams of many independent real events. The nonstationary (nonhomogeneous) Poisson process (NPP) is an enlargement of the basic approach and the main difference lies in a various intensity of events at different times. In the article this tool is presented as a method for modelling and optimization of toll stations on a highway. The car traffic intensity on highways varies, depending on the time of the day. Therefore it is reasonable to apply NPP for its simulation. Results indicate that this approach is promising and can be helpful in determining the most efficient setting for the gates.**

***Index Terms*—nonstationary Poisson process, simulation, modelling, toll station, stochastic process**

I. Introduction

Nowadays computers are extremely useful for humans. They help entrepreneurs in minimizing costs and optimizing production processes. Moreover, they increase the comfort of everyday life. There is a lot of mathematical theories which can be effectively applied in informatics. One of them is the stochastic processes theory, and more precisely the Poisson processes. Many independent streams of events may be approximated with their use which opens the door to their widespread deployment. In this paper, modelling of the toll station on the highway is presented. The simulation is created to help to choose the right number of open toll gates during specific hours of a day. In the article, basic concepts and definitions related to the Poisson processes are presented. These are followed by detailed information about statistical data, the way of method implementation, as well as by description of functions used. Finally, results and conclusions are discussed.

A. Related works

The Poisson process is a subject of many scientific studies. In [8] an improved version of expectation maximization (EM) algorithm designed for nonstationary Poisson processes is presented. In [3] NPP is used for prediction of events like arrivals of patients at Accident and Emergency departments. The interesting research about detecting an anomaly in datasets is inserted in [7]. Authors describe e.g. study about detection of epilepsy. Moreover, this approach can be used in various topics connected with the environment, climate changes (based on such parameter as sea level pressure) [1] and earthquakes [5]. Furthermore, Poisson processes can be very helpful in issues related to wireless networks. For instance, [12] presents a model for sensor deployment and a method based on a division of hexagons in a clustered Wireless Sensor Network. Additionally, authors in [6] show a model of Simple Mail Transfer Protocol session with an application of the Poisson process. An interesting application of NPP in searching parking spaces is described in [9]. This solution is cost-effective and can be helpful for car drivers. Described papers confirm that nonstationary Poisson process may be widely used. The Poisson processes are also connected with queue theory. Some studies about positioning finite-buffer queuing system with an cost-optimization problem are presented in [13].

II. The Poisson Process

Let $X_1, X_2, ..., X_n$ be a sequence of positive independent random variables about the same distribution. Then X_k is the time between $(k-1)$-th and k-th event. Let $N(t)$ be a random variable for fixed t:

$$N(t) = max\{n \geq 0 : \sum_{i=1}^{n} X_i \leq t\}. \tag{1}$$

$N(t)$ is called a counting process [11]. In other words, $N(t)$ gives the number of events appeared until time t. An example trajectory of a counting process is shown in Fig. 1.

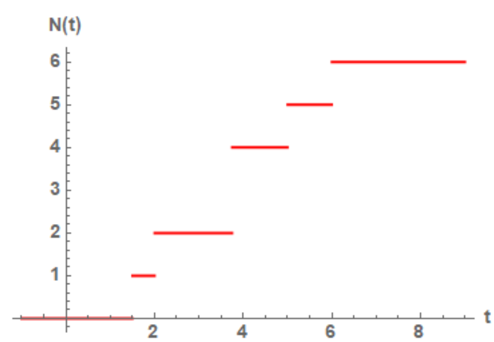


Fig. 1: Chart of a trajectory of a counting process. The horizontal axis represents time, the vertical axis informs about the number of events $\{N(t), t \geq 0\}$.

$N(t)$ is, by definition, a non-negative integer number which fulfils following conditions:

- $t_1 < t_2 \longrightarrow N(t_1) \leq N(t_2)$,
- $N(t_2) - N(t_1)$ is the number of jumps which appeared in the interval $[t_1, t_2], t_1 < t_2$.

Increments are called *independent* if the numbers of jumps (events) of the process in disjoint intervals of time are independent random variables. *Stationarity* of increments consists in the fact that the number of jumps of the process in a given interval is depending only on the length of the interval.

Definition 2.1. A counting process $\{N(t), t \geq 0\}$ such that random variables $X_1, X_2, X_3, ...$ have the same exponential distributions with the rate parameter $\lambda > 0$ is called *the Poisson process*, where λ is the rate of the process.

It can be also proved, that for each $t \geq 0, N(t)$ (i.e. the number of events of the Poisson process having rate λ until time t) has the Poisson distribution with the mean λt:

$$\mathbf{P}\{N(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \tag{2}$$

Assume that $\{N_s(t), t \geq 0\}$ is the superposition of independent counting processes $\{N_{si}(t), t \geq 0\}, i \in \{1, ..., n\}$ which rarely generate the events:

$$N_s(t) = N_{s1}(t) + N_{s2}(t) + ... + N_{sn}(t), \ t \geq 0. \tag{3}$$

By Palm-Khintchine Theorem [4], $N_s(t)$ can be approximated by the Poisson process.

A. Nonstationary Poisson process

Definition 2.2. [11] A counting process $\{N(t), t \geq 0\}$ is called a nonstationary (nonhomogeneous) Poisson process with intensity function $\lambda(t)$, if four following criteria are met:

- $N(0) = 0$; the process in time t (at the beginning) has 0 events.
- The process $\{N(t), t \geq 0\}$ has independent increments.
- $\mathbf{P}\{N(t + \Delta t) - N(t) = 1\} = \lambda(t) \cdot \Delta t + o(\Delta t)$, $\Delta(t) \to 0$ (short time interval).
- $\mathbf{P}\{N(t + \Delta t) - N(t) \geq 2\} = o(\Delta t), \Delta(t) \to 0$.

By definition, $o(\Delta t)$ is such a function that

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0. \tag{4}$$

And finally, it is necessary to describe following theorem:

Theorem 2.1. For each $m, n \geq 0$

$$\mathbf{P}\{N(n+m) - N(n) = l\} = \\ = e^{-(M(n+m)-M(n))} \cdot \frac{(M(n+m) - M(n))^l}{l!}, \tag{5}$$

where $\mathbf{P}\{N(n+m) - N(n) = l\}$ means that in the interval $[n, n+m]$ occurs exactly l events, and $M(n) = \int_0^n \lambda(y)dy$.

III. Modelling of the toll station

A. Description of the task

The main goal of this paper was modelling of the toll station on a highway. During experiments a nonstationary Poisson process was used because events (entering of a single car) occur with varying intensity at different time intervals. Many independent processes which rarely generate events can be approximated by the Poisson process (by Palm-Khintchine Theorem). The model is based on real data from Balice toll station which is located on the highway A4 in Poland (Kraków-Balice). Let $N_1(t), N_2(t)$ be two nonstationary Poisson processes. $N_1(t)$ is to count the cars queuing at toll gates and $N_2(t)$ is responsible for the number of served customers at the toll station (gates throughput). Let $X(t)$ be a random variable expressing the number of cars waiting in the queue at time t:

$$X(t) = max\{N_1(t) - N_2(t), 0\}. \tag{6}$$

Of course, $X(t)$ could not be smaller than 0, and in addition, the number of serviced clients could not be greater than the total number of cars which entered on the highway.

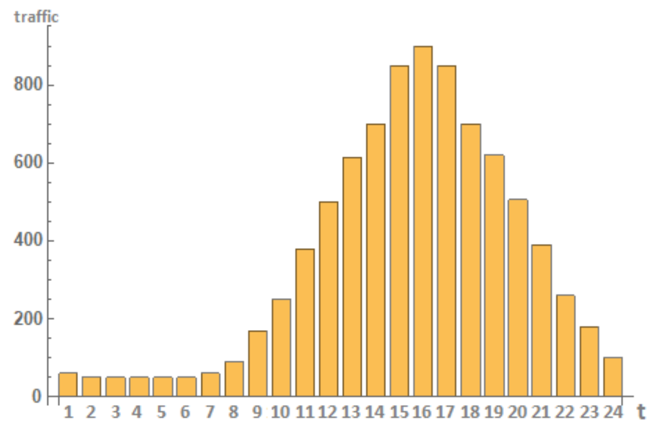


Fig. 2: A bar chart presents traffic in the toll station in Kraków-Balice during Fridays based on Google data. The horizontal axis represents following hours (i.e. first hour means 0:00-1:00, second hour: 1:00-2:00, etc.) and the vertical axis informs about a number of customers during subsequent hours of the day.

B. Implementation

The Algorithm 1 presents the pseudocode of the nonstationary Poisson process dedicated to the described problem. The Exponential Distribution is essential. Its rate parameter depends on time (traffic on a given hour). Because only one day is examined, the process could not have events after 24 hours. $N_1(t)$ function has the following form:

$$N_1(t) = \begin{cases} 0 & 0 \leq t < T_1 \\ 1 & T_1 \leq t < T_2 \\ 2 & T_2 \leq t < T_3 \\ \dots & \\ n-1 & T_{n-1} \leq t < T_n \\ n & T_n \leq t < \infty \end{cases} \tag{7}$$

n is the total number of events and T_i is the time of the i-th event. N_1 is responsible for arrivals of cars in the toll station. The rate parameter of Exponential Distribution in the intensity function during the fixed hour is the same as the number of cars on the highway (Fig. 2). Of course, it is necessary to

Algorithm 1 Pseudocode of the nonstationary Poisson process (responsible for entering cars on the toll station)

Input: range of time: $time$, intensity function: $\lambda(t)$, $t \in [0, time]$.

Generate a pseudorandom value $\mathbf{T}$ from the Exponential Distribution with the rate parameter $\lambda(0)$.
Set $k := 1$.
Create a blank list (called L) intended for the times in which occur events of the process.
while $\mathbf{T} <\ time\ \&\ \mathbf{T} < 24$ **do**
  Append to L the time of the last event.
  Generate a pseudorandom value $\mathbf{T_k}$ from the Exponential Distribution with the rate parameter $\lambda(T)$.
  Set $\mathbf{T} := \mathbf{T} + \mathbf{T_k}$.
  Set $k := k + 1$.
end while
Show the time of consecutive events (L), total number of events (k) and the pattern for the trajectory (N_1).

create the second nonstationary Poisson process (N_2) in charge of customer service at the toll gates. Its principle of operation is almost identical to the first process (N_1) but as mentioned before, the number of served clients cannot be greater than the number of cars on the highway. It is realized by following commands:

- Generate k-th event of 2nd process.
- *if* $(N_2(t) < N_1(t))$
  then append $k-$th event to L_2.

$N_1(t), N_2(t)$ are respectively the numbers of events of the first/second process until time t. L_2 is the list of events of N_2. This time the intensity function is connected to the number of open gates. The remaining part of the pseudocode is identical as in Algorithm 1. It means, that randomly generated jump in N_2 is accepted only if there is a customer (or customers) awaiting in the queue at gates. Now it should be obvious that Eq. 6 notifies about the number of cars in the queue at time t. It is assumed that the maximum number of gates is equal to 8. One client is served in about 30 seconds which corresponds to $\approx$ 0.008h. Similar value was experimentally achieved as a mean value of a random sample from the Exponential Distribution with the rate parameter $\lambda = 120$. This means that the intensity function in $N_2(t)$ is following:

$$\lambda_t^{TC} = \begin{cases} 120 & \text{if 1 toll gate is open at time t} \\ 240 & \text{if 2 toll gates are open at time t} \\ 360 & \text{if 3 toll gates are open at time t} \\ \dots & \\ 960 & \text{if 8 toll gates are open at time t} \end{cases} \tag{8}$$

IV. Results

A. Case 1

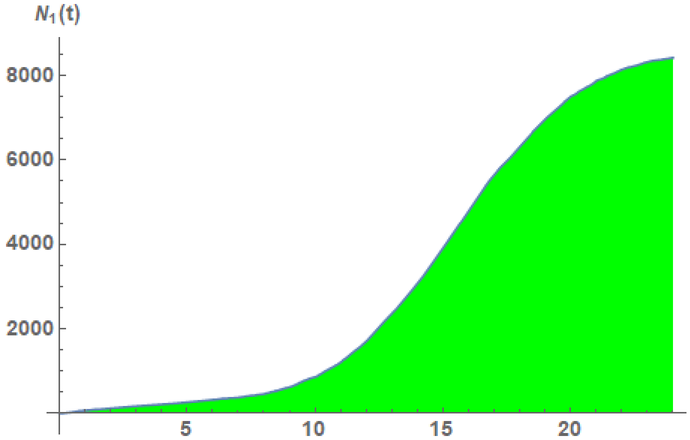
$$\lambda_{N_2}^1(t) = \begin{cases} 120 & 0 < t < 9 \ \ \text{(1 toll gate),} \\ 360 & 9 \le t < 10 \ \ \text{(3 toll gates),} \\ 480 & 10 \le t < 11 \ \ \text{(4 toll gates),} \\ 600 & 11 \le t < 12 \ \ \text{(5 toll gates),} \\ 720 & 12 \le t < 14 \ \ \text{(6 toll gates),} \\ 960 & 14 \le t < 17 \ \ \text{(8 toll gates),} \\ 720 & 17 \le t < 19 \ \ \text{(6 toll gates),} \\ 600 & 19 \le t < 20 \ \ \text{(5 toll gates),} \\ 480 & 20 \le t < 21 \ \ \text{(4 toll gates),} \\ 360 & 21 \le t < 22 \ \ \text{(3 toll gates),} \\ 240 & 22 \le t < 23 \ \ \text{(2 toll gates),} \\ 120 & 23 \le t < 24 \ \ \text{(1 toll gate)} \end{cases} \tag{9}$$

The first experiment was carried out by using the intensity function $\lambda_{N_2}^1$ (Eq. 9). Fig. 3 shows exemplary trajectories for two nonstationary Poisson processes: N_1 and N_2. A trajectory grows rapidly if the number of cars dramatically rises (like in afternoon hours). The choice of a configuration of active stands is the better, the more N_2 is similar to N_1. Fig. 5a gives information about values of $X(t)$ during subsequent hours based on difference between N_1 and N_2 in a given time. It can be concluded that toll gates opening schedule $\lambda_{N_2}^1$ is not efficient and generates a traffic congestion between 8-9 a.m.. A sixty-vehicles-long queue during typical traffic hours is a situation which should not take place. Statistical data presented in Fig. 2 shows that the number of cars increases sharply from 9 a.m. but adopted schedule is not sufficient. In case of other hours, a situation is better (one should remember that is necessary to divide the number of cars by the number of open toll gates) but still not perfect. For instance, about 6. p.m. five cars waited for payment on each available lane. It is possible to improve the throughput by increasing the number of active stands in some hours.

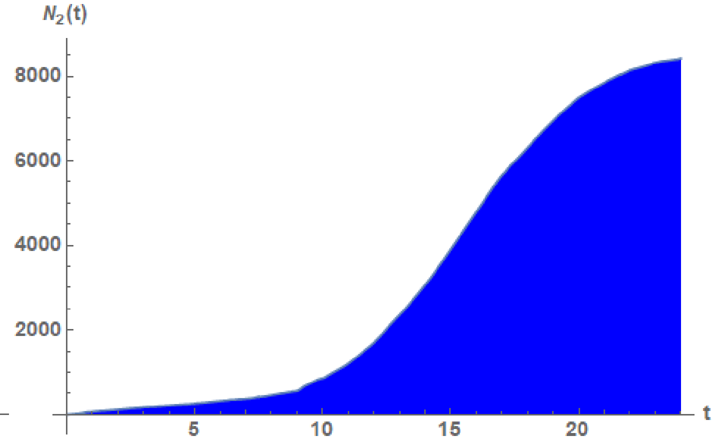
B. Case 2

$$\lambda_{N_2}^2(t) = \begin{cases} 120 & 0 < t < 8 \ \ \text{(1 toll gate),} \\ 240 & 8 \le t < 9 \ \ \text{(2 toll gates),} \\ 360 & 9 \le t < 10 \ \ \text{(3 toll gates),} \\ 480 & 10 \le t < 11 \ \ \text{(4 toll gates),} \\ 600 & 11 \le t < 12 \ \ \text{(5 toll gates),} \\ 720 & 12 \le t < 13 \ \ \text{(6 toll gates),} \\ 960 & 13 \le t < 19 \ \ \text{(8 toll gates),} \\ 600 & 19 \le t < 20 \ \ \text{(5 toll gates),} \\ 480 & 20 \le t < 21 \ \ \text{(4 toll gates),} \\ 360 & 21 \le t < 22 \ \ \text{(3 toll gates),} \\ 240 & 22 \le t < 23 \ \ \text{(2 toll gates),} \\ 120 & 23 \le t < 24 \ \ \text{(1 toll gate)} \end{cases} \tag{10}$$

The second experiment was connected with the intensity function $\lambda_{N_2}^2$ (Eq. 10). Similarly as before, Fig. 4 shows exemplary trajectories for such setting of intensity function. A chart presenting values of $X(t)$ calculated for trajectories

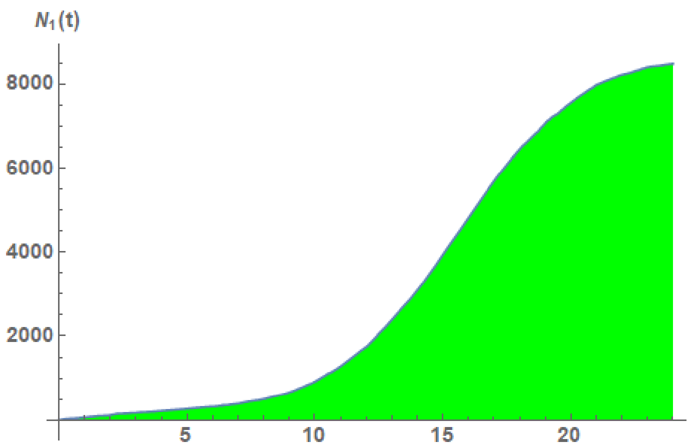


(a) A trajectory of $N_1(t)$

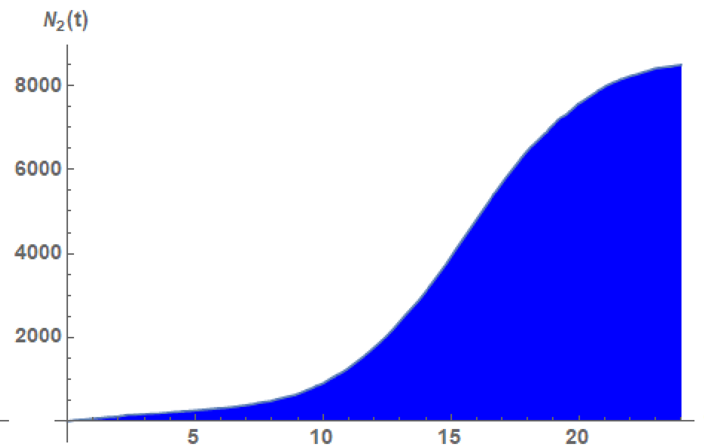


(b) A trajectory of $N_2(t)$

Fig. 3: A sample experiment by using road traffic intensity from Fig. 2 and intensity function $\lambda^1_{N_2}(t)$ (Eq. 9)

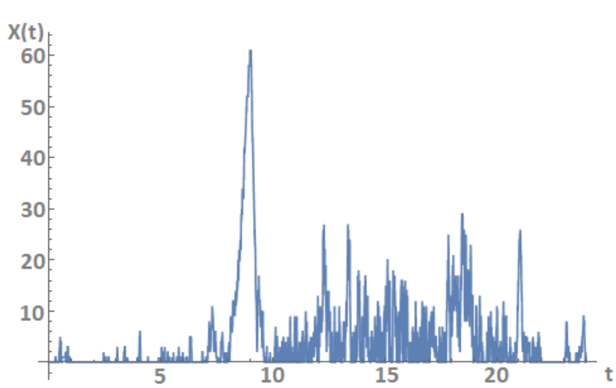


(a) A trajectory of $N_1(t)$

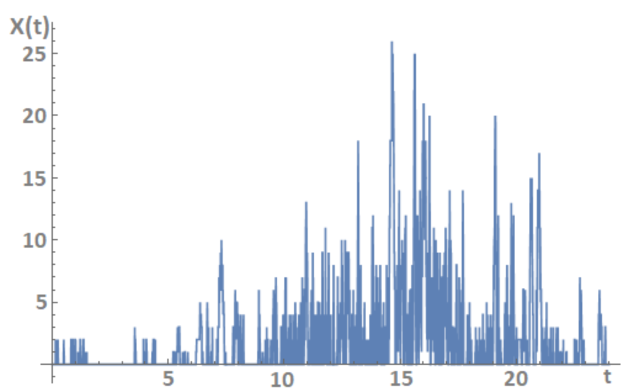


(b) A trajectory of $N_2(t)$

Fig. 4: A sample experiment by using road traffic intensity from Fig. 2 and intensity function $\lambda^2_{N_2}(t)$ (Eq. 10)



(a) $X(t)$ based on trajectories from Fig. 3



(b) $X(t)$ based on trajectories from Fig. 4

Fig. 5: Above functions present how many cars waited for payment in the queue.

from Fig. 4 is exposed in Fig. 5b. In this case, results are satisfactory – the maximum value of $X(t)$ is equal to 26. It means that at most 26 customers (before 3 p.m.) waited in the queue. In this time 8 toll gates were open so this state corresponds to ≈ 3 cars per one stand. Furthermore, opening of the ninth gate is not possible (see 8). It can be concluded that this schedule is right choice for statistical data inserted in Fig. 2.

V. Conclusions and final remarks

Application of nonstationary Poisson process can be helpful in determining the right schedule in toll stations. Obviously, an opening of each new stand is connected with additional costs so one should choose the right setting carefully. It is easier thanks to such branches of mathematics as the theory of stochastic processes. Moreover, it is necessary to adjust the number of open gates with the season (e.g. weekends, vacation). Furthermore, one should consider situations in which traffic is bigger than usual and adjust the number of open gates accordingly.

However, it should be stressed that the described model could be applied in toll stations. It is necessary to collect statistical data connected with each day of week and carry out simulations for all cases. Then, after allowing for possible deviations (season and other factors influence on traffic) one can determine the optimal numbers of open gates during subsequent hours/days based on data coming from trajectories of the Poisson process. It is worth to see that wrong choice of open stands can bring loss of money. Car drivers which waited in the queue too long, can find alternative roads in the future.

During further research one can compare the described method with the other optimization techniques like heuristic algorithms or neural networks and analyze their efficiency.
Measurements and charts were performed by using Wolfram Mathematica 11 software.

References

[1] S. Caires, V. R. Swail, and X. L. Wang. Projection and analysis of extreme wave climate. *Journal of Climate*, 19(21):5581–5605, 2006.

[2] G. Capizzi, G. L. Sciuto, C. Napoli, E. Tramontana, and M. Woźniak. Automatic classification of fruit defects based on co-occurrence matrix and neural networks. In *Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 861–867. IEEE, 2015.

[3] J. Goulding, S. Preston, and G. Smith. Event series prediction via non-homogeneous poisson process modelling. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 161–170, Dec 2016.

[4] D. Heyman and M. Sobel. Stochastic models in operations research, stochastic processes and operating characteristics, vol. 1, 1982.

[5] L.-L. Hong and S.-W. Guo. Nonstationary poisson model for earthquake occurrences. *Bulletin of the Seismological Society of America*, 85(3):814–824, 1995.

[6] Y. Lee and J. S. Kim. Characterization of large-scale smtp traffic: the coexistence of the poisson process and self-similarity. In *2008 IEEE International Symposium on Modeling, Analysis and Simulation of Computers and Telecommunication Systems*, pages 1–10, Sept 2008.

[7] S. Luca, P. Karsmakers, and B. Vanrumste. Anomaly detection using the poisson process limit for extremes. In *2014 IEEE International Conference on Data Mining*, pages 370–379, Dec 2014.

[8] V. Nagaraju, L. Fiondella, P. Zeephongsekul, C. L. Jayasinghe, and T. Wandji. Performance optimized expectation conditional maximization algorithms for nonhomogeneous poisson process software reliability models. *IEEE Transactions on Reliability*, 66(3):722–734, Sept 2017.

[9] L. Peng and H. Li. Searching parking spaces in urban environments based on non-stationary poisson process analysis. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1951–1956, Nov 2016.

[10] D. Połap, M. Woźniak, C. Napoli, E. Tramontana, and R. Damaševičius. Is the colony of ants able to recognize graphic objects? In *International Conference on Information and Software Technologies*, pages 376–387. Springer, 2015.

[11] S. M. Ross. *Introduction to probability models*. Academic press, 2014.

[12] D. Wang. Cluster subdivision towards power savings for randomly deployed wsns - an analysis using 2-d spatial poisson process. In *2016 7th International Conference on Information and Communication Systems (ICICS)*, pages 156–161, April 2016.

[13] M. Woźniak, W. M. Kempa, M. Gabryel, and R. K. Nowicki. A finite-buffer queue with a single vacation policy: An analytical study with evolutionary positioning. *International Journal of Applied Mathematics and Computer Science*, 24(4):887–900, 2014.

[14] M. Wozniak, C. Napoli, E. Tramontana, G. Capizzi, G. L. Sciuto, R. K. Nowicki, and J. T. Starczewski. A multiscale image compressor with rbfnn and discrete wavelet decomposition. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE, 2015.

[15] M. Woźniak, D. Połap, C. Napoli, and E. Tramontana. Graphic object feature extraction system based on cuckoo search algorithm. *Expert Systems with Applications*, 66:20–31, 2016.