

# A HU-GO elektronikus útdíjrendszerből származó adatok felhasználási lehetősége automatikus incidensdetektáló algoritmushoz

Az úthasználati jogosultság meglétének ellenőrzése során, amelyet a Nemzeti Útdíjfizetési Szolgáltató Zrt. (NÚSZ) végez, nagy mennyiségű forgalmi adat keletkezik. Ezen adatok forgalmi menedzsmenthez, illetve automatikus incidensdetektáló (AID) algoritmusokhoz is felhasználhatók. Először egy ARIMA-modell alapú algoritmus kerül bemutatásra, majd ezt követi a hangolási módszertan ismertetése és a valós adatokkal tesztelt, különböző változatok hatékonyságának vizsgálata.

DOI: <https://doi.org/10.24228/KTSZ.2022.6.2>

---

**Gressai Mánuel – Tóth Róbert Péter – Dr. Tettamanti Tamás**

Budapesti Műszaki és Gazdaságtudományi Egyetem

Közlekedésmérnöki és Járműmérnöki Kar

e-mail: tettamanti.tamas@kjk.bme.hu, gressai.manuel@kjk.bme.hu, tothrobertpeter@edu.bme.hu

---

## 1. BEVEZETÉS

2020-ban a Magyar Közlekedési Mérnökképzésért Alapítvány a Nemzeti Útdíjfizetési Szolgáltató Zrt. (NÚSZ) támogatásával ösztöndíjpályázatot hirdetett a Budapesti Műszaki és Gazdaságtudományi Egyetem Közlekedésmérnöki és Járműmérnöki Karán. A NÚSZ a szakmai érdekkörébe tartozó, útdíjfizetéssel összefüggő kutatási tevékenységet és oktatói konzultációt kívánta támogatni általuk meghatározott, specializált tématerületekhez kapcsolódóan. Korábbi kutatásunk folytatásaként, a HU-GO útdíjrendszerből származó adatok automatikus incidensdetektáló algoritmus-hoz történő felhasználási lehetőségével és a megfelelő működéshez szükséges paraméterek hangolásának módszertanával foglalkoztunk.

Általános megfogalmazásban forgalmi incidensnek nevezünk minden olyan váratlan eseményt (legtöbbször baleseteket, műszaki hibás járművek megállását), amely a forgalom zavartalan áramlását akadályozza. Ezen esetek észlelésére automatikus incidensdetektáló (AID – Automatic Incident Detection) algoritmusok alkalmazhatók, amelyek az esemény hatására kialakuló forgalmi állapotok egyedi tulajdonságait használják ki.

Az incidensek detektálása eszközigényük alapján csoportosítható. A legtöbb algoritmus két, egymástól meghatározott távolságban elhelyezett detektáló eszköz adatainak felhasználásával érzékeli a forgalmi viszonyokat. Ezek előnye, hogy a több keresztmetszeten végzett mérésekből adódóan kevésbé hajlamosak ha-

mis riasztást adni, viszont telepítési és üzemeltetési költségük jelentős, valamint a meghibásodás valószínűsége is nagyobb. Bizonyos algoritmusok egy érzékelési pont adatai alapján is alkalmasak az incidensek felismerésére, azonban ezek pontatlanabb eredményt adnak. Az algoritmusok működési elvük alapján a következőképpen csoportosíthatók:

- összehasonlító vagy mintafelismerő,
- statisztikai vagy idősor alapú,
- forgalmi modell alapú,
- mesterséges intelligenciát alkalmazó,
- videófelismerő rendszerek.

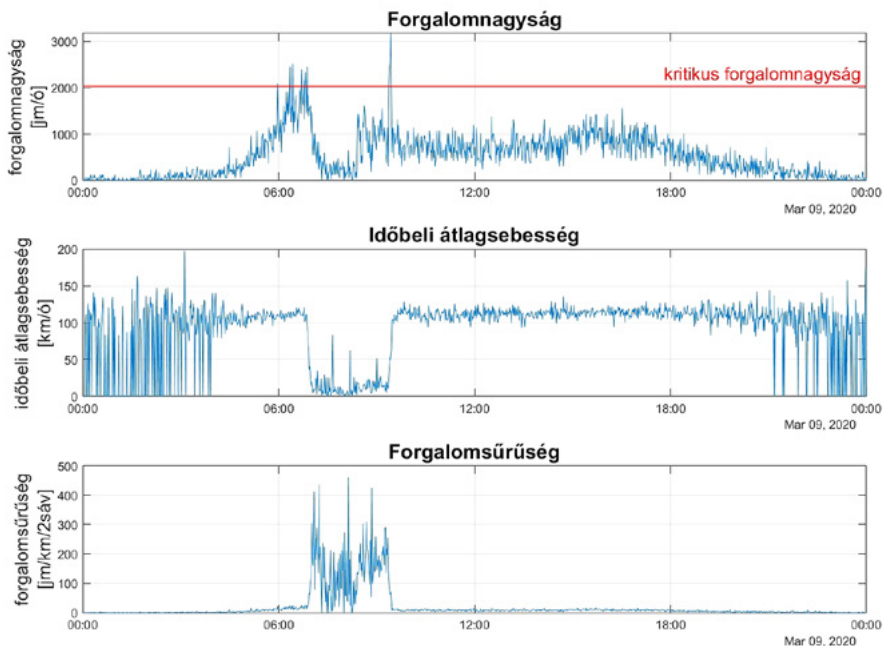
## 2. ARIMA-MODELL ALAPÚ ALGORITMUS

Az ARIMA-modell (Autoregressive Integrated Moving Average) egy statisztikai és idősor alapú incidensfelismerő módszer, amely egyetlen mérési pont - jelen esetben egy NÜSZ ellenőrző portál - adatai alapján is képes a váratlan események detektálására. Az eljárás megbízhatósága és pontossága a modellparaméterek hangolásával tovább nö-

velhető. A hangolás módszere az M6-s autópálya 22+300 km-szelvényében, a Budapest felé tartó oldalon elhelyezkedő portál által az úthasználati jogosultság ellenőrzése során rögzített adatsoron kerül bemutatásra. A helyszínen 2020. március 9-én a reggeli órákban utolérési baleset történt menetirány szerint a detektálási keresztmetszet után nagyjából 400 méterre, így vizsgálhatók a forgalmi áramlat jellemzőiben a baleset hatására bekövetkező változások.

Elsőként az elemzés szempontjából releváns forgalomtechnikai paraméterek (forgalomnagyság, időbeli és térbeli átlagsebesség, forgalomsűrűség) kerültek meghatározásra 1 perces időbeli felbontásban (1. ábra). A diagramokon egyértelműen megmutatkozik az incidens: a baleset ideje, az okozott torlódás időtartama, valamint a kapcsolódó, forgalmi állapotokra vonatkozó paraméterek alakulása is nyomon követhető. A baleset előtt az átlagsebesség változatlansága mellett növekedett a forgalomnagyság, amely a forgalomsűrűség növekedését eredményezte. Az incidens bekövetkezése

1. ábra: Az algoritmus hangolásához kiválasztott adatsor általános vizsgálata



után a forgalomnagyság és az átlagsebesség tekintetében egyaránt hirtelen csökkenés figyelhető meg. Később a járműfolyam némileg rendeződött, nőni kezdett a forgalomnagyság, ekkor kis sebességgel jelentős mennyiségű jármű haladt át a detektálási keresztmetszeten. A torlódott állapot megszűnésének idejében hirtelen ugrás figyelhető meg a forgalomnagyság értékében, az addig araszoló járművek ismét szabad utat kaptak, és feltartóztatás nélkül haladhattak tovább az útszakaszon. A forgalmi állapot rendeződése után ismét a helyszínré jellemző, általánosnak tekinthető paraméter-értékek jelennek meg.

## 2.1. Az ARIMA-modell alapú algoritmus működése

A módszer lényege, hogy egy adott  $t$  mérési időpontban a megelőző három mérés eredményei alapján becslést ad a  $t+1$ -edik mérési időpontban várható értékre, amelyet az előzőleg mért és becsült értékek eltéréséből számít. Amennyiben a  $t+1$ -edik mérési időpontban a ténylegesen mért értékek kiesnek a korábban becsült értékek 95%-os konfidenciaintervallumából, a rendszer incidensriasztást ad le. Az eredeti algoritmus hurokdetektoros mérések foglaltságadataiból dolgozik, amely – a két érték egyenes arányosságából adódóan – az időegységre vetített átlagsebességgel helyettesíthető. Ezen értékek a nyers adatsorokból kinyerhetők, így közvetlenül használhatók az algoritmus működtetéséhez. A logika alap-egyenletei a következők:

$$\begin{aligned}v_{\text{becsült}}(t+1) &= v_{\text{mért}}(t) - e_{t-1} \cdot \theta_1 - e_{t-2} \cdot \theta_2 - e_{t-3} \cdot \theta_3, \\v_{jh}(t+1) &= v_{\text{becsült}}(t+1) + n \cdot \sigma, \\v_{ah}(t+1) &= v_{\text{becsült}}(t+1) - n \cdot \sigma,\end{aligned}$$

ahol az alkalmazott jelölések az alábbiak:

- $v_{\text{becsült}}(t+1)$ : a  $t+1$ -edik időintervallumra becsült átlagsebesség;
- $v_{\text{mért}}(t)$ : a  $t$ -edik időintervallumban mért átlagsebesség;
- $e_{t-1}$ ,  $e_{t-2}$ ,  $e_{t-3}$ : a  $t-1$ -edik,  $t-2$ -edik,  $t-3$ -adik időintervallumra vonatkozó becslési hiba;

- $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ : a becslő algoritmus mozgóátlag paraméterei;
- $v_{jh}(t+1)$ ,  $v_{ah}(t+1)$ : a  $t+1$ -edik időintervallumra vonatkozó felső és alsó intervallumhatárok;
- $n$ : segédparaméter az intervallumhatárok hangolásához;
- $\sigma$ : a mérési zaj által okozott becsült hiba (a mérési helyszín egyedi jellemzője).

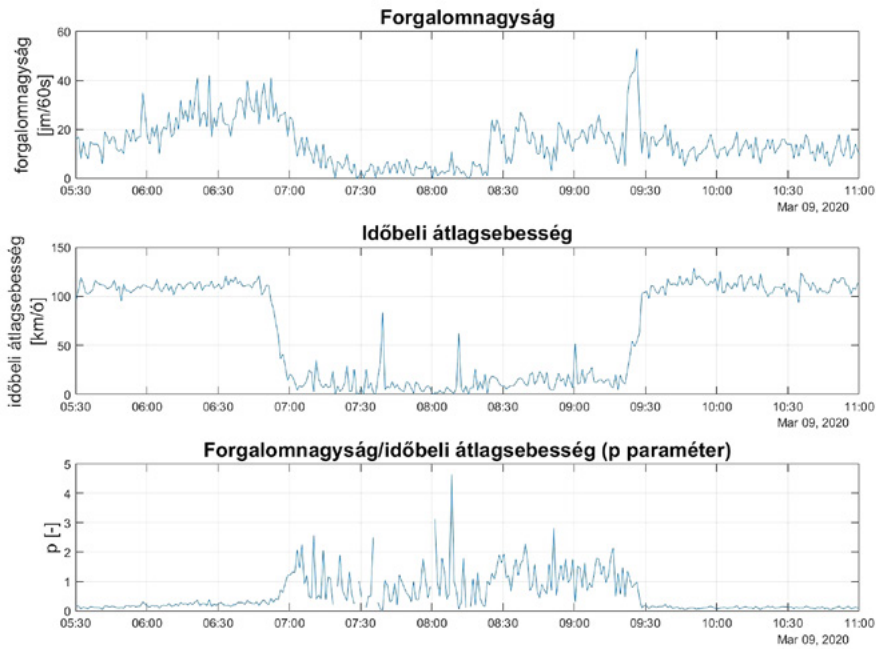
A mérést befolyásoló zaj ( $\sigma$ ) értékének módosításával az intervallumhatárok, a hibák mozgóátlag-paramétereivel ( $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ ) pedig az incidensdetektálás érzékenysége módosítható. Minél hangsúlyosabb az utolsó mérési pont ( $t-1$ ) hibája, annál hamarabb detektál a rendszer incidenst, viszont ez növelheti a hamis riasztások számát. Ha a három méréssel ezelőtti pont ( $t-3$ ) kapja a legnagyobb súlyt, az lassítja a detektálás sebességét, és kisebb eséllyel ad le hamis riasztást, viszont nagyobb valószínűséggel maradhat el egy valódi incidens detektálása.

Az intervallumhatárok hangolásához használt  $n$  paraméter az eredeti egyenletekben konstans értéket vesz fel, viszont ezt dinamikusan változóvá alakítottuk. Amikor egy adott időintervallumban nagyon kevés jármű halad el, ráadásul a sebességeik között is jelentős különbség, az torzulást és hibás incidensjelzést eredményezhet. Ennek megelőzése érdekében vezettük be a  $p$  paramétert, amely információt szolgáltat a hálózat terheltségéről a vizsgált időintervallumban:

$$p = \frac{N}{v_{\text{mért}}}$$

Az adott időintervallumon mért járműszám ( $N$ ) és átlagsebesség ( $v_{\text{mért}}$ ) hányadosa különböző forgalmi viszonyok mellett a következőként alakul:

- kis járműszámnál a sebességtől függetlenül alacsony, nullához közeli értéket ad;

2. ábra: Az ARIMA modellben használt  $p$  paraméter időbeli alakulása

- nagyobb járműszámnál és magasabb sebességértékeknél közepes értéket ad;
- nagy járműszám és alacsony sebességek esetén jelentősen nagy, kiugró értéket ad.

A paraméter várható értéke nem általánosítható, mivel az szoros összefüggésben áll az alkalmazott időbeli felosztással, az intervallumok hosszával. Célunk az, hogy kis járműszám esetén kiterjesszük, nagyobb járműszám esetén pedig szűkítsük az elfogadási intervallumot, így az egyenletekben a  $p$  paraméter reciprokát használjuk fel. A forgalmi áramlat jellemzőit és  $p$  értékének alakulását a 2. ábra szemlélteti. Néhány esetben a paraméter nem számítható a nulla értékű forgalomnagyság és átlagsebesség értékek miatt, viszont ez az algoritmus pontosságát és az incidensek felismerését nem befolyásolja.

A  $p$  paramétert egy olyan időbeli jelként értelmezhetjük, amely felerősíti a kritikus forgalmi állapotokat és az ebből származtatott  $n$  para-

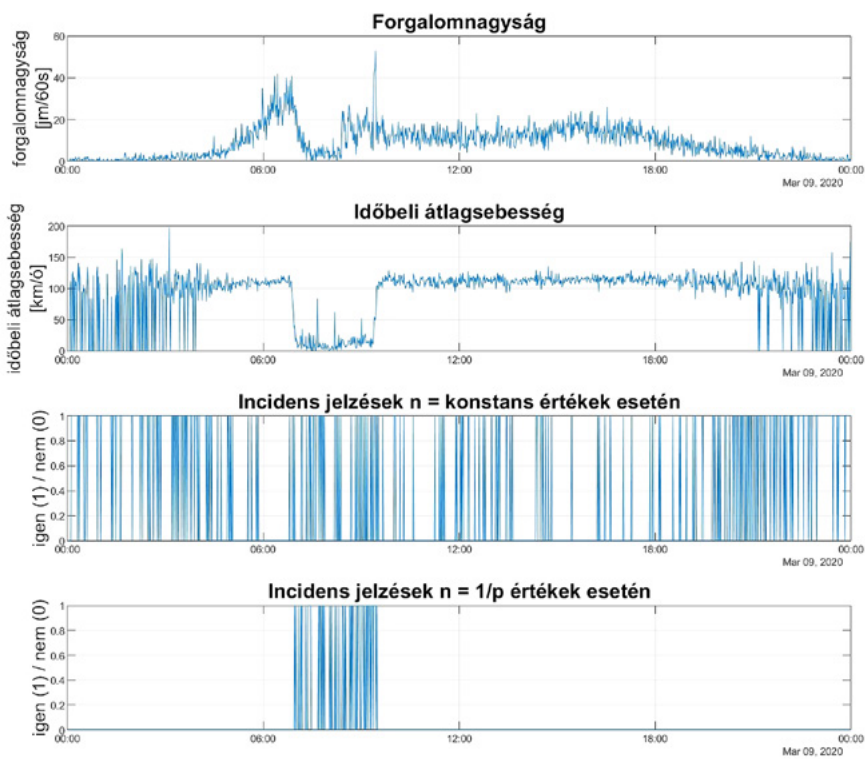
méter szűkebb elfogadási határokat eredményez ezen időintervallumokban, használatával pedig csökkenthető a téves riasztások száma (3. ábra). A modell paramétereinek ismertetésétől jelen pillanatban eltekintünk, kizárólag az  $n$  paraméter megválasztása és az incidensjelzések közötti összefüggést mutatjuk be.

Megfigyelhető, hogy az  $n$  paramétert dinamikus változóként használva az algoritmus csak a ténylegesen kritikus időszakban (ahol  $p$  értéke kiugró) ad le riasztást, ezen kívül nincsenek incidensjelzések. Az elfogadott átlagsebességhez tartozó intervallumhatárok alakulását a 4. ábra és az 5. ábra szemlélteti. Látható, hogy a kritikus forgalmi állapotokhoz a paraméter módosítása következtében jelentősen szűkebb elfogadási sáv tartozik.

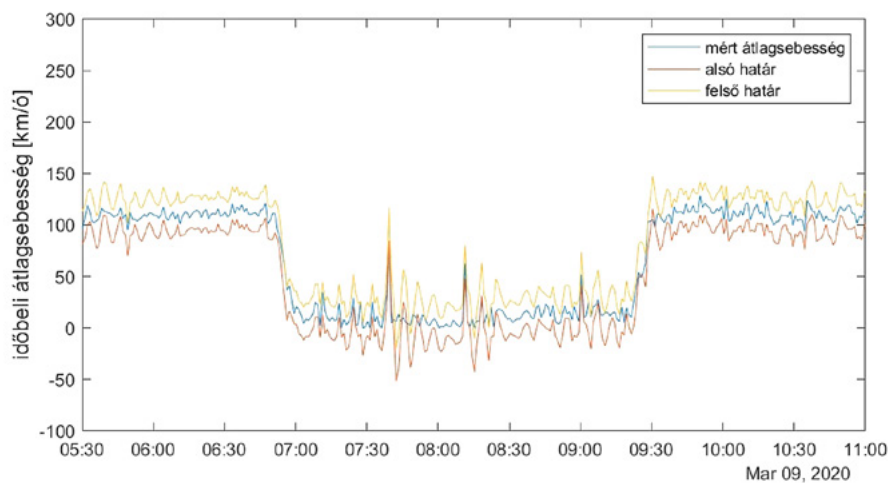
## 2.2. Különböző változatok hatékonyságának vizsgálata

Az ARIMA-modell alapú algoritmus változtatható paramétereit jelentősen befolyá-

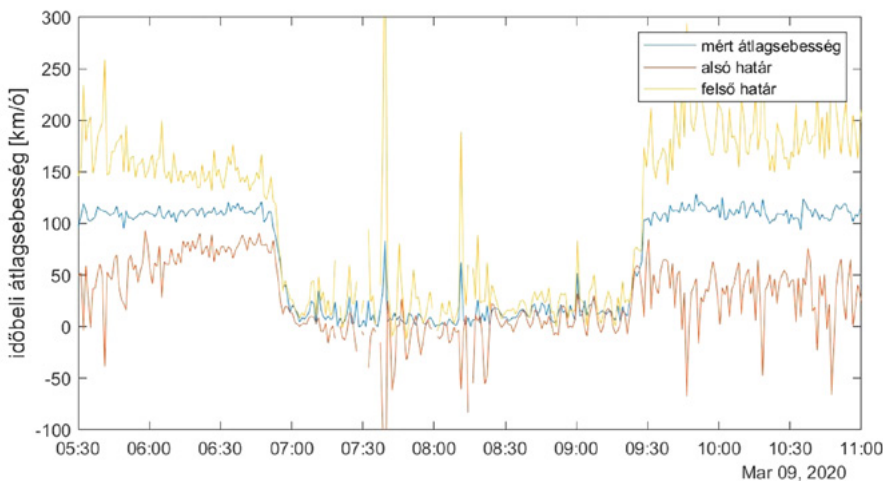
3. ábra: Incidensjelzések



4. ábra: Az n = 2 esethez tartozó elfogadási intervallumok



5. ábra: Az  $n = 1/p$  esethez tartozó elfogadási intervallumok



solják a módszer hatékonyságát, ezért mindenképpen manuális hangolást igényelnek. A paraméterek értékeinek előzetes megválasztásával különböző változatokat (A, B, C, D) állítottunk össze (1. táblázat), a különböző változatokat pedig hatékonysági mutatók alapján vizsgáltuk. A teljesítmény értékelése három általánosan elterjedt mutató alapján történt.

1. táblázat: ARIMA-modell alapú algoritmus változataihoz tartozó paraméterértékek

A					
$\theta_1, \theta_2, \theta_3$	[0,57; 0,285; 0,145]				
$\sigma$	2	3	5	8	15
B					
$\theta_1, \theta_2, \theta_3$	[0,34; 0,33; 0,33]				
$\sigma$	2	3	5	8	15
C					
$\theta_1, \theta_2, \theta_3$	[0,5; 0,25; 0,25]				
$\sigma$	2	3	5	8	15
D					
$\theta_1, \theta_2, \theta_3$	[0,25; 0,5; 0,25]				
$\sigma$	2	3	5	8	15

### 2.2.1. Felismerési arány (DR – Detection Rate)

A felismerési arány a vizsgált helyszínen felismert incidensek számát viszonyítja a szakszon valóban bekövetkezett incidensek számához. A tesztelés során az algoritmus különböző változatainak egyikénél sem fordult elő, hogy a bekövetkezett incidens detektálása elmaradt volna, így ez az érték minden vizsgálatnál 100%-nak tekinthető.

$$DR = \frac{\text{felismert incidensek száma}}{\text{megtörtént incidensek száma}} * 100 [\%].$$

### 2.2.2. Téves riasztási arány (FAR – False Alarm Rate)

Téves riasztásnak tekintünk egy leadott incidensjelzést, ha az olyan időpontban jelentkezett, amikor a valóságban nem következett be a forgalom áramlását akadályozó esemény. Tekintve, hogy a modellben használt  $n$  paramétert dinamikussá alakítottuk, az incidens bekövetkezése és a forgalmi áramlat rendeződése között bekövetkező jelzéseket nem tekintjük téves riasztásnak, mivel a paraméter használatának célja a kritikus forgalmi állapotok ki-



emelése. Esetünkben manuálisan határoztuk meg az incidens bekövetkezésének és a forgalmi állapot rendeződésének időpontját, majd a riasztások számának alakulását vizsgáltuk ezen intervallumon belül és kívül.

$$FAR = \frac{\text{téves riasztások száma}}{\text{összes riasztás száma}} \cdot 100 [\%] .$$

A téves riasztásoknak a valóságban több oka is lehet. Amennyiben a mérőeszköz működése nem megfelelő, az hamis adatokat szolgáltathat az algoritmus számára. Problémát okozhatnak a lassan haladó járművek, amelyek a közlekedés többi résztvevőjét feltartják, és ezáltal a forgalmi áramlatban a balesetekhez hasonló változásokat hoznak létre. Nehézséget okozhat az útpálya vonalvezetésében valamilyen szokatlan elem is (éles kanyar, forgalmi sáv megszűnése), amely a járművek sebességének csökkenését eredményezi. Probléma jelentkezhet a nagy forgalmú autópálya-felhajtók környezetében is, amikor a főpálya forgalma jelentősen megnő, akár annak kapacitását is meghaladva. Ezen tényezők miatt javasolt az algoritmusok egyedi hangolása a különböző mérési helyszíneken.

2.2.3. Felismerési idő  
(TTD – Time To Detect)

A felismerési idő az incidens bekövetkezésének pillanatától annak érzékeléséig eltelt időtartamot jelenti. A mutató számításához a baleset bekövetkezésétől számítva meghatároztuk az azt követő legkorábbi időpillanatot, amikor az algoritmus riasztást adott le. A különböző események detektálási ideje eltérhet, ezért a különböző incidensekre adott reakciók vizsgálata, valamint a paraméterek hangolása kiemelten fontos.

*TTD=incidens detektálásának időpontja –  
incidens bekövetkezésének időpontja.*

2.2.4. Mérőszámok közötti összefüggések

Az incidensfelismerő algoritmus hangolása során arra kell törekedni, hogy a felismerési arány minél nagyobb legyen, míg a téves riasztási arány és a felismerési idő a lehető

legalacsonyabb értéket vegye fel. Ideális esetben a hatékonyságot jellemző paraméterek: DR=100%, FAR=0%, TTD=0 s. A folyamat nehézsége abban rejlik, hogy a mutatók egymástól nem függetlenek. A felismerési ráta növekedésével a téves riasztási arány is nőni fog, hiszen a rendszer sokkal érzékenyebbé válik. Ennek következtében olyan jelenségekre is riasztást ad le, amelyek valójában nem minősülnek a forgalom lefolyását akadályozó eseménynek. Ha a téves riasztások számának minimalizálása az elsődleges szempont, akkor a nem detektált incidensek száma növekedhet. Hasonló probléma jelentkezik a felismerési idővel kapcsolatban is. Ha a logika gyorsan elemzi az adott szituációt, akkor nagyobb lesz a téves riasztási arány, miközben a felismerési arány egyre csökken. A lassabb elemzés viszont magasabb felismerési rátát és kisebb téves riasztási arányt eredményez: minél több ideje van a rendszernek „gondolkodni”, annál pontosabb lesz az eredmény.

2. táblázat: Különböző paraméterezésű változatok teljesítménymutatói					
A					
θ <sub>1</sub> , θ <sub>2</sub> , θ <sub>3</sub>	[0,57; 0,285; 0,145]				
σ	2	3	5	8	15
DR	100%	100%	100%	100%	100%
FAR	35,82%	13,48%	1,92%	0,00%	0,00%
TTD [min]	1	1	2	8	12
B					
θ <sub>1</sub> , θ <sub>2</sub> , θ <sub>3</sub>	[0,34; 0,33; 0,33]				
σ	2	3	5	8	15
DR	100%	100%	100%	100%	100%
FAR	29,35%	9,85%	0,00%	0,00%	0,00%
TTD [min]	1	2	2	6	8
C					
θ <sub>1</sub> , θ <sub>2</sub> , θ <sub>3</sub>	[0,5; 0,25; 0,25]				
σ	2	3	5	8	15
DR	100%	100%	100%	100%	100%
FAR	31,15%	9,77%	1,00%	0,00%	0,00%
TTD [min]	1	1	2	8	8
D					
θ <sub>1</sub> , θ <sub>2</sub> , θ <sub>3</sub>	[0,25; 0,5; 0,25]				
σ	2	3	5	8	15
DR	100%	100%	100%	100%	100%
FAR	32,51%	10,00%	0,00%	0,00%	0,00%
TTD [min]	1	2	2	2	20

3. táblázat: A leghatékonyabb változatok összefoglalása

	A	B	C	D	D
$\theta_1, \theta_2, \theta_3$	[0,57; 0,285; 0,145]	[0,34; 0,33; 0,33]	[0,5; 0,25; 0,25]	[0,25; 0,5; 0,25]	[0,25; 0,5; 0,25]
$\sigma$	5	5	5	5	8
DR	100%	100%	100%	100%	100%
FAR	1,92%	0,00%	1,00%	0,00%	0,00%
TTD [min]	2	2	2	2	2

Az előzetesen paraméterezett algoritmus-változatok teljesítménymutatóit a 2. táblázat foglalja össze. A felismerési arány minden esetben 100%, mivel az algoritmus egyik esetben sem mulasztotta el jelezni a bekövetkezett incidenst. A téves riasztások arányának minimalizálása mellett a felismerési idő is kiemelt szempont, így a teljesítménymutatók egyidejű vizsgálata célszerű. Az A, B és C változat esetén  $\sigma$  értékét 5-nek választva érhető el a legjobb eredmények, a D változatban  $\sigma = 8$  is hasonlóan jól teljesít. Az így megmaradt változatok teljesítménymutatói a 3. táblázatban láthatók.

A második körben elvégzett összehasonlításból látszik, hogy a paraméterértékek különböző kombinációja mellett azonos hatékonysággal teljesített az algoritmus. Tekintve, hogy az eltérő típusú zavarások különböző változásokat idéznek elő a forgalmi áramlat tulajdonságaiban, érdemes több incidenst is megvizsgálni, azokat egymással összehasonlítani és az eredmények alapján megválasztani az optimális hangolási értékeket.

### 3. ÖSSZEFOGLALÁS

Cikkünkben bemutattuk a NÚSZ portálkapuk detektálási adatainak felhasználási lehetőségét egy statisztikai elven működő incidensdetektáló algoritmushoz. Megerősítettük, hogy a NÚSZ detektálási keresztmetszein előálló adatok megfelelő bemeneti adatként szolgálnak az ARIMA-modell alapú incidensdetektáló módszer teszteléséhez, működtetéséhez. Bemutattuk, hogy a gyakorlati alkalmazás és az éles üzem előtt miként lehetséges a különböző paraméterezésű változatok hatékonyságának vizsgálata és ezek alapján a legjobban teljesítő megoldás kiválasztása. Is-

mertettük az összehasonlításhoz használható teljesítménymutatókat, azok számításának elvét és módszerét. Egy kvalitatív megközelítésű módszertant mutattunk be a változatok összehasonlításához, amelyet egy kiválasztott útszakaszon bekövetkezett valós incidenshez tartozó adatsoron sikeresen alkalmaztunk.

A teljes folyamat egy mérési és értékelési utasításként kezelhető, amely az AID algoritmusok jövőbeli gyakorlati alkalmazásba helyezését támogatja.

### KÖSZÖNETNYILVÁNÍTÁS

A cikk szerzői köszönetüket fejezik ki a Nemzeti Útdíjfizetési Szolgáltató Zrt.-nek a szolgáltatott adatokért és a kutatás támogatásáért.

### FELHASZNÁLT IRODALOM

- [1] Horváth, M.T., 2012. Automatikus incidensfelismerő algoritmusok összehasonlítás autópályán, BSc szakdolgozat
- [2] Martin, P.T., Perrin, J., Hansen, B., Kump, R. and Moore, D., 2001. Incident detection algorithm evaluation. Prepared for Utah Department of Transportation.
- [3] Parkany, E. and Xie, C., 2005. A complete review of incident detection algorithms & their deployment: what works and what doesn't.
- [4] Mahmassain, H.S., Haas, C., Zhou, S. and Peterman, J., 1999. Evaluation of incident detection methodologies (No. FHWA/TX-00/1795-1). University of Texas at Austin. Center for Transportation Research.





### Possible uses of data from the HU-GO electronic toll system for an automatic incident detection algorithm

The basic task of the National Toll Payment Services PLC (NÚSZ) is to sell road use authorization and to verify the existence of such authorizations and to provide related services. Traffic monitoring on the toll road network is carried out by means of nearly 130 fixed toll gates equipped with cameras and laser scanners, and video equipment mounted on vehicles. A large amount of detailed data is generated during the verification of the right to use the road, which can be used for traffic management and automatic incident detection (AID) algorithms. In this paper, the different AID methods are described, followed by the tuning methodology of the ARIMA model-based algorithm and the examination of the effectiveness of different versions tested with real data.



### Die Möglichkeit der Nutzung von Daten aus dem elektronischen Mautsystem HU-GO für einen Algorithmus zur automatischen Erkennung von Vorfällen

Die Hauptaufgabe der Nationalen Mauterhebung Geschlossener Dienstleistungsgesellschaft (NÚSZ) ist der Verkauf von Strassenbenutzungsberechtigung, die Überprüfung des Bestehens dieses Rechts und die Erbringung der damit verbundenen Dienstleistungen. Die Verkehrsüberwachung auf dem mautpflichtigen Straßennetz erfolgt mit Hilfe von fast 130 fest installierten Portalschranken, die mit Kameras und Lasermessgeräten ausgestattet sind, sowie mit auf Fahrzeugen montierten Videogeräten. Bei der Überprüfung des Bestehens der Straßenbenutzungsberechtigung wird eine große Menge detaillierter Daten generiert, die auch für das Verkehrsmanagement und AID-Algorithmen (Automatic Incident Detection, automatische Vorfallerkennung) verwendet werden können. In dem Artikel werden zunächst die verschiedenen AID-Methoden beschrieben, gefolgt von der Abstimmungsmethodik des auf dem ARIMA-Modell basierenden Algorithmus und der Wirksamkeit der verschiedenen Versionen, die mit realen Daten getestet wurden.

## E számunk lektorai

Dr. Gulyás András ■ Horváth Gábor

Dr. Katona András ■ Ötvös Viktória ■ Dr. Tóth János