

Oponentní posudek

Habilitační práce

„Datová analýza v bezpečnostní praxi“

Autor: RNDr. Tatiana Hajdúková, PhD.

Vhodnost a aktuálnost tématu

RNDr. Tatiana Hajdúková, PhD. předložila k habilitačnímu řízení práci na téma „*Datová analýza v bezpečnostní praxi*“. Toto téma jako oponent považuji za velmi vhodné, aktuální a velmi významné.

Žijeme v době tzv. „*Informační revoluce*“, která je charakteristická bouřlivým rozvojem informačních a telekomunikačních technologií, které nás dnes obklopují na každém kroku ve všech oblastech lidské činnosti. Pro tuto dobu je význačné, že se objemy vytvořených, přenášených, uchovávaných a zpracovávaných dat neustále navyšují geometrickými způsoby (jedná se o násobky objemů z předchozích období). Tento růst bude neustále pokračovat, protože data v blízké době budou generovat nejenom lidé, ale i nejrůznější zařízení v našich domovech. Jedná se o tzv. fenomén „*Internet věcí*“ (*Internet of things*).

Výše popsany rozvoj se týká i bezpečnostní oblasti, práce Policie SR a dalších evropských zemí, dalších bezpečnostních složek státu. Úkolem bezpečnostních složek je plnit zákony požadované služby a aktivity, chránit zdraví, životy a hodnoty občanů a společnosti.

Zároveň důležitým, nikde nepsaným způsobem je v době informačních a telekomunikačních technologií se dokázat dobře orientovat ve velkém množství informací, neztratit se v něm, nenechat se dezorientovat, ale naopak se dostat včas a ke všem datům, informacím, které mohou změnit průběh negativních událostí, včetně prevence, nebo alespoň rychle a kvalitně, zákonným způsobem vést vyšetřování důsledků, dopadů rozmanitých negativních jevů, které způsobily škody nebo újmy občanům, různým institucím či státu.

Cíl a jeho definice habilitační práce

Cílem habilitační práce je poukázat na možnosti argumentace a vědeckého zdůvodňování prioritně empirických pozorování z bezpečnostní praxe, které jsou vyjádřeny v současné době zejména elektronickými daty (údaji).

Autorka, a její práce důsledně oddělují statistiku od analýzy dat. První pojem – statistika – je pojmem „statickým, pasivním“, při kterém se pouze číselně interpretují různé četnosti výskytu určitých jevů, událostí (zpravidla negativních), zatímco analýza dat je pojmem „dynamickým, aktivním“ jehož cílem je určit předpoklady, vztahy, anomálie, asociace, predikce, prognózy, vazby, dopady, jejich pravděpodobnosti apod. zejména výskytu negativních jevů, dějů a událostí a tím preventivně nebo operativně či při vyšetřování využít získaná data, informace z informačních systémů policie, bezpečnostních složek, veřejné správy a v neposlední řadě i z otevřených zdrojů, zejména z prostředí Internetu, sociálních sítí atd.

Cílem práce je popsat, determinovat různé vědecké způsoby a metody pro zpracování dat (jejich sběr, analýzu a interpretaci, včetně grafické) a názorně, jednoduše je demonstrovat na známých, zřejmých příkladech z bezpečnostní praxe.

Struktura a obsah habilitační práce

Předložená habilitační práce má 146 stran, včetně 8 příloh. V práci nalezneme 16 obrázků, 32 grafů, 32 tabulek. Počet grafů a tabulek z logických důvodů převyšuje počet obrázků, neboť práce se zaměřuje na zpracování dat různými matematickými způsoby (metodami), včetně jejich názorné grafické prezentace.

Práce obsahuje standardní části jako jsou abstrakty ve slovenském a anglickém jazyce, obsah, seznamy tabulek, grafů, obrázků a seznam zkratk, a závěr, seznam literatury. Autorka vhodným způsobem doplnila ještě méně obvyklé části jako jsou výkladový slovník pojmů překladový anglicko-slovenský slovník použitých odborných termínů. Tyto dvě části velice pomáhají pochopení obsahu práce.

Ve struktuře habilitační práce nalezneme v konci i kapitulu, označenou jako diskuzi, která se zaměřuje na logický průnik problematiky zpracování dat a dopravních nehod, které jsou velmi specifickým fenoménem dnešní doby s nemalými ztrátami na majetku, zdraví i životů lidí. Tato oblast je velmi vhodná z hlediska odborného i společenského na zpracování dat (jejich získávání, přenos, analýzu, prezentaci apod.). Autorka zcela přesně vystihla tuto skutečnost a proto ve své práci pro lepší pochopení předmětné problematiky uvádí velké množství názorných, přímo ukázkových příkladů právě z oblasti dopravní problematiky, jako jsou právě dopravní nehody.

Samotný obsah práce je rozdělen do 4 základních kapitol.

První kapitola se zabývá obecnou metodologií a metodami výzkumu a vymezuje základní pojmy, zaměřuje se na úlohu výzkumu ve vědeckém poznání, hypotézy, základní soubory a rozdělení početnosti.

Kapitola druhá s názvem „*Základní vědecké metody*“ klade akcent na pozorování, porovnání (komparaci), klasifikaci a tzv. Paretovu analýzu.

Kapitola třetí je „*Statistická analýza dat*“, kde autorka se primárně zaměřuje na oblasti analýzy rozptylu, analýzy časových řad, měsíční sezónnost časový řad, kontingenční tabulky a grafického zobrazení (interpretaci dat).

Kapitola čtvrtá je věnována moderním přístupům, které se objevují až s výpočetní technikou, s moderními technologiemi, s vyspělými SW prostředky, s databázemi s velkými objemy dat a je nazvána „*Dataminingová analýza dat*“, obsahující podkapitoly „Big data“, „Data mining“, „Deskriptivní modelování dat“, „Prognózování a predikce“ a v neposlední řadě „Základy umělé inteligence“.

Logičnost a přístup ke zpracování

Kapitoly na sebe harmonicky navazují, doplňují se, jsou obsahově, strukturálně i objemově vyvážené z pohledu stanoveného cíle. Velké pozornosti je věnováno statistické analýze dat a dataminingové analýze dat, jež jsou jádrem práce a tyto kapitoly co do počtu stran jsou rozsáhlejší. Všechny 4 kapitoly jsou číslovány pouze do 2 úrovní nadpisů, což pomáhá dobré orientaci v habilitační práci a jejímu pochopení.

Novátorství zpracovávaného tématu

Většinou lidem se zdá, že bezpečnost, a matematika nejdou moc k sobě. Bezpečnostní praxe je však zcela jiná. Přístupy například kriminalistiky při vyhledávání stop a zpracování dat v oblasti obecné informatiky jsou velmi podobné. Stopy jsou svým způsobem též data, informace. (Kriminalistické) stopy nejprve vyhledáváme, dokumentujeme, zajišťujeme, přenášíme do laboratoře, zpracováváme (popisujeme, analyzujeme), prezentujeme, uchováváme. V obecné informatice data jsou nějakým způsobem generována (jako výsledek interakce určitých objektů, procesů v konkrétním prostředí), získávána, přenášena, zpracovávána (výpočetní technikou pomocí daných algoritmů odpovídajících určitým procesům, prezentována a rovněž uchovávána určitou dobu pro určité účely pozdějšího zpracování.

Autorka ve své habilitační práci ukazuje právě cesty, způsoby, vědecké metody a postupy, jak obecně přistupovat ke zpracování dat tak, aby výsledky byly využitelné v praxi i pro bezpečnostní účely. Policejní, bezpečnostní praxe je zvyklá používat statistické výstupy nebo ad hoc jednoúčelové dotazy do různých informačních zdrojů. Tento postup je z mého pohledu „statický, reaktivní, pasivní“, jednoúčelový.

Autorka ve své poslední kapitole poukazuje na problematiku dataminingového zpracování zejména velkého objemu dat. Tento přístup však přináší mnohem více možností využití, podstatně širší záběr, než pouhé statistické přístupy. Je to právě především o nalezení, určení rozmanitých předpokladů, vztahů, anomálií, asociací, predikcí, prognóz, vazeb, dopadů, jejich pravděpodobností atd., které lze využít při stanovení různých vyšetřovacích hypotéz, scénářů postupů atd. Všechny tyto závislosti, vztahy atd. kromě operativního využití mají obrovský význam pro prevenci, stejně tak i prognózování různých zejména negativních událostí s různě dalekým výhledem „do budoucna“. Umělá inteligence pak už nabízí i možnosti určitých automatických procesů, které budou využitelné pro řešení bezpečnosti a zajištění bezpečí různých subjektů.

Práce s literaturou a dalšími informačními zdroji

Bibliografický seznam použité literatury má 5 stran, uvádí 66 převážně knižních slovenských, českých a anglických zdrojů. Všechny zdroje v bibliografickém seznamu jsou uvedeny korektně podle platných norem.

V průběhu textu habilitační práce autorka uvádí minimální počet odkazů na výše uvedený seznam. Svědčí to o tom, že výše uvedené zdroje použila pouze pro svůj celkový přehled o problematice, pro inspiraci. Práce je pak psána dále samostatně, autorsky.

Jazyková a grafická stránka

Z hlediska jazyka, způsobu vyjadřování se, či grafické úpravy jsem v textu habilitační práce nenašel žádné pochybení. Práce je velmi kvalitně, profesionálně zpracována. Je přehledná, srozumitelná, dobře čtivá i pro laiky.

Přínosnost tématu

Mohu konstatovat, že autorka zcela splnila cíl své práce. Předložila systematicky zpracovaný základní dokument, který popisuje jakým způsobem, jakými vědeckými metodami využít velké objemy dat pro různé bezpečnostní úlohy. Cenné jsou i jednoduché, vhodně vybrané a demonstrovány příklady, postupy, které ukazují, že to není zase až tak v praxi složité.

Práce je přínosná nejen pro policejní sbory, ale obecně pro jakékoliv bezpečnostní složky státního, veřejného či privátního charakteru.

Otázky k předložené práci pro habilitanta

Na straně 137 habilitační práce uvádíte systémy z rozhraní CLK. Které tyto informační systémy jsou využitelné pro aplikaci metod a technologií popsanych v kapitole 3 a 4. Uveďte současné, již existující řešení a navrhněte, co dalšího a jakým způsobem by bylo možné pro bezpečnostní praxi vytvořit. Zkuste uvést nějaké konkrétní návrhy a jaké by měly výstupy, funkčnost a jak by pomohly nové funkcionality řešit různé bezpečnostní situace.

Závěrečné doporučení oponenta

Jako oponent habilitační práce mohu potvrdit, že tato splňuje všechna kritéria na habilitační práci a doporučuji habilitační komisi v pokračování celého procesu, jehož cílem je udělení pedagogické hodnosti docent paní RNDr. Tatianě Hajdúková, PhD. Jsem hluboce přesvědčen, že habilitantka je velmi dobře připravena na to, aby se stala docentem v oboru Ochrana osob a majetku.

Rovněž doporučuji, aby předložená habilitační práce po doplnění a vhodných úpravách byla publikována jako odborná monografie.

Prof. Ing. Roman Rak, PhD.

V Praze dne 30.4.2019