TECHNICKÁ UNIVERZITA V LIBERCI
Ekonomická fakulta

VČASNÉ VAROVÁNÍ PŘED ZVÝŠENÝM RIZIKEM VZNIKU DOPRAVNÍ NEHODY S VYUŽITÍM DATA MININGU

Autoreferát disertační práce

Studijní program: P6209 – Systémové inženýrství a informatika
Studijní obor: 6209V003 – Ekonomická informatika

Autor práce: **Ing. Bc. Marián Lamr**
Školitel: doc. Ing. Jan Skrbek, Dr.

Disertační práce byla vypracována v prezenční formě doktorského studia na Katedře informatiky Ekonomické fakulty Technické univerzity v Liberci.

Uchazeč: Ing. Bc. Marián Lamr
Ekonomická fakulta Technické univerzity v Liberci
Katedra informatiky
Voroněžská 13
461 17 Liberec 1

Školitel: doc. Ing. Jan Skrbek, Dr.
Ekonomická fakulta Technické univerzity v Liberci
Katedra informatiky
Voroněžská 13
461 17 Liberec 1

Autoreferát byl rozeslán dne:

Obhajoba disertační práce se koná dne 22. 11. 2018 před komisí na Ekonomické fakultě Technické univerzity v Liberci, Voroněžská 13, Liberec 1 v zasedací místnosti děkanátu Ekonomické fakulty.

S disertační prací je možno se seznámit na katedře informatiky Ekonomické fakulty Technické univerzity v Liberci.

doc. Ing. Jan Skrbek, Dr.
předseda dílčí (fakultní) oborové rady
DSP Systémové inženýrství a informatika

Čestné prohlášení:

Prohlašuji, že jsem disertační práci vypracoval samostatně a na základě literatury a pramenů uvedených v seznamu použité literatury.

V Liberci dne 19. září 2018 ……

Ing. Bc. Marián Lamr

Anotace a klíčová slova

V poslední dekádě dochází v České republice každoročně k nárůstu počtu dopravních nehod. Obdobný trend prezentují i statistiky počtu lehkých i těžkých zranění vzniklých na silničních komunikacích. Navzdory očekávání neklesá významněji ani výše materiálních škod. V reakci na tento trend policie pravidelně realizuje represivní kroky vůči řidičům, porušujícím pravidla silničního provozu, které však ne vždy splňují původní očekávání. Následující text se snaží přispět ke zvýšení bezpečnosti silničního provozu formou prevence vzniku dopravních nehod díky využití progresivních ICT prostředků.

Předložená disertační práce se zabývá především včasným varováním řidičů před zvýšeným rizikem vzniku dopravní nehody s využitím data miningu. Nejprve je v práci popisován aktuální stav preventivních prvků a systémů zvyšujících bezpečnost účastníků silničního provozu v ČR. Dále jsou analyzovány a vyhodnocovány nejvhodnější data miningové metody, algoritmy a přístupy s ohledem na potřeby řešení realizovaného v rámci disertační práce.

Stěžejním tématem disertační práce je konceptuální návrh systému včasného varování před zvýšeným rizikem dopravní nehody. Navrhovaný systém by měl v reálném čase a místě predikovat riziko vzniku dopravní nehody na základě modelů vytvořených pomocí data miningových algoritmů a poskytovat adekvátní upozornění pro řidiče. Součástí práce je implementace databáze dopravních nehod a vyhodnocení možností využití těchto dat pro predikční účely. V práci je dále řešeno hledání skrytých závislostí v datech o dopravních nehodách, které je nedílnou součástí řídicí části navrhovaného řešení. V neposlední řadě jsou v práci představeny modelové situace a ekonomické aspekty a přínosy celého řešení.

Klíčová slova: včasné varování, data mining, predikce, systém, dopravní nehody, data, analýza, shluky

Annotation and keywords

Recently, the number of traffic accidents has been regularly increasing. This unflattering phenomenon unevitably resulted in a growing amount of ligh and havy injuries. Despite the anticipation, material demages are not declining either. In response to this trend, the police is regularly implementing repressive actions against drivers, but these do not always meet initial expectations. This doctoral thesis deals with the use of modern ICT tools to prevent the occurrence of traffic accidents.

This thesis deals with the possibility to generate early warnings in order to inform the drivers about the increased risk of traffic accidents using data mining techniques. The first section describes the current state of existing features and systems increasing the safety of road users in the Czech Republic. Further, the most suitable data mining methods, algorithms and approaches are analyzed and evaluated with regard to the needs and specifications the solution proposed in this dissertation. The main objective of the thesis is the conceptual design of an early warning system informing the drivers about the increased risk of a traffic accident. The proposed system should predict the risk of accidents in real-time and on the basis of the models created using data mining algorithms, and provide adequate warnings to the driver. The implementation of the traffic accidents database and the evaluation of the possibilities of using this data for prediction purposes represents a significant part of the dissertation. The examination of the hidden dependencies in traffic accidents data is described as an integral proces of the proposed solution. Last but not least, the thesis presents model situations, economic aspects and benefits of the potential implemention of such concept.

Keywords: early warning, data mining, prediction, system, traffic accidents, data, analysis, clusters

Obsah

Seznam zkratek 8

Úvod 9

1 Aktuální stav systémů zvyšujících bezpečnost dopravy v ČR 11

1.1 Pasivní prvky 11

1.2 Aktivní prvky 11

1.3 Telematické systémy v ČR 12

1.3.1 RDS-TMC a informační tabule 12

1.3.2 eCall 12

2 Data mining: metodologie a algoritmy 13

2.1 Metodologie v data miningu 13

2.2 Použité DM nástroje 14

2.2.1 Vyhodnocení možností DM nástrojů pro využití v systému včasného varování 15

3 Vybrané algoritmy data miningu 16

3.1 Asociační pravidla a algoritmus Apriori 16

3.2 Shluková analýza 16

4 Cíle práce, metody a doporučení pro další práci 18

4.1 Metody a fáze výzkumu 18

5 Princip systému včasného varování před dopravní nehodou 20

5.1 Řídicí část 20

5.2 Uživatelská část 22

5.3 Spolupráce řídící a uživatelské části 23

5.4 Systémový návrh řešení a pohled z hlediska obecné teorie systémů 24

5.4.1 Místo systému v hierarchii dopravně bezpečnostních orgánů 24

5.4.2 Prvky a vazby systému včasného varování 25

5.5 Ekonomické zhodnocení systému včasného varování 27

5.6 Modelové situace 28

Závěr 32

Seznam použité literatury 35

Publikace autora související s tématem disertační práce 38

Seznam zkratek

ABS - Anti-lock Brake System

ACC - Autonomous Cruise Control

API - Application Programming Interface

ASR - Anti-Slip Regulation

CRISP-DM - Cross-Industry Standard Process for Data Mining

CARMA - Continuous Association Rule Mining Algorithm

CURL - Client for URLs

DBSCAN - Density-Based Spatial Clustering of Applications with Noise

DENCLUE - DENsity-based CLUstEring

ECALL – Emergency Call

ELKI - Environment for Developing KDD-Applications Supported by Index-Structures

ESP - Electronic Stability Program,

FTP - File Transfer Protocol

HeERO – Harmonised eCall European Pilot

HTTP - Hypertext Transfer Protocol

KNIME - Konstanz Information Miner

MinPts- Minimum of Points

LTE - Long Term Evolution

OPTICS - Ordering points to identify the clustering structure

PČR – Policie české republiky

RDS-TMC - Radio Data System - Traffic Message Channel

SEMMA - Sample, Explore, Modify, Model, and Assess

SVM -Ssupport Vector Machine

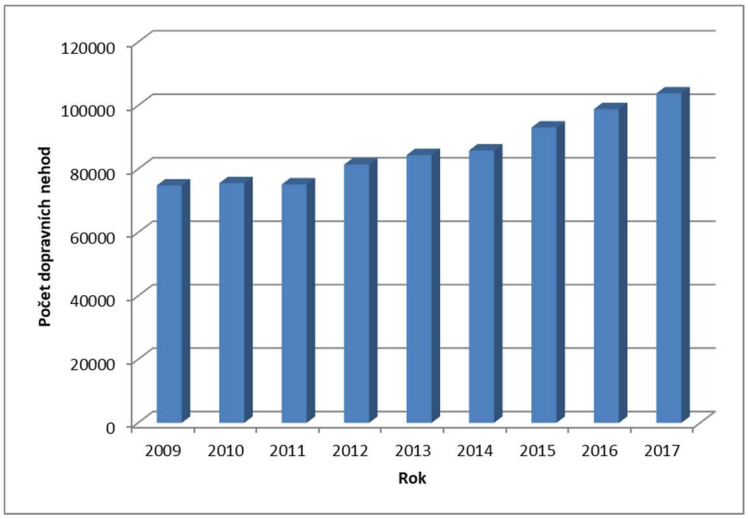
WEKA - Waikato Environment for Knowledge Analysis

Úvod

Dopravní nehody jsou bohužel každodenní a neoddělitelnou součástí silničního provozu. I přesto, že se policie snaží různými způsoby zvýšit bezpečnost silničního provozu, počty dopravních nehod, ztráty na životech, majetku a počty zranění se nedaří snížit podle očekávání.

Z aktuálních statistik dopravní nehodovosti vyplývá, že od roku 2011 dochází k neustálému nárůstu dopravních situací končících materiálními škodami či zdravotní újmou. V Grafu 1 je pro relevantní porovnání zobrazen vývoj počtu dopravních nehod od roku 2009 do roku 2017.

Graf 1: Vývoj počtu dopravních nehod



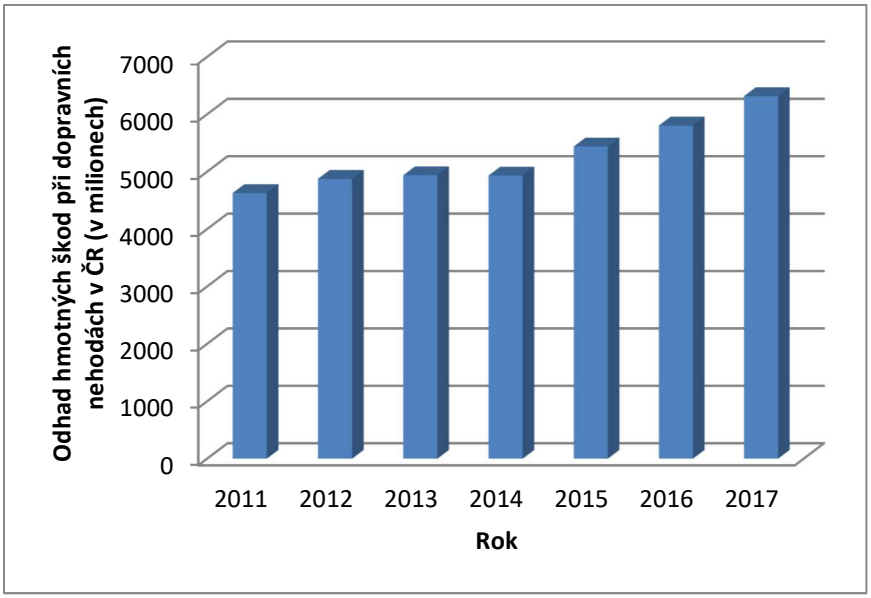
Zdroj: vlastní zpracování, data: Policie ČR

Vzhledem k již zmiňovanému narůstajícímu množství dopravních nehod se policie opakovaně snaží hledat způsoby zvýšení bezpečnosti provozu. Mezi tato opatření lze zařadit zejména frekventovanější silniční kontroly, zvýšeni finančních pokut za dopravní přestupky, zavedení bodového systému a řadu regionálních či celostátních dopravně bezpečnostních akcí. Ministerstvo dopravy plánuje další zvyšování sankcí za určité druhy přestupků a to i navzdory prokázané krátkodobé účinnosti tohoto druhu postihu. („Ministerstvo dopravy chce zvýšit pokuty za rychlost. Až trojnásobně", 2015). I přes všechna zmiňovaná opatření neklesají ani počty lehce a těžce zraněných osob.

S rostoucím počtem dopravních nehod souvisí i zvyšování hmotných škod. Vývoj trendu v této oblasti v posledních letech je znázorněn v Grafu 2, kde jsou zobrazeny odhady škod při dopravních nehodách podle policie ČR od roku 2011 do roku 2017. I z toho důvodu je třeba hledat nové přístupy předcházení dopravním nehodám. Jedním z možných způsobů zlepšení

stávající situace na silničních komunikacích, je cílená implementace moderních technologií za účelem zvýšení bezpečnost dopravy díky lepší informovanosti řidičů o potenciálních nebezpečích na trase.

Graf 2: Hmotné škody při dopravních nehodách 2011-2017



Zdroj: vlastní zpracování, Policie ČR

Cílem této práce je vytvořit konceptuální návrh systému umožňujícího v reálném čase a místě predikovat riziko dopravní nehody. Systém je založený na predikčních modelech, které jsou vytvářeny pomocí data miningových technik a nástrojů využívajících především algoritmy shlukové analýzy. V následujících kapitolách budou diskutována východiska a předpoklady pro existenci systému včasného varování před zvýšeným rizikem dopravní nehody. Nejprve bude provedena rešerše aktuálního stavu technických zařízení a systémů zvyšujících bezpečnost účastníků dopravního provozu. Dále bude věnována pozornost základním i pokročilým poznatkům z oboru data mining. Budou diskutovány vybrané postupy, programové vybavení a algoritmy použité při návrhu řešení popsaného v této práci. V dalších kapitolách se čtenář dozví, jak je v současné době možné získávat, zpracovávat a uchovávat data o dopravních nehodách a jak tato data využít při tvorbě systému včasného varování. Následně bude představen princip systému včasného varování se vzorovými úlohami demonstrujícími jeho využití. V neposlední řadě jsou v práci představeny modelové situace užití navrhovaného řešení, systémový pohled, ekonomické aspekty a společenské přínosy celého řešení.

1 Aktuální stav systémů zvyšujících bezpečnost dopravy v ČR

Zvyšování bezpečnosti účastníků silničního provozu je možné provádět pomocí celé řady prvků. V úvodu je nutné upozornit, že v této kapitole nebude věnována pozornost preventivním opatřením typu dopravně bezpečnostních akcí prováděných policií, či zásahům do bodového systému. Následující odstavce se zaměří především na vybrané typické aktivní a pasivní prvky umístěné v běžných automobilech dnešní doby, asistenční systémy a nástroje, které by měly předcházet dopravním nehodám, nebo umožnit zmírnění jejich dopadu.

1.1 Pasivní prvky

Nejprve budou popsány vlastnosti vozu, zařízení a systémy, které ke své funkci nepoužívají žádný externí zdroj informací a jsou integrovány ve vozidlech. Tato zařízení ke své činnosti buď nepotřebují žádná čidla, anebo využívají informace, jež jim dodávají pouze čidla dostupná v daném automobilu.

Prvky pasivní bezpečnosti lze definovat jako konstrukční zařízení, která jsou aktivována až ve chvílích, kdy nastane nebezpečná situace nebo dopravní nehoda. Cílem těchto zařízení je tudíž minimalizace následků dopravní nehody. Vzniklou situaci tedy řeší ex post. Nejběžnějšími prvky pasivní bezpečnosti v dnešních automobilech jsou především zádržné systémy (bezpečnostní pás, přepínač bezpečnostního pásu, dětská sedačka), opěrka hlavy, bezpečná konstrukce karoserie a airbag (Vlk, 2006).

1.2 Aktivní prvky

Aktivní prvky bezpečnosti ve vozidle se na rozdíl od prvků pasivní bezpečnosti snaží dopravním nehodám a krizovým situacím předcházet. Jedná se především o nejrůznější technická zařízení, systémy a vlastnosti vozu využívající data z čidel ve vozidle. Základními prvky aktivní ochrany jsou kvalitní brzdy a přesné řízení. V současných vozidlech však existuje celá řada sofistikovaných systémů a řešení jako například elektronický protiblokovací systém ABS (Anti-lock Braking Systems), protiprokluzový systém ASR (Antriebs-Schlupf-Regelung) a elektronický stabilizační systém ESP (Electronic Stability Program) a další.

1.3 Telematické systémy v ČR

1.3.1 RDS-TMC a informační tabule

RDS-TMC (Radio Data System - Traffic Message Channel) označuje kanál umožňující přenos předzpracovaných dopravních informací do vozidla, kde jsou tyto údaje dále upraveny a distribuovány řidiči. RDS-TMC je součástí dopravně informačního systému a vždy je pevně spjat s konkrétní rozhlasovou stanicí, a proto se vysílaná data mohou stanici od stanice lišit. RDS-TMC mohou využívat navigace i mobilní telefony s TMC dekodérem. I přesto, že RDS-TMC informuje o problémech na komunikaci, které se již staly, lze jej zařadit k aktivním systémům přispívajícím k prevenci nehod. S obdobně nastíněnými problémy je nutné počítat u dalšího prostředku pro šíření informací pro řidiče, a to u informačních tabulí. Projekt jednotného systému dopravních informací je společným dílem Ministerstva dopravy, Ředitelství silnic a dálnic a několika dalších organizací (Lamr, 2015a), (Lamr, 2015c).

1.3.2 eCall

Systém eCall řeší především následky dopravní nehody, a proto patří mezi pasivní systémy. Jde o službu navrženou pro umožnění rychlého pohotovostního zásahu v případě dopravní nehody, a to kdekoli na území EU. Cílem systému eCall je zvýšit bezpečnost dopravy v Evropě a snížit počet úmrtí zapříčiněných dopravními nehodami, a také zamezit zraněním a ztrátě na majetku s nimi spojených (European Commission MEMO 13/547, 2013).

Tento projekt spolufinancovaný Evropskou unií předpokládá, že po plném nasazení dokáže v EU každý rok zachránit 2500 životů. eCall by měl urychlit dobu reakce při mimořádné události o 40 % v městských oblastech a o 50 % na venkově. Pilotní testování a zavádění systému eCall v Evropě provádí HeERO consorcium. Do projektu HeERO (Harmonised eCall European Pilot) je zapojena i Česká republika. Podle odhadů stojí ročně dopravní nehody okolo 160 miliard eur, v případě plného nasazení systému eCall by se mělo ušetřit až 20 miliard eur ročně (ECall deployment - Publication by the European Commission of the Delegated Regulation No 305/2013, 2013).

Dle posledních informací Evropská komise schválila, že systém eCall se automaticky aktivuje v případě, kdy senzory v automobilu detekují nehodu. Více o nasazení systému ecall do reálného provozu lze nalézt například v (Ecall becomes a reality!, 2015).

2 Data mining: metodologie a algoritmy

V současném moderním světě většina společností shromažďuje data o svých zákaznících. Tato data obsahují mnohdy i velice citlivé a osobní informace, které lze využívat pro nejrůznější účely. Obrovské datové základny ale nebudují jen velké korporace, ale například i státní organizace. Jestliže v dobách minulých znamenal data mining konkurenční výhodu, dnes se pomalu stává nepostradatelnou nutností a nachází se ve všech komerčních i nekomerčních sférách a nejrůznějších oborech. I samotný vznik data miningu byl podmíněn tím, že organizace shromažďovaly s vynaložením značných prostředků velké množství dat, a právě data mining umožňuje zhodnotit tyto prostředky.

Data mining můžeme chápat jako cílené prozkoumávání údajů, kdy na začátku často čerpáme z velkých objemů dat. Postupným zaměřováním se na konkrétní cíl se objem dat zužuje až po fázi modelování, která pracuje již s významně menší datovou množinou. Gartner group definuje data mining jak proces objevování významných netriviálních závislostí, vzorů a trendů prozkoumáváním velkých objemů dat pomocí algoritmů pro odhalování pravidel a pomocí matematických a statistických algoritmů.

Data mining využívá celá řada oborů. Ať už je obor, na který aplikujeme data miningové nástroje a postupy jakýkoliv, vždy je při řešení data miningových úloh kladen důraz na konkrétní nasazení do praxe. V praxi není podstatné, jaké modely při řešení problému použijeme, je však nutné, aby řešení jako celek fungovalo a přineslo něco užitečného.

Díky možnostem, které data miningové postupy nabízejí v případě zpracovávání velikých objemů dat, se využití data miningu při hledání skrytých závislostí u nehod, které tvoří v nějakém konkrétním místě shluk, jeví jako ideální prostředek.

2.1 Metodologie v data miningu

S postupem času začaly v data miningu vznikat metodologie, který byly následně všeobecně aplikovány na různé data miningové projekty. Zpočátku byly metodologie utvářeny především velkými společnostmi zabývajícími se tímto oborem. Příkladem takovéhoto postupu jsou např. metodologie SEMMA či 5A, které vytvořily společnosti SAS resp. SPSS. Tyto metodologie si byly velmi podobné a v podstatě řešily stejný problém. V roce 1996 byl evropskou komisí financován grant, který měl zajistit vytvoření jednotné univerzální metodologie pro data mining. Díky tomuto grantu vznikla v roce 1999 první verze volně dostupné metodologie

CRISP DM (Cross-Industry Standard Process for Data Mining). Tuto metodologii zaštiťovalo konsorcium společností SPSS, NCR, DAIMLER CHRYSLER a OHRA.

První dvě fáze metodologie jsou analytické. První fáze se zaměřuje na definici problému a rozplánování celého projektu. Jsou rovněž stanovena kritéria úspěšnosti, tj. očekávání přínosů daného projektu. Druhá fáze je věnována porozumění datům, která jsou v tomto kroku analyzována, je hodnocena jejich dostupnost a kvalita. Ve třetí fázi označované jako *Data Preparatio*n je vytvářena modelovací matice (tabulka), jež je dále využívána pro DM modely. Jde o časově nejnáročnější fázi celého projektu. Dále následuje modelování spočívající ve vytváření predikčních modelů. V případě používání sofistikovaných softwarových (sw) řešení je nutné tyto modely pouze správně nastavit. Po vytváření modelů následuje evaluace reprezentující kontrolu a vyhodnocení celého vlastního řešení. V rámci této fáze jsou odhadovány budoucí přínosy a je hodnoceno naplnění počátečních kritérií, která byla stanovena ve fázi *Business Understanding*.

Standardizace a otevřenost metodologie CRISP DM představují značnou výhodu, a proto lze tuto metodologii považovat za nejvhodnější rámec pro zpracování data miningové části této práce.

2.2 Použité DM nástroje

Vytváření modelů pro systém včasného varování před vysokým rizikem dopravní nehody je možné realizovat několika cestami. Jednou z možností je využití nejrůznějších softwarových produktů, více či méně univerzálních, které jsou vhodné pro naše potřeby. Cílem této kapitoly je ukázat možnosti vybraných data miningových nástrojů pro hledání shluků dopravních nehod s podobnými vlastnostmi. Nástroje budou porovnávány vzhledem k existenci jednotlivých potřebných algoritmů, rozšiřitelnosti softwarového nástroje a možnostem vizualizace výsledků v mapovém podkladu.

Pro porovnání možností při hledání shluků nehod s podobnými vlastnostmi byly vybrány následující softwarové nástroje a utility: IBM SPSS Modeler, KNIME, RapidMiner, Weka, Orange, ELKI. Tyto nástroje byly vybrány, jelikož se jedná o nejběžnější a uživatelsky nejoblíbenější data miningové nástroje.

Pro analýzu a vytváření prvotních modelů byl použit DM nástroj IBM SPSS modeler a open source nástroj KNIME. Postupně byly vyzkoušeny nejběžnější algoritmy shlukové analýzy, které by měly umožnit vytvořit mapu nebezpečných úseků na vozovkách.

Každý z těchto nástrojů má vzhledem k vyhledávání shluků v geografických datech a jejich následné reprezentaci pomocí mapových podkladů, své pozitivní, ale také negativní vlastnosti.

2.2.1 Vyhodnocení možností DM nástrojů pro využití v systému včasného varování

Analýza vybraných data miningových nástrojů ukázala, že pro potřeby vyhledávání shluků dopravních nehod s podobnými vlastnostmi není žádný z uvedených nástrojů plně vybaven. Ideálním nástrojem pro přípravu datové matice se zdá být IBM SPSS Modeler. Pro vyhledávání shluků na základě geografických dat jsou důležité optimalizované algoritmy jako DBSCAN či OPTICS. Ty jsou k dispozici pouze v nástroji ELKI, který ale nelze použít na žádnou další námi vyžadovanou činnost. Optimalizovanou verzi DBSCANu lze nalézt také v nástroji RapidMiner studio. Neoptimalizované verze se nacházejí v softwaru Weka a KNIME. Algoritmy generující asociační pravidla nabízejí všechny nástroje. Vizualizace dat do mapových podkladů je snadno realizovatelná v nástroji KNIME a při doinstalování potřebného balíčku je dostupná i v IBM SPSS Modeler. Většina vybraných nástrojů je dále rozšiřitelná o další balíčky a algoritmy. Všechny aplikace umožňují vytváření vlastních skriptů ve vybraných jazycích jako je JAVA, jazyk R či Python. Hodně možností při psaní vlastních skriptů poskytuje KNIME. Každý nástroj byl bodově ohodnocen pomocí sedmi kritérií. Každému kritériu byla přidělena určitá bodová maximální hranice dle jeho důležitosti

Jelikož v žádném nástroji nelze proces detekce shluků podobných vlastností zcela automatizovat, je nejlepší cestou implementace vlastního uzlu. I přesto, že například nástroj KNIME obsahuje jak uzly pro detekci shluků v geo datech, tak uzel pro hledání asociačních pravidel, nelze v prostředí KNIME automaticky a hromadně v detekovaných shlucích následně hledat asociační pravidla. Nejschůdnějším řešením je implementace vlastního upraveného algoritmu APRIORI do aplikace KNIME. Řešení, které za tímto účelem vzniklo, může využívat pro vyhledávání shluků vestavěné algoritmy DBSCAN a OPTICS a následně automaticky prohledávat všechny shluky a detekovat asociační pravidla v jednotlivých shlucích. Upravený uzel Apriori umožnuje kromě volby základních charakteristik asociačních pravidel i volbu proměnné, podle které se budou pravidla tzv. grupovat.

3 Vybrané algoritmy data miningu

V následujících podkapitolách jsou rozebrány vybrané algoritmy, jež je vhodné využít v řídící části navrhovaného systému včasného varování. Pozornost je věnována nejprve generování asociačních pravidel pomocí algoritmu Apriori, poté následuje kapitol zabývající se algoritmy shlukové analýzy.

3.1 Asociační pravidla a algoritmus Apriori

V běžném životě lidé často využívají konstrukty typu IF-THEN, aniž by si to mnohdy uvědomovali. Tyto zmíněné konstrukty lze proto najít ve všech běžných programovacích jazycích. Počátkem 90. let byl termín asociační pravidla zpopularizován v souvislosti s analýzou nákupního koše. Při analýze nákupního košíku je zjišťováno, jaké zboží zákazníci supermarketu nakupují společně. Jde tedy o hledání skrytých vzájemných vazeb mezi položkami, např. zboží v obchodě. Žádné zboží nemusí být v tomto případě na straně závěru pravidla. Asociační pravidla zapisujeme ve formě implikací ve tvaru: X <- A & B & C, přičemž A, B, a C jsou předpoklady a X je závěr. Pro vyhledávání asociačních pravidel je všeobecně nejpoužívanější algoritmus Apriori (Berka, 2005).

3.2 Shluková analýza

Jedním ze způsobů, jak vyhledávat nehody s podobnými vlastnostmi, je shluková analýza. Jak samotný název napovídá, jde v případě této metody využívané nejen v data miningu o vytváření shluků ze vstupních dat. Shluk je možné definovat jako množinu objektů mající podobné vlastnosti.

Shlukovací metody můžeme podle cíle členit na hierarchické a nehierarchické. Dále lze hierarchické metody dělit na aglomerativní (metoda nejbližšího souseda, metoda nejvzdálenějšího souseda, metoda centroidní) a divizní (např: MacNaughton-Smithův algoritmus).

Aglomerativní metody vycházejí ze shluků obsahujících pouze jeden objekt. Divizní metody mají zpočátku jediný shluk obsahující všechny objekty. V každém kroku rozkládání se jeden shluk rozdvojí, přičemž takové dělení pokračuje, dokud není počet shluků roven počtu objektů (Řezanková, 2009). V nehierarchických metodách lze ještě od sebe odlišit metody zachovávající počet shluků a metody optimalizující počet shluků. Někdy jsou tyto algoritmy

označovány jako rozdělovací. Mezi metody zachovávající počet shluků lze zařadit algoritmus K-Means, jehož modifikovanou variantou disponuje i nástroj IBM SPSS Modeler. Jako nejvhodnější algoritmy pro hledání shluků v geografických datech se jeví algoritmy DBSCAN a OPTICS patřící do skupiny algoritmů založených na hustotě. Metody založené na hustotě umožňují hledat shluky libovolného tvaru. Výhodou těchto metod je, že narozdíl od metod založených na rozdělování umí nalézt i shluky jiného než kulovitého tvaru. Shluky algoritmů založených na hustotě jsou chápány jako oblasti s velkou hustotou a jsou od sebe odděleny šumem, tj. oblastmi s hustotou nízkou. To znamená, že některé objekty nemusí být přiřazeny do žádného shluku (Řezanková, 2009).

Přehled shlukovacích metod

Rozdělovací metody:

- Nacházejí vzájemně vylučující se kulové shluky
- Principiálně jsou založeny na základě vzájemné vzdálenosti bodů
- Mohou používat nejbližší středy nebo medoidy pro reprezentaci středu shluku
- Jsou efektivní v případech malých až středně velkých souborů dat

Hierarchické metody:

- Shlukování je provedeno přes hierarchický rozklad (několika-úrovňový)
- Nedokáží opravit chybné sloučení či rozdělení
- Mohou zapojovat další techniky jako například mikro-shlukování či brát v potaz "propojení" objektů

Metody založené na hustotě:

- Dokáží nalézt shluky libovolných tvarů
- Shluky jsou oblasti s velkou hustotou objektů v prostoru, které jsou odděleny oblastmi o nízké hustotě(šumem)
- Hustota shluku: Každý bod musí mít ve svém "sousedství" alespoň dané minimum ostatních bodů
- Mohou odfiltrovat šum

Mřížkové metody:

- Využívají mřížkové struktury dat o několika rozlišeních
- Jsou pro ně typické rychlé výpočty (typicky nezávislé na počtu datových objektů, však závislé na velikosti mřížky)

4 Cíle práce, metody a doporučení pro další práci

Hlavním cílem této práce je vytvořit konceptuální návrh systému umožňujícího v reálném čase a místě predikovat riziko dopravní nehody. Systém je založený na využívání predikčních modelů, které jsou vytvářeny pomocí data miningových technik a nástrojů využívajících především algoritmy shlukové analýzy a asociačních pravidel. Dílčími cíli práce jsou analýza současného stavu systémů zvyšujících bezpečnost účastníků silničního provozu, analýza využitelnosti technik dobývání znalostí z databází a určení vhodných algoritmů a postupů v souvislosti s realizací daného systému. Dalšími výsledky by mělo být vytvoření modelů schopných identifikovat nebezpečná místa na silničních komunikacích.

4.1 Metody a fáze výzkumu

Při zpracovávání práce je vycházeno z rešerše dostupných publikací v databázích, knižních zdrojů a zdrojů dostupných online. V první fázi byla provedena rešerše stávajících systémů zvyšujících bezpečnost účastníků silničního provozu. V současné době existuje řada systémů zvyšujících bezpečnost v dopravě, ale doposud nebyla realizována myšlenka systému včasného varování popsaného v této disertační práci.

V druhé fázi bylo nutné získat data, za pomoci nichž by bylo možné postavit systém včasného varování. Jednou variantou je simulace dat užívaných pro predikci. Od této možnosti bylo prozatím upuštěno. Aby byla práce postavena na skutečných datech, byla provedena analýza současných možností pro získání a zpracování relevantních dat o dopravních nehodách.

Současně bylo nutné provést výběr vhodných nástrojů a algoritmů. Za tímto účelem byla provedena rešerše a využita forma analytického zhodnocení umožňující výběr algoritmů a nástrojů pro vyhledávání shluků a následné dolování informací s využitím asociačních pravidel. Na základě analytického zhodnocení vhodných data miningových nástrojů jsou syntetizovány poznatky a vytvořena doporučení pro optimální řešení zkoumané problematiky.

V současné chvíli je vyřešen návrh tzv. řídicí části systému včetně popisu funkce jeho jednotlivých částí. V dalších etapách se výzkum zaměří na rozšíření funkcí implementovaného systému. V tomto kontextu je nutné se zabývat automatizovanou distribucí dat o dopravních nehodách a dat o hustotě dopravy na jednotlivých komunikacích do systému včasného varování. V době vzniku této práce byla k dispozici pouze data o sčítání dopravy z roku 2010. Jelikož se

sčítání dopravy provádí většinou jednou za 5 let, je nutné zajistit aktualizaci a analýzu těchto nových dat.

Na základě analýzy datové matice lze vytvářet doporučení například pro Policii ČR nejen z hlediska prevence, ale také doporučení pro změnu struktury dat, která policie sbírá. V době vzniku předem definované struktury sbíraných dat, kterou policie používá, nebylo počítáno s tím, že by data šla využít pro predikci rizika vzniku nehody v reálném čase a místě. V případě, že by policie realizovala tyto změny, došlo by jistě k významnému kvalitativnímu skoku v predikci rizika vzniku nehody. Zvětšení datové matice o další atributy a zpřesnění hodnot, jichž atributy nabývají, by mohlo přinést také zavedení systému eCall, jehož jednotky v nově vyrobených automobilech zasílají řadu dalších dat do tísňového centra v blízkosti nehody. Další fází výzkumu bude tedy i návrh propojení systému eCall se systémem včasného varování před zvýšeným rizikem dopravní nehody popisovaným v této práci.

V neposlední řadě byl v rámci této doktorské disertační práce vytvořen systémový návrh konceptu systému včasného varování tak, aby řešení bylo co nejuniverzálnější a dále rozšiřitelné.

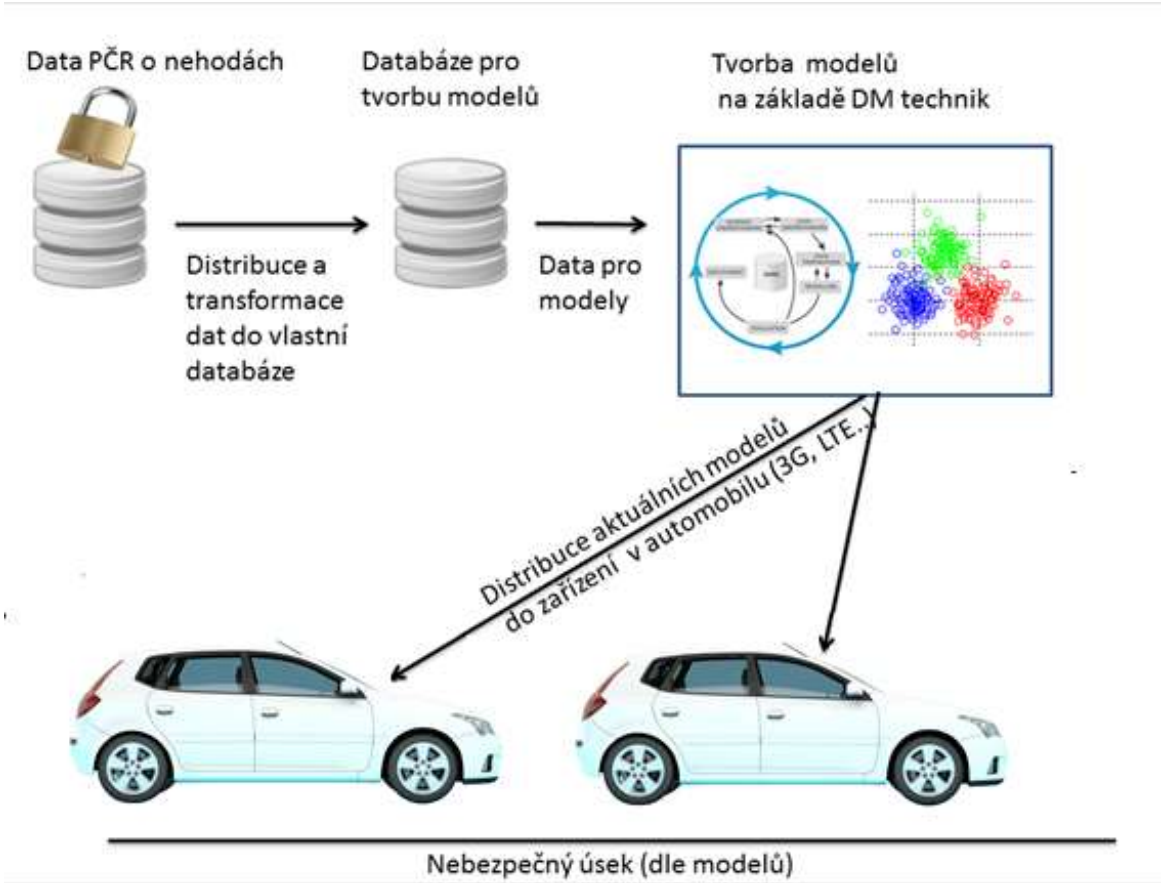
Navrhované postupy, techniky a samotný princip systému včasného varování je možné využít i na zcela odlišné typy úloh s typem dat stejného charakteru týkajících se však zcela jiné oblasti zájmu. Příkladem využití navrhovaného řešení může být vyhledávání skrytých závislostí v datech týkajících se kriminální činnosti.

Budoucí vývoj navrhovaného systému včasného varování by se měl zabývat především samotnou implementací jednotlivých bloků řídicí a uživatelské části. Z hlediska uživatelské části, je dále vyvíjena mobilní aplikace realizující samotné varování. Dále je počítáno s dalším vývojem vlastní aplikace umožňující provádět hromadnou a automatizovanou realizaci činností řídicí části systému

5 Princip systému včasného varování před dopravní nehodou

Ve zkratce lze konceptuální návrh systému včasného varování před dopravní nehodou popsat jako komplexní systém využívající vytvořených modelů předpovídajících na základě aktuální polohy a dalších atributů riziko nehody v reálném čase. Systém včasného varování před vysokým rizikem dopravní nehody se skládá ze dvou základních částí. První řídicí část zajišťuje shromažďování, zpracovávání a distribuci dat o dopravních nehodách. Druhým úkolem řídicí části je tvorba a distribuce predikčních modelů. Uživatelská část systému v reálném čase vyhodnocuje situaci a vhodným způsobem informuje řidiče o vysokém riziku dopravní nehody. Schéma vystihující nejdůležitější části systému je znázorněno na obrázku 1.

Obrázek 1: Princip systému

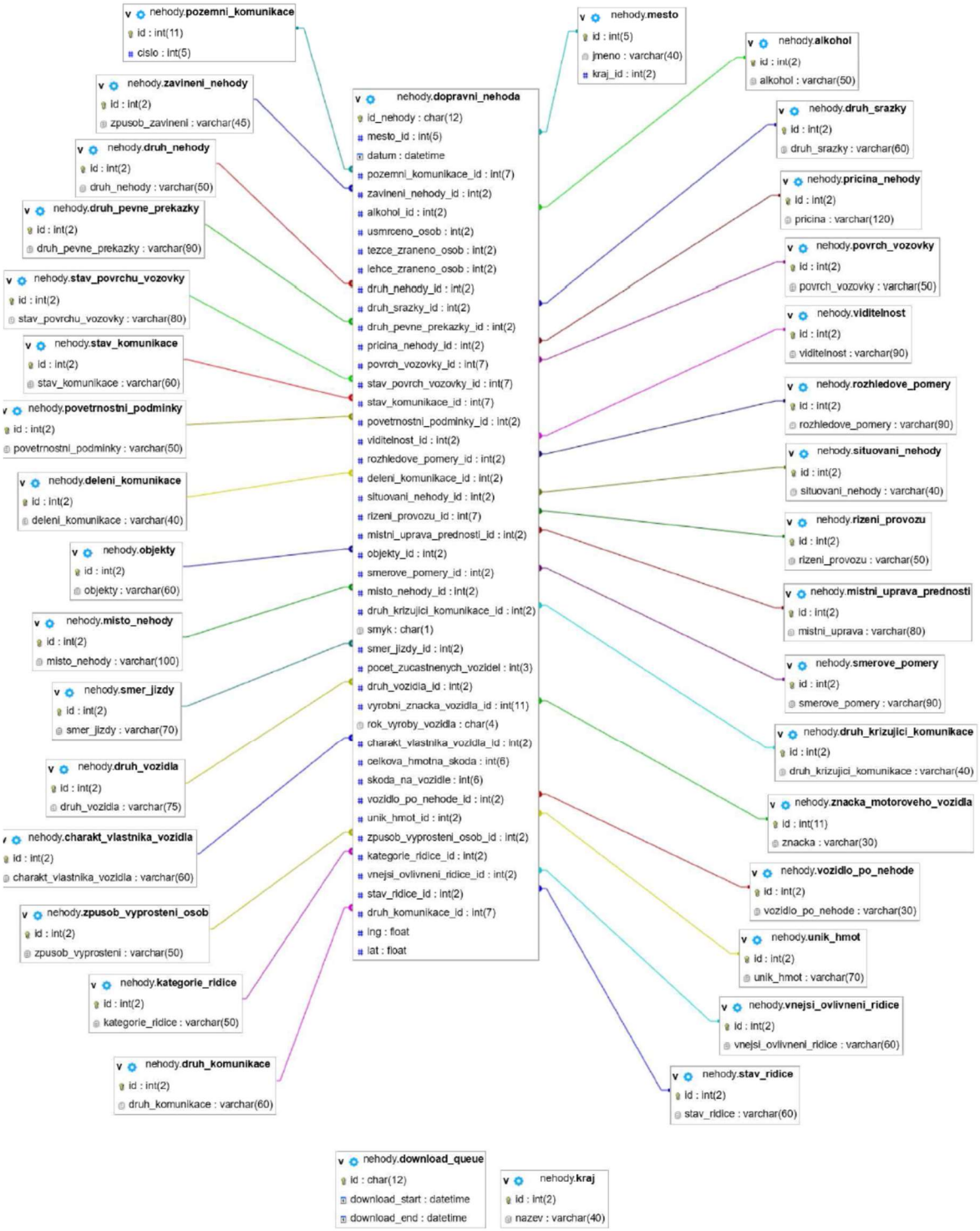


Zdroj: vlastní zpracování

5.1 Řídicí část

Řídicí část sytému slouží pro získávání a zpracovávání heterogenních dat o dopravních nehodách od Policie České republiky, resp. Ministerstva dopravy.

Obrázek 2: Konceptuální návrh vlastní databáze dopravních nehod (MySQL)



Zdroj: vlastní zpracování

Získávání dat z jednotné dopravní mapy PČR je v tuto chvíli možné pomocí skriptů napsaných v linuxového programu cURL, což je nástroj pro přenos dat po protokolech jako HTTP, FTP a dalších. Takto získaná data je třeba ještě pomocí dalších skriptů importovat do databáze pro

pozdější zpracovávání. Součástí řídicího systému je tedy databáze dopravních nehod sloužící nejen pro potřeby data miningu. Prozatím jsou data ukládána do relační databáze MySQL běžící na linuxovém stroji. Za účelem ukládání dat do databáze byl navržen vlastní konceptuální model databáze, respektující okolnosti zaznamenávané Policií ČR při dopravní nehodě. Aktuální konceptuální model databáze je zobrazen na obrázku 2. Pomocí vhodných skriptů jsou získaná data transformována do potřebné podoby a následně ukládána do databáze.

Zmiňovaná databáze slouží pro potřeby vývoje při vytváření modelů popisujících aspekty nehod. Ve druhé fázi je databáze využívána pro vytvoření komplexních modelů předpovídající riziko nehody v daném čase a místě.

5.2 Uživatelská část

Uživatelská část systému předpokládá využívání predikčních modelů ve speciálním zařízení automobilu (klient), jež v reálném čase vyhodnocuje riziko nehody v závislosti na čase, aktuální poloze, stavu vozovky, stavu automobilu a počasí a dalších atributů reflektujících aktuální situaci. V dnešních automobilech je většina těchto informací běžně dostupná z čidel, kterými vozidla disponují. Zařízení v automobilu totiž porovnává aktuální situaci s výsledkem predikce a v případě vysoké míry podobnosti aktuální situace s predikcí řidiče upozorní. Řidič vozidla je upozorněn v případě, že se v daném místě za obdobných podmínek stalo více podobných dopravních nehod. V praxi se však setkáme s nebezpečnými místy, která nebudou specifická (např. časem, povětrnostními podmínkami a dalšími atributy). Taková místa jsou označována jako obecné shluky a jsou vytvořena pouze na základě častého výskytu dopravních nehod v dané lokalitě.

Po hardwarové stránce může být uživatelská část realizována velmi obdobným způsobem jako zařízení typu tablet, či mobilní telefon. Nejdůležitějším faktorem u těchto zařízení je samotný výkon mikroprocesoru a velikost paměti. Velikost paměti a výkon procesoru jsou důležitými parametry z hlediska vytváření rozhodování a varování řidiče na základě modelů. Druhým faktorem, který je důležitý především z hlediska bezpečnosti, je dostatečná velikost a čitelnost dotykové obrazovky ovládající uživatelskou část systému v automobilu. Dalším prvkem uživatelské části zvyšujícím uživatelský komfort a s tím velmi úzce související bezpečnost provozu celého systému je hlasové ovládání, které by mělo umožňovat jednoduchými příkazy měnit základní uživatelská nastavení. Propojení zařízení realizujícího uživatelskou část

systému včasného varování s automobilem by mělo být realizováno pomocí standartních komunikačních protokolů používaných pro diagnostiku vozů.

Jednou z variant implementace uživatelské části popisovaného systému včasného varování je také její zabudování přímo do infotainmentu vozidla. To by však vyžadovalo spolupráci jednotlivých výrobců automobilů.

5.3 Spolupráce řídící a uživatelské části

V principu může být fungování uživatelské části realizováno třemi způsoby. Veškerá komunikace predikčního zařízení v automobilu s řídicí částí by měla probíhat přes wifi či LTE sítě čtvrté generace. (Lamr, 2015a)

První přístup (varianta A) předpokládá, že již v řídicí části budou vybudovány a do mapového podkladu přidány tzv. heat mapy, které budou představovat jakási „horká místa“, kde se vyskytlo více dopravních nehod podobného charakteru. Analýza takových shluků by měla ukázat, zda jsou i pro takové shluky nějaká další společná kritéria jako je například roční období, resp. měsíc, denní doba, teplota, či povětrnostní podmínky. V případě existence dalších aspektů ovlivňujících dané místo nehody, bude takovýto shluk reprezentován jako jeden bod se specifickými vlastnostmi. V případě, že nebude prokázána specifická charakteristika pro dané místo, bude takový shluk označen všeobecně jako „místo častých dopravních nehod“. Zařízení v automobilu v reálném čase kontroluje pohyb vozidla a v případě, že se vozidlo bude blížit k nějakému shluku reprezentujícímu častější výskyt nehod označené jako „místo častých dopravních nehod“ bude řidič pouze upozorněn, že projíždí místem častých dopravních nehod. Pokud se bude vozidlo blížit ke shluku, který bude charakteristický specifickými podmínkami, zařízení v automobilu porovná realitu se specifickými podmínkami blížícího se shluku a řidiče upozorní na vysoké riziko nehody vzhledem k tomu, že se blíží k místu, kde se za podobných podmínek stalo vícero nehod. S tímto přístupem by bylo nutné pravidelně aktualizovat celý mapový podklad s vygenerovanými shluky do uživatelských zařízení v automobilu, zařízení by však pro korektní fungování nemuselo být stále online.

Druhý možný přístup (varianta B) využívá řídicí část spíše jako výkonnou databázi shluků dopravních nehod a více operací je přesunuto na klienta v uživatelské části. Analýzou dat v řídicí části by byly vyřazeny nevýznamné atributy nepotřebné k predikcím, a tak by byla zredukována datová náročnost při přenosu informací o nehodách ke klientovi. Shluky v řídicí části by byly vytvářeny pouze na základě počtu výskytů dopravních nehod a jejich

geografických souřadnic. Zařízení v automobilu by, stejně jako v předchozím přístupu, monitorovalo pohyb automobilu, ale ze serveru by stahovalo s předstihem data o nehodách z míst, do kterých se automobil blíží. Každý takto stažený shluk by byl testován z hlediska ročního období (měsíce), data, času, povětrnostních podmínek či příčiny nehody (např.: nedodržení bezpečné vzdálenosti za vozidlem). Takový přístup by odstranil nutnost pravidelné masivní aktualizace mapových podkladů klienta, jelikož by se v zařízení tato data neudržovala, ale stahovala se vždy jen data na krátkodobé použití. Nevýhodou je nutnost stálého připojení klienta k internetu, což by mohlo znamenat v případě nedostupnosti datového připojení určitá úskalí. Tím, že by se všechny shluky z databáze testovaly na svoji specifičnost (např.: místo s nehodami za deště, místo s nehodami večer) online v zařízení by vzrostly i hardwarové nároky na klienta.

Třetí možný přístup (varianta C) by mohl být kombinací prvních dvou. Stejně jako v případě varianty A by v řídicí části byly v databázi uloženy nejen shluky všeobecné vytvořené na základě počtu nehod v určitém místě, ale byly by zde i shluky specifické. Stahování dat do klienta by probíhalo obdobně jako v případě varianty B, tedy stahovala by se vždy data potřebná pro aktuální potřebu predikce v daném místě na trase vozidla. Varianta C se v tuto chvíli jeví jako nejoptimálnější varianta pro reálné použití.

5.4 Systémový návrh řešení a pohled z hlediska obecné teorie systémů

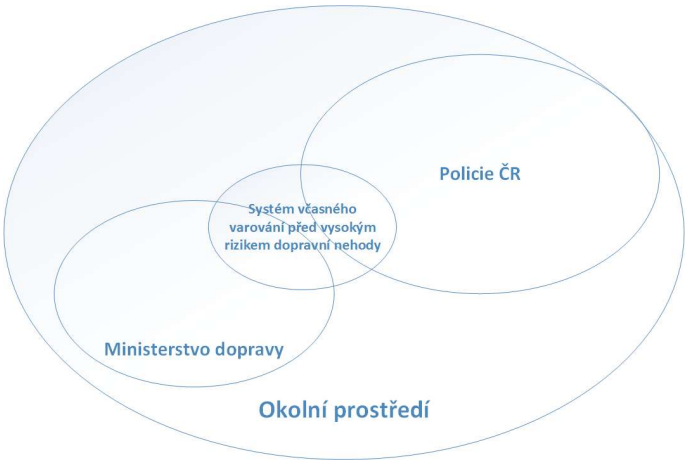
Systém lze definovat jako účelově uspořádanou množinu prvků a množinu vazeb mezi nimi, s dynamickým chováním, které společně určují vlastnosti celku. V rámci dekompozice daného systému je možné vyčlenit podsystém. Podsystémem rozumíme podmnožinu systémových prvků a vazeb, která je z nějakého důvodu vyčleněna ze systému a je chápana jako nový systém nebo jako prvek. Každý systém tedy obsahuje prvky a vazby, které tvoří základ struktury systému.

5.4.1 Místo systému v hierarchii dopravně bezpečnostních orgánů

Z hlediska provozování systému a jeho dalšího vývoje by měl systém včasného varování před vysokým rizikem dopravní nehody patřit (být podsystémem) pod vyšší autoritu jako je například Ministerstvo dopravy, Ministerstvo vnitra či Policie ČR. Vzhledem k tomu, že nedílnou součástí systému jsou data o dopravních nehodách, kterými disponuje Ministerstvo dopravy, se jeví jako logické, aby systém z větší části spravovalo Ministerstvo dopravy. To však veškerá data dostává od Policie ČR, která data o každé dopravní nehodě (od roku 2008 o

každé nehodě nad 100 000 Kč) zaznamenává. Policie by do systému včasného varování zasahovala nejen poskytováním dat o nehodách, ale také je zodpovědná za jejich kvalitu, formát a zjišťované atributy při šetření dopravní nehody a není jednoduše možné z vlivu na systém včasného varování před vysokým rizikem dopravní nehody policii vynechat. Grafické vyjádření výše popsaného je zobrazeno na obrázku 3. Ministerstvo dopravy však provozuje i další systémy, které mají zvýšit bezpečnost na komunikacích. Mezi nejvýznamnější telematické systémy, které mají zvyšovat bezpečnost na silničních komunikacích v ČR patří „systém informačních tabulí“, a systém šíření dopravních informací pomocí technologie RDS-TMC. Systém včasného varování před vysokým rizikem dopravní nehody se od zmíněných systémů liší především tím, že se snaží dopravním nehodám aktivně předcházet, kdežto zmíněné systémy řeší situaci ve chvíli, kdy se nehoda stane. Existovat vedle sebe mohou paralelně všechny zmíněné systémy.

Obrázek 3: Systém včasného varování jeho místo v systémové hierarchii dopravně bezpečnostních orgánů



Zdroj: vlastní zpracování

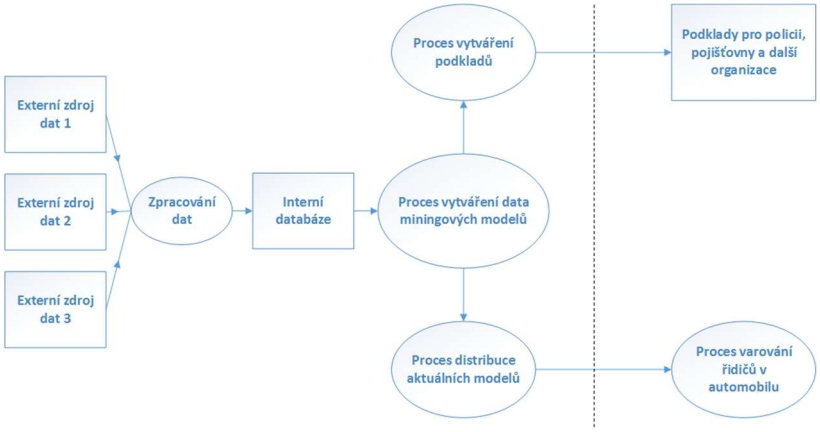
5.4.2 Prvky a vazby systému včasného varování

Jak již bylo řečeno, systém včasného varování před zvýšeným rizikem dopravní nehody podléhá jistým suprasystémům, ale na druhou stranu i on se skládá z určitých prvků a vazeb mezi nimi, přičemž některé procesy lze definovat jako podsystémy systému včasného varování. Hranice systému jsou na jedné straně tvořeny vstupy v podobě externích zdrojů dat a na straně druhé končí varováním řidiče v reálném čase a místě.

Pomyslná hranice mezi řídicí částí a uživatelskou částí systému je na obrázku 4 znázorněna čárkovanou čárou. Data z externích zdrojů (Obrázek 4) v systému včasného varování

reprezentují například databázi dopravních nehod Policie ČR, která však nemusí být jediným datovým zdrojem navrhovaného systému. Pro vytváření predikčních modelů je k dispozici v tuto chvíli více jak 700 000 záznamů o dopravních nehodách na území ČR, přičemž každý záznam obsahuje 44 atributů s číselnými a kategoriálními daty.

Obrázek 4: Prvky systému včasného varování



Zdroj: vlastní zpracování

Díky různorodosti potenciálních dalších datových zdrojů je nutné data zpracovat (proces zpracování dat) a uložit do interní databáze systému. Jak již bylo řečeno, data o dopravních nehodách od policie nemusí být jediným datovým zdrojem, který systém využívá.

Dalším příkladem externích zdrojů dat může být databáze výsledků celostátního sčítání dopravy. Základní výsledky sčítání dopravy včetně metodiky sběru dat lze nalézt na webové prezentaci projektu (Prezentace výsledků sčítání dopravy 2010, 2011). Jednotlivé komunikace jsou rozděleny na úseky a na těchto úsecích jsou počítány četnosti různých druhu dopravních prostředků (osobní automobily, autobusy, nákladní vozy atd..) a také je k dispozici již vypočtená relativní hustota provozu v daném místě. Data však nejsou jednoduše přístupná a je nutné je spojit s jednotlivými záznamy o dopravních nehodách.

Stěžejním je pro systém včasného varování subsystém vytváření data miningových modelů, které budou sloužit pro predikci nebezpečných míst za daných okolností. I když by bylo vhodné co nejvíce tento subsystém automatizovat, aby charakter co nejvíce odpovídal systému tvrdému, nelze to v případě vytváření modelů zcela realizovat a vždy bude nutné, aby do procesu vstupoval lidský faktor. V procesu vytváření data minigových modelů je třeba načítat z databáze informace o dopravních nehodách a vytvořit tzv. modelovací matici (data denormalizovat). Data v modelovací matici procházejí analýzou například pomocí data

miningového nástroje a následně jsou v datech pomocí algoritmů hledány vzory sloužící pro budoucí predikci.

Vytvořené modely je třeba následně distribuovat do uživatelské části systému, což zajišťuje proces distribuce aktuálních modelů. Přenos informací by měl být možný pomocí několika alternativních přenosových kanálů.

Proces varování řidičů v automobilu (uživatelská část systému) předpokládá využívání predikčních modelů ve speciálním zařízení automobilu (klient), které v reálném čase vyhodnocuje riziko nehody v závislosti na čase, aktuální poloze, stavu vozovky, stavu automobilu a počasí a dalších atributů reflektujících aktuální situaci.

5.5 Ekonomické zhodnocení systému včasného varování

V souvislosti se zvyšujícím se počtem dopravních nehod je nutné zamyslet se nad ekonomicko-společenskými důsledky souvisejícími nejen s hmotnými škodami, které jenom v roce 2017 dosáhly 6, 316 miliardy Kč, ale především nelze opomenout újmy na lidských životech. Zatímco materiální škody je velmi snadné vyčíslit, v případě lidských životů lze uvést pouze přibližné hodnoty vyjádřené pomocí současných plnění poskytovaných českými pojišťovnami.

Nejčastější a také nejbezpečnější důsledek nárazu při autonehodách bývá zranění hlavy. Ke zraněním hlavy dochází při více než 70% dopravních nehod a jsou také nejfrekventovanější příčinou úmrtí. Mezi další častá zranění patří také poranění hrudníku či páteře, poranění kostí a zlomeniny, popáleniny a šokové stavy.

Ekonomická náročnost samotné implementace systému včasného varování do automobilů závisí na konstrukční náročnosti řídicí a uživatelské části systému.

V řídící i uživatelské části je kromě vývoje softwaru třeba počítat i s náklady na hardwarové části. Vzhledem k tomu, že vývoj aplikace řídící části a vývoj aplikace jedné z variant uživatelské části systému včasného varování je tvořen v rámci disertační práce, a bude dále řešen v rámci budoucího výzkumu, je nutné podotknout, že náklady by připadly pouze na pořízení dostatečně robustního hardwaru.

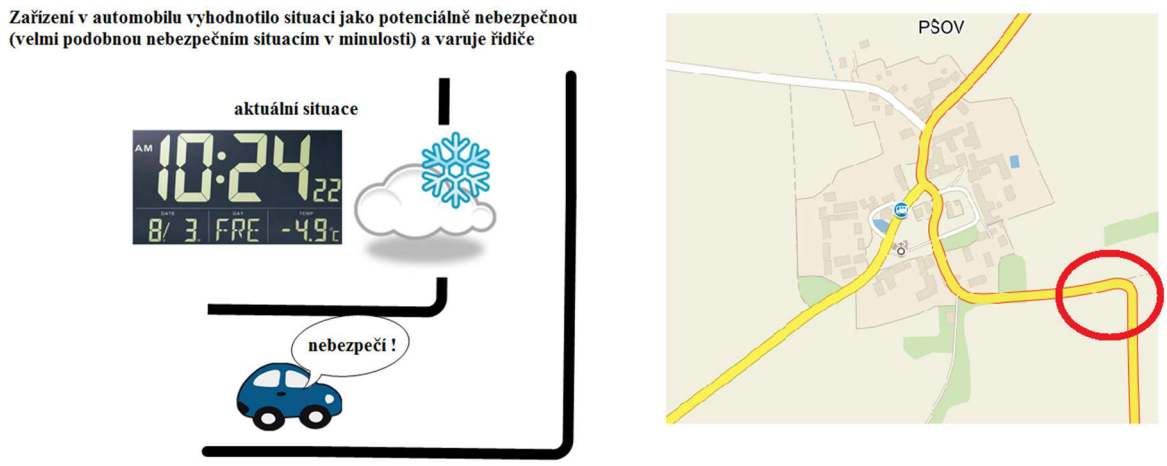
V řídící části systému je nutné disponovat serverem, na němž bude probíhat konverze heterogenních dat do databáze, dále serverem, kde bude prováděno vytváření modelů a servery zajištujícími distribuci dat do klientských zařízení.

Uživatelská část může být realizována třemi variantami. První varianta předpokládá vývoj speciálního zařízení v automobilu, které by však mělo být ze své podstaty velmi podobné zařízení tabletového typu a bylo by třeba vyřešit především implementaci konektoru pro spojení s automobilem. Druhá varianta předpokládá řešení uživatelské části formou mobilní aplikace, která je vyvíjena v rámci výzkumu autora této práce a pro uživatele by byla dostupná volně ke stažení. Uživatel by byl nucen pořídit si, dle konkrétního typu automobilu, konektor umožňující komunikaci automobilu s telefonem prostřednictvím bluetooth. Třetí varianta předpokládá implementaci uživatelské části systému do infotainmentu automobilů. Náklady s tím spojené by připadly na úkor jednotlivých výrobců automobilů.

5.6 Modelové situace

Princip systému včasného varování před dopravní nehodou lze ilustrovat na zatáčce, která je označována Centrem dopravního výzkumu ČR jako nejnebezpečnější zatáčka z celé sítě silnic první třídy a dálnic v ČR. Danou situaci ilustruje obrázek 5.

Obrázek 5: Ilustrace fungování systému v nebezpečném místě



Zdroj: vlastní zpracování

Tato zatáčka se nachází se na silnici 1. třídy s číslem 27 u obce Pšov na Podbořansku. Silnici, která je významným tahem mezi Ústeckým krajem a Plzní, ve velké míře využívají kamiony. Často se při projíždění ostré zatáčky dostávají minimálně částí vozidla do protisměru a ohrožují tak řidiče, kteří přijíždějí ve směru od Pšova. Mnohá auta jedou ještě pár desítek metrů před záhybem silnice výrazně vyšší rychlostí, než je zde povolená třicítka. Pokud je pak silnice

mokrá, nebo je na ní dokonce sníh či náledí, auto se snadno dostane do smyku a havárie je téměř neodvratná. Je evidentní, že i přes několikanásobné značení nebezpečného místa řidiči toto značení podceňují. Riziko nehody se v tomto místě rapidně zvyšuje v případě sněhové pokrývky nebo náledí. Ve chvíli, kdy bude automobil vybaven systémem včasného varování před vysokým nebezpečím nehody, který porovnává aktuální povětrnostní situaci s výsledkem predikčních modelů, je možné řidiče varovat pouze v situacích, které jsou opravdu nebezpečné. Věříme, že v případě, že budou řidiči dostávat pouze adekvátní informace, které jsou z hlediska bezpečnosti opravdu zásadní, budou mít v tento systém důvěru a přizpůsobí styl jízdy v případě potřeby. (Lamr, 2016c)

Pro ilustraci vyhledávání shluků jsme z databáze dopravních nehod vybrali geografickou oblast spadající dle policie ČR pod město Podbořany. Tato množina obsahuje 430 záznamů o dopravních nehodách za období let 2007-2013.

Oblast nebezpečné zatáčky je označena algoritmem DBSACN jako Cluster 55 a obsahuje 71 případů. Podívejme se nyní detailněji na tuto množinu. Nejvíce nehod se stává mezi druhou a třetí hodinou odpoledne. Většina nehod, které se zde odehrály, naštěstí není tragických, není zde výskyt smrtelného zranění, a jen ve dvou případech došlo k těžkému zranění. Zhruba ve třiceti procentech nehod však dochází k lehkým zraněním, přičemž ve většině případů jde o zranění pouze jedné osoby.

Zajímavé je, že celá polovina ze všech nehod se tu stává během června, července a srpna. Budeme-li dále zkoumat tuto skupinu nehod, zjistíme, že v 90 procentech nehod bylo jako příčina nehody označeno nepřizpůsobení rychlosti dopravně technickému stavu vozovky. V 63 procentech případů nehod situaci zkomplikovaly povětrnostní podmínky v podobě deště. Ve více jak třech čtvrtinách případů byl atribut *Stav povrchu vozovky* označen jako mokrý. U 55 procent případů byla Policií ČR označena viditelnost jako snížená právě vlivem povětrnostních podmínek.

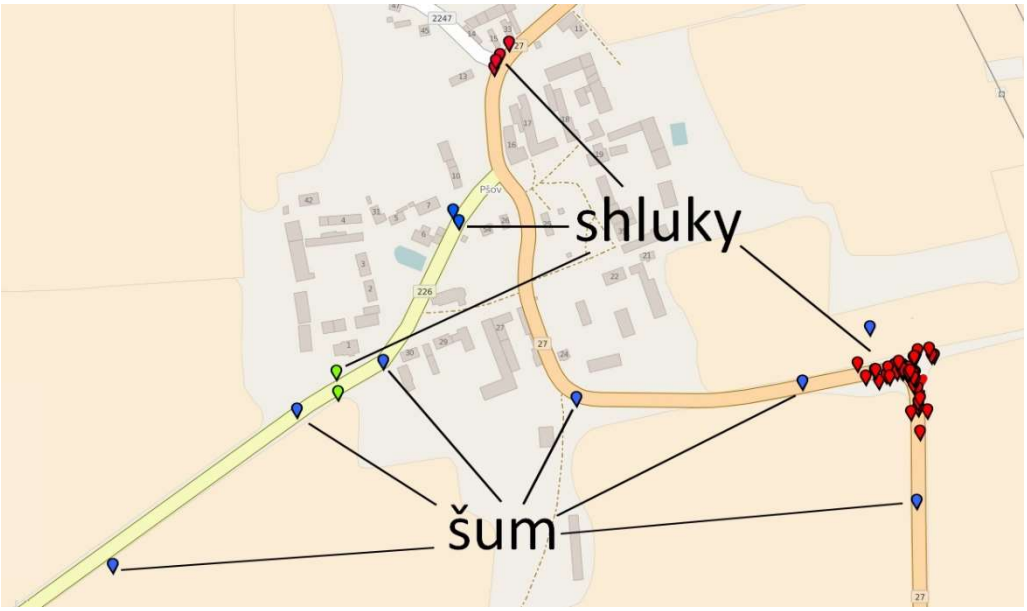
Porovnat výsledky algoritmů založených na hustotě (DBSCAN) a algoritmu K-means lze v obrázcích 6 a 7. Na obrázku 6 jsou shluky vytvořeny pomocí algoritmu DBSCAN v software KNIME. Shluky byly vytvářeny na datové matici obsahující 431 záznamů o dopravních nehodách v okolí “nejnebezpečnější” zatáčky v ČR, která se nachází u obce Pšov. V případě algoritmu DBSCAN se jako nejoptimálnější volba jeho parametrů s ohledem na počet záznamů jeví eps=0,025 a MinPts=2. Při daném nastavení nalezl algoritmus v dané oblasti 87 clusterů, resp. 86 clusterů a posledním clusterem byl šum. Jako šum jsou ve většině případů označeny

odlehlé hodnoty, což je v pořádku. V následujících obrázcích jsou jednotlivé záznamy o dopravních nehodách patřící do stejného clusteru vyznačeny stejnou barvou.

Pro srovnání jsou v obrázku 7 zobrazeny shluky dopravních nehod vytvořené pomocí shlukovacího algoritmu K-Means. Abychom dosáhli stejného počtu shluků jako v předchozím případě, nastavili jsme k=87. A lgoritmus K-means však nepočítá s šumem a shluky které vytvořil, nejsou určeny tak dobře jako pomocí algoritmu DBSCAN. Z nehod v inkriminované zatáčce, které zjevně patří do jednoho stejného shluku, vytvořil shluků hned několik. V místě zmiňované zatáčky s větší koncentrací dopravních nehod tedy algoritmus selhal. V oblastech s menší koncentrací nehod algoritmus vytvořil některé shluky dle očekávání. Dále jsou v obrázku 7 vyznačena dvě místa, která algoritmus K-means označil jako jeden cluster, i když spolu evidentně nesouvisí. Hlavní nevýhodou algoritmu K-means v případě použití pro hledání shluků v geografických datech je nutnost zadat předpokládaný počet shluků. (Lamr, 2016c)

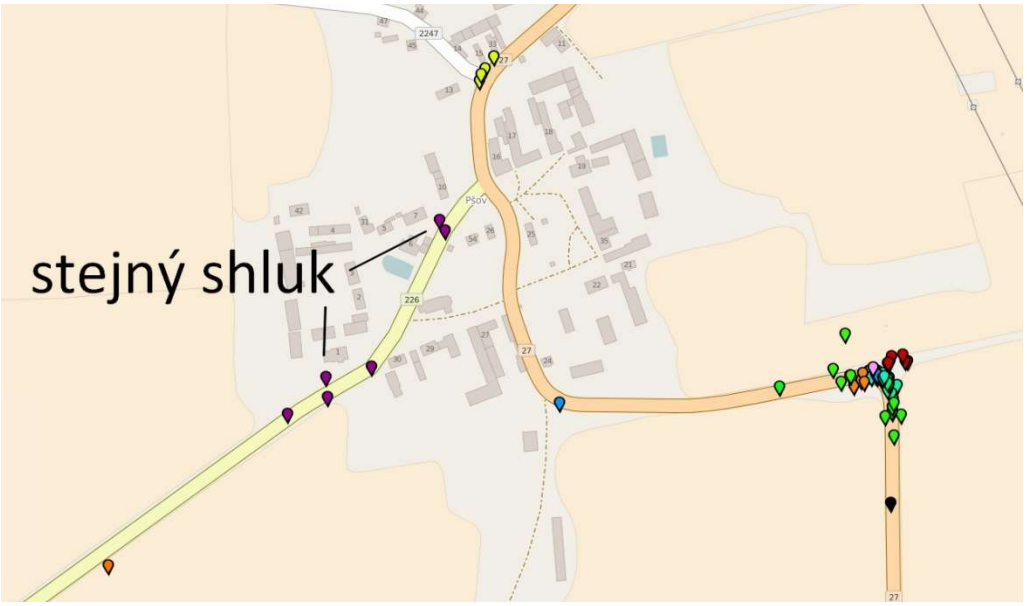
Systém včasného varování by měl být užitečný v případech, kdy řidič jede danou trasou poprvé. V případech, kdy jede řidič takovým místem prvně, může situaci na první pohled vyhodnotit zdánlivě za méně zákeřnou a díky nepřiměřené rychlosti, kterou vjíždí do zatáčky, nezvládne řízení. Ve chvíli, kdy bude řidič přijíždějící do daného místa včas upozorněn na vysoké riziko nebezpečí nehody, by mohlo dojít ke snížení počtu podobných nehod. (Lamr, 2016c)

Obrázek 6: Shluky nehod v okolí nejnebezpečnější zatáčky v ČR (DBSACN eps=0,025, MinPts=2)



Zdroj: vlastní zpracování

Obrázek 7: Shluky nehod v okolí nejnebezpečnější zatáčky v ČR (K-means k=87)



Zdroj: vlastní zpracování

Závěr

Se zvyšujícím se počtem nehod se policie opakovaně snaží zvrátit nepříznivý stav na dnešních silničních komunikacích pomocí represivních zásahů. S vědomím tohoto nepříznivého trendu by se měla společnost snažit přicházet především s novými nápady řešícími nastíněné problémy. Myslím si, že příznivé výsledky by měla přinášet především systémová řešení a využívání nových informačních a telekomunikačních technologií. V dnešní době disponujeme prostředky, které nám pomáhají rychle budovat obrovské datové základny. Využívání dat o dopravních nehodách společně s algoritmy pro získávání užitečných vzorů je dle mého názoru jednou z cest, jak by se silniční komunikace mohly stát bezpečnějším místem pro řidiče.

Datová základna využívaná v rámci této práce disponuje obrovským množstvím dat, která se dají využívat pro hledání skrytých vzorů a trendů. Databáze dopravních nehod, kterou vytváří Policie České republiky však ne vždy postihuje všechny okolnosti, za kterých se nehoda stala. Tento fakt stěžuje detekci nebezpečných míst na vozovkách, která jsou něčím specifická.

Hledání shluků dopravních nehod je nedílnou součástí vývoje systému včasného varování před zvýšeným rizikem dopravní nehody. Po důkladné analýze byly vybrány a v práci popsány nejvhodnější metody a postupy shlukové analýzy pro řešenou problematiku. Shlukovací algoritmy založené na hustotě (DBSCAN, OPTICS) přinášejí dobré výsledky při vytváření shluků dopravních nehod založených na geografických datech, jelikož předpokládají existenci šumu (nehod, které nepatří do žádného shluku). Pro získávání lepších výsledků při detekci shluků je třeba do datové matice o nehodách získat informaci o hustotě dopravy. Tak je možné vyhledávat shluky odděleně v místech s rozdílnou dopravní hustotou. Sběr informací o hustotě dopravy provádí každých 5 let Ředitelství silnic a dálnic. Reálně je ale problém datové matice spojit, jelikož neexistuje jednoznačný klíč, pomocí kterého je to možné provést. V rámci této práce byl navržen způsob, který spojení obou databází řeší.

Hledání skrytých závislostí je možné provádět pomocí řady data miningových nástrojů. V rámci této práce byla provedena analýza nejvhodnějších data miningových nástrojů s ohledem na specifika řešené problematiky. Pro řešení byly využívány především IBM SPSS Modeler a software KNIME. Pro porozumění datům a jejich přípravu je efektivnější využít IBM SPSS Modeler. Při hledání skrytých závislostí byly používány oba nástroje. Výhodou open source nástroje KNIME je snadnější implementace vlastního algoritmu, čehož bylo při práci využito.

V rámci práce jsou testovány dva přístupy k detekci nebezpečných míst.

První přístup spočívá v prvotní detekci míst častých dopravních nehod. Vstupními parametry pro algoritmy, které jsou přímo určené na vyhledávání shluků v geo datech, jsou GPS souřadnice nehod. Odděleně je možné vyhledávat shluky v místech s velmi odlišnou hustotou dopravy díky informacím o počtu vozidel projíždějících daným místem za 24 hodin a typu silniční komunikace. Druhým krokem tohoto přístupu je detekce relevantních asociačních pravidel v rámci každého detekovaného shluku. Za tímto účelem je možné využít vlastního uzlu například v prostředí KNIME, který hromadně prochází jednotlivé shluky a hledá v nich asociační pravidla.

Druhý přístup řešený v rámci této práce je vyhledávání shluků podobných dopravních nehod s využitím většiny atributů popisujících aspekty dopravní nehody bez využívání GPS souřadnic. Tento přístup by mohl být využíván v případě budování nových komunikací, tak aby se předcházelo vzniku míst, která jsou z hlediska svého charakteru nebezpečná. V tomto případě je nutné důkladně analyzovat data tak, aby do algoritmů nevstupovaly atributy, které spolu korelují. Pro vyhledávání podobných nehod bylo využíváno algoritmů automatické klasifikace (IBM SPSS Modeler), který dokáže kombinovat několik shlukovačích algoritmů najednou. Samostatně bylo testováno vyhledávání skupin pomocí algoritmu TwoStep, který nevyžaduje na vstupu pevný počet shluků, jelikož částečně využívá hierarchického shlukování. Při vytváření modelů s tímto přístupem nebylo dosaženo prozatím uspokojivých výsledků.

Zásadním problémem druhého přístupu detekce podobných dopravních nehod je to, že data o dopravních nehodách nepostihují všechny aspekty, za kterých se nehoda stala. Tak nemohou být odhalena místa častých dopravních nehod, u kterých nejsou dostatečné informace popisující okolnosti nehody.

První přístup dovoluje vyhledat i místa, jež jsou nebezpečná počtem výskytu dopravních nehod relativně k hustotě dopravy v daném místě. Taková místa nemusí být na první pohled ničím typická (nemáme o nich prozatím dostatečné informace). V konceptu systému včasného varování definujeme tato místa jako obecné shluky.

Modely založené na datech o dopravních nehodách a algoritmech pro odhalování skrytých závislostí by mohly najít využití i v autonomních automobilech.

V práci je popsána řídicí a uživatelská část systému včasného varování před zvýšeným rizikem dopravní nehody. Diskutovány jsou i možnosti kooperace mezi řídicí a uživatelskou částí systému.

Nastíněný princip systému včasného varování by mohl přispět k prevenci častých dopravních nehod i předcházet tragickým nehodám. Navrhovaný systém by měl být užitečný především v situacích, kdy řidič jede konkrétní lokalitou poprvé. Hlavní výhodou navrhovaného systému včasného varování by mělo být především to, že se nehodám snaží předcházet, tj. neřeší situaci ve chvíli, kdy nehoda nastala, ale zaměřuje se na to, aby k nehodě vůbec nedošlo.

Nastíněné postupy, techniky a samotný princip systému včasného varování je možné využít díky zobecnění, které je v práci taktéž prezentováno i na zdánlivě zcela odlišné typy úloh s jiným typem dat. Příkladem využití navrhovaného řešení může být vyhledávání skrytých závislostí v datech týkajících se kriminální činnosti. Budoucí vývoj navrhovaného systému včasného varování by se měl zabývat především samotnou implementací jednotlivých bloků řídicí a uživatelské části. Z hlediska uživatelské části, je dále vyvíjena mobilní aplikace realizující samotné varování. Dále se také počítá s dalším vývojem a prací na vylepšování vlastní aplikace umocňující provádět hromadnou a automatizovanou realizaci činností řídicí části systému.

Seznam použité literatury

Citace

Aktivní a pasivní prvky bezpečnosti motorových vozidel, 2015. Observatoř bezpečnosti silničního provozu [online]. Praha: Centrum dopravního výzkumu [cit. 2017-09-12]. Dostupné z: http://www.czrso.cz/clanky/aktivni-a-pasivni-prvky-bezpecnosti-motorovych-vozidel/

AGRAWAL, Rakesh a Ramakrishnan SRIKANT, 1994. Fast Algorithms for Mining Association Rules. In: Proceeding VLDB '94 Proceedings of the 20th International Conference on Very Large Data Bases [online]. San Francisco: Morgan Kaufmann Publishers, s. 487-499 [cit. 2017-04-29]. ISBN 1-55860-153-8.

ANKERST, Mihael, Markus M BREUNIG, Hans-Peter KRIEGEL a Jorg SANDER, 1999. OPTICS: Ordering Points To Identify the Clustering Structure. In: Proceedings of the 1999 ACM SIGMOD international conference on Management of dat [online]. Philadelphia: ACM [cit. 2017-04-29]. ISBN 1-58113-084-8. Dostupné z: http://www.dbs.ifi.lmu.de/Publikationen/Papers/OPTICS.pdf

BARNSLEY, Les, Susan LORD a Nikolai BOGDUK, 1994. Whiplash injury. Pain [online]. 58(3), 283-307 [cit. 2017-09-14]. DOI: 10.1016/0304-3959(94)90123-6. ISSN 0304-3959. Dostupné z: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00006396-199409000-00001

BENGLER, Klaus, Klaus DIETMAYER, Berthold FARBER, Markus MAURER, Christoph STILLER a Hermann WINNER, 2014. Three Decades of Driver Assistance Systems: Review and Future Perspectives. IEEE Intelligent Transportation Systems Magazine [online]. 6(4), 6-22 [cit. 2017-09-14]. DOI: 10.1109/MITS.2014.2336271. ISSN 1939-1390. Dostupné z: http://ieeexplore.ieee.org/document/6936444/

BEN-HUR, Asa, David HORN, Hava SIEGELMANN a Vladimir VAPNIK, 2002. Support vector clustering. The Journal of Machine Learning Research [online]. 2002, 2002(2), 125-137 [cit. 2018-09-01]. Dostupné z: http://delivery.acm.org/10.1145/950000/944807/2-125-horn.pdf?ip=147.230.232.216&id=944807&acc=OPEN&key=D6C3EEB3AD96C931%2E981AA4EB460CF8D4%2E4D4702B0C3E38B35%2E6D218144511F3437&__acm__=1537872235_3ff7c5a148e61c4fe4eb51ba6a104e5e

BERKA, Petr, 2005. *Dobývání znalostí z databází*. Praha: Academia. ISBN 8020010629.

BERTALANFFY, Ludwig, 1968. General system theory: foundations, development, applications. New York: George Braziller.

City Safety, 2014. Volvo Cars Support [online]. Praha: Volvo Cars Support [cit. 2017-09-14]. Dostupné z: http://support.volvocars.com/cz/cars/Pages/owners-manual.aspx?mc=v526&my=2016&sw=15w46&article=dedd207c897617ddc0a801514bfcc7b4

CRISP DM: Data Mining Session 2, 2013. STATSOFT [online]. Palo Alto: Statsoft [cit. 2016-03-13]. Dostupné z: http://www.statsoft.com/support/blog/entryid/540/crisp-data-mining-session-2

DBSCAN pseudocode, 2011. GITHUB [online]. Kanwar Bhajneek [cit. 2017-08-01]. Dostupné z: https://github.com/KanwarBhajneek/DBSCAN

Dopravní nehody a jejich následky v roce 2017, 2018. Centrum služeb pro silniční dopravu [online]. Praha: Centrum služeb pro silniční dopravu [cit. 2018-09-19]. Dostupné z: https://www.cspsd.cz/707-dopravni-nehody-a-jejich-nasledky-v-roce-2017

Ecall becomes a reality!, 2015. European Emergency Number Association [online]. Brusel: European Emergency Number Association [cit. 2017-09-12]. Dostupné z: http://www.eena.org/press-releases/ecall-becomes-a-reality#.Wbg2A8hJZaS

ECall deployment - Publication by the European Commission of the Delegated Regulation No 305/2013, 2013. Harmonised eCall European Pilot - HeERO [online]. Brusel: Evropská komise [cit. 2017-09-12]. Dostupné z: http://www.heero-pilot.eu/view/en/media/news/20130415.html

ELKI [online], 2018. Germany: Ludwig Maximilian University of Munich [cit. 2018-07-29]. Dostupné z: https://elki-project.github.io/

ESTER, Martin, Hans KRIEGEL, Jorg SANDER a Xiaowei XU, 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING [online]. AAAI Press, s. 1-6 [cit. 2017-04-29]. ISBN 978-1-57735-004-0.
Dostupné z: http://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf

Euro NCAP Advanced Reward 2011, 2011. Euro NCAP [online]. Leuven: Euro NCAP [cit. 2017-09-14]. Dostupné z: https://www.euroncap.com/en/ratings-rewards/euro-ncap-advanced-rewards/2011-ford-active-city-stop/

European Commission MEMO 13/547, 2013. European Commission [online]. Brusel: European Commission [cit. 2015-09-12]. Dostupné z: http://europa.eu/rapid/press-release_MEMO-13-547_en.htm

EVANS, Leonard, 2004. Traffic Safety. 1. Michigan: Science Serving Society. ISBN 978-0975487105.

HAN, Jiawei. a Micheline. KAMBER, c2011. Data mining: concepts and techniques. 3rd ed. Burlington, MA: Elsevier. ISBN 9780123814791.

IBM SPSS Modeler, 2016. IBM [online]. New York: IBM Corporation [cit. 2017-04-29]. Dostupné z: http://www-03.ibm.com/software/products/cs/spss-modeler

JONES, W.D., 2001. Keeping cars from crashing. IEEE Spectrum [online]. 38(9), 40-45 [cit. 2017-09-13]. DOI: 10.1109/6.946636. ISSN 00189235. Dostupné z: http://ieeexplore.ieee.org/document/946636/

CHAPMAN, Pete, Julian CLINTON, Randy KERBER, Thomas KHABAZA, Thomas REINARTZ, Colin SHEARER a Rüdiger WIRTH, 2000. CRISP-DM 1.0: Step-by-step data mining guide [online]. 1. CRISP-DM consortium: [cit. 2018-09-15]. Dostupné z: https://www.the-modeling-agency.com/crisp-dm.pdf

KNIME, 2017. KNIME [online]. Konstanz: University of Konstanz [cit. 2017-04-29]. Dostupné z: https://www.knime.org/

KOVÁČIK, Milan, 2016. Hledání shluků podobných dopravních nehod. Liberec. Diplomová práce. Technická univerzita v Liberci. Vedoucí práce Marián Lamr.

MADSEN, Henrik, 2008. Time series analysis. 1. Boca Raton: Chapman & Hall/CRC. Texts in statistical science. ISBN 978-1-4200-5967-0.

MAURER, Markus, 2012. Forward Collision Warning and Avoidance. MAURER, Markus. Handbook of Intelligent Vehicles. 1. London: Springer London, s. 657-687. DOI: 10.1007/978-0-85729-085-4_25. ISBN 978-0-85729-084-7. Dostupné také z: http://link.springer.com/10.1007/978-0-85729-085-4_25

Ministerstvo dopravy chce zvýšit pokuty za rychlost, 2015. ČTK. DENÍK.CZ. DENÍK.CZ [online]. ČTK, ČTK. Praha: VLTAVA-LABE-PRESS, 17.8.2015, s. 2 [cit. 2015-11-13]. Dostupné z: http://www.denik.cz/automoto-denik/ministerstvo-dopravy-chce-zvysit-pokuty-za-rychlost-az-trojnasobne-20150817.html

Orange [online], 2018. Ljubljana: University of Ljubljana [cit. 2018-08-29]. Dostupné z: https://orange.biolab.si/

PETR, Pavel, 2010-. *Data Mining*. Vyd. 3. Pardubice: Univerzita Pardubice. ISBN 9788073953256.

PETR, Pavel, 2014. Metody Data Miningu. Pardubice: Univerzita Pardubice. ISBN 9788073958732.

Prezentace výsledků sčítání dopravy 2010, 2011. ŘSD [online]. Praha: ŘSD [cit. 2016-09-20]. Dostupné z: http://scitani2010.rsd.cz/pages/informations/default.aspx

RapidMiner, 2017. RapidMiner [online]. Boston [cit. 2017-04-29]. Dostupné z: https://rapidminer.com/

ŘEZANKOVÁ, Hana, Dušan HÚSEK a Václav SNÁŠEL, 2009. Shluková analýza dat. 2., rozš. vyd. Praha: Professional Publishing, 218 s. ISBN 9788086946818.

Statistika nehodovosti, 2014. POLICIE ČR [online]. Praha: [cit. 2015-11-13]. Dostupné z: http://www.policie.cz/clanek/statistika-nehodovosti-900835.aspx

SIEGEL, Eric, 2016. Predictive analytics: the power to predict who will click, buy, lie, or die. Revised and Updated Edition. Hoboken, New Jersey: Wiley. ISBN 9781119145677.

SHMUELI, Galit, Peter C BRUCE, Inbal YAHAV, Nitin R PATEL a Kenneth C LICHTENDAHL, 2018. Data mining for business analytics: concepts, techniques, and applications in R. 1. Hoboken, New Jersey: John Wiley. ISBN 9781118879368.

Statistika nehodovosti, 2018. Statistika nehodovosti - Policie České republiky [online]. Praha: Policie ČR [cit. 2018-09-21]. Dostupné z: Statistika nehodovosti na pozemních komunikacích v ČR 2016, 2017. Autoklub ČR [online]. Praha: Autoklub ČR [cit. 2017-09-08]. Dostupné z: http://www.autoklub.cz/dokument/12022-statistika-nehodovosti-za-rok-2016.html

Statistika nehodovosti na pozemních komunikacích v ČR 2015, 2016. Autoklub ČR [online]. Praha: Autoklub ČR [cit. 2017-09-08]. Dostupné z: http://www.autoklub.cz/dokument/9698-statistika-nehodovosti-za-rok-2015.html

Statistika nehodovosti na pozemních komunikacích v ČR 2016, 2017. Autoklub ČR [online]. Praha: Autoklub ČR [cit. 2017-09-08]. Dostupné z: http://www.autoklub.cz/dokument/12022-statistika-nehodovosti-za-rok-2016.html

Toto je nejnebezpečnější zatáčka v zemi, 2016. PECÁK, Radek. Aktuálně.cz [online]. ČR: Economia [cit. 2016-02-18]. Dostupné z: http://zpravy.aktualne.cz/ekonomika/auto/toto-je-nejnebezpecnejsi-zatacka-v-cr-pred-rizikem-ted-u-pso/r~2977752cc9bc11e593630025900fea04/

TwoStep Cluster Node, 2012. IBM® IBM Knowledge Center [online]. USA: IBM [cit. 2017-08-31]. Dostupné z: https://www.ibm.com/support/knowledgecenter/en/SS3RA7_15.0.0/com.ibm.spss.modeler.help/clusternode_general.htm

V srpnu zemřelo na silnicích 70 lidí, o 12 víc než loni, 2015. Aktuálně.cz [online]. ČTK. Praha: Economia, 1.9.2015 [cit. 2015-11-13]. Dostupné z: http://zpravy.aktualne.cz/domaci/v-srpnu-zemrelo-na-silnicich-70-lidi-o-12-vice-nez-loni/r~9d183314509e11e5adcb0025900fea04/

Vláda: Cena lidského života je 120 mil. Kč, 2016. Banky cz [online]. Praha: Top-in.cz, a.s [cit. 2018-09-01]. Dostupné z: https://www.banky.cz/clanky/vlada-cena-lidskeho-zivota-je-120-mil-korun/

VLK, František, 2006. Automobilová elektronika: systémy řízení podvozku a komfortní systémy. 1. Brno: Prof.Ing.František Vlk,DrSc., nakladatelství a vydavatelství, 308 s. ISBN 80-239-7062-3.

WEKA [online], 2018. Waikato: University of Waikato [cit. 2018-07-29]. Dostupné z: http://www.cs.waikato.ac.nz/ml/weka/

WINNER, Hermann, Bernd DANNER a Joachim STEINLE, 2009. Adaptive Cruise Control. Handbuch Fahrerassistenzsysteme [online]. 1. Wiesbaden: Vieweg+Teubner, s. 478. DOI: 10.1007/978-3-8348-9977-4_33. ISBN 978-3-8348-0287-3. Dostupné také z: http://www.springerlink.com/index/10.1007/978-3-8348-9977-4_33

WINNER, Hermann, 2012. Adaptive Cruise Control. Handbook of Intelligent Vehicles [online]. 1. London: Springer London, s. 613. DOI: 10.1007/978-0-85729-085-4_24. ISBN 978-0-85729-084-7. Dostupné také z: http://link.springer.com/10.1007/978-0-85729-085-4_24

Základní informativní výpis o nehodě, 2014. POLICIE ČR. Jednotná dopravní vektorová mapa [online]. Praha: PČR [cit. 2015-09-27]. Dostupné z: http://pcr.jdvm.cz/pcr/Reports.aspx?S_Type=01&S_LID=41aa962a-f5bb-4e2b-953d-c56b6ba94b63&S_IdNehoda=002100070013

Zdravý člověk má cenu 10 milionů korun, stanovili experti, 2014. Idnes.cz [online]. Praha: Mafra [cit. 2018-09-01]. Dostupné z:
https://zpravy.idnes.cz/cena-lidskeho-zivota-je-10-milionu-d4b-/domaci.aspx?c=A140414_131414_domaci_hv

Publikace autora související s tématem disertační práce

LAMR, Marián a Jan SKRBEK, 2017a. A Systems Approach to Designing a Traffic Collision Avoidance Early Warning System. Journal of Systemics, Cybernetics and Informatics [online]. Orlando, USA, 15(Volume 15 - Number 3), 55-59 [cit. 2017-09-13]. ISSN 1690-4524. Dostupné z: http://www.iiisci.org/journal/sci/Contents.asp?Previous=ISS1703 (100 %)

LAMR, Marián a Jan SKRBEK, 2017b. Using Data Mining Tools for Retrieving Information from Databases of Traffic Accidents. In: IDIMT-2017 Digitalization in Management, Society and Economy: 25th Interdisciplinary Information Management Talks,. Austria: Trauner Verlag, s. 391-400. ISBN 978-3-99062-119-6. (100 %)

LAMR, Marián a Robin DVOŘÁK, 2016a. Application for visualization and analysis of traffic accident information. In: Proceedings of the 28th International Business Information Management Association Conference – Vision 2020: Innovation Management, Development Sustainability, and Competitive Economic Growth. Madrid: International Business Information Management Association, s. 1703-1709. ISBN 978-0-9860419-8-3. (70 %)

LAMR, Marián a Jan SKRBEK, 2016b. Real-time approaches to improving traffic safety. In: 2016 5th Mediterranean Conference on Embedded Computing (MECO) [online]. Bar: IEEE, s. 452-455 [cit. 2017-09-14]. DOI: 10.1109/MECO.2016.7525690. ISBN 978-1-5090-2222-9. Dostupné z: http://ieeexplore.ieee.org/document/7525690/ (100 %)

LAMR, Marián a Jan SKRBEK, 2016c. Searching for Traffic Accident Clusters to Increase Road Traffic Safety. In: IDIMT-2016 Information Technology, Society and Economy - Strategic Cross-Influence - 24th Interdisciplinary Information Management Talks. Linz, Österreich: Johannes Kepler Universitat Linz., s. 425-432. ISBN 978-3-99033-869-8. (100 %)

LAMR, Marián a Jan SKRBEK, 2016d. Traffic Data and Possibilities of their Utilization for Safer Traffic. In: Proceedings of the International Conference: Liberec Informatics Forum 2016. Liberec: Technical University of Liberec, s. 61-73. ISBN 978-80-7494-303-4. (100 %)

LAMR, Marián a Jan SKRBEK, 2015a. Advanced Approaches to Traffic Accident Prevention. In: IBIMA 2015 - Proceedings of the 26th International Business Information Management Association Conference. Madrid: International Business Information Management Association, s. 10. ISBN 978-0-9860419-5-2. (100 %)

SKRBEK, Jan a Marián LAMR, 2015b. Increasing effectiveness of early warning through smart ICT. In: Information and Communication Technology (ICoICT), 2015 3rd International Conference. 2015. Indonesia: IEEE, s. 160-165. ISBN 978-1-4799-7751-2. (100 %)

LAMR, Marián, David KUBÁT a Jan SKRBEK, 2015c. New Approaches to Smart Solutions for eliminating Car Accidents. In: Proceedings of the 12th International Conference Liberec Economic Forum 2015. 1. Liberec: Vysokoškolský podnik, spol., s. 392-401. ISBN 978-80-7494-225-9. (50 %)

LAMR, Marián a Jan SKRBEK, 2015d. System Approach to Increasing Safety of Road Traffic. In: System approaches‘15 Interaction of soft and hard systems. Prague: University of Economics, Prague Publishing Oeconomica, 54 – 58. ISBN 978-80-245-2125-1. (100 %)

LAMR, Marián a Jan SKRBEK, 2015e. The Options for Actively Increasing Road Safety. In: IDIMT 2015: Information Technology and Society – Interaction and Interdependence – 23rd Interdisciplinary Information Management Talks. Linz, Österreich: Johannes Kepler Universitat Linz, 487 – 494. ISBN 978-399033395-2. (100 %)

Ostatní publikace autora

ANTLOVÁ, Klára, Petra RYDVALOVÁ a Marián LAMR, 2017. Motivation in the students´ start–ups. In: DIMT-2017 Digitalization in Management, Society and Economy: 25th Interdisciplinary Information Management Talks. Austria: Trauner Verlag, s. 79-84. ISBN 978-3-99062-119-6. (30 %)

CÍSAŘOVÁ, Klára, Marián LAMR, Přemysl SVOBODA a Pavel TYL, 2016. Attitudes of students to the lecture streaming and new elements of e-learning portal. In: Proceedings of the 28th International Business Information Management Association Conference – Vision 2020: Innovation Management, Development Sustainability, and

Competitive Economic Growth. Seville: International Business Information Management Association, s. 2080-2087. ISBN 978-0-9860419-8-3. (25 %)

CÍSAŘOVÁ, Klára, Marián LAMR, Jan LOUFEK a Pavel TYL, 2015. Stačí technické prvky a nové technologie, aby byly nastavené rovné šance na vzdělávání pro studenty s SVP. In: Sborník příspěvků z X. ročníku z konference „vysokoškolské studium bez bariér". 1. Liberec: Vysokoškolský podnik, spol., s.r.o 59-69. ISBN 978-80-7494-228-0. (25 %)

HYBLEROVÁ, Šárka, Martina ČERNÍKOVÁ, Olga MALÍKOVÁ, Klára CÍSAŘOVÁ a Marián LAMR, 2017. Indicating Financial Health of Czech Companies with the Support of Modern Methods of Multidimensional Data Processing. In: SGEM 2017: 4th international multidisciplinary scientific conference on social sciences & arts, Albena, 24-30 August 2017 : [recenzovaný zborník]. Book 3 Science and Society, Vol. 5. Sofia: STEF92 Technology, 353 – 360. ISBN 978-619-7408-15-7. ISSN 2367-5659. (10 %)

LAMR, Marián, Klára CÍSAŘOVÁ a Jana VITVAROVÁ, 2015. Advanced Learning Space as an Asset for Students with Disabilities. Turkish Online Journal of Educational Technology [online]. 14(2), 10-14 [cit. 2015-06-03]. ISSN 2146 - 7242. Dostupné z: http://www.tojet.net/volumes/v14i2.pdf (50 %)

LAMR, Marián, Pavel TYL a Alena GREGOVÁ, 2014. Nabídka nových sw řešení na elearningovém portále ALS na TUL v Liberci. In: Sborník příspěvků z VIII. ročníku mezinárodní konference "Vysokoškolské studium bez bariér". 1. Liberec: Vysokoškolský podnik, spol., s.r.o, s. 58-66. ISBN 978-80-7494-066-8. (33 %)

LOUFEK, Jan, Marián LAMR a Klára CÍSAŘOVÁ, 2014. Využití richmedií pro studenty s poruchou zraku. In: Sborník příspěvků ze IX. ročníku konference „Vysokoškolské studium bez bariér“. Liberec: Vysokoškolský podnik spol., s.r.o, s. 36-44. ISBN 978-80-7494-169-6. (33 %)

RYDVALOVÁ, Petra, Klára ANTLOVÁ a Marián LAMR, 2017. Vztah studentů k podnikání: celosvětový průzkum GUESSS 2016 z pohledu studentů českých vysokých škol. Liberec: Technická univerzita v Liberci. ISBN 978-80-7494-379-9. (33 %)

Strukturovaný životopis

Jméno	**ING. BC. MARIÁN LAMR**

PRACOVNÍ ZKUŠENOSTI

• Období (od – do)	10/2009- současnost
• Jméno a adresa zaměstnavatele	Technická univerzita v Liberci, Fakulta mechatroniky, informatiky a mezioborových studií
• Dosažené pozice	Cvičící/přednášející odborných předmětů, komunikátor posilující feedback vzdělávacího procesu, Expert pro multimediální uživatelský návod, Postprocesing streamů
• Hlavní pracovní náplň a odpovědnost	Příprava materiálů pro výuku, výklad a procvičování látky daného předmětu. Technická podpora uživatelů při streamování a užívání elearningového portálu, post processing přednášek
• Období (od – do)	7/2008 – 11/2012
• Jméno a adresa zaměstnavatele	PCSTORM, Přerušená 357, Trutnov
• Oblast podnikání nebo název odvětví	Počítačová školící střediska
• Dosažená pozice	Lektor, předseda zkušební komise
• Hlavní pracovní náplň a odpovědnost	Školení a přezkušování firemních a soukromých klientů v oblasti užívání software (zejména MS Office), příprava podkladů pro výuku, správa místní počítačové sítě

VZDĚLÁNÍ A KURZY

• Období (od – do)	9/2006 - 6/2009
• Název a typ organizace poskytující vzdělání či kurzy	Technická univerzita v Liberci, Fakulta mechatroniky, informatiky a mezioborových studií, obor Informační technologie
• Získaný titul	Ing.
• Období (od – do)	9/2003 - 6/2006
• Název a typ organizace poskytující vzdělání či kurzy	Technická univerzita v Liberci, Fakulta mechatroniky, informatiky a mezioborových studií, obor Elektronické, informační a řídící systémy
• Získaný titul	Bc.
• Období (od – do)	9/2006 - 9/2010
• Název a typ organizace poskytující vzdělání či kurzy	Technická univerzita v Liberci, Fakulta přírodovědně-humanitní a pedagogická, obor pedagogika
• Získaný titul	Bc.
• Období (od – do)	9/1999 - 6/2003

STÁŽE A VÝZNAMNÉ ODBORNÉ KONFERENCE	2017 -Fakulta informatiky, Masarykova univerzita v Brně 2018 konference ES 2018 (Kypr) 2017 konference IDIMT 2017(ČR), konference WMSCI 2017 (USA) 2016 konference IDIMT 2016(ČR), konference LIF2016(ČR) 2016konference IBIMA 2016(Španělsko), MECO 2016(Černá hora) 2015 LEF 2015(ČR), konference SP 2015(ČR), konference IDIMT 2015(ČR)