

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Odkrivanje zavarovalniških goljufij z uporabo analize omrežij

Lovro Šubelj

Univerza v Ljubljani,
Fakulteta za računalništvo in informatiko

FMF, 30.5.2013

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Goljufije v zavarovalništvu

Odkrivanje zavarovalniških goljufij

Ekspertni sistem

Rezultati in doprinos

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ izsiljene ali uprizorjene prometne nesreče
- ▶ neupravičena pridobitev sredstev od zavarovalnice
- ▶ nevarnost za udeležence v cestnem prometu
- ▶ višje zavarovalnine za ostale zavarovance:
 - ▶ $\approx 15\text{-}30\%$ škodnih zahtevkov vsaj pretiranih
 - ▶ $\approx 10\%$ odhodkov za škode na račun goljufij
 - ▶ v Sloveniji ≈ 100 milijonov € izgub na leto



Vir: <http://www.insurancefraud.org/>

Odkrivanje
goljufov preko
analize omrežij

Lovro Šubelj

Goljufige v
zavarovalništvu

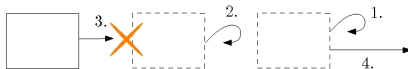
Odkrivanje
zavarovalniških
goljufov

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ posebej pereče organizirane skupine goljufov (vozniki, avtomehaniki, zdravniki, odvetniki, policisti itd.)
- ▶ znane sheme za uprizarjanje prometnih nesreč:



ZVCP: "[V]arnostna razdalja mora ne glede na vozne razmere omogočati, da lahko voznik zmanjša hitrost ali ustavi in s tem prepreči trčenje, če voznik, ki vozi pred njim, zmanjša hitrost ali ustavi".

- ▶ skupne značilnosti uprizorjenih prometnih nesreč (mlajši moški, pozno zvečer, izven naselja, ni alkohola itd.)
- ▶ **cilj**: ekspertni sistem za odkrivanje skupin goljufov

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ občutljivi podatki (oteženo pridobivanje)
- ▶ pomanjkanje označenih podatkov (nove goljufije)
- ▶ neuravnotežena ciljna spremenljivka

Podatki v policijskih zapisnikih (in registrih):

- ▶ entitete: nesreče, udeleženci, vozila, policisti
- ▶ atributi: čas in kraj, gmotna škoda, poškodbe, spol in starost udeležencev itd.
- ▶ relacije: voznik-sopotnik-nesreča-policist



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Goljufije v zavarovalništvu

Odkrivanje zavarovalniških goljufij

Ekspertni sistem

Rezultati in doprinos

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

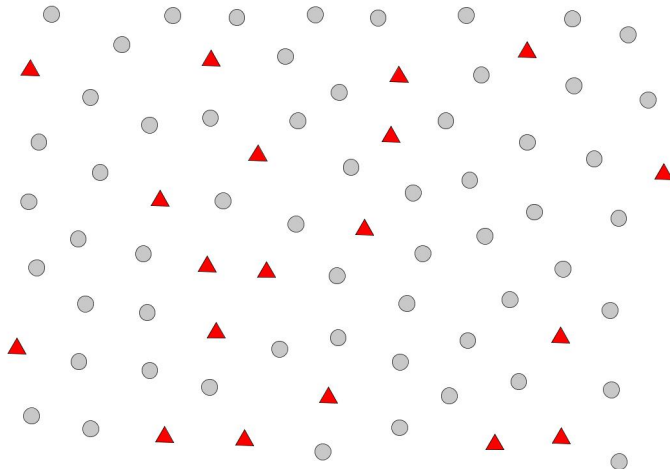
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



- ▶ strojno učenje, podatkovno rudarjenje, statistika, eksperti

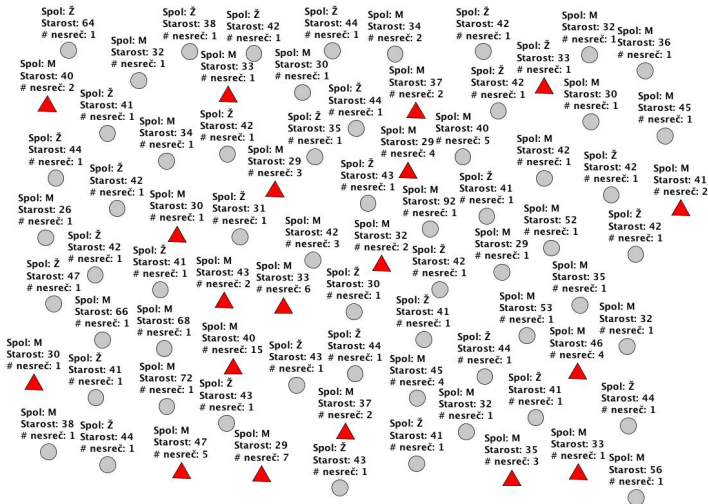
Goljufije v zavarovalništvu

Odkrivanje zavarovalniških goljufij

Eksperтни
sistem

Rezultati in doprinos

Zaključek



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

- ▶ strojno učenje, podatkovno rudarjenje, statistika, eksperti
- ▶ v praksi ni razlik, številne goljufije ostanejo neodkrita

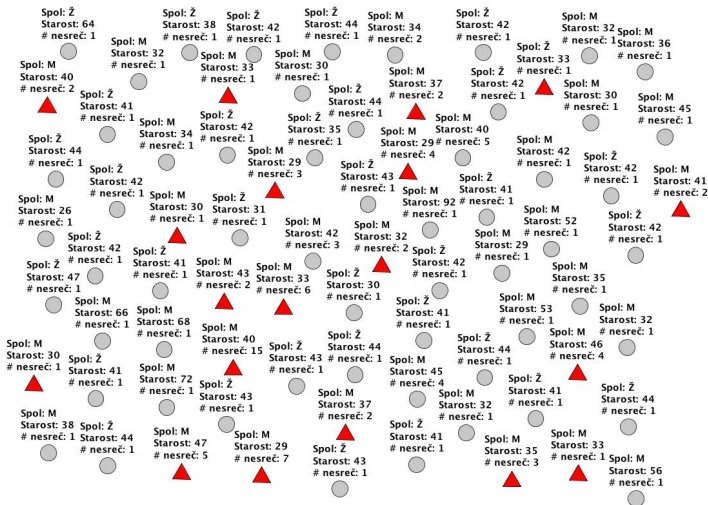
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Slabosti pristopov odkrivanja goljufij:

- ▶ večji nabor označenih podatkov
- ▶ odkrivanje znanih goljufij (ne anomalij)
- ▶ neprimerna predstavitev podatkov (ni relacij)



Vir: <http://http://www.reevewrites.com/>

Odkrivanje
goljufij preko
analize omrežij

- udeležence nesreč povežemo v družbena omrežja

Lovro Šubelj

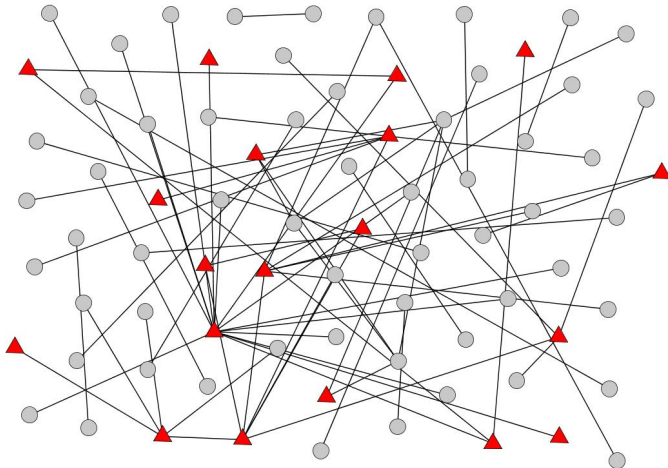
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

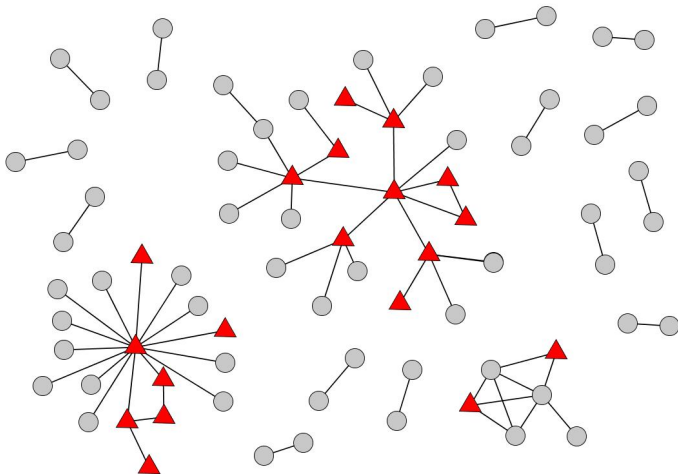
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ udeležence nesreč povežemo v družbena omrežja
- ▶ goljufe lahko odkrijemo že s prostim očesom



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Goljufije v zavarovalništvu

Odkrivanje zavarovalniških goljufij

Ekspertni sistem

Rezultati in doprinos

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

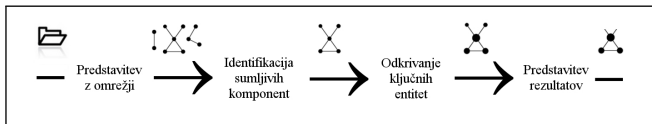
Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Sistem za odkrivanje goljufij:

1. predstavitev nesreč z družbenimi omrežji
2. odkrivanje sumljivih skupin nesreč
3. odkrivanje ključnih udeležencev
4. prikaz in interpretacija rezultatov



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

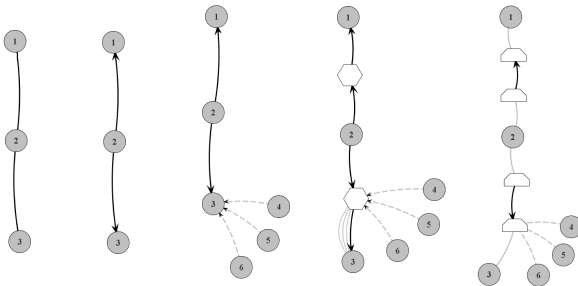
Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Predstavitev z omrežji:

- ▶ vozlišča predstavljajo udeležence, nesreče in vozila
- ▶ povezave predstavljajo relacije med njimi (6 načel gradnje)
- ▶ omogoča formulacijo kompleksnih vzorcev povezovanja



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

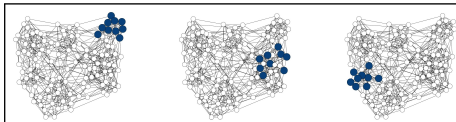
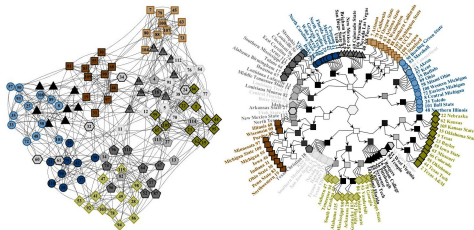
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinosi

Zaključek

Omrežja poenostavimo (tj. delimo) glede na skupnosti vozlišč.



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

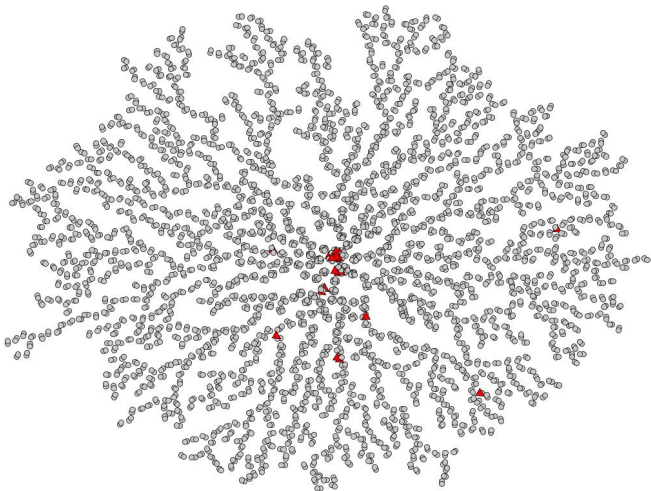
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

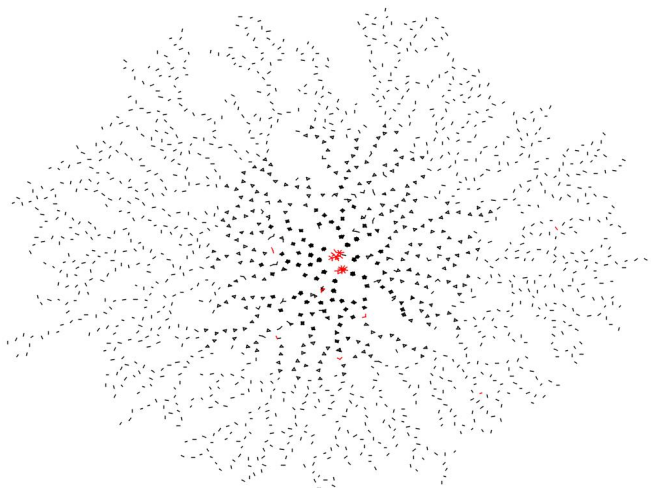
Ekspertni
sistem

Rezultati in
doprinos

Zaključek



- omrežje razpade na več povezanih komponent



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

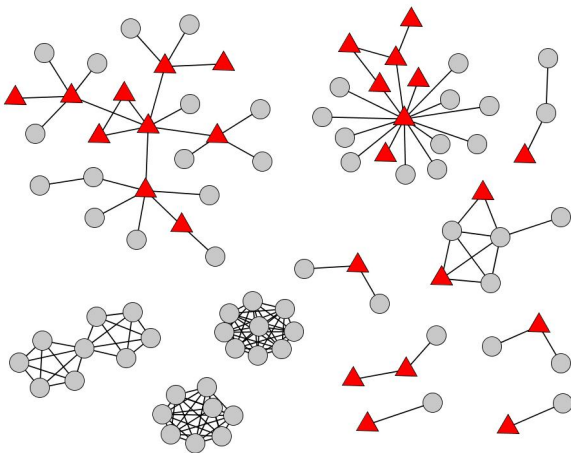
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ omrežje razpade na več povezanih komponent
- ▶ skupne lastnosti “goljufivih” komponent (npr. gostota, premer, nakopičenost, največja stopnja in centralnost itd.)



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

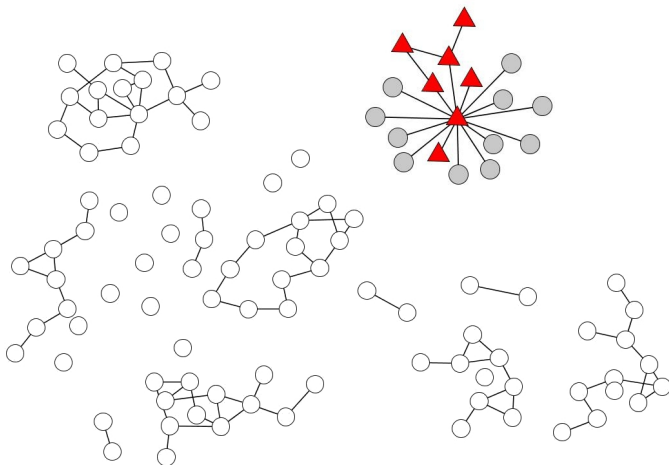
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinosa

Zaključek

- ▶ omrežje razpade na več povezanih komponent
- ▶ skupne lastnosti “goljufivih” komponent (npr. gostota, premer, nakopičenost, največja stopnja in centralnost itd.)



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

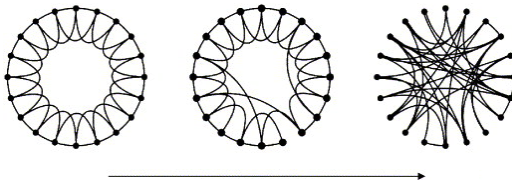
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ vsaki lastnosti priredimo indikator (0-1 atribut)
- ▶ sumljive komponente razkrijemo s pomočjo simulacije naključnih omrežij in preverjanja hipotez
- ▶ možnosti: posamezni indikator, *majority*, *(P)RIDIT* itd.



Vir: <http://http://www.nature.com/>

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

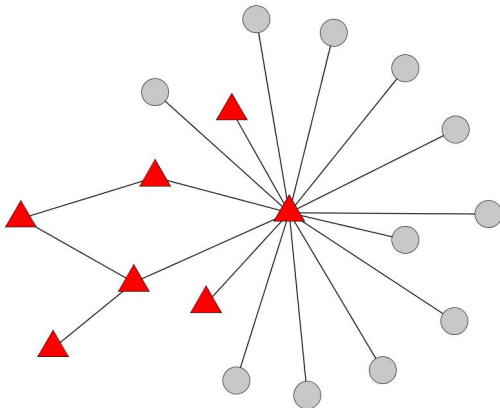
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



- ▶ sumljivost udeležencev določimo na podlagi neposredno povezanih entitet, in obratno (vrsta omrežja)

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

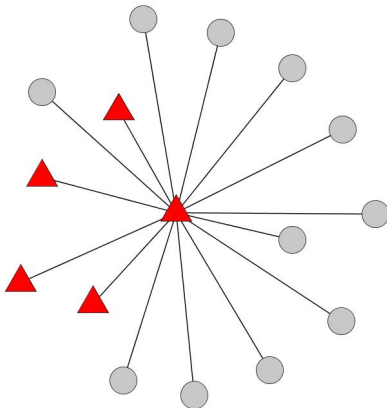
Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

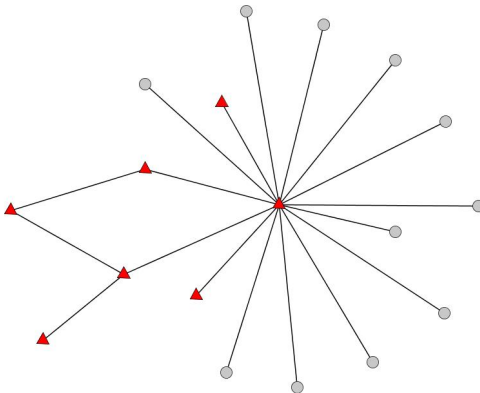
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinosi

Zaključek

- ▶ sumljivost udeležencev določimo na podlagi neposredno povezanih entitet, in obratno (vrsta omrežja)
- ▶ modeli: utežena linearna kombinacija (domenski ekspert)



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

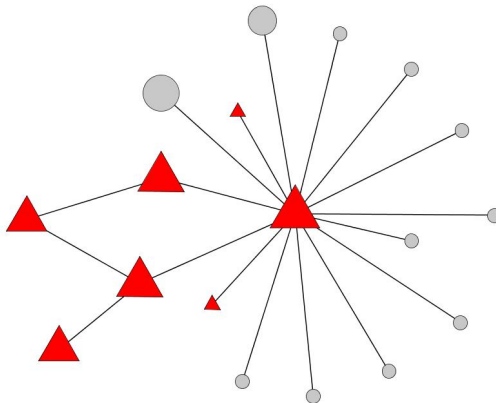
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinosi

Zaključek

- ▶ sumljivost udeležencev določimo na podlagi neposredno povezanih entitet, in obratno (vrsta omrežja)
- ▶ modeli: utežena linearna kombinacija (domenski ekspert)
- ▶ iterativni algoritem za ocenjevanje sumljivosti



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

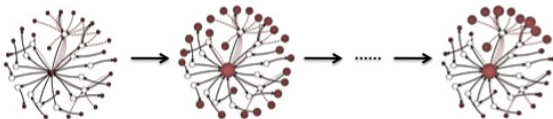
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

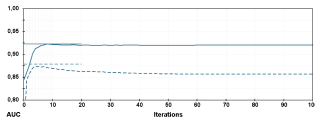
Rezultati in
doprinosi

Zaključek

- ▶ z iterativnim ocenjevanjem premagamo lokalnost (propagacija sumljivosti po omrežju)



- ▶ ocenimo zgolj sumljivost udeležencev (posredno za ostale)
- ▶ povezani (sumljivi) udeleženci → skupina goljufov
- ▶ rangiranje udeležencev, ne klasifikacija!



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

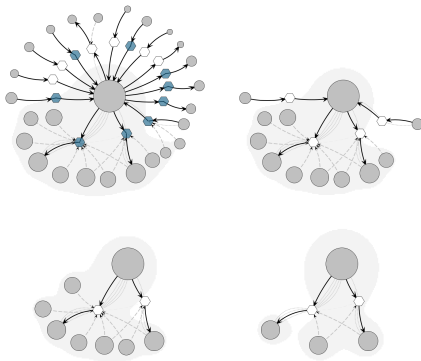
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ popolnoma avtomatski sistem v praksi ni mogoč
- ▶ predstavitev rezultatov domenskemu ekspertu, analitiku
- ▶ uporaba znanja pridobljenega skozi analizo



Odkrivanje
goljufij preko
analize omrežij

Goljufije v zavarovalništvu

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Odkrivanje zavarovalniških goljufij

Ekspertni
sistem

Ekspertni sistem

Rezultati in
doprinos

Rezultati in doprinos

Zaključek

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Testni nabor podatkov:

- ▶ policijski zapisniki o nesrečah v Sloveniji (1999-2008)
- ▶ vsi (označeni) podatki:
 - ▶ 1561 (91) prometnih nesreč
 - ▶ 3451 (211) udeležencev v nesrečah
 - ▶ 46 predhodno odkritih goljufov
- ▶ pazljivost pri interpretaciji rezultatov:
 - ▶ podatki niso izbrani naključno
 - ▶ majhen (nereprezentativen) vzorec
 - ▶ podatke uporabljamo pri zasnovi
 - ▶ vprašljivost oznak podatkov

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Sistem doseže $\approx 93\%$, drugi pristopi brez omrežij 80-85%.

Rangiranje

AUC = **0,9288** (ploščina pod krivuljo)

Klasifikacija (minimiziramo skupno ceno)

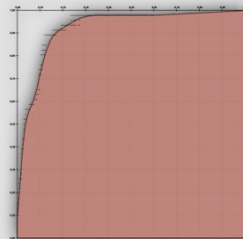
CA = 0,8720 (točnost)

Recall = 0,8913 (priklic, TPR)

Precision = 0,6508 (natančnost)

Specificity = 0,8667 (specifičnost, 1-FPR)

F1 ocena = 0,7523



	Sumljivi	Nesumljivi
Goljufi	41	5
Negoljufi	22	143
Neznano	0	3240

Znanstveni prispevki:

- ▶ revija → izjemni znanstveni dosežek v Sloveniji (ARRS)
- ▶ diploma → fakultetna Prešernova nagrada (FRI)
- ▶ konferenca → nagrada za prispevek (DSI)

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

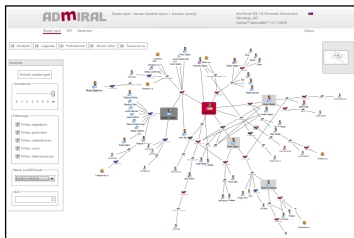
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ podjetje *Optilab* razvija orodje *Admiral* (Motor)
→ odkrivanje goljufij na podlagi analize omrežij
- ▶ *Admiral* uporablja Slovensko zavarovalno združenje



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Goljufije v zavarovalništvu

Odkrivanje zavarovalniških goljufij

Ekspertni sistem

Rezultati in doprinos

Zaključek

Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

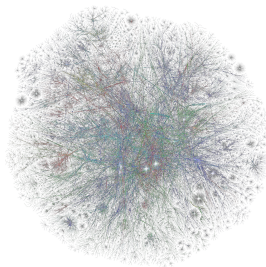
Ekspertni
sistem

Rezultati in
doprinos

Zaključek

- ▶ sistem za odkrivanje goljufij uporaben v praksi
- ▶ možnosti uporabe druge (npr. zavarovanje plovil itd.)
- ▶ omrežja so danes prisotna na vsakem koraku
- ▶ definicije, formule, algoritmi:

Šubelj, L., Furlan, Š. & Bajec, M. (2011). An expert system for detecting automobile insurance fraud using social network analysis, *Expert Systems with Applications* 38(1), 1039-1052.



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

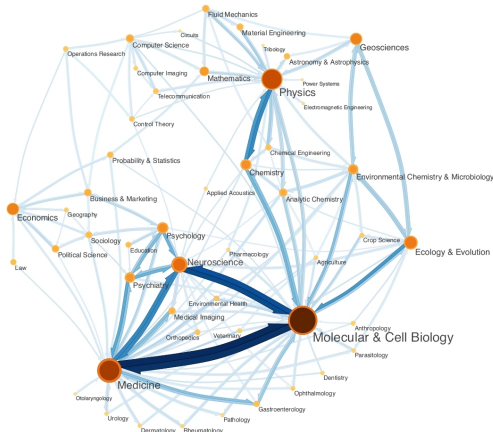
Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Analiza omrežij ponuja številne pristope:
vozlišča → skupine vozlišč → omrežje, povezave itd.



Odkrivanje
goljufij preko
analize omrežij

Lovro Šubelj

Goljufije v
zavarovalništvu

Odkrivanje
zavarovalniških
goljufij

Ekspertni
sistem

Rezultati in
doprinos

Zaključek

Hvala za pozornost



Laboratorij za podatkovne tehnologije

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani

<http://lovro.lpt.fri.uni-lj.si/>

lovro.subelj@fri.uni-lj.si

Soavorji in drugi:

izr. prof. dr. Marko Bajec (Univerza v Ljubljani)

doc. dr. Matjaž Kukar (Univerza v Ljubljani)

doc. dr. Jure Leskovec (Univerza Stanford)

dr. Štefan Furlan (Optilab, d.o.o.)