Elektrotehniški vestnik 75(5): 293-298, 2008
Electrotechnical Review: Ljubljana, Slovenija

Sistem za avtomatsko detekcijo goljufij pri elektronskem poslovanju

Štefan Furlan, Marko Bajec

Laboratorij za informatiko, Fakulteta za računalništvo in informatiko, Tržaška cesta 25, 1000 Ljubljana
E-pošta: stefan.furlan@fri.uni-lj.si

Povzetek. V času elektronskega poslovanja in pojava poslovnih modelov deljenja dobička (ang. revenue sharing) goljufije postajajo pereč problem. Zaradi goljufije klikanja na oglase podjetje Google izgubi več kot milijardo dolarjev na leto, saj je od 10 do 15 odstotkov klikov na njihove oglase goljufivih. Proti goljufijam se lahko uspešno borimo le z uporabo učinkovitih sistemov za detekcijo goljufij. Kot vsako področje ima tudi področje detekcije goljufij določene posebnosti, ki jih moramo pri graditvi takih sistemov upoštevati, če želimo biti uspešni. Specifike je mogoče strniti v šest razredov: specifike podatkov, problem ocenjevanja uspešnosti detekcije, problem razlage sklepanja, problem kombiniranja več metod, zahteva po adaptivnosti in zahteva po zmogljivosti. Članek poda predlog zasnove sistema za detekcijo goljufij in skozi opis statičnega in dinamičnega vidika tega sistema pojasni, kako le-ta uspešno naslovi omenjene specifike. Statični vidik sistema predstavljata opisa strukture jedra za detekcijo goljufij, ki izvaja samo detekcijo goljufij, in strukture celotnega sistema, ki poleg jedra za detekcijo goljufij vsebuje tudi podatke ter uporabniški vmesnik. Dinamični vidik zajema dva pomembna procesa, ki ju mora sistem podpreti. To sta proces detekcije in razreševanja goljufij ter proces učenja in izboljšave sistema.

Ključne besede: detekcija goljufij, detekcija anomalij, elektronsko poslovanje, goljufija s klikanjem

A System for Automatic Fraud Detection in e-Business

Extended abstract. Nowadays e-business is in its full swing and the providers of e-services are undertaking new business models. The newest and most promising business model is based on revenue sharing and relies heavily on advertising. In such circumstances, fraud is becoming an immense issue, as fraudsters no longer just inflict damage but can actually earn a lot of money. Google alone is facing billion dollar losses due to fraud annually, since between 10 % and 15 % clicks on their advertisements are fraudulent. Because fraud represents a big problem, companies have started forming departments that deal merely with the problem of fraud, monitoring and analyzing every business transaction. A great contribution to this effort are systems that help the fraud analytics to focus their efforts on the transactions, providers, etc. that are more suspicious. The fraud detection domain has its own specificities which can be organized into the following sections: (A) specific data, (B) the problem of assessing the detection, (C) the necessity to explain the system's inference, (D) the need to combine more than one method, (E) the need to adapt and (F) good system performance. The paper proposes a concept of a system which successfully addresses the aforementioned issues. It turns out that in order to effectively address these matters, the fraud detection engine is not enough. We need a system which of course includes an efficient fraud detection engine, but also supports two key processes and in order to do so, provides other crucial components, such as user-friendly graphical interface and an efficient access to the company's transactional, and other data. The paper provides a description of a system which is given through both a dynamic and static perspective. The dynamic perspective is represented by two key processes: (1) the fraud detection and resolution process and (2) the learning and system adaptation process. The static view on the system is provided via the system architecture and the fraud detection engine architecture.

Keywords: fraud detection, anomaly detection, e-business, click fraud

Prejet 19. november 2007
Odobren 4. oktober, 2008

1 Uvod

Goljufije so dandanes pereč problem, saj so v različnih sektorjih razlog za izgubo približno 10 odstotkov prihodkov. Že nekaj časa se z detekcijo goljufij ukvarjajo v sektorjih, kot so na primer zavarovalništvo, telekomunikacije in banke, kjer so goljufije razlog za izgubo od 3 do 20 odstotkov prihodkov [1, 2, 3, 4].

Od nastanka elektronskega poslovanja [5] je uporabnikom in podjetjem čedalje več storitev na voljo na spletu. Odpiranje je šlo do te mere, da so določeni ponudniki storitev omogočili vsem uporabnikom sooblikovanje njihovih storitev (primer Wikipedia) in so pripravljeni z njimi celo deliti svoj dobiček (primer Google s programom AdWords). S to odprtostjo pa so podjetja postala odprta tudi za goljufe.

Ena zelo znanih goljufij je goljufija s klikanjem na oglase (ang. click fraud), ki se je razpasla predvsem z razcvetom spletnega oglaševanja [6, 7]. Prek ponudnikov, kot sta na primer Yahoo in Google, lastnik spletne strani na svojo stran zelo preprosto doda oglase

in dobi del dobička od oglaševanja ob vsakem kliku na te oglase. Goljufi ta poslovni model izkoriščajo tako, da napišejo programe, ki nenehno klikajo na te oglase, s čimer posledično zaslužijo velikanske količine denarja. Business Week poroča, da je pri Googlu od 10 do 15 odstotkov klikov na oglase goljufivih, kar na letni ravni pomeni več kot milijardo dolarjev (podatki za leto 2006) [6].

Zaradi obsežnosti problema so si tako raziskovalci kot tudi praksa enotni, da je mogoče goljufije odkrivati in preprečevati le z nenehnim spremljanjem in nadzorovanjem poslovanja. Pri tem so nam lahko v veliko pomoč sistemi za detekcijo goljufij, ki nam omogočajo, da prizadevanja pri odkrivanju goljufij osredotočimo na primere, ki so bolj sumljivi in s tem povečamo uspešnost odkrivanja goljufij.

V nadaljevanju članka je v treh sklopih predstavljen sistem za detekcijo goljufij. Sekcija 2 podrobno opiše specifike oziroma probleme področja odkrivanja goljufij. V poglavju 3 je predstavljen predlog zasnove sistema za detekcijo goljufij. Zasnova je predstavljena skozi statični (arhitektura) in dinamični vidik (procesi), pri čemer je poudarek na tem, kako so v okviru teh vidikov naslovljeni omenjeni problemi. Zadnje poglavje povzame in s tem logično zaključi prispevek.

2 Karakteristike sistemov za detekcijo goljufij

Izkušnje tako praktikov kot tudi teoretikov pri odkrivanju goljufij kažejo, da ima problemska domena nekatere specifike oziroma probleme. Le-ti bistveno vplivajo na učinkovitost sistemov za detekcijo goljufij in kot taki pomenijo zahteve, ki jih morajo sistemi za detekcijo goljufij uspešno nasloviti. Te specifike lahko razporedimo v šest razredov:

A) specifičnost podatkov;
B) problem ocenjevanja uspešnosti detekcije;
C) pomembnost razlage sklepanja;
D) problem kombiniranja več metod;
E) zahteva po adaptivnosti;
F) zahteva po zmogljivosti sistema.

Prva posebnost (A) so podatki, nad katerimi se detekcija izvaja. Ker gre za velike količine podatkov, je v njih tudi veliko šuma. Dodaten problem je dejstvo, da je zelo težko pridobiti kakovostno učno množico. Eden od razlogov, da podjetja te množice skrivajo, je, da bi z objavo le-teh pomagala tudi goljufom, ki bi tako posredno ugotovili, kakšne vrste goljufij je podjetje sposobno odkriti. Dodaten problem pri učnih podatkih je, da ne moremo biti nikoli popolnoma prepričani o njihovi kakovosti. Nikoli namreč ne moremo biti prepričani, ali podatki, deklarirani kot pošteni, dejansko ne vsebujejo goljufivih podatkov, medtem ko nasprotno ne velja – podatki, ki so deklarirani kot goljufije, so vedno dejansko goljufije. S tujko ta pojav poimenujemo omission error. Dodaten problem je nesimetrična distribucija podatkov, ki otežuje uporabo klasičnih klasifikacijskih algoritmov [8, 9, 10].

Druga posebnost (B) je, da zaradi narave podatkov, kot tudi domene, ni smiselno uporabljati klasičnih metod za ocenjevanje uspešnosti detekcije, kot na primer klasifikacijske točnosti. Namesto tega se priporoča opiranje na ekonomska merila, kot so na primer čim večji denarni prihranki [1, 11].

Tretja posebnost (C) je posledica dejstva, da gre pri razreševanju goljufij za delikatne aktivnosti, kjer imamo opravka s strankami. Ker si ne želimo, da bi bil učinek razreševanja goljufije izguba strank, je zelo pomembno, da zna sistem utemeljiti svoje sklepanje. Zato se tehnike, kot je na primer metoda podpornih vektorjev (ang. support vector machines), ne uporabljajo veliko, čeprav so tradicionalno zelo uspešne. Nekateri raziskovalci so poskušali tudi nasprotno – torej da bi v ozki domeni detekcije goljufij poskušali bolje razložiti rezultate tehnik, kot na primer nevronskih mrež [12, 13].

Četrta posebnost (D) je, da se različne vrste goljufij v podatkih odražajo na različne načine. Zato potrebujemo različne metode za detekcijo in različne vhodne podatke. Za odkrivanje nekaterih goljufij potrebujemo podatke daljšega časa, medtem ko je za druge dovolj ena sama transakcija. Zelo pomembni lastnosti sistema za detekcijo goljufij sta zmožnost uporabe več različnih načinov detekcije in učinkovita kombinacija rezultatov le-teh [14].

Peta posebnost (E) domene ja izrazita potreba po adaptivnosti. Poslovno okolje je dinamično in prav tako storitve, ki jih ponujajo poslovni sistemi. Tem spremembam se prilagajajo tudi goljufi, ki nenehno iščejo nove načine zaslužka in s tem nove oblike goljufij. Ker so goljufi vedno korak pred nami, mora biti tudi sistem zasnovan adaptivno, kar pomeni, da se mora znati prilagajati novim oblikam goljufij [1].

Šesta posebnost (F) so visoke performančne zahteve. Pri detekciji goljufij je treba preiskovati velike količine podatkov. V domenah, kot so na primer telekomunikacije, bančništvo in nekateri segmenti e-poslovanja, je potrebno hitro ukrepanje, zato je treba goljufije detektirati v realnem času. Zato je potrebno, da je arhitektura sistema zasnovana učinkovito in razširljivo ter da je mogoče hitro dostopati do podrobnejših podatkov [8, 9].

3 Sistem za avtomatsko detekcijo goljufij pri e-poslovanju

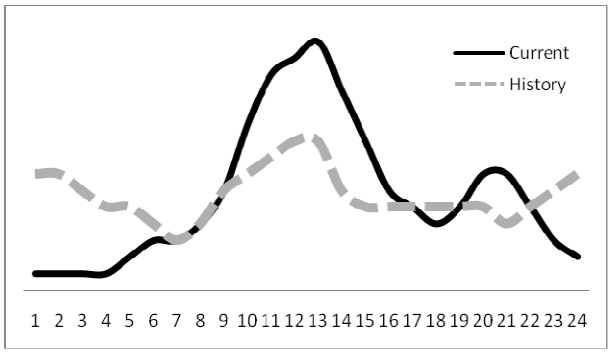
Pred nadaljevanjem je treba razjasniti distinkcijo med pojmoma sistem za detekcijo goljufij (ang. fraud detection system, v nadaljevanju včasih tudi samo sistem) in jedra za detekcijo goljufij (ang. fraud detection engine, v nadaljevanju FDE). Medtem ko FDE izvaja detekcijo goljufij, je sistem širši pojem. Sistem

vsebuje poleg FDE tudi uporabniški vmesnik, implementacijo učinkovitega dostopa do podatkov, podporo procesoma razreševanja goljufij, učenja in prilagajanja.
Zagovarjamo tezo, da lahko vse posebnosti in zahteve, naštete v prejšnjem poglavju, uspešno naslovimo zgolj na ravni sistema; sam FDE ni dovolj.

3.1 Arhitektura FDE

Komponenta FDE je z arhitekturnega vidika razdeljena na tri večje logične sklope oziroma module (Slika 2).
Množica modulov za odkrivanje znanih oblik goljufij omogoča učinkovito odkrivanje znanih oblik goljufij na različne medsebojno komplementarne načine. Primeri posameznih modulov so modul, ki deluje na podlagi pravil, ki jih poda domenski strokovnjak. Tak pristop že tradicionalno zelo dobre rezultate [19]. Drugi komplementarni moduli vključujejo predvsem različne klasifikacijske modele, ki so zgrajeni na podlagi označenih podatkov. Pri tem se lahko uporabljajo vse klasifikacijske tehnike s področja umetne inteligence, z izjemo tistih, ki so tradicionalno slabe pri razlagi sklepanja, kot na primer metoda podpornih vektorjev in nevronske mreže.
Moduli za odkrivanje novih oblik goljufij temeljijo na predpostavki, da so goljufije ena od oblik anomalij. Ta predpostavka se je že večkrat [1, 15, 16, 18] izkazala kot uporabna, saj dovoljuje uporabo različnih statistično-verjetnostnih metod in metod umetne inteligence za odkrivanje anomalij. V našem primeru se uporabljajo trije različni moduli.

- Modul za iskanje odstopanja od obnašanja v preteklosti. Modul deluje tako, da obnašanje danega ponudnika opišemo z večdimenzionalno podatkovno strukturo, ki jo na področju telekomunikacij imenujejo prstni odtis (ang. fingerprint) [17,18]. Ta podatkovna struktura povzema porazdelitve pomembnih atributov obnašanja ponudnika in tako ponuja kompaktno in medsebojno primerljivo reprezentacijo delovanja ponudnika v danem času. Tipični predstavniki dimenzij so čas transakcije in lokacija uporabnika, ki je uporabil storitev. Tako dobimo različne porazdelitve, ki nam pomenijo osnovo za iskanje odstopanja od tipičnega obnašanja (Slika 1).

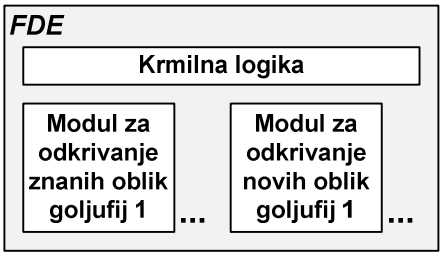


Slika 1: Porazdelitev količine storitev čez dan

Figure 1: Distribution of services during a day.

- Modul za iskanje odstopanja od obnašanja sorodnih ponudnikov e-storitev. Modul deluje enako kot iskanje odstopanja od preteklega obnašanja, le da v tem primeru ponudnika ne primerjamo s samim seboj, temveč s sorodnimi ponudniki.
- Modul za iskanje odstopanja od obnašanja sorodnih ponudnikov na podlagi hevristik obnašanja. Hevristike obnašanja so merljive metrike, ki naj se ne bi bistveno razlikovale med sorodnimi ponudniki. Primer take hevristike je odstotek obiskovalcev strani, ki prihajajo iz druge države kot ponudnik. Modul deluje tako, da išče ponudnike, ki odstopajo po več hevristikah. Za iskanje odstopanja se lahko uporabijo različne statistično-verjetnostne metode. Tak pristop je bil že uspešno uporabljen pri iskanju sumljivih ponudnikov zdravstvenih storitev na področju zdravstvenega zavarovanja [15, 16].

Tretja logična enota FDE je krmilna logika, katere glavna naloga je, da na učinkovit način združuje alarme, ki jih prožijo različni moduli. Za vsak posamezni alarm se odloči, ali ga je smiselno poslati analitiku v razreševanje. Pri tem je treba upoštevati različne vidike, kot na primer razmerje med stroški razreševanja in obsegom potencialne izgube; verjetnost, da gre dejansko za goljufijo; ali gre za znano ali novo obliko goljufije itd. Zadnji vidik je pomemben zato, ker (1) je verjetnost, da gre za goljufijo, pri alarmu, sproženem s strani modula za odkrivanje znanih oblik goljufij, veliko večja kot pri drugi vrsti modulov; (2) potencialen strošek znane oblike goljufij je zgolj strošek ene skupine transakcij, pri novih oblikah goljufij pa je treba upoštevati tudi potencialne prihranke, ki jih bomo imeli, če bomo znali učinkoviteje detektirati to novo obliko goljufij. Z vprašanjem, kdaj prožiti alarm, se je med drugim ukvarjal Viaene [11], s problemom kombiniranja pa npr. Brockett [14].



Slika 2: Arhitektura jedra za detekcijo goljufij

Figure 2. Fraud-detection engine architecture.

Z enoto FDE smo naslovili vse probleme, naštete v prejšnjem poglavju, z izjemo zahteve po visokih zmogljivostih.

3.2 Arhitektura sistema za detekcijo goljufij

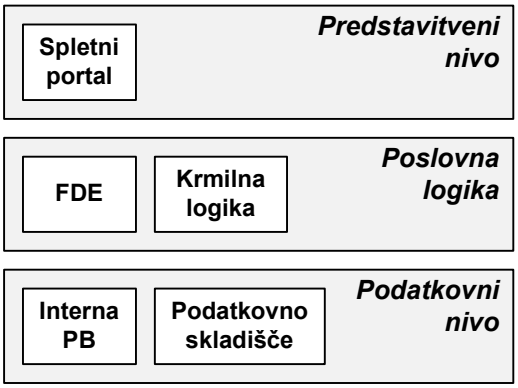
Sistem za odkrivanje goljufij je glede na priporočila učinkovite gradnje informacijskih sistemov dandanes zasnovan večnivojsko (Slika 3). Sestavljajo ga trije nivoji: (1) predstavitveni nivo, (2) poslovna logika in (3) podatkovni nivo.

Predstavitveni nivo vsebuje uporabniški vmesnik, ki uporabniku omogoča, da na učinkovit in uporabniku prijazen način komunicira s sistemom. Pomembne funkcionalnosti uporabniškega vmesnika so:

- učinkovit prikaz vseh sumljivih dogodkov v sistemu, ki analitiku omogočajo, da se pri veliki količini le-teh osredotoči na pomembnejše;
- učinkovit prikaz sklepanja, podprt z različnimi grafičnimi prikazi in možnostjo vpogleda v podrobne poslovne in transakcijske podatke;
- podpora procesoma odkrivanja in razreševanja goljufij (predstavljen v poglavju 3.3) ter procesu učenja in izboljšave (predstavljen v poglavju 3.4);

Nivo poslovne logika vsebuje vso logiko, ki omogoča delovanje sistema. Pomemben del tako poslovne logike kot tudi celotnega sistema je FDE, ki je bil podrobneje predstavljen v prejšnjem poglavju. Poleg FDE pa krmilna logika, katere naloga je, da učinkovito krmili druge module.

Podatkovni nivo zagotavlja persistenco podatkov, potrebnih za delovanje sistema, ter učinkovit in predvsem hiter dostop do teh. Naš predlog je, da se za ta namen uporabi kombinacija podatkovne baze in podatkovnega skladišča. Podatkovna baza hrani podatke, ki so potrebni za delovanje samega sistema za detekcijo goljufij. Podatkovno skladišče pa v večdimenzionalnih shemah hrani transakcijske podatke o poslovanju in tako zgornjim nivojem omogoča zelo hiter dostop do teh.



Slika 3: Arhitekturni diagram sistema za detekcijo goljufij

Figure 3. Architecture diagram of the proposed fraud detection system.

S predlagano arhitekturno zasnovo, predvsem večnivojsko zasnovo in uporabo podatkovnega skladišča, je mogoče zadostiti visokim performančnim zahtevam sistemov za detekcijo goljufij (F). Z uporabo učinkovitega uporabniškega vmesnika in učinkovitega dostopa do podatkov o poslovanju lahko ponudimo primerno utemeljitev vsakega posameznega alarma (C).

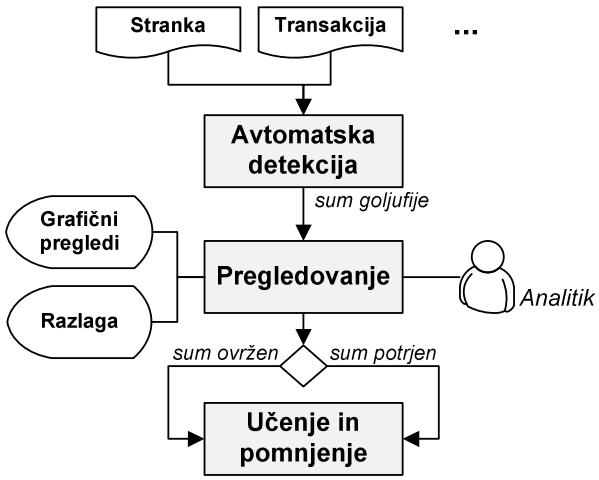
Sama struktura oziroma arhitektura sistema je zgolj statični vidik in z učinkovito arhitekturo lahko zagotavljamo učinkovitost sistema le do določene mere. Medtem ko na primer lahko zagotovimo, da so moduli za detekcijo goljufij adaptivni, s tem še ne zagotovimo, da bodo ti moduli dejansko dobili ustrezne podatke, ki so potrebni za proces adaptacije. Zato je treba s sistemom za odkrivanje goljufij podpreti tudi procese razreševanja goljufij.

3.3 Proces detekcije in razreševanja goljufij

Proces detekcije in razreševanja goljufij poteka v treh korakih. Prvi korak je avtomatska detekcija. V tem koraku FDE na podlagi različnih podatkov identificira sumljive vzorce obnašanja in če presodi, da je to smiselno, javi alarm.

Sledi drugi korak – pregledovanje –, v okviru katerega analitik pregleda posamezen alarm, pri čemer si pomaga z razlago sklepanja, ki jo poda FDE, ter različnimi grafičnimi prikazi in vpogledi v podatke poslovanja. Aktivnost pregledovanja posameznega alarma se logično konča z odločitvijo, ali gre v določenem primeru za goljufijo ali ne.

Alarm, opremljen z oznako, ali gre za goljufijo ali ne, je označen podatek, ki si ga moramo zapomniti, saj pomeni vhod za ponovno učenje in izboljševanje oziroma adaptacijo modulov za detekcijo goljufij.



Slika 4: Diagram procesa odkrivanja in razreševanja goljufije

Figure 4. Fraud-detection and resolution-process diagram.

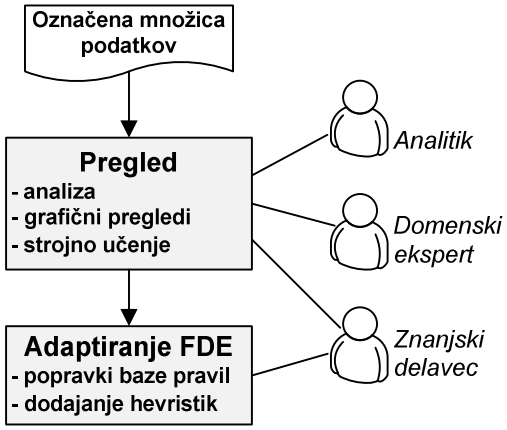
Proces odkrivanja in razreševanja goljufije poteka vsakodnevno in zagotavlja, da moduli dobivajo ustrezen odziv od analitika, na podlagi katerega se lahko izboljšujejo (E). Poleg tega proces zagotavlja beleženje aktivnosti uporabnika ter s tem povečuje učinkovitost ocenjevanja smiselnosti proženja alarma (B). Poleg tega je treba v okviru procesa analitiku prikazati učinkovito razlago sklepanja ter mu ponuditi dostop do poslovnih in transakcijskih podatkov, ki so bili vhod v posamezne module za detekcijo goljufij (C in D).

3.4 Proces učenja in izboljšave

Drug proces, ki ga mora podpirati sistem, je dolgotrajnejši in ga je smiselno izvesti ob daljših periodah, pri čemer se velja opreti predvsem na periodo, v kateri se izoblikuje dovolj velika množica označenih podatkov.

Glavni vhod v proces je množica označenih podatkov, ki se v okviru pregleda podrobno analizira. Pri tem se uporabijo različni grafični pregledi in orodja za strojno učenje. Pri tej aktivnosti morajo nujno sodelovati analitik, ki dnevno analizira goljufije in pozna posamezne primere; domenski strokovnjak, ki dobro pozna samo področje, in znanjski delavec, ki se spozna na strojno učenje in zna formalizirati znanje v tako obliko, ki je primerna za nadaljnjo uporabo s strani FDE.

Druga aktivnost procesa je adaptiranje FDE, pri čemer gre v glavnem za polavtomatsko popravljanje baze pravil, ki se uporablja pri odkrivanju znanih oblik goljufij. Pri tem se znanje, ki se je akumuliralo pri odkrivanju novih oblik goljufij, formalizira, s čimer se oblike goljufij, ki do tega trenutka niso bile identificirane, jasno definira in s tem neprimerno poveča natančnost in uspešnost njihovega detektiranja. Če nismo odkrili nobenih novih oblik goljufij, v tej aktivnosti zgolj dodajamo in popravljamo hevristike, ki naj izboljšajo natančnost odkrivanja goljufij.



Slika 5: Diagram procesa učenja in izboljševanja sistema

Figure 5. System-learning and adaptation-process diagram.

S procesom učenja in izboljševanja zagotovimo periodično izboljšavo predvsem modulov za detekcijo znanih oblik goljufij in s tem nenehno izboljšujemo uspešnost detekcije le-teh.

4 Sklep

Članek kot celoto predstavlja sistem za detekcijo goljufij v domeni e-poslovanja. Tabela 1 prikazuje, kako vidiki predlaganega sistema rešujejo specifike področja detekcije goljufij.

	Statični vidiki		Dinamični vidiki	
	Arhitektura FDE	Arhitektura sistema	Proces detekcije	Proces učenja
A	X	X		
B	X		X	X
C	X	X	X	X
D	X		X	X
E	X		X	X
F		X		

Tabela 1: Pokrivanje specifik domene odkrivanja goljufij, opisanih v poglavju 2, z različnimi vidiki sistema za detekcijo goljufij, predstavljenimi v poglavju 3

Table 1. Coverage of specifics, presented in Chapter 2 by different aspects given in Chapter 3.

Kot pojasnitev naj navedemo primer pomembnosti razlage sklepanja (C). Kot je razvidno iz tabele, je razlaga sklepanja pomembna v okviru obeh procesov. To, ali je mogoče sklepanje razložiti, je odvisno predvsem od algoritmov, ki so implementirani v modulih FDE. Poleg tega je učinkovita razlaga mogoča samo s sodelovanjem drugih komponent sistema, tj. s prijaznim uporabniškim vmesnikom s podporo vizualizaciji in možnostjo dostopa do podatkov o

poslovanju. Analogno je tudi večino drugih specifik mogoče nasloviti zgolj na ravni sistema za detekcijo goljufij kot celote.

5 Literatura

[1] R. J. Bolton, D. J. Hand, Statistical Fraud Detection: A Review. *Statistical Science*, 17, p.p. 235–249, 2002.

[2] M. H. Cahill, D. Lambert, H. C. Pinheiro, D. X. Sun, Detecting Fraud in Real World. *Handbook of Massive Data Sets*, p.p. 913–930, Kluwer Academic Publishers, 2002.

[3] NHCAA, *The Problem of Health Care Fraud: A Serious and Costly Reality for All Americans*. http://www.nhcaa.org/eweb/dynamicPage.aspx?webcode=anti_fraud_resource_centr&wpscode=TheProblemOfHCFraud, 2004.

[4] D. A. Hyman, HIPAA and Health Care Fraud: An Empirical Perspective. *Cato Journal*, 1, pp. 151–178, 2002.

[5] R. Kalakota, M. Robinson., E-Business, Addison-Wesley, New York, 1999.

[6] B. Grow, B. Elgin, M. Herbst, Click Fraud, The Dark Side of Online Advertising. *Business Week*, 6. 10. 2006.

[7] A. Tuzhilin, The Lane's Gift v. Google Report. http://googleblog.blogspot.com/pdf/Tuzhilin_Report.pdf, 7. 12. 2006.

[8] K. Tuyls, S. Maes, B. Vanschoenwinkel, *Machine Learning Techniques for Fraud Detection*. Master's thesis, Vrije Universiteit Brussel, Belgium, 2000.

[9] P. K. Chan, W. Fan, A. L. Prodromidis, S. J. Stolfo, Distributed Data Mining in Credit Card Fraud Detection. *IEEE Intelligent Systems*, 14, p.p. 67–74, 1999.

[10] M. Artis, M. Ayuso, M. Guillen, Detection of Automobile Insurance Fraud With Discrete Choice Models and Misclassified Claims. *The Journal of Risk and Insurance*, vol. 63, no. 3, pp. 325–340.

[11] S. Viaene, M. Ayuso, M. Guillen, D. V. Gheel, G. Dedene, Strategies for Detecting Fraudulent Claims in the Automobile Insurance Industry. *European Journal of Operational Research*, vol. 176, pp. 565–583, 2007.

[12] R. Wheeler, S. Aitken, Multiple Algorithms for Fraud Detection. *Knowledge-Based Systems*, 13, p.p. 93–99, 2000.

[13] S. Viaene, G. Dedene, R. A. Derrig, Auto Claim Fraud Detection Using Bayesian Learning Neural Networks. *Expert Systems with Applications*, vol. 29, pp. 653–666, 2005.

[14] P. L. Brockett, R. A. Derrig, L. L. Golden, A. Levine, M. Alpert, Fraud Classification Using Principal Component Analysis of RIDIT. *The Journal of Risk and Insurance*, vol. 69, no. 3, pp. 341–371, 2002.

[15] E. Cox, A Fuzzy System for Detecting Anomalous Behaviors in Healthcare Provider Claims. *Intelligent Systems in Finance and Business*, Wiley, 2005.

[16] J. A. Major, D. R. Riedinger, EFD: Heuristic Statistics for Insurance Fraud Detection. *Intelligent Systems in Finance and Business*, Wiley, 2005.

[17] T. Fawcett, F. Provost, Adaptive Fraud Detection. *Data Mining and Knowledge Discovery*, Kluwer, 1997.

[18] D. Xing, M. Girolami, Empoying Latent Dirichlet Allocation for Fraud Detection in Telecommunications. *Pattern Recognition Letters*, doi: 10.1016/j.patrec.2007.04.015, 2007.

[19] J. Shawe-Taylor, K. Howker, P. Burge, R. Holloway, Detection of Fraud in Mobile Telecommunications. *Information Security Technical Report*, vol. 4, no. 1, pp. 16–26, 1999.

Štefan Furlan je doktorski študent na Fakulteti za računalništvo in informatiko, kjer je tudi zaposlen kot mladi raziskovalec. Dodiplomski študij je leta 2006 predčasno končal in za diplomsko delo prejel Prešernovo nagrado. Med študijem je sodeloval s številnimi laboratoriji in raziskovalnimi institucijami, tako s področja računalništva in informatike kot tudi bioinformatike, prava in ekonomije. Poleg tega je delal v več slovenskih podjetjih na področju informatike. Trenutno se poklicno ukvarja s področjem odkrivanja goljufij. Na tem področju se udejstvuje tako raziskovalno kot tudi s sodelovanjem pri različnih gospodarskih projektih. Je član Slovenskega društva INFORMATIKA ter ustanovni in častni član Skupine GI.

Dr. Marko Bajec je docent na Fakulteti za računalništvo in informatiko Univerze v Ljubljani, kjer predava predmete s področja informacijskih sistemov. Raziskovalno in v praksi se ukvarja predvsem s področji, kot so: načrtovanje in uvajanje metodologij razvoja informacijskih sistemov, strateško planiranje informatike, poslovno modeliranje; elektronsko poslovanje ipd. Je izvoljeni predsednik društva Association for Information Systems, podpredsednik Slovenskega društva INFORMATIKA ter član več strokovnih in znanstvenih združenj. Svoje prispevke objavlja v domačem in mednarodnem prostoru.