ISKANJE DOBRIH NASTAVITEV MODULA V INTELIGENTNEM SISTEMU ZA NADZOR PRISTOPA

Tea Tušar, Matjaž Gams
Odsek za inteligentne sisteme
Institut "Jožef Stefan"
Jamova cesta 39, 1000 Ljubljana, Slovenija
tea.tusar@ijs.si, matjaz.gams@ijs.si

POVZETEK

Predstavljamo modul mikro učenja v inteligentnem sistemu za nadzor pristopa, ki opazuje obnašanje uporabnikov pred posameznimi točkami vstopa in s pomočjo algoritma za odkrivanje izjem opozarja na nenavadne vstope. Naša naloga je nastaviti parametre modula tako, da bo delal čim bolje tj. čim bolj pravilno prepoznaval neobičajne vstope od običajnih. V ta namen na množici poskusnih vstopov definiramo želene oznake teh vstopov in preiskujemo prostor parametrov modula, da bi dobili oznake, ki so čim bolj podobne želenim. Najboljša nastavitev parametrov modula mikro učenja doseže 92.5% točnost na poskusnih vstopih.

1 UVOD

Sisteme za nadzor pristopa običajno sestavlja nekaj (biometričnih) senzorjev, ki se uporabljajo za identifikacijo in verifikacijo uporabnikov. Če uporabnik pristopi k vsem senzorjem pravilno, ima omogočen vstop v varovani prostor. Z uporabo inteligentnih metod želimo obstoječi sistem za nadzor pristopa nadgraditi tako, da bo omogočal prepoznavanje ustaljenih vzorcev obnašanja za vsakega posameznega uporabnika in posledično odkrivanje izjem, ki lahko predstavljajo poskus vstopa neavtorizirane osebe. Na neobičajne vstope se bo odzival s posredovanjem *opozorila* ali *alarma* nadzorniku, ki bo lahko nato ustrezno ukrepal.

Obravnavani inteligentni sistem za nadzor pristopa je sestavljen iz štirih modulov:

- modula, ki za odkrivanje prepovedanih vstopov uporablja *ekspertna pravila*,
- modula *mikro učenja*, ki opazuje obnašanje pred posamezno točko vstopa,
- modula *makro učenja*, ki nadzoruje gibanje med različnimi točkami vstopa, ter
- *kamere*, ki snema vse vstope in pri tem spremlja gibe uporabnika.

Po vstopu uporabnika bo vsak modul sporočil ali je pri vstopu opazil kakšne posebnosti. Možni izidi vsakega modula so: *OK* (✓), *opozorilo* (?) in *alarm* (X). Ko sistem dobi izhode vseh modulov, iz njih sestavi skupen rezultat, ki je odvisen od izbrane občutljivosti. Računanje skupnega izida si lahko ogledamo s pomočjo primera s slike 1. Tu modul z ekspertnimi pravili ni zaznal težav, medtem ko je modul makro učenja vrnil opozorilo, modula mikro učenja in kamere pa alarm. Občutljivost je nastavljena na 3/4, kar pomeni da bo skupni rezultat enak tretjemu najboljšemu izidu – v našem primeru je to alarm.

Slika 1: Izidi posameznih modulov in skupni rezultat na primeru vstopa nekega uporabnika.

V tem prispevku se bomo omejili le na modul mikro učenja – natančneje, na iskanje nastavitev tega modula tako, da bodo rezultati modula in posledično inteligentnega sistema za nadzor pristopa čim boljši.

2 MODUL MIKRO UČENJA

Obravnavali bomo sistem za nadzor pristopa, ki ga sestavljajo čitalnik brezkontaktnih identifikacijskih kartic, biometrični čitalnik prstnih odtisov in senzorji na vratih. V našem primeru modul mikro učenja spremlja običajno

obnašanje uporabnika pred posameznim vstopom, ki ga merimo s pomočjo naslednjih treh časov:

- čas med identifikacijo z brezkontaktno kartico in verifikacijo s prstnim odtisom,
- čas med verifikacijo s prstnim odtisom in odprtjem vrat,
- čas med odprtjem in zaprtjem vrat.

S spremljanjem teh treh časov se modul s pomočjo algoritma za odkrivanje izjem lahko nauči kakšno je običajno vstopanje uporabnika in opozori, ko pride do odstopanj. Pri tem je treba v modulu nastaviti nekaj parametrov, ki vplivajo na njegovo delovanje.

2.1 Algoritem LOF

Med množico algoritmov, ki se lahko uporabijo za odkrivanje izjem [3], smo za naš konkreten primer izbrali algoritem LOF[2, 1]. Zanj smo se odločili, ker nam ne vrne zgolj podatka, ali je nek vstop izjemen ali ne, ampak definira oceno izjemnosti kot realno število, imenovano *lokalni koeficient izjemnosti*, ali krajše LOF. To je zelo zaželena lastnost, saj moramo za vsak vstop posebej določiti ali je običajen ali ne – v slednjem primeru pa rabimo tudi informacijo o tem, kolikšno je odstopanje od običajnih vstopov, da se lahko odločimo ali bomo sprožili alarm ali samo opozorilo.

Dodatna prednost algoritma LOF je v tem, da izjemnost opazovanega vstopa definira glede na lokalno gostoto tega vstopa in lokalne gostote njegovih najbližjih sosedov kot:

$$LOF(p) = \frac{1}{|\{sosedi(p)\}|} \sum_{o \in \{sosedi(p)\}} \frac{lokalna\ gostota(o)}{lokalna\ gostota(p)}.$$

Zato dela dobro tudi v primerih, ko so časi vstopov neenakomerno distribuirani po prostoru. Število najbližjih sosedov, ki jih upoštevamo v tem izračunu, je parameter algoritma. Za običajne vstope je $LOF \leq 1$, medtem ko lahko vstope, pri katerih je $LOF > 1$ štejemo za izjemne – večji kot je LOF, večja je izjemnost vstopa.

2.2 Nastavitve modula mikro učenja

Za praktično uporabo modula mikro učenja na našem primeru moramo določiti vrednosti naslednjih parametrov:

- število najbližjih sosedov v algoritmu LOF (n_s),
- mejno vrednost LOF med običajnimi vstopi in tistimi, ki sprožijo opozorilo (m_o),
- mejno vrednost LOF med vstopi, ki sprožijo opozorilo, in vstopi, ki sprožijo alarm (m_a).

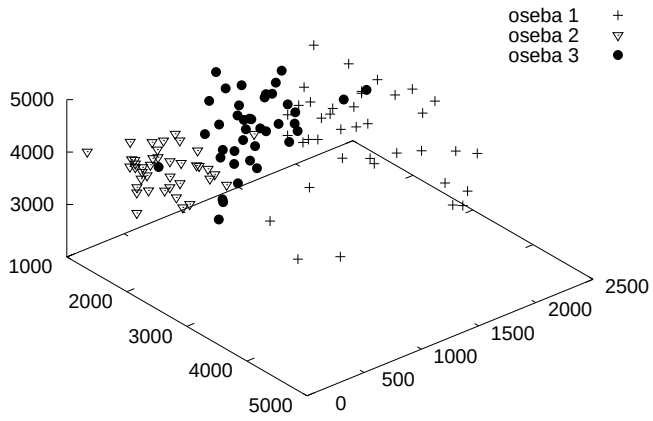
Ker je od izbire vrednosti teh parametrov močno odvisno delovanje modula za mikro učenje, smo bomo vrednosti parametrov določili na podlagi rezultatov poskusov, ki so opisani v naslednjem razdelku.

3 POSKUSI

3.1 Podatki

Za potrebe naše raziskave smo beležili čase vstopov treh različnih oseb. Za vsako smo zabeležili 40 regularnih vstopov, poleg tega pa še skupno 17 neregularnih vstopov vseh oseb, ki so posnemali "nenavadne" vstope. Med slednje spadajo igrane ugrabitve, poskusi vstopa z nenavadnimi predmeti ter t. i. smukanje (vstop več oseb naenkrat pri identifikacije ene same osebe).

Kot že rečeno, je vsak vstop definiran s tremi časi, ki se beležijo med vstopom. Tako lahko vsak vstop predstavimo s točko v tridimenzionalnem prostoru. Na sliki 2 predstavljamo regularne vstope vseh treh oseb. Hitro lahko opazimo dvoje: (1) časi vstopov treh oseb se mestoma prekrivajo ter (2) medtem ko so razlike med časi posameznih vstopov osebe 2 zelo majhne, to ne velja za osebi 1 in 3.



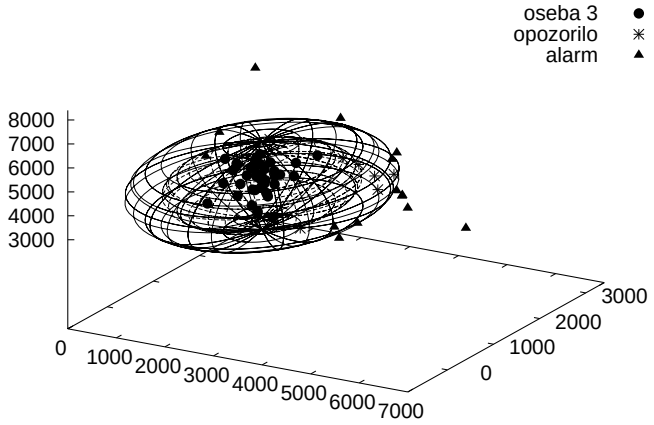
Slika 2: Vstopi treh različnih oseb kot točke v tridimenzionalnem prostoru (na oseh so predstavljeni posamezni časi vstopov v milisekundah).

Za vsako osebo smo zbrane vstope razdelili na učno in testno množico na dva načina:

1. Vse regularne vstope obravnavane osebe smo (po vzoru 10-kratnega prečnega preverjanja) razdelili na 10 učnih in 10 testnih množic tako, da vsak vstop nastopa natanko enkrat v testni množici. Želimo si, da bi algoritem LOF za vse vstope iz testnih množic pravilno ugotovil, da niso izjemni.
2. Vse regularne vstope obravnavane osebe smo združili v učno množico, vse preostale vstope (80 regularnih vstopov drugih dveh oseb in vseh 17 neregularnih vstopov) pa v testno množico. Tu si želimo, da bi algoritem LOF znal pravilno razločevati med vstopi, ki so izjemni, in tistimi, ki (s stališča navad obravnavane osebe) to niso.

Torej moramo za vse vstope določiti želen izid (OK, opozorilo ali alarm) in parametre modula mikro učenja nastaviti tako, da se bodo tem izidom čim bolj približali.

Da bi se izognili ročnemu označevanju vstopov, smo si pomagali z elipsoidi. Okrog regularnih vstopov obravnavane osebe smo naredili osnovni elipsoid, ki ima središče v centru vseh teh vstopov in katerega polos r_i v smeri dimenzije i je enaka polovici razdalje med dvema najbolj skrajnima vstopoma v dimenziji i. Nato smo naredili še dva druga elipsoida – enega s polosmi enakimi $r_i^1 = 1.6r_i$ in drugega, katerega polosi so enake $r_i^2 = 2.2r_i$. Konstanti 1.6 in 2.2 smo določili eksperimentalno tako, da so v manjšem elipsoidu vsebovani vsi regularni vstopi osebe, med obema elipsoidoma vstopi, ki bi morali sprožiti opozorilo, ter zunaj večjega elipsoida vstopi, ki bi morali sprožiti alarm. Elipsoida torej označujeta mejo med vstopi OK/opozorilo ter opozorilo/alarm. Na sliki 3 lahko vidimo kako izgledata takšna elipsoida za osebo 3.



Slika 3: Elipsoida, ki označujeta mejo med vstopi OK/opozorilo ter opozorilo/alarm za osebo 3.

3.2 Prostor parametrov

S pomočjo elipsoidov smo označili vse vstope z želenim izidom. Želimo najti tiste nastavitve modula za mikro učenje, ki bodo vstope označile čim bolj podobno želenemu izidu. Za vsak parameter smo preizkusili naslednje nastavitve:

- število najbližjih sosedov $n_s \in \{3, 4, \ldots, 30\}$,
- mejna vrednost LOF za opozorilo $m_o \in \{1.1, 1.2, 1.3, 1.4, 1.5, 1.6\}$,
- mejna vrednost LOF za alarm $m_a \in \{1.6, 1.7, 1.8, 1.9, 2.0, 2.1\}$.

Vseh možnih nastavitev je 28 224.

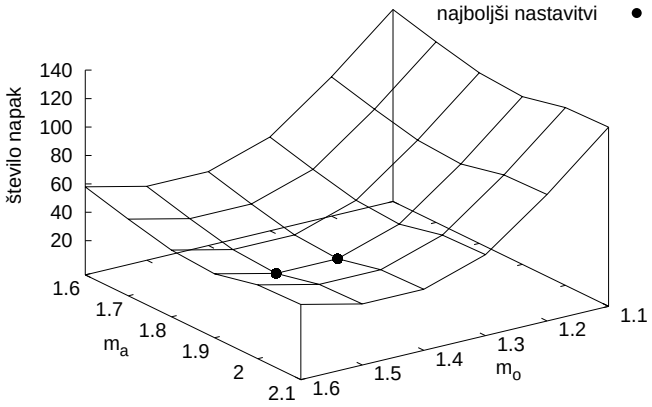
3.3 Rezultati in diskusija

Za vsako nastavitev smo pognali modul mikro učenja za vsako osebo na obeh nalogah učenja in beležili, koliko napak je naredil modul glede na predhodno določene želene oznake. Vse napake smo sešteli in iskali nastavitev, pri kateri je skupna napaka najmanjša. Najmanjšo skupno napako, enako 31, smo dobili v dveh primerih:

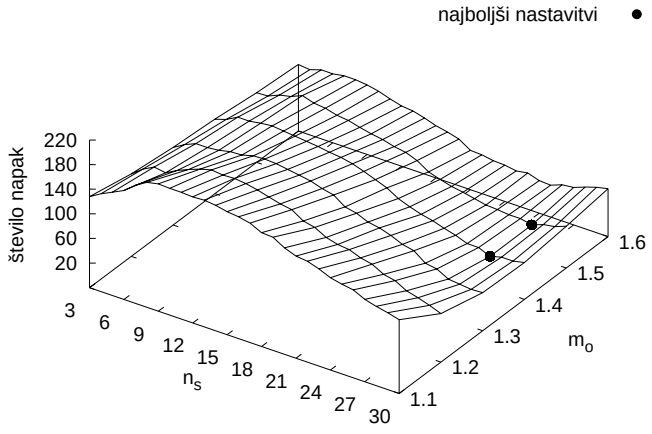
- $n_s = 27$, $m_o = 1.4$, $m_a = 1.9$ in
- $n_s = 27$, $m_o = 1.5$, $m_a = 1.9$.

V prvem primeru naredi modul mikro učenja eno napako na prvi nalogi in 30 napak na drugi nalogi, medtem ko so v drugem primeru vse napake narejene na drugi nalogi učenja. Ker si želimo, da bi bili vsi regularni vstopi označeni kot OK, bomo za praktično uporabo izbrali drugo nastavitev.

Vpliv posameznih parametrov na izvajanje modula mikro učenja si lahko pogledamo s pomočjo slik 4 in 5. Prva slika prikazuje kako se spreminja skupna napaka glede na izbrani mejni vrednosti LOF za opozorilo in alarm pri določenem številu najbližjih sosedov ($n_s = 27$), medtem ko druga kaže skupno napako pri izbrani mejni vrednosti LOF za alarm ($m_a = 1.9$).



Slika 4: Skupno število napak za $n_s = 27$ pri različnih vrednostih parametrov m_o in m_a.



Slika 5: Skupno število napak za $m_a = 1.9$ pri različnih vrednostih parametrov n_s in m_o.

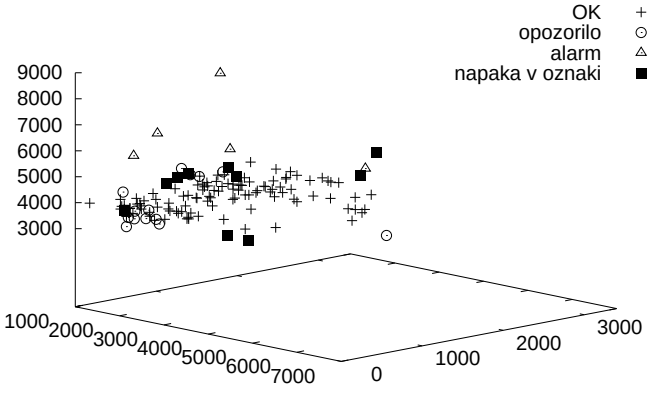
Kot priznavajo tudi avtorji algoritma LOF v [2], je zelo težko določiti pravo vrednost za število najbližjih

sosedov n_s. Kot lahko vidimo s slike 5, v našem primeru prinesejo najboljše rezultate vrednosti okrog 27. Vendar pa je pri tem parametru potrebna pazljivost, saj je močno odvisen od števila obravnavanih točk. V našem primeru imamo v učni množici bodisi 36 bodisi 40 točk. Pri (bistveno) drugačnem številu točk, bi morali poskus ponoviti, saj bi vrednost $n_s = 27$ verjetno ne dala najboljših rezultatov. Preostala dva parametra modula mikro učenja sta skoraj neobčutljiva na velikost učne množice. Odvisna sta predvsem od tega, kako občutljiv modul želimo imeti.

Za konec si poglejmo še kako deluje modul pri izbranih vrednostih parametrov: $n_s = 27$, $m_o = 1.5$, $m_a = 1.9$. Kot že rečeno, je vseh napak 31, pri čemer ni nobena storjena na prvi nalogi učenja. Pri drugi nalogi pa so napake razdeljene po osebah tako, kot prikazuje tabela 1. Za posamezne osebe prikazujemo napake tudi na slikah 6, 7 in 8, iz katerih je razvidno, kateri vstopi so modulu povzročili največ težav. Poudariti je treba, da modul ni storil nobenih "dvakratnih" napak, pri katerih bi OK vstop označil za alarm oz. alarm označil kot OK. To je še posebej pomembno s praktičnega vidika, saj veča zaupanje nadzornika v pravilno delovanje modula.

napoved	pravilno	oseba 1	oseba 2	oseba 3	$\sum$
OK	opozorilo	6	4	3	13
OK	alarm	0	0	0	0
opozorilo	OK	3	1	0	4
opozorilo	alarm	1	2	8	11
alarm	OK	0	0	0	0
alarm	opozorilo	0	3	0	3
	$\sum$	10	10	11	31

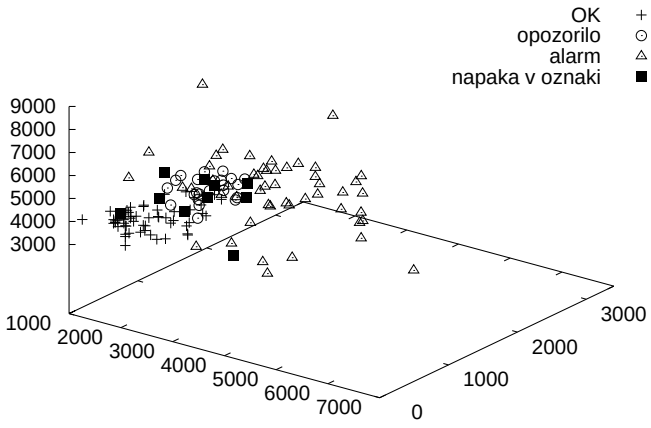
Tabela 1: Tipi napak za vsako osebo.



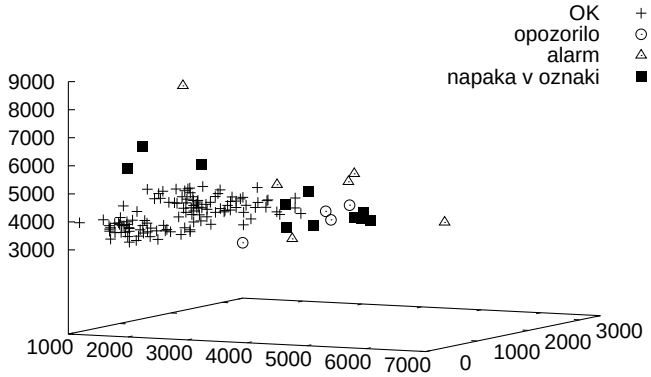
Slika 6: Izidi modula mikro učenja (s poudarjenimi napakami) za osebo 1.

4 ZAKLJUČEK

V prispevku smo predstavili modul za mikro učenje obnašanja uporabnika pred točkami vstopa, ki z uporabo algoritma za odkrivanje izjem označuje vstope z oznakami OK, opozorilo ali alarm. Modul smo preiskusili na testni množici podatkov, ki smo jih prej označili z želenimi izidi. Na teh podatkih modul doseže 92.5% točnost. Posebej pomembno pa je to, da ne dela "dvakratnih" napak med vstopi OK/alarm.



Slika 7: Izidi modula mikro učenja (s poudarjenimi napakami) za osebo 2.



Slika 8: Izidi modula mikro učenja (s poudarjenimi napakami) za osebo 3.

Literatura

[1] M. M. Breunig. *Quality Driven Database Mining.* PhD thesis, University of Munich, 2001.

[2] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. LOF: Identifying density-based local outliers. V *Proceedings of the International Conference on Management of Data (SIGMOD'00)*, strani 93–104, 2000.

[3] T. Tušar and M. Gams. Odkrivanje izjem na primeru inteligentnega sistema za kontrolo pristopa. V *Zbornik devete mednarodne multiconference Informacijska družba (IS 2006)*, zvezek B, strani 136–139, 2006.