

*Obvladovanje zaupanja v e-okoljih*

Eva Zupančič

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2014



## IZJAVA

*Izjavljam, da sem avtor dela in da slednje ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo za pridobitev naziva na univerzi ali na drugem visokošolskem zavodu, razen v primerih, kjer so navedeni viri.*

— Eva Zupančič —  
november 2014

ODDAJO SO ODOBRILI

dr. Denis Trček

*redni profesor za računalništvo in informatiko*

MENTOR IN ČLAN OCENJEVALNE KOMISIJE

dr. Marjan Krisper

*izredni profesor za računalništvo in informatiko*

PRESEDNIK OCENJEVALNE KOMISIJE

dr. Matjaž Debevc

*izredni profesor za računalništvo in informatiko*

ZUNANJI ČLAN OCENJEVALNE KOMISIJE

Univerza v Mariboru



## PREDHODNA OBJAVA

Izjavljam, da so bili rezultati obravnavane raziskave predhodno objavljeni/sprejeti za objavo v recenzirani reviji ali javno predstavljeni v naslednjih primerih:

- [1] E. Zupančič, D. Trček. QADE: A Novel Trust and Reputation Model for Handling False Trust Values in e-Commerce Environments with Subjectivity Consideration. *Technological and Economic Development of Economy*, sprejeto za objavo.
- [2] E. Zupančič, D. Trček. Deployment of Trust Management System in Environment of e-commerce. V: *Proceedings of Agent and multi-agent systems: technologies and applications (AMSTA), Lecture Notes in Artificial Intelligence*, 7327, str. 484–493, Berlin, Heidelberg, 2012. Springer-Verlag. doi: [10.1007/978-3-642-30947-2\\_53](https://doi.org/10.1007/978-3-642-30947-2_53)
- [3] E. Zupančič, D. Trček, K. Taveter. Agent-Oriented Engineering of Trust Management Systems. V: *Proceedings of the First International Conference on Agreement Technologies (AT 2012)*, str. 349–350, Dubrovnik, Croatia, 2012. ACTA Press.
- [4] E. Zupančič, D. Trček. The evaluation of Qualitative Assessment Dynamics (QAD) methodology for managing trust in pervasive computing environments. V: *Proceedings of ICPCA/SWS 2011*, str. 67–73, Port Elizabeth, 2011. IEEE.  
doi: [10.1109/ICPCA.2011.6106480](https://doi.org/10.1109/ICPCA.2011.6106480)
- [5] E. Zupančič, D. Trček. Pregled nekaterih simulacijskih tehnik za modeliranje človeškega dejavnika. V: *Zbornik devetnajste mednarodne Elektrotehniške in računalniške konference ERK 2010, Zvezek A*, str. 433–436, Portorož, Slovenija, 2010. Slovenska sekcija IEEE.
- [6] E. Zupančič, D. Trček. Pregled metodologij za računalniško podprto obvladovanje zaupanja in ugleda. V: *Zbornik osemnajste mednarodne elektrotehniške in računalniške konference ERK 2009, Zvezek B*, str. 81–84, Portorož, Slovenija, 2009. Slovenska sekcija IEEE.

Potrjujem, da sem pridobila pisna dovoljenja vseh lastnikov avtorskih pravic, ki mi dovoljujejo vključitev zgoraj navedenega materiala v pričujočo disertacijo. Potrjujem,

*f*

E. Zupančič

da zgoraj navedeni material opisuje rezultate raziskav, izvedenih v času mojega podiplomskega študija na Univerzi v Ljubljani.

## POVZETEK

Zaupanje je bistvenega pomena za učinkovito sodelovanje, kar pomeni, da medsebojna izbira (poslovnih) partnerjev in posledično sprejemanje odločitev temelji na tem, koliko ti zaupajo drug drugemu. Način vrednotenja zaupanja v e-okolju se razlikuje od tistega v fizičnem svetu, saj je spletno okolje v primerjavi s fizičnim svetom precej omejeno glede razpoložljivosti in prisotnosti indikatorjev zaupanja. V e-okoljih zaupanje vrednotimo z uporabo sistemov za obvladovanje zaupanja in ugleda, ki temeljijo na različnih matematičnih modelih zaupanja in ugleda. Predpogoj za učinkovit sistem za obvladovanje zaupanja in ugleda je, da implementira razumno in smiselno formalizacijo zaupanja in ugleda v e-okoljih.

Doktorska disertacija temelji na modelu kvalitativne dinamike ocenjevanja (angl. Qualitative Assessment Dynamics, QAD), ki v definicijo procesa dojemanja in vrednotenja zaupanja vpelje človeške dejavnike. S tem predstavlja preudaren in prepričljiv model zaupanja in ugleda. Na podlagi računalniških simulacij smo model QAD ovrednotili in analizirali. Načrtovali smo različne scenarije, v katerih imajo agenti v e-okolju vloge iskalcev in ponudnikov storitev, pri čemer nekateri ponudniki storitev priskrbijo storitve z nezadovoljivo kakovostjo. Izvedli smo dva niza simulacijskih eksperimentov. Prvi niz prikaže učinek uvedbe preprostega sistema za obvladovanje zaupanja in ugleda v e-okolje ter prikaže osnovne lastnosti vrednotenja zaupanja, kot ga definira model QAD. Model QAD definira t. i. *operatorje QAD*, ki modelirajo vedenje agentov in postopek vrednotenja zaupanja. Z drugim sklopom poskusov demonstriramo dodatne lastnosti in učinke operatorjev QAD. Analizirali smo vedenje agentov, kot ga definirajo operatorji QAD, in s primerjavo odločitev, ki so jih sprejeli agenti s pripisanimi različnimi operatorji, pojasnili lastnosti operatorjev QAD.

Ena od izrazitih težav v sistemih za obvladovanje zaupanja in ugleda je prisotnost neresničnih ocen zaupanja, ki jih v sistem posredujejo zlonamerni agenti. Za rešitev

navedenega problema smo predlagali model zaupanja QADE (angl. Qualitative Assessment Dynamics Extended), ki osnovni model zaupanja QAD razširja z elementi, kot so *zasebni vektor ocen zaupanja*, *zgodovinske matrike zaupanja*, *zgodovinski zasebni vektorji ocen*, *splošna miselnost agenta*, *operatorji zaupanja QADE*, *agent napadalec in funkcija podobnosti sim*. S tem model zaupanja in ugleda QADE predpostavlja možnost obstoja neresničnih ocen zaupanja. Hkrati model QADE upošteva človeško naravo zaupanja in predpostavlja, da različne ocene zaupanja lahko pomenijo neresnične (zlagane) ocene, lahko pa izkazujejo razlike v odnosih do zaupanja različnih agentov. Na podlagi modela zaupanja QADE razvijemo metodo za obvladovanje zlaganih ocen zaupanja. Metoda je sestavljena iz dveh delov. Najprej agent ocenjevalec poišče ostale agente v danem e-okolju s podobnim odnosom do zaupanja. Podobnost med agenti se izračuna s primerjavo ocen zaupanja, ki sta jih podala primerjana agenta, in primerjavo t. i. *splošne miselnosti* primerjanih agentov. Agent ocenjevalec primerja lastne ocene zaupanja z ocenami, ki so jih podali podobni agenti, in s t. i. *črno piko* oceni tiste, ki so bistveno različne. Ocene zaupanja, ki jih s črno piko označi določeno število agentov, veljajo za nepošteno in so izključene iz nadaljnjega postopka izračuna vrednosti zaupanja. Učinkovitost predlagane metode smo ovrednotili na podlagi večagentnih simulacij. Izvedli smo simulacije v simulacijskih skupnostih z različnim številom agentov napadalcev in napadenih agentov kot tudi z različnimi vrstami napadalcev – skupinskimi in samostojnimi. Rezultati so pokazali, da je naša predlagana metoda v povprečju boljša za 28 % do 58 % v primerjavi z reprezentativno eksogeno metodo, ki jo je predlagal Teacy s soavtorji, in z reprezentativno endogeno metodo, ki so jo predlagali Whitby in soavtorji.

Ocene zaupanja, ki jih posredujejo agenti v e-okolju, so subjektivne, saj je zaupanje odvisno tudi od agentovega *nagnjenja k zaupanju*. Izračun ugleda agentov z združevanjem posredovanih ocen zaupanja je lahko vprašljiv, če dejavnik subjektivnosti pri vrednotenju zaupanja ni ustrezno upoštevan. Da bi vsi uporabniki enako razumeli vrednosti ugleda agentov, je treba posredovane ocene zaupanja ustrezno prilagoditi vsakemu agentu posebej. V zadnjem delu doktorske disertacije obravnavamo ustrezno združevanje in prilagajanje posredovanih ocen zaupanja. Kot rešitev predlagamo razširitve in modifikacije modela zaupanja QAD ter razvijemo metodo HOMRA (angl. Human-Oriented Method for Ratings Adaptation), ki izračuna posameznemu agentu prilagojene vrednosti ugleda drugih agentov z izločitvijo dejavnika subjektivnosti. Z uporabo metode za prilagoditev ocen zaupanja HOMRA imajo vsi agenti primerljive

možnosti pri izbiri agentov v nadaljnjih transakcijah, ne glede na njihovo nagnjenje k zaupanju. Predlagana metoda HOMRA z uporabo neparametričnih statističnih testov poišče agente s podobnimi nagnjenji k zaupanju. V naslednjem koraku izračuna posameznemu agentu prilagojeno vrednost ugleda drugega (ocenjevanega) agenta v e-okolju, tako da združi ocene zaupanja do ocenjevanega agenta, ki so jih posredovali agenti s podobnimi nagnjenji k zaupanju. Lastnosti agentovega nagnjenja k zaupanju so izpeljane iz preteklih ocen zaupanja, s katerimi je dani agent ocenil druge agente. Metoda HOMRA od agentov ne zahteva razkritja procesov vrednotenja zaupanja kot tudi ne kriterijev, prepričanj, meril, motivacijskih ali drugih faktorjev, ki vplivajo na proces evalvacije zaupanja. Uspešnost naše metode smo ocenili na podlagi obsežnih simulacij z različnim številom agentov v e-okolju, različno porazdelitvijo njihovih tipov osebnosti in različnim odstotkom razpoložljivih ocen zaupanja. Rezultati so pokazali bistvene izboljšave naše metode v primerjavi z drugimi metodami za prilagoditev ocen zaupanja. V povprečju je metoda HOMRA 50 % učinkovitejša od metode, ki sta jo predlagala Abdul-Rahman in Hailes, in 73 % učinkovitejša od metode, ki so jo predlagali Hasan in soavtorji, upoštevajoč konfiguracije z različnim številom agentov v e-okolju, konfiguracije z različnimi porazdelitvami osebnostnih tipov agentov kot tudi konfiguracije z različnimi odstotki razpoložljivih ocen zaupanja.

*Ključne besede:* zaupanje, ugled, modeliranje zaupanja in ugleda, sistemi za obvladovanje zaupanja in ugleda, zlagane ocene, subjektivnost, človeški dejavniki, e-okolje



University of Ljubljana  
Faculty of Computer and Information Science

Eva Zupančič  
*Managing trust in e-environments*

## ABSTRACT

Trust is essential to effective collaboration, meaning that (business) partners choose each other and make decisions based on how much they trust one another. The way how to assess trust in e-environments is different from those in physical world, as there are limited indicators available in online environment. In e-environments, trust is evaluated with the usage of trust and reputation management systems (TRMS) that are based on different mathematical trust and reputation models. The underlying mathematical trust and reputation model should be reasonable in order to implement effective TRMS.

The thesis is based on the Qualitative Assessment Dynamics (QAD) trust model, which considers human factors involved in trust reasoning and as such presents sound mathematical trust and reputation model. At first, we evaluated the QAD trust model based on simulations. We proposed different scenarios with agents that played service requester and service provider roles, where some of the service providers provide unsatisfying services. Further, we performed two set of experiments. The first set of experiments demonstrates the effect of introducing trust and reputation management system into environment and demonstrates the basic properties of trust evaluation defined within the QAD trust model. The QAD trust model defines different QAD operators that model an agent's behavior and process of trust evaluation. The second set of experiments shows the properties and the effects of the QAD operators. We analyzed the behavior of the agents modeled with different QAD operators and explained the QAD operators properties based on comparison of the decisions they made.

One of the prominent problems within trust and reputation management systems is the presence of unfair ratings, which has not been sufficiently addressed so far. To address the problem of unfair trust ratings we present the QADE trust model, which extends the QAD trust model with elements such as *private trust vector*, *historical trust*

*matrices, historical private trust vectors, agents' general mindset, QADE operators, attacker agent and similarity function.* As such, the QADE trust and reputation model assumes and models the existence of unfairly reported trust assessments. It considers human-centric nature of trust, where differently reported trust values do not necessarily mean false value propagation but can also imply differences in agents' trust attitudes. Based on the QADE trust model, we provide the method to identify and filter out the presumably unfair trust values. The method is two-part. Firstly, a trust evaluator finds other agents in society that are similar to him. Similarity between agents takes into account pairwise similarity of trust values and similarity of agents' *general mindsets*. Secondly, the trust evaluator black-marks reported trust values if they differ with theirs. Trust values that are black-marked by certain amount of agents are considered to be unfair and excluded from trust computation. We compared the effectiveness of methods to decrease the effect of unfair ratings based on simulations. We made the simulations in environments with varying number of attackers and targeted agents, as well as different kind of attackers — individuals and collaborative attackers. The results showed significant improvements of our proposed method with average improvements from 28% to 58% compared to the other most representative filtering methods by Teacy and Whitby.

Trust ratings shared by agents in e-environments are subjective as trust evaluation depends on evaluator's personal disposition to trust. As such, aggregation of shared trust ratings to compute an agent's reputation may be questionable without proper consideration of rating subjectivity. So that all agents understand the ratings in the same way, the reported trust ratings should be adjusted to each agent individually. In the last part of the thesis, we address the problem of proper trust rating analysis and aggregation. We propose an extension of the QAD trust model, named QADES. We propose a novel Human-Oriented Method for Ratings Adaptation (HOMRA) that derives adjusted reputations compliant with the behavioral patterns of the evaluators. With HOMRA method, all participants have comparable opportunities to choose trustworthy agents in future transactions, regardless of their trust dispositions. The proposed HOMRA method finds the agents with similar trust dispositions, taking advantages of non-parametric statistical methods. After that, it computes the personalized reputation scores of other agents with the aggregation of trust values shared by agents with similar trust dispositions. The method derives the characteristics of agents' trust dispositions implicitly from their past ratings and does not request them to disclose

any part of their trust evaluation process, such as motivating criteria for trust assessments, underlying beliefs, or criteria preferences. We evaluated the performance of our method with extensive simulations with varying number of agents, different distributions of agents' personality types, as well as simulations with different number of available trust ratings. The results showed significant improvements of our HOMRA method with average improvement of 50% over the Abdul-Rahman and 73% over the Hasan method, when comparing the efficiency of the methods depending on the number of agents in certain e-environment, the efficiency of the methods depending on the distributions of agents with different personality types and the efficiency of the methods depending on the availability of trust ratings.

*Key words:* trust, reputation, trust and reputation modeling, trust and reputation management systems, false ratings, subjectivity, human factors, e-environment



## ZAHVALA

*Iskreno se zahvaljujem vsem, ki ste spremljali moje delo, mi pomagali in me pri tem podpirali. Najprej hvala mentorju prof. dr. Denisu Trčku za vso pomoč, nasvete in usmeritve v času doktorskega študija ter pri izdelavi doktorske disertacije. Hvala vsem tistim, ki ste z mano razpravljali o zaupanju. Za maratonske debate o obvladovanju zaupanja v e-okoljih še posebej hvala dr. Davidu Jelencu.*

*Zahvaljujem se celotni družini, ki mi je ves čas stala ob strani. Miha, Nina, mami in ati, vsakemu posebej hvala za vso podporo in potrpljenje. Prav posebna zahvala gre Matjažu. Hvala ti za spodbudo, razumevanje, podporo in predvsem ljubezen. In najlepša hvala moji Eli. S svojo iskrivostjo me dan za dnem osrečuješ!*

— Eva Zupančič, Ljubljana, november 2014.



# KAZALO

<i>Povzetek</i>	<i>i</i>
<i>Abstract</i>	<i>v</i>
<i>Zahvala</i>	<i>ix</i>
<b>1</b> <i>Uvod</i>	<b>1</b>
1.1 Motivacija . . . . .	2
1.2 Cilji . . . . .	3
1.3 Pregled vsebine . . . . .	3
<b>2</b> <i>Pregled področja</i>	<b>5</b>
2.1 Varnost, zaupanje in ugled . . . . .	6
2.1.1 Varnost . . . . .	6
2.1.2 Zaupanje . . . . .	6
2.1.3 Ugled . . . . .	8
2.2 Modeli zaupanja in ugleda v e-okoljih . . . . .	9
2.2.1 Marshev model . . . . .	9
2.2.2 Abdul-Rahmanov in Hailesev model . . . . .	11
2.2.3 Sistem BRS . . . . .	12
2.2.4 Yujev in Singhov model . . . . .	13
2.2.5 Jøsangova subjektivna logika . . . . .	15
2.2.6 REGRET . . . . .	18
2.2.7 SPORAS in HISTOS . . . . .	19
2.2.8 Spletni modeli ugleda . . . . .	20

2.3	Sistemi za obvladovanje zaupanja in ugleda . . . . .	21
3	<i>Temeljna formalizacija zaupanja v e-okolju</i> . . . . .	25
3.1	Model kvalitativne dinamike ocenjevanja (QAD) . . . . .	26
3.2	Osnovna notacija . . . . .	27
3.3	Operatorji zaupanja QAD . . . . .	29
3.4	Analiza modela zaupanja QAD . . . . .	32
3.4.1	Opis simulacij . . . . .	32
3.4.2	Splošna analiza lastnosti modela zaupanja QAD . . . . .	35
3.4.3	Poglobljena analiza lastnosti operatorjev QAD . . . . .	39
4	<i>Metoda za obvladovanje zlaganih ocen zaupanja</i> . . . . .	63
4.1	Predstavitve problema . . . . .	64
4.2	Predlagana rešitev . . . . .	65
4.2.1	Model zaupanja in ugleda QADE . . . . .	65
4.2.2	Metoda za izračun zaupanja z obrambo pred napadi QADE-filter . . . . .	75
4.3	Ovrednotenje metode QADE-filter . . . . .	80
4.3.1	Opis simulacij . . . . .	82
4.3.2	Ovrednotenje metod za izločanje zlaganih vrednosti pri napadih z očrnitvijo . . . . .	85
4.3.3	Ovrednotenje metod za izločanje zlaganih vrednosti pri napadih s hvaljenjem . . . . .	92
4.4	Kvalitativna primerjava metode QADE-filter z alternativnimi pristopi . . . . .	96
5	<i>Metoda za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti</i> . . . . .	99
5.1	Predstavitve problema . . . . .	100
5.2	Predlagana rešitev . . . . .	102
5.2.1	Model zaupanja in ugleda QADES . . . . .	102
5.2.2	Metoda za prilagoditev ocen zaupanja HOMRA . . . . .	113
5.3	Ovrednotenje metode HOMRA . . . . .	115
5.3.1	Opis simulacij . . . . .	115
5.3.2	Analiza učinkovitosti metode HOMRA glede na vrednost parametra <i>tdsTh</i> . . . . .	118
5.3.3	Analiza učinkovitosti metode HOMRA glede na število agentov . . . . .	120

5.3.4	Analiza učinkovitosti metode HOMRA glede na število transakcij . . . . .	121
5.3.5	Analiza učinkovitosti metode HOMRA glede na porazdelitev agentov z različnimi tipi osebnosti . . . . .	126
5.4	Kvalitativna primerjava metode HOMRA z alternativnimi pristopi . . . . .	132
6	<i>Zaključek</i> . . . . .	137
6.1	Sklepne ugotovitve . . . . .	138
6.2	Prispevki k znanosti . . . . .	141
A	<i>Rezultati simulacij pri poglobljeni analizi lastnosti operatorjev QAD</i> . . . . .	145
B	<i>Rezultati simulacij pri ovrednotenju metode QADE-filter</i> . . . . .	177
	<i>Literatura</i> . . . . .	183



*Uvod*

## 1.1 Motivacija

Zaupanje je ključ do uspešnega sodelovanja tako v fizičnem kot virtualnem svetu. V fizičnem svetu se ljudje tipično pozanimajo pri svojih (zaupanja vrednih) znancih, kakšne izkušnje imajo z nekom. Tako se lažje odločijo, s kom sodelovati oziroma poslovati [1]. Če je prisotna tudi osebna izkušnja, odločitev, ali bodo nekomu zaupali ali ne, temelji tudi na podlagi (nezavedne) mimike in kretenj partnerja, kot so umikanje pogleda, nagib glave, stisk roke ipd. [2, 3].

V e-okolju je treba k vrednotenju zaupanja pristopiti drugače, saj se oddaljeno (angl. on-line) komuniciranje med entitetami bistveno razlikuje od interakcij v fizičnem okolju. Pri fizičnih oblikah komunikacije je na voljo velik spekter raznovrstnih pokazateljev zaupanja (ali nezaupanja), za katere je treba poiskati nadomestke v e-okolju. Poleg tega gre v e-okolju velikokrat za enkratne interakcije med entitetami in brez neposredne izkušnje, zato je izziv smotrnega ocenjevanja zaupanja še toliko večji.

Zaupanje v e-okolju ocenimo z uporabo *sistemov za obvladovanje zaupanja in ugleda* (angl. *trust and reputation management systems, TRMS*) [4, 5]. Osnovna ideja sistemov za obvladovanje zaupanja in ugleda je to, da entitete v danem e-okolju podajo mnenje oziroma ocene zaupanja do drugih entitet, s katerimi so imele izkušnje, nato pa sistem združi podane ocene in tako izpelje vrednost ugleda za posamezno entiteto (tj. uporabnika, agenta, storitev ipd.) [6, 7]. Sistemi za obvladovanje zaupanja in ugleda implementirajo različne matematične modele zaupanja in ugleda, ki definirajo proces evalvacije zaupanja med entitetami v virtualnih skupnostih. Matematični modeli zaupanja in ugleda temeljijo na različnih pristopih, kot so verjetnostni račun, mehka logika, teorija grafov in različne ad hoc metode [6, 8–11].

Internet je postal pomemben medij za poslovne transakcije in dandanes se velik del poslovanja odvija na elektronski način, ki temelji na uporabi informacijskih tehnologij (IT) oz. omrežju internet [12, 13]. Raziskave so pokazale, da je k velikemu uspehu elektronskih tržišč pripomogla tudi uvedba sistemov za obvladovanje zaupanja in ugleda [14]. Dellarocas [14] trdi, da so ti sistemi dosegli svoj prvotni namen: ustvariti zadostno zaupanje med kupci, na podlagi katerega ti prevzemajo tveganje in opravijo transakcije s popolnimi neznanci.

## 1.2 Cilji

Na podlagi predstavljenih izhodišč lahko ugotovimo, da je obvladovanje zaupanja in ugleda v e-sistemih relevantno in pomembno raziskovalno področje. Doslej predlagani sistemi za obvladovanje zaupanja in ugleda temeljijo na različnih matematičnih modelih. Učinkovit sistem za obvladovanje zaupanja in ugleda mora biti podprt z ustreznim matematičnim modelom zaupanja in ugleda, zato bomo v pričujoči doktorski disertaciji najprej ustrezno formalizirali zaupanje in ugled v e-okoljih. Pri tem bomo izhajali iz modela kvalitativne dinamike ocenjevanja (angl. Qualitative Assessment Dynamics, QAD) [15, 16], ki ga bomo ustrezno modificirali in razširili. Predlagani matematični model zaupanja in ugleda bo služil kot osnova za reševanje dveh ključnih problemov sistemov za obvladovanje zaupanja in ugleda: (1) obvladovanje zlaganih ocen zaupanja, in (2) izločitev dejavnika subjektivnosti iz ocen zaupanja.

V ta namen bomo v disertaciji predlagali model zaupanja, ki bo upošteval možnost prisotnosti neresničnih (zlaganih) ocen zaupanja in s tem omogočal razvoj metode za obvladovanje teh. Predlagana metoda bo zlagane ocene obravnavala bolj učinkovito od reprezentativnih metod, kar bomo dokazali z uporabo večagentnih simulacij.

Naslednji pomemben cilj disertacije je razširitev matematičnega modela zaupanja in ugleda s konstrukti, ki bodo omogočali razvoj metode za prilagoditev posredovanih ocen zaupanja na podlagi izločitve dejavnika subjektivnosti. Cilj razvite metode je učinkovito izločiti dejavnik subjektivnosti iz posredovanih ocen zaupanja in izboljšati do sedaj predlagane metode na tem področju. Učinkovitost predlagane metode v primerjavi z drugimi metodami bomo preverili z uporabo večagentnih simulacij.

Z realizacijo zastavljenih ciljev bo disertacija pomembno doprinesla k nadaljnjemu razvoju modelov zaupanja in pristopov za obvladovanje zaupanja in ugleda v e-okoljih. Obenem bodo predstavljeni izvirni doprinosi, ki do sedaj še niso bili naslovljeni na način, kot ga obravnava pričujoča doktorska disertacija.

## 1.3 Pregled vsebine

Doktorska disertacija je sestavljena iz šestih poglavij. V drugem poglavju predstavimo raziskovalno področje zaupanja in ugleda v e-okoljih. Opišemo pojme varnosti, zaupanja in ugleda v e-okoljih ter navedemo razlike in povezave med posameznimi pojmi. V nadaljevanju poglavja predstavimo ključne matematične modele zaupanja in ugleda ter poglavje zaključimo s predstavitvijo in opisom lastnosti sistemov za računalniško

podprto obvladovanje zaupanja in ugleda.

V tretjem poglavju predstavimo matematični model zaupanja in ugleda v e-okolju, ki bo služil kot osnova za formalizacijo modelov zaupanja in ugleda ter ki bo (z dodatnimi razširitvami) upošteval obstoj neresničnih ocen zaupanja in prisotnost dejavnika subjektivnosti v ocenah zaupanja. Delovanje in lastnosti temeljnega modela podrobno opišemo z interpretacijo simulacij, ki so bile izvedene v ta namen. Posebej se osredotočimo na analizo t. i. operatorjev zaupanja, ki definirajo proces vrednotenja zaupanja in s tem predstavljajo ključni element predlaganega modela zaupanja in ugleda.

Sledi četrto poglavje, v katerem najprej opišemo problem prisotnosti neresničnih (zlaganih) ocen zaupanja v sistemih za obvladovanje zaupanja in ugleda ter za dani problem predlagamo rešitev. Najprej predstavimo razširitev temeljnega modela zaupanja in ugleda, nato pa na podlagi razširjenega modela razvijemo metodo za izračun zaupanja z izločanjem zlaganih ocen. Predlagano metodo ovrednotimo in primerjamo z drugimi metodami za izločanje domnevno zlaganih ocen s simulacijami, izvedenimi s simulacijskim orodjem Zaupnik. Poleg numeričnih rezultatov simulacij, ki ponazarjajo učinkovitost naše predlagane metode, podamo tudi kvalitativno primerjavo predlagane metode z alternativnimi pristopi.

V petem poglavju podrobno opišemo problematiko subjektivnosti, ki je prisotna v ocenah zaupanja, in opišemo predlagano rešitev. Navedemo modifikacije in nadgradnje temeljne formalizacije zaupanja in ugleda, nato pa opišemo predlog metode za prilagoditev ocen zaupanja, ki je osnovana na predlaganem razširjenem modelu. Nadaljujemo z opisom večagentnih simulacij, izvedenih s simulacijskim orodjem Zaupnik, na podlagi katerih smo ovrednotili učinkovitost predlagane metode, in predstavimo rezultate vrednotenja. Poleg tega podamo še kvalitativno primerjavo predlagane metode z drugimi metodami za izločitev dejavnika subjektivnosti, definiranimi v literaturi.

Delo zaključimo s šestim poglavjem, v katerem podamo sklepne ugotovitve prejšnjih poglavij in navedemo izvirne prispevke tega dela k znanosti. Doktorska disertacija vsebuje tudi dodatke, v katerih so predstavljeni dodatni rezultati izvedenih simulacij.

## *Pregled područja*

## 2.1 Varnost, zaupanje in ugled

### 2.1.1 Varnost

V računalništvu se pojem zaupanja (angl. trust) pogosto pojavlja skupaj s pojmom varnosti (angl. security). Čeprav sta oba pojma v povezavi drug z drugim, ju ne smemo enačiti ali zamenjevati. Varnost je v ISO-standardih [17] definirana kot “mehanizem za zagotavljanje zaupnosti, celovitosti in razpoložljivosti informacij” (angl. confidentiality, integrity, availability, CIA). Pri tem se navadno uporabljajo overjanje, avtorizacija in druge metode za omejevanje in kontrolo dostopa, ki temeljijo na kriptografskih protokolih. Tovrstne varnostne storitve imenujemo tradicionalni varnostni mehanizmi in govorimo o *trdi varnosti* (angl. *hard security*) [18]. Mehanizmi za zagotavljanje trde varnosti ne zagotavljajo zaščite pred zlonamernim vedenjem drugih entitet v sistemu. Entiteta lahko na legitimen način vstopi v sistem, a se po vstopu v sistem ne obnaša zaupanja vredno in na primer zavedno (namenoma) podaja napačne ali zavajajoče informacije. V tem primeru govorimo o ogrožanju *mehke varnosti* (angl. *soft security*) [18]. Trenutno najbolj učinkovito orodje za zaščito pred grožnjami mehke varnosti so sistemi za obvladovanje zaupanja in ugleda (angl. trust and reputation management systems, TRMS) [19], ki bodo opisani v nadaljevanju.

### 2.1.2 Zaupanje

Po SSKJ [20] je zaupanje “prepričanje, da je kdo sposoben, voljen narediti, kar se pričakuje”. Avtorji, ki se ukvarjajo z raziskovanjem zaupanja v e-okolju, so podali podobne definicije. Li in Singhal [21] sta zaupanje definirala kot “prepričanje, da se je entiteta v določenem primeru sposobna vesti zanesljivo in odgovorno”. Gambetta [22] je zapisal: “Zaupanje je določena stopnja subjektivne verjetnosti, da bo agent opravil določeno nalogo, in sicer preden lahko spremljamo opravljanje naloge in v kontekstu, v katerem opravljena naloga vpliva na naše dejavnosti.” Na slednjo definicijo se sklicujejo predvsem avtorji, ki so podali verjetnostne modele zaupanja in ugleda (glej poglavje 2.2). Mui in soavtorji [23, 24] so zaupanje definirali kot “subjektivno prepričanje agenta o prihodnjem obnašanju drugega agenta, ki temelji na zgodovini njunih srečanj”. Jøsang [6] je zaupanje klasificiral v dve skupini. *Zaupanje v zanesljivost* (angl. *reliability trust*) je definirano kot “subjektivna verjetnost, s katero posameznik *A* pričakuje, da bo drugi posameznik *B* opravil določeno dejanje, od katerega je odvisna blaginja posameznika *A*”. Definicija izhaja iz že omenjene definicije, ki jo je podal

Gambetta [22]. Drugo skupino predstavlja *odločitveno zaupanje* (*angl. decision trust*), ki ga Jøsang definira kot “obseg, v katerem je posameznik pripravljen biti odvisen od nekoga (ali nečesa) v določeni situaciji z občutkom relativne varnosti, čeprav so možne tudi negativne posledice”.

Definicije in poglobljeno obravnavanje pojma zaupanja večinoma izvirajo iz družbenih ved, ki preučujejo dinamiko človeške družbe [19]. Značilnosti zaupanja, ki veljajo v človeških združbah, je treba ustrezno prenesti v e-okolje. Osnovne lastnosti zaupanja so naslednje [6, 8, 19, 22]:

- Zaupanje je potrebno le v okoljih, kjer je prisotna negotovost in so udeleženci za doseg svojih ciljev odvisni drug od drugega.
- Zaupanje je odvisno od konteksta. V literaturi se zaupanje pogosto opisuje s 3-delno relacijo “*A* zaupa *B* za *X*”. Del relacije “za *X*” se nanaša na opravljanje različnih nalog.
- Zaupanje je subjektivno. Oblikovanje mnenja o nekom ni odvisno samo od njegovih ravnanj in obnašanja, ampak tudi od tega, kako ocenjevalec zaupanja to dojema.
- Zaupanje je enosmerno. Ocenjevalec zaupanja lahko zaupa komu, pri čemer ustvarjeno zaupanje temelji na osebni izkušnji, priporočilih ali drugih informacijah o ocenjevanem agentu. Pri tem je možno, da ocenjevana entiteta ocenjevalca sploh ne pozna ali pa kljub enakim izkušnjam ne vzpostavi zaupanja do ocenjevalca. Zato o recipročnosti zaupanja ne moremo govoriti.
- Zaupanje ni nujno tranzitivno. Denimo, da *A* zaupa *B* in *B* zaupa *C*. Ali bo *A* na podlagi tega zaupal *C*, je odvisno od konteksta, v katerem *A* zaupa *B*. Zaupanje v priporočila je pomemben koncept v računalniškem zaupanju, pogosto obravnavan pod pojmom *verodostojnost priporočevalca* (*angl. credibility of witness*) [19, 25].

V pričujočem delu bomo zaupanje definirali kot “odnos (ali razmerje) med izvorno in ciljno entiteto” [15]. Izvorna entiteta oceni stopnjo zaupanja v ciljno entiteto, da lahko izvede določeno nalogo/storitev v danem kontekstu. Ob tem predpostavljamo, da je izvorna entiteta *razumna* (*angl. cognitive*) in je torej zmožna ocenjevanja in odločanja o ciljni entiteti na podlagi prejetih informacij iz okolja in preteklih izkušenj,

ciljna entiteta pa je lahko tudi abstraktni pojem [13]. Entitete lahko predstavljajo osebe, programske agente, storitve, organizacije ipd. V nadaljevanju dela bomo (razumne) entitete označevali z *agenti*. Formalna definicija zaupanja sledi v poglavju 3.2.

### 2.1.3 Ugled

S pojmom zaupanja je tesno povezan tudi pojem ugleda (angl. reputation). SSKJ [20] ugled definira kot “zelo dobro mnenje, ki ga ima več ljudi o kom zaradi njegovih lastnosti, dejanj”. Sabater in soavtorji [26, 27] ugled definirajo na zelo poenostavljen način, in sicer kot “mnenje ali stališče nekoga o nečem”. Abdul-Rahman in Hailes [28] sta ugled definirala kot “pričakovanje glede obnašanja agenta, ki temelji na informacijah in opazovanju preteklega obnašanja”. Definicijo ugleda, precej prilagojeno delovanju sistemov za obvladovanje zaupanja in ugleda v e-okoljih, sta podala Wang in Vassileva [29], ki sta ugled definirala kot “prepričanje agenta v zmogljivost, poštenost in zanesljivost drugega agenta, ki temelji na priporočilih drugih agentov”. Podobno so ugled opisali tudi Jøsang in soavtorji [6], ki ugled definirajo kot “tisto, kar je splošno povedano ali se verjame o lastnostih osebe oziroma stvari”.

V disertaciji bomo ugled obravnavali kot “mnenje skupnosti o vrednosti zaupanja v ciljno entiteto”. Formalna definicija ugleda je navedena v poglavju 5.2.1.

Kot je razvidno iz definicij, med pojmom ugleda in zaupanja kljub njuni tesni povezanosti obstaja pomembna razlika. Zaupanje je subjektiven fenomen, ki temelji na različnih dejavnikih, med katerimi so različnim agentom pomembni različni faktorji. Zapleteno relacijo med zaupanjem in ugledom lahko ponazorimo z naslednjima povedma [6]: (1) “Zaupam ti, ker imaš dober ugled,” in (2) “Zaupam ti, čeprav imaš slab ugled.” V prvem primeru je zaupanje pogojeno z ugledom ciljnega agenta. Drugi primer nakazuje možnost, da ima izvorni agent določene informacije o ciljnem agentu, npr. neposredne izkušnje, na podlagi katerih ovrednoti zaupanje in pretehtajo splošno mnenje o ciljni entiteti. Lahko pa gre zgolj za različno dojemanje zaupanja v prvem in drugem primeru. Dejstvo je, da je pomanjkanje osebne/neposredne izkušnje v e-okolju pogosto, saj gre večinoma za enkratne transakcije med agenti. Ugled agenta, ki je izražen preko mnenja skupnosti o tem agentu, je zato pogosto pomemben (v nekaterih primerih edini) faktor pri modeliranju zaupanja v e-okoljih (glej poglavje 2.2).

## 2.2 Modeli zaupanja in ugleda v e-okoljih

Zaupanje in ugled agentov v danem sistemu na formalen način opišemo z matematičnimi modeli. Formalizacija zaupanja in ugleda v e-okoljih se zaradi drugačnih in pogosto tudi zelo omejenih virov podatkov razlikuje od splošne formulacije zaupanja in ugleda v človeških družbah, vendar mora ustrezno upoštevati lastnosti zaupanja in ugleda, ki veljajo v fizičnem svetu. Dober model zaupanja in ugleda mora upoštevati relacije med agenti, subjektivnost, faktorje okolja, tveganje in številne druge dejavnike ter jih definirati tako, da čim bolj odsevajo resnično situacijo.

Matematični modeli zaupanja so temelj za implementacijo sistemov za obvladovanje zaupanja in ugleda, ki bodo opisani v poglavju 2.3. Za postavitve dobrega sistema za obvladovanje zaupanja in ugleda morata biti vrednosti zaupanja in ugleda ustrezno modelirani in izračunani. Torej, definiran mora biti ustrezen matematični model zaupanja in ugleda.

Doslej predlagani modeli zaupanja in ugleda temeljijo na različnih pristopih, kot so verjetnosti račun in Bayesove mreže [23, 29–31], Dempster-Shaferjeva teorija dokazov [32, 33], teorija iger [34, 35], mehka logika [26, 27, 36, 37], in na različnih ad hoc metodah [28, 38, 39].

V nadaljevanju poglavja bomo na kratko predstavili najvidnejše formalizacije zaupanja in ugleda. Obsežnejši pregled računskih modelov zaupanja in ugleda najdemo v [6, 8–10, 40–44].

### 2.2.1 Marshev model

Stephen Marsh je bil eden prvih, ki je definiral matematični model zaupanja v računalniškem okolju [38, 45]. Marsh je vrednost zaupanja med agentoma opisal s skalarjem iz intervala  $[-1, 1]$ , pri čemer je zaupanje razdelil v tri skupine:

- *Osnovno zaupanje* (*angl. basic trust*) predstavlja splošno nagnjenje k zaupanju agenta  $x$  ob času  $t$  in je označeno s  $T_x^t$ . Visoka vrednost osnovnega zaupanja agenta pomeni, da je bolj zaupljive narave. Model definira tri osnovne stopnje nagnjenja k zaupanju: optimizem, pesimizem in realizem. Optimistični agenti so zelo nagnjeni k zaupanju, ne glede na pretekle (slabe) izkušnje. Zaupanja do drugih agentov ne morejo zmanjšati, lahko ga samo povečajo. Nasprotno velja za pesimiste, medtem ko je osnovno zaupanje realističnih agentov bolj fleksibilne narave.

- *Splošno zaupanje* (*angl. general trust*) predstavlja zaupanje med agenti. Notacija  $T_x(y)^t$  označuje zaupanje agenta  $x$  do drugega agenta  $y$  ob času  $t$ . Kot pove ime samo, predstavlja splošno zaupanje do drugega agenta in *ni* odvisno od specifične situacije.
- *Situacijsko zaupanje* (*angl. situational trust*) definira zaupanje agenta  $x$  do drugega agenta  $y$  v dani situaciji  $\alpha$  ob času  $t$  in je označeno z oznako  $T_x(y, \alpha)^t$ . Formula za izračun vrednosti situacijskega zaupanja je naslednja:  $T_x(y, \alpha)^t = U_x(\alpha)^t \times I_x(\alpha)^t \times \overline{T_x(y)}^t$ , kjer  $U_x(\alpha)^t$  predstavlja koristnost situacije  $\alpha$  za agenta  $x$ ,  $I_x(\alpha)^t$  predstavlja pomembnost situacije, faktor  $\overline{T_x(y)}^t$  pa oceno splošnega zaupanja med  $x$  in  $y$ , ki upošteva vse informacije, ki jih ima agent  $x$  o agentu  $y$ . Ocena splošnega zaupanja  $\overline{T_x(y)}^t$  je odvisna od osnovnega zaupanja agenta  $x$ , (tj. nagnjenja agenta k zaupanju) in se lahko izračuna kot povprečna, minimalna ali maksimalna vrednost ocen zaupanja iz preteklih izkušenj. Koristnost situacije  $U_x$  je vrednost iz intervala  $[-1, 1]$ , pomembnost situacije  $I_x$  pa vrednost iz intervala  $[0, 1]$ .

Predlagani model formalizira zaupanje med agenti, medtem ko pojem ugleda ni obravnavan in definiran. Formalizacija pri izračunu zaupanja na primer ne upošteva priporočil drugih agentov. S tem je količina informacij za ovrednotenje zaupanja omejena, predvsem kadar ima agent malo neposrednih interakcij z ostalimi agenti.

### Primer

Z Marshevim modelom bomo izračunali vrednost zaupanja t. i. realističnega agenta  $x$  do agenta  $y$  v dani situaciji  $\alpha$  ob času  $t$ . Ker gre za realističnega agenta, bomo za vrednost splošnega zaupanja uporabili povprečno oceno zaupanja iz preteklih izkušenj, izračunano po formuli:  $\overline{T_x(y)}^t = \frac{1}{|A|} \sum_{\alpha \in A} T_x(y)^t$ , kjer je  $A$  množica vseh situacij  $\alpha$ , v katerih je imel agent  $x$  interakcijo z agentom  $y$ . Agent  $x$  je imel v preteklosti z agentom  $y$  tri interakcije, na podlagi katerih je ocenil zaupanje do  $y$  z vrednostmi 0, 62, 0, 71 in 0, 59. Iz tega sledi, da je vrednost povprečne ocene zaupanja glede na pretekle izkušnje enaka  $\overline{T_x(y)}^t = 0, 64$ . Agent  $x$  ocenjuje, da je dana situacija  $\alpha$  "srednje" pomembna, zato jo ovrednoti z  $I_x(\alpha)^t = 0, 5$ , medtem ko koristnost situacije oceni z vrednostjo  $U_x(\alpha)^t = 0, 8$ . Iz danih vrednosti po Marshevi formalizaciji izračunamo vrednost zaupanja  $x$  do agenta  $y$  v dani situaciji  $\alpha$  ob času  $t$ , kot sledi:  $T_x(y, \alpha)^t = U_x(\alpha)^t \times I_x(\alpha)^t \times \overline{T_x(y)}^t = 0, 8 \times 0, 5 \times 0, 64 = 0, 256$ .

### 2.2.2 Abdul-Rahmanov in Hailesev model

V modelu, ki sta ga predlagala Abdul-Rahman in Hailes [28], so možne vrednosti zaupanja podane s štiristopenjsko lestvico, ki vsebuje vrednosti *vt* (*very trustworthy*), *t* (*trustworthy*), *u* (*untrustworthy*) in *vu* (*very untrustworthy*). Podobna lestvica je uporabljena tudi za kategorizacijo izkušenj: *vg* (*very good*), *g* (*good*), *b* (*bad*) in *vb* (*very bad*). Vsak agent si zapomni število zelo dobrih, dobrih, slabih in zelo slabih neposrednih izkušenj z vsakim drugim agentom. Kategorija, v kateri je največ izkušenj, odraža vrednost zaupanja med agentoma na podlagi neposrednih izkušenj. Ugled agenta je izpeljan iz priporočil drugih agentov, ki so regulirana glede na semantično razdaljo med agenti. Semantična razdalja meri podobnost dojemanja zaupanja in ocenjevanja izkušenj med agentom  $X$ , ki vrednoti zaupanje oziroma ugled, in agentom  $Y$ , ki mu posreduje svoje izkušnje in priporočila. Semantična razdalja se izračuna s primerjavo neposrednih izkušenj agentov  $X$  in  $Y$  z drugimi agenti. Semantična razdalja je shranjena v obliki  $(c, Y, \{T_{vg}, T_g, T_b, T_{vb}\})$ , kjer  $c$  predstavlja kontekst,  $Y$  je agent, ki je posredoval oceno, in  $T_e$  je prilagoditvena vrednost za posredovane ocene z vrednostjo  $e$ . Če sta agenta drugače ocenjevala iste agente, bo agent  $X$  pred izračunom končne vrednosti zaupanja ustrezno zvišal/znižal priporočila, posredovana od agenta  $Y$ . Predlagani model ugled izračuna na podlagi obteženih priporočil drugih agentov; uteži so definirane v vpogledni tabeli. Ocene osebnih izkušenj se pri modeliranju ugleda ne upoštevajo neposredno, uporabijo se zgolj za primerjavo med agenti in morebitno prilagajanje posredovanih informacij.

#### Primer

Agent  $x$  želi oceniti zaupanje do agentov  $y$  in  $z$  za isti kontekst  $c$ . Za priporočila o agentih  $y$  in  $z$  vpraša agente  $r_1$ ,  $r_2$  in  $r_3$ , ki podajo naslednje ocene: izkušnja agenta  $r_1$  z agentom  $y$  je *vg* (*very good*), izkušnja agenta  $r_2$  z agentom  $y$  je *b* (*bad*) in prav tako je izkušnja agenta  $r_3$  z agentom  $z$  ocenjena z *b* (*bad*). Agenti  $r_1$ ,  $r_2$  in  $r_3$  so v preteklosti že posredovali svoje ocene agentu  $x$ , na podlagi katerih je agent  $x$  izračunal semantične razdalje z njimi. V danem primeru naj bodo vrednosti semantičnih razdalj v danem kontekstu naslednje:  $(c, r_1, (\{0\}, \{-1\}, \{-1\}, \{2\}))$ ,  $(c, r_2, (\{1\}, \{1\}, \{1\}, \{2\}))$  in  $(c, r_3, (\{1\}, \{0\}, \{0\}, \{1\}))$ . Semantične razdalje agenta  $x$  z agentom  $r_1$  so:  $T_{vg} = \{0\}$ ,  $T_g = \{-1\}$ ,  $T_b = \{-1\}$  in  $T_{vb} = \{2\}$ . To pomeni, da bo agent  $x$  ocene, ki jih bo posredoval agent  $r_1$  in bodo imele vrednost *vg*, upošteval take, kot so, ocene z vre-

dnostjo  $g$  pomanjšal za eno stopnjo, ocene z vrednostjo  $b$  pomanjšal za eno stopnjo in ocene z vrednostjo  $vb$  povečal za dve stopnji. Na enak način interpretiramo tudi semantične razdalje z agentoma  $r_2$  in  $r_3$ . Na podlagi semantičnih razdalj agent  $x$  posredovane ocene izkušenj interpretira na naslednji način: prilagojena ocena izkušnje agenta  $r_1$  z agentom  $y$  je  $vg + T_{vg} = vg + 0 = vg$ , prilagojena ocena izkušnje agenta  $r_2$  z agentom  $y$  je  $b + T_b = b + 1 = g$  in prilagojena ocena izkušnje agenta  $r_3$  z agentom  $z$  je  $b + T_b = b + 0 = b$ . Nadalje, Abdul-Rahman in Hailes sta definirala vpogledno tabelo z utežmi za obtežitev priporočil drugih agentov, ki določa večjo utež za tiste agente z manjšo semantično razdaljo in manjšo utež za tiste z večjo. Pri oceni zaupanja agenta  $x$  do agenta  $y$  ima večjo utež ocena, ki jo je podal agent  $r_1$ , saj prilagajanje ni bilo potrebno zaradi semantične razdalje enake 0. Končna vrednost ugleda agenta  $y$  je tako ocenjena z vrednostjo  $vg$  (*very good*), ugled agenta  $z$  pa je ocenjen z vrednostjo  $b$  (*bad*). Denimo, da je agent  $x$  nato naredil transakcijo z agentoma  $y$  in  $z$  ter s tem ocenil neposredno vrednost izkušenj. Izkušnjo z agentom  $y$  je ocenil kot  $g$  (*good*) in izkušnjo z agentom  $z$  kot  $b$  (*bad*). Na podlagi lastne izkušnje agent  $x$  posodobi semantične razdalje z ostalimi agenti, kot sledi:  $(c, r_1, (\{0, -1\}, \{-1\}, \{-1\}, \{2\}))$ ,  $(c, r_2, (\{1\}, \{1\}, \{1, 1\}, \{2\}))$  in  $(c, r_3, (\{1\}, \{0\}, \{0, 0\}, \{1\}))$ . Nove vrednosti se uporabijo pri naslednjem vrednotenju ugleda.

### 2.2.3 Sistem BRS

Model BRS (Beta Reputation System) [30] je osnovni model iz skupine verjetnostnih modelov zaupanja. Temelji na uporabi družine beta funkcij porazdelitev za izračun vrednosti ugleda agenta. Sistem BRS ugled agenta izrazi z verjetnostjo, ki predvideva določeno (slabo ali dobro) obnašanje agenta v prihodnjih interakcijah. Nova vrednost ugleda se izračuna z uporabo beta porazdelitvene funkcije na podlagi informacij o preteklem obnašanju agenta in trenutne ocene. Funkcijo gostote verjetnosti za beta porazdelitev  $beta(p|\alpha, \beta)$  izrazimo s funkcijo  $\Gamma$ , kot sledi:  $beta(p|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}$ , kjer je  $0 \leq p \leq 1$ ,  $\alpha > 0$ ,  $\beta > 0$ . Pričakovana vrednost beta porazdelitve je  $E(p) = \frac{\alpha}{\alpha + \beta}$ .

V sistemu BRS agent  $X$  oceni drugega agenta  $T$  z oceno  $(r_T^X, s_T^X)$ , kjer  $r_T^X$  predstavlja stopnjo zadovoljstva in  $s_T^X$  stopnjo nezadovoljstva agenta  $X$  z agentom  $T$  po opravljeni transakciji. V preprostih binarnih sistemih  $r_T^X$  predstavlja število dobrih izkušenj in  $s_T^X$  število slabih izkušenj z agentom  $T$ . Zaupanje agenta  $X$  do agenta  $T$  je definirano

kot  $\rho_T^X = E(\text{beta}(p|\alpha, \beta))$ , kjer je  $\alpha = r_T^X + 1$  in  $\beta = s_T^X + 1$ . Sistem BRS definira funkcijo združevanja vrednosti zaupanja različnih agentov do ciljnega agenta  $T$ , s čimer se izračuna splošni ugled ciljnega agenta  $T$ . Ugled ciljnega agenta  $T$  je definiran s funkcijo  $\rho_T^{\mathcal{X}} = E(\text{beta}(p|\alpha, \beta))$ , kjer  $\alpha = r_T^{\mathcal{X}} + 1$  predstavlja seštevek vrednosti stopnje zadovoljstva  $\sum_{x \in \mathcal{X}} r_T^x$  vseh agentov, ki so imeli interakcijo z agentom  $T$  in  $\beta = s_T^{\mathcal{X}} + 1$  predstavlja akumulirano vrednost stopenj nezadovoljstva. Sistem BRS vpelje tudi faktor staranja in novejše ocene zaupanja bolj obteži:  $r_{T,t_R}^{X,t} = \lambda^{t-t_R} * r_{T,t_R}^X$  in  $s_{T,t_R}^{X,t} = \lambda^{t-t_R} * s_{T,t_R}^X$ , kjer je  $0 \leq \lambda \leq 1$ ,  $t$  predstavlja trenutni čas in  $t_R$  čas, v katerem je bila podana ocena (čas opravljene transakcije).

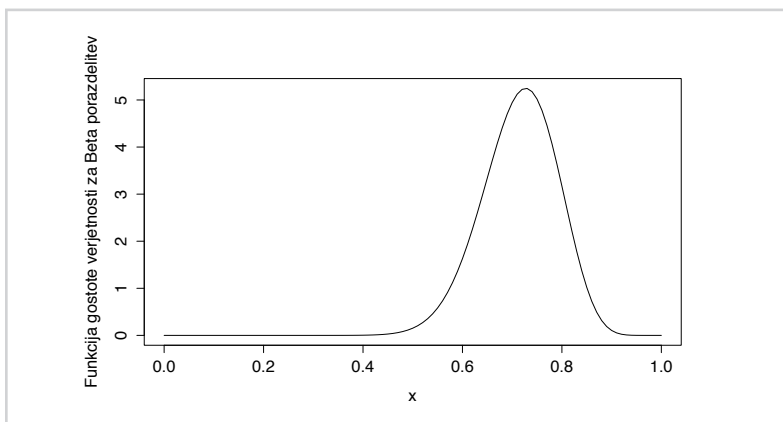
Model zaupanja BRS so uporabili ali nadgradili številni drugi avtorji. Whitby in drugi [46] so sistem BRS nagradili s konstrukti za prepoznavanje zlaganih ocen izkušenj. Pristop za detekcijo zlaganih ocen so predstavili tudi Teacy in soavtorji [31] v modelu TRAVOS ter Zhang in Cohen [47, 48] – vsi so predstavili alternativne modifikacije in nadgradnje modela BRS za filtriranje neresničnih ocen izkušenj. Podoben pristop modeliranja zaupanja in ugleda z uporabo beta porazdelitvenih funkcij je predlagal tudi Mui s soavtorji [23, 24].

### Primer

Agent  $X$  želi s sistemom BRS izračunati splošni ugled agenta  $T$ . Trenutna ocena agenta  $X$  za agenta  $T$  je (7, 3), tj. agent  $X$  je imel z agentom  $T$  v preteklosti 7 dobrih in 3 slabe izkušnje. Z agentom  $T$  so imeli v preteklosti transakcije tudi agenti  $A$ ,  $B$  in  $C$ . Njihove ocene za agenta  $T$  so naslednje: (5, 2), (8, 3) in (4, 1). V skupnem so agenti izkušnjo z agentom  $T$  ocenili kot dobro v 24 primerih in 9-krat kot slabo, iz česar izpeljemo parametra za izračun ugleda agenta:  $\alpha = 24 + 1 = 25$  in  $\beta = 9 + 1 = 10$ . V danem primeru predvidevamo, da agenti vse ocene upoštevajo z enako mero, ne glede na starost ocene. Ugled ciljnega agenta  $T$  je  $\rho_T^{\mathcal{X}} = E(\text{beta}(p|\alpha, \beta)) = E(\text{beta}(p|25, 10)) = \frac{25}{25+10} = 0,71$ . Za boljšo ponazoritev smo dodali tudi grafičen prikaz funkcije gostote verjetnosti za porazdelitev  $\text{beta}(p|25, 10)$ . Funkcija je prikazana na sliki 2.1.

### 2.2.4 Yujev in Singhov model

Formalizacija zaupanja in ugleda, ki sta jo predlagala Yu in Singh [32, 49, 50], temelji na Dempster-Shaferjevi teoriji dokazov [51]. Osnovni koncept Dempster-Shaferjeve teorije dokazov je *okvir razločevalnosti* (angl. *frame of discernment*)  $\Theta$ , ki predstavlja



Slika 2.1

Funkcija gostote verjetnosti za porazdelitev beta za  $\alpha = 25$  in  $\beta = 10$ .

množico možnih stanj. Znotraj okvira je ob vsakem času lahko resnično samo eno stanje. *Osnovna obtežitev prepričanja* (angl. *basic probability assignment*) je podana s funkcijo  $m : 2^{\Theta} \mapsto [0, 1]$ , kjer je  $m(\emptyset) = 0$  in  $\sum_{A \subseteq \Theta} m(A) = 1$ . Iz formule sledi, da za stanje  $T$  velja  $m(\{T\}) + m(\{-T\}) + m(\{T, -T\}) = 1$ . Obtežitev prepričanja  $m_{\Theta}(X)$  izraža obtežitev prepričanja za celo množico stanj  $X$  in ne za posamezna podstanja  $X$ . Dempster-Shaferjeva teorija dokazov definira tudi *funkcijo prepričanja* (angl. *belief function*)  $Bel(A)$ , ki je vsota vseh obteženih prepričanj za stanja v dani množici stanj  $A$ . Na primer:  $Bel(\{T, -T\}) = m(\{T\}) + m(\{-T\}) + m(\{T, -T\}) = 1$ .

Nadalje, Dempsterjevo združevalno pravilo definira izračun nove vrednosti funkcije prepričanja na podlagi združevanja različnih dokazov. Za vsako podmnožico stanj  $A$  združevalno pravilo definira  $m_1(A) \oplus m_2(A)$  kot vsoto vseh produktov v obliki  $m_1(X) * m_2(Y)$  po vseh množicah stanj  $X$  in  $Y$ , ki so v preseku z  $A$ .

Yu in Singh sta predlagala model zaupanja, ki temelji na uporabi funkcije prepričanja. V modelu predlagata dve možni stanji za vsakega agenta: agent  $A$  je zaupanja vreden (stanje  $T_A$ ) in agent  $A$  ni zaupanja vreden (stanje  $\neg T_A$ ); in obtežitvi prepričanj za obe stanji:  $m(T_A)$  in  $m(\neg T_A)$ . Ugled agenta  $A$  je definiran kot  $\Gamma(A) = m(T_A) - m(\neg T_A)$ , kjer je  $m(T_A), m(\neg T_A) \in [0, 1]$  in  $\Gamma(A) \in [-1, 1]$ . Obtežitvi prepričanj za stanji  $T$  in  $\neg T$  sta izpeljani iz števila negativnih in pozitivnih preteklih izkušenj.

### Primer

Želimo izračunati zaupanje do agenta  $X$ , o katerem sta agenta 1 in 2 podala naslednje obtežitve prepričanj za stanje  $T = \text{“vreden zaupanja”}$ :

- $m_1(\{T_X\}) = 0,8, m_1(\{\neg T_X\}) = 0, m_1(\{T_X, \neg T_X\}) = 0,2$
- $m_2(\{T_X\}) = 0,9, m_2(\{\neg T_X\}) = 0, m_2(\{T_X, \neg T_X\}) = 0,1$

Dokaze obeh agentov združimo po Dempsterjevem združevalnem pravilu:

$$m_{12}(\{T_X\}) = m_1(\{T_X\}) * m_2(\{T_X\}) + m_1(\{T_X\}) * m_2(\{T_X, \neg T_X\}) \\ + m_2(\{T_X\}) * m_1(\{T_X, \neg T_X\}) = 0,72 + 0,08 + 0,18 = 0,98.$$

Na enak način združimo tudi verjetnosti za množici stanj  $\{\neg T_X\}$  in  $\{T_X, \neg T_X\}$  ter dobimo  $m_{12}(\{\neg T_X\}) = 0$  in  $m_{12}(\{T_X, \neg T_X\}) = 0,02$ . Na podlagi združenih dokazov je verjetnost, da je izbrani agent zaupanja vreden, enaka 0,98, medtem ko je verjetnost, da ni zaupanja vreden, enaka 0. Iz izpeljanih vrednosti prepričanj za stanji  $T_X$  in  $\neg T_X$  izračunamo ugled agenta:  $\Gamma(X) = m(T_X) - m(\neg T_X) = 0,98$ .

### 2.2.5 Jøsangova subjektivna logika

Jøsangovo subjektivno logiko [33, 52] uvrščamo v skupino modelov prepričanja (angl. belief models). Subjektivna logika je razširitev verjetnostnega modela in prav tako kot model, ki sta ga predlagala Yu in Singh [32, 50] (glej poglavje 2.2.4), temelji na uporabi elementov iz Dempster-Shaferjeve teorije dokazov [51].

V Jøsangovi subjektivni logiki je zaupanje agenta  $A$  v trditev  $p$  modelirano z *mnenjem* (angl. *opinion*) in je označeno z notacijo  $\omega_p^A$ . Mnenje je definirano s komponentami  $\{b, d, u, a\}$ , za katere velja  $b, d, u, a \in [0, 1]$  in  $b + d + u = 1$ . Komponenta  $b$  predstavlja stopnjo zaupanja (tj. verjetnost, da trditev drži), komponenta  $d$  stopnjo nezaupanja (tj. verjetnost, da trditev ne drži) in komponenta  $u$  stopnjo negotovosti (tj. ne vemo, ali trditev drži ali ne drži) glede resničnosti trditve  $p$ . Element  $a$  je *relativna atomarnost* (angl. *relative atomicity*), ki določa, v kolikšni meri stopnja negotovosti  $u$  vpliva na pričakovano vrednost zaupanja v  $p$ :  $E(\omega_p^A) = b + au$ . V splošnem je vrednost elementa  $a = 0,5$  in ga lahko izpustimo iz nadaljnjih izračunov.

V subjektivni logiki je trenutno definiranih 15 različnih operatorjev za operacije nad mnenji [52]. Med pomembnejšimi so (skupaj s pripadajočimi definicijami):

- *Konjunkcija*: Naj bosta  $\omega_p^A = \{b_p^A, d_p^A, u_p^A\}$  in  $\omega_q^A = \{b_q^A, d_q^A, u_q^A\}$  mnenji subjekta  $A$  o dveh neodvisnih trditvah  $p$  in  $q$ . Potem konjunkcija mnenj  $\omega_p^A$  in  $\omega_q^A$  predstavlja konjunktno resničnost trditvev  $p$  in  $q$ , ki jo definiramo kot

$$\omega_{p \wedge q}^A = \omega_p^A \wedge \omega_q^A = \{b_{p \wedge q}^A, d_{p \wedge q}^A, u_{p \wedge q}^A\},$$

kjer

$$\begin{cases} b_{p \wedge q}^A = b_p^A b_q^A, \\ d_{p \wedge q}^A = d_p^A + d_q^A - d_p^A d_q^A, \\ u_{p \wedge q}^A = b_p^A u_q^A + u_p^A b_q^A + u_p^A u_q^A. \end{cases}$$

- *Splošno mnenje*: Naj bosta  $\omega_p^A = \{b_p^A, d_p^A, u_p^A\}$  in  $\omega_p^B = \{b_p^B, d_p^B, u_p^B\}$  mnenji subjektov  $A$  in  $B$  o trditvi  $p$ . Potem splošno mnenje  $\omega_p^{A,B}$  definiramo kot

$$\omega_p^{A,B} = \omega_p^A \oplus \omega_p^B = \{b_p^{A,B}, d_p^{A,B}, u_p^{A,B}\},$$

kjer

$$\begin{cases} b_p^{A,B} = (b_p^A u_p^B + b_p^B u_p^A) / (u_p^A + u_p^B - u_p^A u_p^B), \\ d_p^{A,B} = (d_p^A u_p^B + d_p^B u_p^A) / (u_p^A + u_p^B - u_p^A u_p^B), \\ u_p^{A,B} = (u_p^A u_p^B) / (u_p^A + u_p^B - u_p^A u_p^B). \end{cases}$$

- *Priporočilo*: Naj bosta  $A$  in  $B$  agenta. Mnenje agenta  $B$  o trditvi  $p$  naj bo  $\omega_p^B = \{b_p^B, d_p^B, u_p^B\}$ . Mnenje agenta  $A$  o priporočilu agenta  $B$  naj bo  $\omega_B^A = \{b_B^A, d_B^A, u_B^A\}$ . Mnenje agenta  $A$  o trditvi  $p$  po priporočilu agenta  $B$  definiramo kot

$$\omega_p^{AB} = \omega_B^A \otimes \omega_p^B = \{b_p^{AB}, d_p^{AB}, u_p^{AB}\},$$

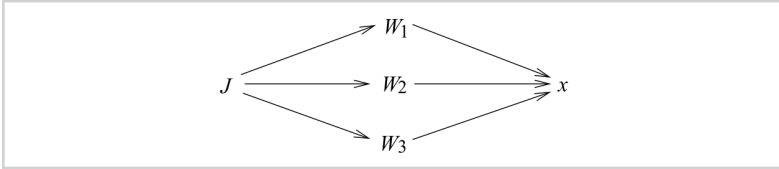
kjer

$$\begin{cases} b_p^{AB} = b_B^A b_p^B, \\ d_p^{AB} = b_B^A d_p^B, \\ u_p^{AB} = d_B^A + u_B^A + b_B^A u_p^B. \end{cases}$$

Subjektivna algebra ima tako kot ostali verjetnostni modeli dobro matematično ozadje. Mnenje  $\omega$  lahko izrazimo s porazdelitveno funkcijo *beta* ( $B$ ), saj med prostoro  $\Omega$  (množico mnenj) in  $\Phi$  (množico porazdelitvenih funkcij  $B$ ) obstaja bijektivna preslikava. Slaba stran modelov prepričanja je, da predvideva racionalne entitete, ki bodo podale ustrezne vrednosti  $b$ ,  $d$  in  $u$ , kar je težka naloga za povprečnega uporabnika e-okolja (potrebno je namreč dobro razumevanje verjetnostnega računa in logike).

### Primer

Agent  $J$  želi oceniti zaupanje do agenta  $X$  in bo zato ovrednotil trditev  $x =$  "agent  $X$  je vreden zaupanja". Za priporočila vpraša agente  $W_1$ ,  $W_2$  in  $W_3$ . Opisana situacija je prikazana na sliki 2.2.



Slika 2.2

Zaupanje v trditve  $x$  glede na priporočila agentov  $W_1$ ,  $W_2$  in  $W_3$ . Vir: [33].

Vsi agenti  $W_1$ ,  $W_2$  in  $W_3$  imajo enako mnenje o tem, da je  $X$  zaupanja vreden agent, in sicer:  $\omega_x^{W_1} = \omega_x^{W_2} = \omega_x^{W_3} = \{0, 90, 0, 00, 0, 10\}$ . Mnenja agenta  $J$  o priporočilih agentov  $W_1$ ,  $W_2$  in  $W_3$  pa so različna in imajo naslednje vrednosti:  $\omega_{W_1}^J = \{0, 90, 0, 00, 0, 10\}$ ,  $\omega_{W_2}^J = \{0, 00, 0, 90, 0, 10\}$  in  $\omega_{W_3}^J = \{0, 10, 0, 00, 0, 90\}$ . Kot vidimo, ima  $J$  visoko stopnjo zaupanja za priporočilo od agenta  $W_1$ , visoko stopnjo nezaupanja za priporočilo od agenta  $W_2$  in visoko stopnjo negotovosti glede agenta  $W_3$ . Mnenje agenta  $J$  o trditvi  $x$  po priporočilu agenta  $W_1$  izračunamo kot  $\omega_x^{JW_1} = \omega_{W_1}^J \otimes \omega_x^{W_1} = \{b_x^{JW_1}, d_x^{JW_1}, u_x^{JW_1}\} = \{b_{W_1}^J b_x^{W_1}, b_{W_1}^J d_x^{W_1}, d_{W_1}^J + u_{W_1}^J + b_{W_1}^J u_x^{W_1}\} = \{(0, 90 * 0, 90), (0, 90 * 0, 00), (0, 00 + 0, 10 + 0, 90 * 0, 10)\} = \{0, 81, 0, 00, 0, 19\}$ . Po enakem postopku izračunamo še mnenje agenta  $J$  o trditvi  $x$  po priporočilih agentov  $W_2$  in  $W_3$ . Dobljeni vrednosti sta  $\omega_x^{JW_2} = \{0, 00, 0, 00, 1, 00\}$  in  $\omega_x^{JW_3} = \{0, 09, 0, 00, 0, 91\}$ . Agent  $J$  z operatorjem "splošno mnenje" združi mnenja agentov  $W_1$ ,  $W_2$  in  $W_3$ , upoštevajoč njegovo mnenje o priporočilih teh agentov:

$$\begin{aligned} \omega_x^{J(W_1, W_2, W_3)} &= (\omega_{W_1}^J \otimes \omega_x^{W_1}) \oplus (\omega_{W_2}^J \otimes \omega_x^{W_2}) \oplus (\omega_{W_3}^J \otimes \omega_x^{W_3}) = \\ &= \{0, 81, 0, 00, 0, 19\} \oplus \{0, 00, 0, 00, 1, 00\} \oplus \{0, 09, 0, 00, 0, 91\} = \\ &= \left\{ \frac{(0,81*1,00)+(0,00*0,19)}{0,19+1,00-(0,19*1,00)}, \frac{(0,00*1,00)+(0,00*0,19)}{0,19+1,00-(0,19*1,00)}, \frac{0,19*1,00}{0,19+1,00-(0,19*1,00)} \right\} \\ &\quad \oplus \{0, 09, 0, 00, 0, 91\} = \\ &= \{0, 81, 0, 00, 0, 19\} \oplus \{0, 09, 0, 00, 0, 91\} = \{0, 8135, 0, 0000, 0, 1865\}. \end{aligned}$$

Glede na izračun po Jøsangovi subjektivni logiki je stopnja zaupanja agenta  $J$  do agenta  $X$  v danem primeru enaka 0,8135.

### 2.2.6 REGRET

Model REGRET [26, 27, 53] pri formalizaciji ugleda agentov upošteva tri vire informacij: neposredne izkušnje, informacije drugih agentov in družbeno strukturo mreže agentov. Model definira tri dimenzije ugleda:

- *Individualna dimenzija* (angl. *individual dimension*) modelira neposredno interakcijo med dvema agentoma. *Neposredni ugled* (angl. *outcome reputation*) je izračunan na podlagi izidov neposrednih interakcij in je obravnavan kot najbolj zanesljiv. Označen je z  $R_{O_{a \rightarrow b}}(\varphi)$ , kjer je  $a$  agent ocenjevalec,  $b$  je ocenjevani agent in  $\varphi$  predstavlja aspekt ocenjevanja.<sup>1</sup>
- *Družbena dimenzija* (angl. *social dimension*) modelira ugled na podlagi informacij, ki jih posredujejo drugi agenti. Kljub temu, da velja neposredni ugled za najbolj zanesljiv vir informacij, ne more biti vedno uporabljen, saj vsak agent nima neposredne izkušnje z vsemi drugimi agenti in zato za izračun uporabi druge informacije. Družbena dimenzija razlikuje med *priporočenim ugledom* (angl. *witness reputation*)  $R_{W_{a \rightarrow b}}(\varphi)$ , ki se izračuna na podlagi mnenj drugih agentov iz skupnosti, *ugledom poznanstev* (angl. *neighbourhood reputation*)  $R_{N_{a \rightarrow b}}(\varphi)$ , izpeljanim glede na družbene relacije agentov v omrežju, ki se razvijajo preko interakcij agentov, in *sistemskim ugledom* (angl. *system reputation*)  $R_{S_{a \rightarrow b}}(\varphi)$ , ki izhaja iz vlog in lastnosti posameznih agentov.
- *Ontološka dimenzija* (angl. *ontological dimension*) pri modeliranju ugleda upošteva in združi različne aspekte ocenjevanja.

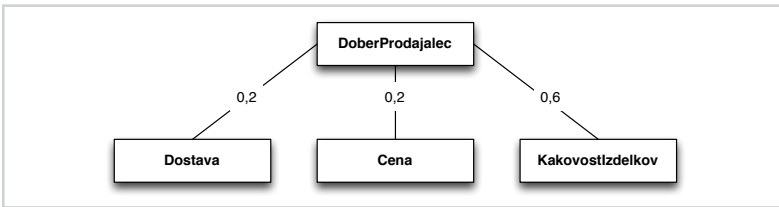
Poleg različnih tipov ugleda model REGRET definira tudi zanesljivost posameznega tipa ugleda. Definicija končne vrednosti ugleda agenta upošteva vse tri dimenzije ugleda in izračunane vrednosti zanesljivosti vseh vrst ugleda.

#### Primer

Agent  $A$  je kupec v spletni trgovini in želi z modelom REGRET izračunati vrednost zaupanja do prodajalca  $B$ . Aspekt ocenjevanja je  $\varphi = \text{DoberProdajalec}$ . Agent  $A$  še ni imel preteklih izkušenj z agentom  $B$ , zato neposrednega ugleda  $R_{O_{A \rightarrow B}}(\varphi)$  ni mogoče izračunati. Za izračun zaupanja zato uporabi t. i. družbene dimenzije ugleda.

<sup>1</sup>Primeri aspektov pri ocenjevanju spletnega trgovca so: cena izdelkov, čas dostave, kakovost ipd.

Ker agent  $A$  prvič kupuje v dani spletni trgovini in še nima relacij z drugimi agenti, tudi ugleda poznanstev ne moremo izračunati. Poleg tega tudi nihče od ostalih agentov v sistemu še ni podal mnenja o prodajalcu  $B$ , zato tudi izračun priporočenega ugleda ni mogoč. V takem primeru model REGRET definira izračun ugleda na podlagi sistemskega ugleda. Agent  $B$  je zaposlen v podjetju *Company*. Za delavce tega podjetja veljajo naslednje vrednosti sistemskega ugleda za posamezne aspekte:  $R_{S_{A \rightarrow B}}(\text{Dostava}) = 0,7$ ,  $R_{S_{A \rightarrow B}}(\text{KakovostIzdelkov}) = 0,6$  in  $R_{S_{A \rightarrow B}}(\text{Cena}) = 0,9$ . Od teh aspektov je odvisen aspekt ocenjevanja  $\varphi = \text{DoberProdajalec}$  z ontološko strukturo, kot je prikazano na sliki 2.3. Glede na dane vrednosti sistemskega ugleda in ontološko strukturo aspekta *DoberProdajalec* izračunamo stopnjo zaupanja, da je agent  $B$  dober prodajalec, kot sledi:  $R_{A \rightarrow B}(\text{DoberProdajalec}) = 0,2 * R_{S_{A \rightarrow B}}(\text{Dostava}) + 0,6 * R_{S_{A \rightarrow B}}(\text{KakovostIzdelkov}) + 0,2 * R_{S_{A \rightarrow B}}(\text{Cena}) = 0,2 * 0,7 + 0,6 * 0,6 + 0,2 * 0,9 = 0,68$ .



Slika 2.3

Ontološka struktura aspektov pri ocenjevanju prodajalca. Vir: [27].

### 2.2.7 SPORAS in HISTOS

SPORAS [54] modelira ugled agenta na podlagi ocen, ki jih o tem agentu posredujejo drugi agenti v skupnosti. Če imata agenta več interakcij in po vsaki interakciji posredujeta nove ocene, se pri izračunu ugleda upoštevajo samo najnovejše ocene. Pri izračunu ugleda se upošteva tudi ugled agenta, ki je podal oceno za drugega uporabnika. Agent, ki na novo vstopi v sistem, ima najmanjšo možno vrednost ugleda, ki jo lahko postopoma poveča. Agentom, ki so že dosegli visoko stopnjo ugleda, se vrednost ugleda po vsaki nadaljnji interakciji zelo malo spremeni, in sicer v skladu s funkcijo shranjevanja ugleda. Ugled agenta po  $i$ -ti interakciji je definiran kot  $R_i = R_{i-1} + \frac{1}{\theta} * \Phi(R_i) * R_i^{other} * (W_i - E(W_i))$ . V dani enačbi  $R_{i-1}$  predstavlja trenutno vrednost ugleda agenta (ki je bila izračunana po prejšnji interakciji s tem agentom). Parameter  $\theta$  je konstantno celo število večje od 1, ki določa, kako hitro se lahko spremeni

vrednost ugleda.  $\Phi(R_i) = 1 - \frac{1}{1 + e^{\frac{-(R_{i-1}-D)}{\sigma}}}$  predstavlja funkcijo shranjevanja ugleda, pri čemer je  $D$  razpon vrednosti ugleda (v sistemu SPORAS ima ugled vrednost med 0 in 3000),  $\sigma$  pa je pospeševalni faktor funkcije shranjevanja ugleda.  $W_i$  je ocena ugleda agenta, ki ima vrednost med 0, 1 (najslabše) in 1, 0 (najboljše),  $R_i^{other}$  je ugled agenta, ki je to oceno posredoval.  $E(W_i) = \frac{R_{i-1}}{D}$  predstavlja pričakovano vrednost ocene ugleda.

Model HISTOS [54] je nadgradnja modela SPORAS, ki v formalizacijo ugleda vpele še personalizem. Agenti bolj upoštevajo ocene tistih, ki jim zaupajo in so že sodelovali z njimi, kot pa nepoznanim agentom. HISTOS definira družbeno omrežje agentov, v katerem vozlišča predstavljajo agente, uteži na povezavah med vozlišči pa najnovije ocene ugleda, s katerimi se agenta ocenita. Ugled agenta, ki poda oceno za drugega uporabnika, ni več splošen ugled, ampak vsak agent izračuna subjektivno vrednost ugleda na podlagi poznanstev (tj. tranzitivnih povezav) med agenti.

### Primer

Zanima nas vrednost ugleda agenta po dani interakciji  $i$ , ki ga izračunamo z modelom SPORAS. Pospeševalni faktor funkcije shranjevanja ugleda ima vrednost  $\sigma = 10$ . Trenutna vrednost ugleda agenta je  $R_{i-1} = 300$ , na podlagi česar izračunamo tudi pričakovano oceno ugleda  $E(W_i) = \frac{R_{i-1}}{D} = \frac{300}{3000} = 0,1$ . Agent z vrednostjo ugleda 2000 (tj.  $R_i^{other} = 2000$ ) je ugled ocenjevanega agenta ocenil z vrednostjo  $W_i = 0,5$ . Če za izračun uporabimo vrednost parametra  $\theta = 1000$ , je vrednost ugleda agenta enaka  $R_i = R_{i-1} + \frac{1}{\theta} * \Phi(R_i) * R_i^{other} * (W_i - E(W_i)) = 300 + \frac{1}{1000} * (1 - \frac{1}{1 + e^{\frac{-(300-3000)}{10}}}) * 2000 * (0,5 - 0,1) = 301$ . Nova vrednost ugleda je minimalno višja od trenutne vrednosti ugleda. Hitrost, kako hitro se lahko spremeni vrednost ugleda, določamo s parametrom  $\theta$ . Če uporabimo nižjo vrednost za parameter  $\theta$ , se spremembe vrednosti ugleda odrazijo hitreje. Na primer, za vrednost parametra  $\theta = 10$  je nova vrednost ugleda po interakciji  $i$  (po enakem izračunu) enaka  $R_i = 380$ .

### 2.2.8 Spletni modeli ugleda

Amazon [55] in eBay [56] sta primera komercialnih spletišč z implementacijami sistemov za obvladovanje zaupanja in ugleda, ki temeljijo na preprostih modelih zaupanja in ugleda. V praksi implementirani in razširjeni sistemi za obvladovanje zaupanja in ugleda večinoma temeljijo na formalizacijah, ki zaupanje in ugled modelirajo z uporabo osnovnih matematičnih operacij seštevanja in deljenja. Uporabniki sistema eBay drug

drugega ocenjujejo z vrednostmi  $-1$  (negativna izkušnja),  $1$  (pozitivna izkušnja) in  $0$  (nevtralna izkušnja). Ugled uporabnika je formaliziran kot odstotek pozitivnih ocen, ki jih je uporabnik dobil v zadnjih dvanajstih mesecih:  $R(u) = \frac{\text{total positive}}{\text{total positive} + \text{total negative}}$ . Nov uporabnik sistema ima začetno vrednost ugleda enako  $0$ . Pri Amazonu je ugled agenta izračunan kot povprečje vseh prejetih ocen. Kot rečeno, gre za preproste modele zaupanja in ugleda, ki ne upoštevajo konteksta, ugleda ocenjevalca, psiholoških in drugih dejavnikov, ki vplivajo na zaupanje in ugled agentov. Kljub enostavnosti in mnogim pomanjkljivostim pa zasedajo spletni modeli ugleda zaradi svoje razširjenosti pomembno mesto med modeli zaupanja in ugleda.

### Primer

Izračunali bomo vrednost ugleda uporabnika v sistemu, ki uporablja eBay. Uporabnika  $A$  je v zadnjem letu ocenilo  $15492$  drugih uporabnikov, od tega je  $15026$  izkušnjo z njim ocenilo kot pozitivno,  $182$  kot nevtralno in  $284$  kot negativno. Ugled uporabnika  $A$  zato izračunamo kot  $R(A) = \frac{15026}{15026 + 284} = 98,15\%$ .

## 2.3 Sistemi za obvladovanje zaupanja in ugleda

*Sisteme za obvladovanje zaupanja in ugleda* (angl. *trust and reputation management systems*) definiramo kot "sisteme, ki zbirajo in obdelujejo informacije, potrebne za vzpostavitev in regulacijo zaupanja med agenti v e-okolju". Z vpeljavo teh sistemov agentom v danem e-okolju zagotavljamo mehko varnost. Za razliko od klasičnih varnostnih mehanizmov, ki zagotavljajo trdo varnost, sistemi za obvladovanje zaupanja in ugleda upoštevajo tudi psihološko in sociološko komponento zaupanja. Implementirajo različne matematične modele – ključne smo opisali v poglavju 2.2. Poudariti velja, da sistemi za obvladovanje zaupanja in ugleda *niso* odločitveni sistemi, a so s slednjimi lahko povezani [57]. Vrednost zaupanja, ki jo vrne sistem za obvladovanje zaupanja in ugleda, je namreč eden izmed parametrov odločanja v mnogih odločitvenih modelih, npr. [19, 58–60].

Obvladovanje zaupanja in ugleda v e-okoljih se seveda razlikuje od obvladovanja zaupanja in ugleda v fizičnem svetu. Tradicionalnih iztočnic za določanje zaupanja, kot so npr. obrazna mimika in telesne kretnje, v spletnem okolju ne moremo direktno opazovati in jih je treba nadomestiti z drugimi dosegljivimi elektronskimi viri [4, 6]. Po drugi strani pa internet omogoča izmenjavo in širjenje informacij, povezanih z zaupanjem in ugledom na globalni ravni, kar je v fizičnem svetu sicer omejeno na lokalne

skupnosti. Glavni namen in cilj raziskav na področju sistemov za obvladovanje zaupanja in ugleda sta tako [6]:

1. Poiskati ustrezne spletne elemente, ki bodo nadomestili tradicionalne iztočnice za določanje zaupanja in ugleda, ki smo jih vajeni v fizičnem svetu; in opredeliti elektronske vire (specifične za določeno e-okolje), s katerimi lahko ustrezno določimo stopnjo zaupanja in ugleda.
2. Izkoristiti prednosti interneta in informacijske tehnologije za zbiranje teh podatkov ter izračun zaupanja in ugleda, s čimer bi podprli proces odločanja in izboljšali kakovost interakcij v e-okolju.

Osnovna ideja sistemov za obvladovanje zaupanja in ugleda je, da agenti podajo mnenje oziroma oceno zaupanja do drugih agentov, s katerimi so imeli izkušnje, nato pa sistem za obvladovanje zaupanja in ugleda obdela podane ocene in druge vire informacij ter s tem izpelje vrednost zaupanja do posameznega agenta. Za uspešno delovanje sistemov za obvladovanje zaupanja in ugleda morajo veljati naslednje predpostavke [1]:

- Agenti morajo biti v sistemu prisotni daljši čas, tako da lahko pričakujemo transakcije agenta v prihodnosti in napovemo kakovost teh.
- Ocene oziroma mnenja o izvršenih transakcijah so zajete in vidne ostalim agentom v e-okolju.
- Ocene preteklih interakcij so vodilo za odločanje glede prihodnjih interakcij.

Področje uporabe, kjer so sistemi za obvladovanje zaupanja in ugleda potrebni oziroma zaželeni, je zelo široko. Uporaba preprostih sistemov za obvladovanje zaupanja in ugleda je zelo razširjena na spletnih mestih za elektronske licitacije in trgovanje, kot je na primer spletno mesto eBay [56]. V tovrstnih e-okoljih je obvladovanje zaupanja in ugleda še kako pomembno, saj gre za denarne transakcije med nepoznanimi agenti. Asimetrija informacij (kupec ne ve, kaj bo dobil, in prodajalec ne ve, ali bo kupec res plačal blago) še dodatno povečuje tveganje, ki pa ga je moč omiliti z uporabo tovrstnih sistemov. S spletnimi tržišči je tesno povezana tudi uporaba sistemov za obvladovanje zaupanja in ugleda na spletiščih za recenzije izdelkov in storitev, kot so na primer spletišča Epinions [61], Yelp [62], Amazon [55] ali BizRate [63]. Sistemi za obvladovanje zaupanja in ugleda se uporabljajo za učinkovitejše delovanje rešitev, ki tečejo

na vzporednih računalnikih, in sicer za izračun zaupanja, ali bo vsak avtonomno delujoči računalnik v mreži deloval po pričakovanjih [64, 65]. S sistemi za obvladovanje zaupanja in ugleda prav tako izboljšamo strukture, ki so namenjene komunikaciji in skupinskemu delu na daljavo [66] ter portale za izmenjavo oziroma pridobivanje strokovnih mnenj, kot sta AllExperts [67] in Advogato [68]. Poudariti velja, da sistemi za obvladovanje zaupanja in ugleda s prepoznavanjem (in morebitnim izločanjem) slabih/zlonamernih agentov v danem spletnem okolju obenem predstavljajo vzpodbudo za dobro obnašanje agentov, s čimer se poveča kakovost ponujenih storitev v e-okoljih.

Na področju računalniško podprtega obvladovanja zaupanja in ugleda obstaja še veliko nerešenih ali nezadovoljivo rešenih problemov. Glavna odprta vprašanja so naslednja [6, 69, 70]: Kako agente vzpodbuditi k podajanju mnenj o drugih agentih in deljenju osebnih izkušenj? Kako preprečiti namerno podajanje napačnih ocen? Kako obravnavati nagnjenost k pozitivnemu ocenjevanju?<sup>2</sup> Kako identificirati agente z zlonamernim obnašanjem? Kako preprečiti menjave identitet agentov? Kako prepoznati in preprečiti organizirano (naročeno) skupinsko posredovanje mnenj in ocen? Kako obravnavati subjektivnost ocen? Kako prepoznati diskriminatorno obnašanje agentov do posameznih agentov?

V sklopu doktorske disertacije bomo obravnavali in predlagali (1) rešitev za identifikacijo in izločanje zlaganih ocen, ki so posredovane v sistem za obvladovanje zaupanja in ugleda, in (2) rešitev za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti v ocenah. Ostala odprta vprašanja so izven obsega doktorske disertacije in zgolj nakazujejo kompleksnost problematike na področju sistemov za obvladovanje zaupanja in ugleda.

---

<sup>2</sup>Analiza eBay-evega sistema za obvladovanje ugleda je pokazala, da je negativnih samo 0,6 % vseh ocen, ki so jih posredovali kupci, oziroma 1,6 % vseh ocen, posredovanih s strani prodajalcev [1]. Možne razlage za nagnjenost k pozitivnemu ocenjevanju so: (1) do takega ocenjevanje pride zaradi vljudnosti agentov [1], (2) uporabniki podajo pozitivno oceno v upanju, da bodo tudi sami kot povračilo dobili pozitivno oceno [71], oziroma (3) uporabniki ne podajo negativnih ocen zaradi strahu, da jim ocenjeni uporabnik ne bi "iz maščevanja" vrnil slabe ocene [1].



*Temeljna formalizacija  
zaupanja v e-okolju*

### 3.1 Model kvalitativne dinamike ocenjevanja (QAD)

Naše delo temelji na modelu kvalitativne dinamike ocenjevanja (angl. Qualitative Assessment Dynamics, QAD) [15, 16], ki dopolnjuje obstoječe modele zaupanja in ugleda.

Za razvoj dobrega modela zaupanja in ugleda je potrebna identifikacija dejavnikov, ki posredno ali neposredno vplivajo na vrednost zaupanja med agenti in posledično na stopnjo ugleda agentov. Model QAD med drugim upošteva raziskave, ki jih je predstavil švicarski razvojni psiholog Jean Piaget [72]. Zaupanje je izraz procesov razmišljanja in presojanja, ki imajo izvor v psihologiji [15]. Pri formalizaciji zaupanja je zato treba upoštevati določena dejstva in ugotovitve, ki so domena psihologije. Na podlagi teh dejstev model QAD definira naslednje ključne dejavnike zaupanja [13, 15]:

- *Časovna dinamika.* Odnos zaupanja med agentoma se dinamično spreminja s časom.
- *Racionalnost in iracionalnost.* Treba je upoštevati, da se agent lahko pod določenimi pogoji obnaša racionalno, pod drugimi pa iracionalno.
- *Interakcije z okoljem.* Zaupanje predstavlja osnovo za dejanja oziroma vrsto dejanj agenta z okoljem.
- *Zunanji dejavniki.* Zaupanje ni zgolj produkt neodvisnega mišljenja agenta, ampak nanj vpliva tudi okolje. Na zaupanje vpliva mnenje družbe, ki se v e-okoljih v največji meri odraža prek priporočil drugih agentov.
- *Razlikovanje zaupanja.* Zaupanje se odraža v različnih oblikah. Razlogi za to so različne jezikovne sposobnosti agentov za izražanje zaupanja in različno zaznavanje sposobnosti ocenjenih agentov. Govorimo tudi o subjektivnosti zaupanja.

Zaupanje temelji na zaznavanju in ocenjevanju, ki ga vodijo različni dejavniki (racionalni, neracionalni, okoljski ipd.), ki jih je pri modeliranju zaupanja treba upoštevati [40]. Model QAD upošteva nekatere prepoznane družbeno-kognitivne faktorje zaupanja ter modelira vrednosti zaupanja in ugleda s kvalitativnimi vrednostmi, kot so *zaupanja vreden odnos*, *delno zaupanja vreden odnos* itd. V tem poglavju bomo podrobno predstavili model zaupanja QAD.

### 3.2 Osnovna notacija

V formalizaciji zaupanja, ki temelji na modelu QAD, virtualna skupnost sestoji iz množice agentov  $\mathcal{A} = \{a_1, a_2, a_3, \dots, a_n\}$  in odnosov zaupanja med agenti. Kot že opisano v poglavju 2.1.2, *agenti* predstavljajo entitete v sistemu, kot so ljudje, programski agenti, ki si jih ljudje lastijo in jih upravljajo, storitve ipd. Končne vrednosti zaupanja med agenti so odvisne od agentu pripisanega *operatorja*, ki predstavlja agentov “značaj”. Vrednosti zaupanja med agenti se lahko skozi čas spreminjajo.

Za ustrezno opredelitev zaupanja v e-okoljih je potrebna definicija, ki je dovolj ozka, da omogoča formalno obravnavo in jo posledično lahko ustrezno umestimo v okolja z računalniško podprtim obvladovanjem zaupanja [40]. Pojmovanje zaupanja v modelu QAD temelji na definiciji, ki jo je podala D. E. Denning: “Zaupanje je ocena, ki je posledica izkušenj, ki si jih ljudje prek interakcij izmenjujejo, ter se posodobi ob vsaki uporabi sistema [73].” Izhajajoč iz predlagane definicije, model QAD definira zaupanje, kot sledi:

*Definicija 1:* Zaupanje je relacija med agentoma  $a_i$  and  $a_j$ . Oznaka  $\omega_{i,j}$  opredeljuje oceno zaupanja agenta  $a_i$  do agenta  $a_j$ . Ocena zaupanja je vrednost iz množice ocen zaupanja  $\Omega = \{-2, -1, 0, 1, 2\}$ , kjer elementi množice simbolizirajo zaupanja nevreden, delno zaupanja nevreden, neodločen, delno zaupanja vreden in zaupanja vreden odnos. Če odnos zaupanja med dvema agentoma ni definiran, to označimo z notacijo “/”.

Pri vrednotenju zaupanja sodelujeta dva agenta: AGENT OCENJEVALEC in OCENJEVANI AGENT. Velja predpostavka, da je agent ocenjevalec zmožen ocenjevanja, medtem ko ocenjevani agent lahko predstavlja tudi abstraktno (in ne nujno razumno) vrsto agenta. Agent ocenjevalec vrednoti zaupanje (tj. izračuna oceno zaupanja) do ocenjevanega agenta. V definiciji 1 je  $a_i$  agent ocenjevalec in  $a_j$  je ocenjevani agent.

Posebej velja poudariti razliko med “zaupanja nevrednim odnosom” in “nedefiniranim odnosom zaupanja” med agentoma:

- Odnos med agentoma  $a_i$  in  $a_j$  je *nedefiniran*, če agent ocenjevalec še ni ocenil zaupanja do ocenjevanega agenta. Agent ocenjevalec bodisi še ni imel osebnih izkušenj z ocenjevanim agentom bodisi ocenjevanega agenta še ni primerjal z ostalimi na podlagi mnenj drugih agentov. Pravimo, da agent ocenjevalec ne

pozna ocenjevanega agenta. Odnos med agentom je nedefiniran tudi, če agent ocenjevalec ne razkrije vrednosti zaupanja do ocenjevanega agenta. Nedefiniran odnos označimo z  $\omega_{i,j} = \text{"/}$ ". Odnos zaupanja je enosmerna relacija in  $\omega_{i,j} = \text{"/}$ " ne pomeni tudi  $\omega_{j,i} = \text{"/}$ ".

- *Zaupanja nevreden* odnos med agentoma nakazuje to, da je imel agent ocenjevalec slabe izkušnje z ocenjevanim agentom oziroma ga je kot zaupanja nevrednega ocenil na podlagi mnenj drugih agentov in njegovega (negativnega) nagnjenja k zaupanju. Agent ocenjevalec torej ve za ocenjevanega agenta in je na podlagi znanja o ocenjevanem agentu ocenil, da je ta zaupanja nevreden. Tako kot v prejšnjem primeru, tudi v tem vrednost zaupanja  $\omega_{i,j} = -2$  ne pomeni nujno  $\omega_{j,i} = -2$ .

*Definicija 2:* Zaupanje med agenti v danem kontekstu  $\Gamma$  predstavimo z matriko zaupanja  $\mathcal{M}_\Gamma$ . Element matrike  $\omega_{i,j}$  označuje oceno zaupanja agenta  $a_i$  do agenta  $a_j$  v kontekstu  $\Gamma$ .

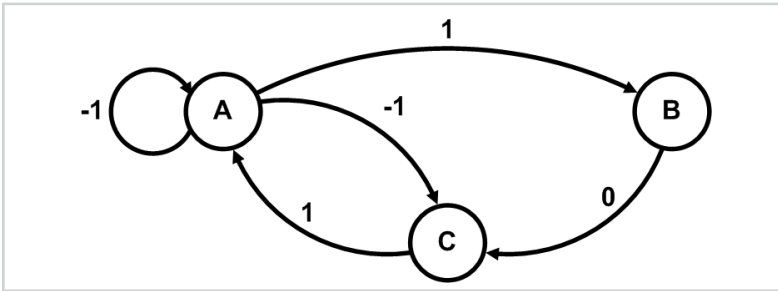
Vrednost zaupanja med agenti je odvisna od konteksta. Ocena zaupanja agenta  $a_i$  do agenta  $a_j$  namreč ni nujno enaka za različne kontekste. Na primer, agent  $a_i$  oceni, da je agent  $a_j$  zaupanja vreden za popravilo računalnika, a hkrati oceni, da je zaupanja nevreden za vožnjo letala. Kakorkoli, v nadaljevanju dela bomo predpostavljali, da ima agent v določenem sistemu samo eno vlogo, in tako upoštevali le en kontekst. Oznaka za kontekst je zato iz vseh nadaljnjih definicij izpuščena.

Splošna oblika matrike zaupanja v okolju z  $n$  agenti je naslednja:

$$\mathcal{M} = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \dots & \omega_{1,n} \\ \omega_{2,1} & \omega_{2,2} & \dots & \omega_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{n,1} & \omega_{n,2} & \dots & \omega_{n,n} \end{bmatrix}.$$

Širjenje zaupanja v e-okolju lahko ekvivalentno predstavimo tudi z *grafom zaupanja*. Povezave v grafu so utežene s kvalitativnimi utežmi in usmerjene. Oceno zaupanja agenta  $A$  do agenta  $B$  podamo s povezavo, ki je usmerjena od  $A$  proti  $B$ . Primer grafa zaupanja je prikazan na sliki 3.1.

Matrika in graf zaupanja se skozi čas spreminjata, in sicer po vsaki transakciji, ki je spremenila vrednost zaupanja med dvema agentoma.



Slika 3.1

Primer grafa zaupanja s tremi agenti: A, B in C.

*Definicija 3:* V matriki zaupanja  $\mathcal{M}$  vrstica  $k$  predstavlja vektor ocen zaupanja agenta  $a_k$ . Vsebuje ocene zaupanja agenta  $a_k$  do ostalih agentov v skupnosti. Označimo ga z  $\mathcal{M}_{k,n} = [\omega_{k,1}, \omega_{k,2}, \dots, \omega_{k,n}]^T$ . Nadalje,  $\overline{\mathcal{M}}_{k,n_k} = [\omega_{k,1}, \omega_{k,2}, \dots, \omega_{k,n_k}]^T$  označuje podvektor ocen zaupanja agenta  $a_k$ , v katerem so izpuščene vrednosti “/”.

V (pod)vektorju ocen zaupanja agenta so shranjene vrednosti, ki so odraz agentovega subjektivnega pogleda na stopnjo zaupanja do ocenjevanih agentov. Odraz celotne skupnosti o vrednosti zaupanja do ocenjevanega agenta je prikazan z vektorjem ugleda:

*Definicija 4:* V matriki zaupanja  $\mathcal{M}$  stolpec  $k$  predstavlja vektor ugleda agenta  $a_k$ . Vsebuje ocene zaupanja do agenta  $a_k$ , ki so jih podali ostali agenti v skupnosti. Označimo ga z  $\mathcal{M}_{n,k} = [\omega_{1,k}, \omega_{2,k}, \dots, \omega_{n,k}]$ . Nadalje,  $\overline{\mathcal{M}}_{n,k} = [\omega_{1,k}, \omega_{2,k}, \dots, \omega_{n_k,k}]$  označuje podvektor ugleda agenta  $a_k$ , v katerem so izpuščene vrednosti “/”.

V definicijah 3 in 4 indeks  $n_k$  označuje število definiranih ocen zaupanja v podvektorju ocen zaupanja in podvektorju ugleda. Če so vse vrednosti definirane, potem sta vektor in njegov podvektor enaka.

### 3.3 Operatorji zaupanja QAD

V resničnem svetu ljudje zaupajo na različne načine, tudi v popolnoma enakih okoliščinah. Kot smo že omenili, vrednotenje zaupanja med agenti v modelu QAD temelji na “značajih” agentov. Model zaupanja QAD definira množico operatorjev QAD, ki

so pripisani agentom in modelirajo njihov odnos zaupanja.

*Definicija 5:* Operatorji zaupanja QAD so elementi množice  $\Psi = \{\uparrow, \Downarrow, \rightsquigarrow, \leftrightarrow, \uparrow, \downarrow\}$ , kjer simboli predstavljajo ekstremno optimistični, ekstremno pesimistični, centralistični, oportunistični, zmerno optimistični in zmerno pesimistični operator. Operatorji so funkcije  $f_i \in \Psi$ , tako da  $f_i : \underline{\mathcal{M}}_{n_i,j} = [\omega_{1,j}^-, \omega_{2,j}^-, \dots, \omega_{n_i,j}^-] \rightarrow \omega_{i,j}^+$ , kjer nadpisani “-” označuje vrednost pred operacijo in nadpisani “+” označuje vrednost po operaciji izračuna zaupanja. Indeks  $n_a$  označuje število vseh elementov, ki so vzeti v izračun vrednosti zaupanja. Preslikave za določene operatorje QAD so definirane, kot sledi:

a)  $\omega_{i,j}^- \neq \text{“} \infty \text{”}$ :

▪

$$\uparrow_i: \quad \max(\underline{\mathcal{M}}_{n_i,j}) \rightarrow \omega_{i,j}^+$$

▪

$$\Downarrow_i: \quad \min(\underline{\mathcal{M}}_{n_i,j}) \rightarrow \omega_{i,j}^+$$

▪

$$\rightsquigarrow_i: \quad \begin{cases} \lceil \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- \rceil \rightarrow \omega_{i,j}^+ & \text{če } \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- < 0 \\ \lfloor \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- \rfloor \rightarrow \omega_{i,j}^+ & \text{sicer} \end{cases}$$

▪

$$\leftrightarrow_i: \quad \begin{cases} \lceil \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega_{i,j}^- \rceil \rightarrow \omega_{i,j}^+ & \text{če } \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- > 0 \\ \lfloor \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega_{k,j}^- \rfloor \rightarrow \omega_{i,j}^+ & \text{sicer} \end{cases}$$

▪

$$\uparrow_i: \quad \begin{cases} \omega_{i,j}^- \rightarrow \omega_{i,j}^+ & \text{če } \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- \leq \omega_{i,j}^- \\ \omega_{i,j}^- + 1 \rightarrow \omega_{i,j}^+ & \text{sicer} \end{cases}$$

▪

$$\downarrow_i: \quad \begin{cases} \omega_{i,j}^- \rightarrow \omega_{i,j}^+ & \text{če } \frac{1}{n_a} \sum_{\omega^- \in \underline{\mathcal{M}}_{n_i,j}} \omega^- \geq \omega_{i,j}^- \\ \omega_{i,j}^- - 1 \rightarrow \omega_{i,j}^+ & \text{sicer} \end{cases}$$

$$\text{b) } \omega_{ij}^- = "f": \quad "f" \rightarrow \omega_{ij}^+$$

Za model zaupanja QAD velja, da je vrednost  $n_a = |\mathcal{M}_{n_j}|$ .

Agent ocenjevalec v izračun zaupanja do ocenjevanega agenta vzame NEPOSREDNO OCENO ZAUPANJA do ocenjevanega agenta ( $\omega_{ij}^-$ ) in vrednosti drugih agentov v skupnosti do ocenjevanega agenta ter jih preslika glede na predpis funkcije operatorja zaupanja QAD. *Druge* (neodvisne) agente v nadaljevanju poimenujemo AGENTI SVETOVALCI, agregirano vrednost zaupanja ( $\omega_{ij}^+$ ), ki predstavlja rezultat funkcije operatorja, pa KONČNA OCENA ZAUPANJA.

V matriki zaupanja  $\mathcal{M}$  so shranjene končne ocene zaupanja  $\omega_{ij}^+$ . V nadaljevanju dela velja predpostavka, da  $\omega_{ij} = \omega_{ij}^+$ . Nadpisani "+" bo s tem praviloma izpuščen iz notacije.

Lastnosti operatorjev lahko neformalno opišemo, kot sledi. Če je ocena zaupanja nedefinirana, potem ostane nedefinirana tudi po operaciji vrednotenja zaupanja, sicer se spremeni glede na preslikavo, ki jo določa operator. Ekstremno optimistični ocenjevalni operator ( $\uparrow$ ) za novo oceno zaupanja vzame najbolj pozitivno vrednost iz podvektorja ugleda ocenjevanega agenta. Ravno nasprotno velja za ekstremno pesimistični ocenjevalni operator ( $\downarrow$ ), ki za novo vrednost vzame najbolj negativno vrednost. Operatorji  $\sim$ ,  $\leftrightarrow$ ,  $\uparrow$  in  $\downarrow$  vrednost zaupanja modelirajo na podlagi "povprečne" vrednosti ocen iz podvektorja ugleda ocenjevanega agenta. Ker je zaupanje v modelu QAD definirano s kvalitativnimi ocenami, ki simbolizirajo zaupanja nevreden (-2), delno zaupanja nevreden (-1), nedoločen (0), delno zaupanja vreden (1) in zaupanja vreden (2) odnos, se ocene za potrebe izračuna povprečne vrednosti pretvorijo v pripadajoče numerične vrednosti. Centralistični operator ( $\sim$ ) tako izračuna povprečno vrednost ocen iz podvektorja ugleda ocenjevanega agenta in dobljeno vrednost zaokroži na celo število, ki je bližje vrednosti 0. Povprečna vrednost ocen iz podvektorja ugleda se izračuna tudi pri uporabi oportunističnega ocenjevalnega operatorja ( $\leftrightarrow$ ), ki dobljeno vrednost (v nasprotju s prejšnjim operatorjem) zaokroži na celo število, ki je dlje od vrednosti 0. Nova ocena zaupanja, izračunana z zmernim optimističnim operatorjem ( $\uparrow$ ), je enaka stari oceni zaupanja, če je povprečne vrednosti elementov iz podvektorja ugleda ocenjevanega agenta nižje od neposredne ocene zaupanja agenta ocenjevalca do ocenjevanega agenta. Če imajo agenti svetovalci v povprečju višjo oceno zaupanja, tj. povprečne elementov iz podvektorja ugleda je višje od neposredne ocene zaupanja, je nova ocena zaupanja za eno stopnjo višja od stare vrednosti ocene zaupanja. Na zmerni pesimi-

stični ocenjevalni operator ( $\Downarrow$ ) (ravno obratno kot pri operatorju  $\Uparrow$ ) vplivajo negativne ocene zaupanja v skupnosti. Nova ocena zaupanja, izračunana z operatorjem  $\Downarrow$ , je za eno stopnjo nižja, če je povprečje elementov iz podvektorja ugleda ocenjevanega agenta nižje od neposredne vrednosti zaupanja. S tem se zmanjša razlika med neposredno oceno zaupanja agenta ocenjevalca do ocenjevanega agenta in splošnim mnenjem drugih agentov v skupnosti do ocenjevanega agenta. Če je povprečje elementov iz podvektorja ugleda ocenjevanega agenta višje od neposredne ocene zaupanja, je nova neposredna ocena zaupanja enaka stari.

Neformalnemu opisu lastnosti operatorjev sledi eksperimentalna analiza delovanja operatorjev v naslednjem poglavju.

### 3.4 Analiza modela zaupanja QAD

Lastnosti operatorjev QAD in ostale splošne lastnosti modela zaupanja QAD bomo opisali z rezultati simulacij, ki so bile izvedene s simulacijskim orodjem Zaupnik, ki smo ga implementirali v ta namen. Za potrebe analize modela zaupanja QAD smo izvedli dva niza demonstracijskih eksperimentov. Prvi niz eksperimentov je namenjen splošni analizi lastnosti modela QAD, drugi niz pa poglobljeni analizi lastnosti operatorjev QAD in procesu vrednotenja zaupanja z operatorji.

#### 3.4.1 Opis simulacij

Model zaupanja QAD definira izračun vrednosti zaupanja med agenti neodvisno od vrste transakcij med agenti. Za namene ovrednotenja modela QAD s simulacijami smo definirali način izvedbe transakcij in možne izide teh, kot sledi: Vsak agent v skupnosti je lahko bodisi ponudnik storitve bodisi iskalec storitve, vendar ne more imeti obeh vlog v isti transakciji. Simulacija poteka v diskretnih časovnih korakih in v vsakem koraku se lahko izvrši ena transakcija. Za izvedbo vsake transakcije sta naključno izbrana dva agenta iz skupnosti, pri čemer je eden od njiju iskalec storitve ( $a_R$ ) in drugi ponudnik storitve ( $a_P$ ). V danem simulacijskem koraku se izvrši transakcija med  $a_R$  in  $a_P$ , medtem ko ostali agenti v danem simulacijskem koraku ne izvajajo transakcij. Nekateri agenti v skupnosti so slabi ponudniki storitve, teh je  $u$ . Preostali so dobri ponudniki storitve, takih je skupno  $1 - u$ . Dobri ponudniki storitve iskalcu storitve vedno zagotovijo storitev z zadovoljivo kakovostjo. Slabi ponudniki storitve *včasih* (samo v nekaterih transakcijah) delujejo zlonamerno in zagotovijo storitev z nezadovoljivo

kakovostjo. Pogostost zlonamernega obnašanja slabih ponudnikov storitev je definirana s frekvenco zlonamernosti  $r_{mal}$ . Dalje, vsak agent  $a_x$  hrani zgodovino kakovosti transakcij (oznaka  $\bullet$  za storitev z nezadovoljivo kakovostjo in oznaka  $*$  za storitev z zadovoljivo kakovostjo), in sicer za vsakega agenta, s katerim je agent  $a_x$  imel transakcijo. Zgodovina vključuje ocene za zadnje štiri transakcije med agentoma  $a_R$  in  $a_P$ .

Za potrebe analize lastnosti modela zaupanja QAD smo določili dva različna simulacijska scenarija:

### Scenarij Z

V simulirani skupnosti sodeluje  $n$  agentov. V vsakem koraku simulacije sta naključno izbrana dva agenta: agent  $A$  (iskalec storitve) in agent  $B$  (ponudnik storitve). Če je to prva interakcija med  $A$  in  $B$ , potem  $A$  v vsakem primeru naredi transakcijo (tj. zahteva storitev) z agentom  $B$ . Če sta pred tem že imela interakcijo, potem  $A$  z verjetnostjo  $1 - p$  zahteva storitev, ki jo ponuja agent  $B$ . Po zahtevku agent  $B$  ponudi storitev, ki ima lahko zadovoljivo ali nezadovoljivo kakovost.

Verjetnost  $p$  se nanaša na količino zlonamernega obnašanja agentov v skupnosti in jo izračunamo kot  $p = u * r_{mal}$ . Če je v skupnosti 30 % slabih agentov, ki se v 80 % transakcij obnašajo zlonamerno, tj. ponudijo storitev z nezadovoljivo kakovostjo, potem je  $p = 0,3 * 0,8 = 0,24$ . To pomeni, da bo iskalec storitve naredil transakcijo z naključno izbranim ponudnikom storitev v 76 od 100 primerov.

Agenti v scenariju Z ne vrednotijo vrednosti zaupanja med njimi in zato tudi nimajo pripisanih operatorjev QAD. Sprejemanje oziroma zavračanje ponudnika storitve agenta  $B$  ne temelji na ovrednotenju vrednosti zaupanja do agenta  $B$ , pač pa je odločanje naključno.

Potek transakcije v posameznem koraku za opisani scenarij Z je s psevdokodo prikazan v algoritmu 1.

### Scenarij Y

V scenariju Y agenti sprejmejo oziroma zavrnejo ponudnika storitve glede na oceno zaupanja do ponudnika storitve, ovrednotenega po modelu zaupanja QAD. V simulirani skupnosti sodeluje  $n$  agentov s pripisanimi operatorji QAD in odnosi zaupanja med njimi. Na začetku je odnos zaupanja med vsemi agenti nedefiniran, tj.  $\omega_{i,j} = \text{''}l\text{''}$ ,  $\forall i,j$ .

*Algoritem 1*

Korak simulacije po scenariju Z

---

*Vhod:* naključno izbran agent iskalca storitve  $a_R$ , naključno izbran agent ponudnik storitve  $a_P$

$st\_transakcij(a_R, a_P) \leftarrow zgod\_transakcij(a_R, a_P).prestej$

$p \leftarrow odstotek\_slabih\_agentov * frekvenca\_zlonamernosti$

*if* ( $st\_transakcij = 0$ ) *then*

$izid\_transakcije \leftarrow naredi\_transakcijo(a_R, a_P)$

$zgod\_transakcij(a_R, a_P).dodaj(izid\_transakcije)$

*else*

$vrednost\_kovanca \leftarrow vrni\_nakljucno\_stevilo(0, 1)$

*if* ( $vrednost\_kovanca < (1 - p)$ ) *then*

$izid\_transakcije \leftarrow naredi\_transakcijo(a_R, a_P)$

$zgod\_transakcij(a_R, a_P).dodaj(izid\_transakcije)$

*end if*

*end if*

---

Enako kot v scenariju Z tudi v scenariju Y v vsakem koraku simulacije naključno izberemo iskalca storitve  $A$  in ponudnika storitve  $B$ . Če gre za prvo interakcijo med  $A$  in  $B$ , potem agent  $A$  v vsakem primeru zahteva in sprejme storitev od agenta  $B$ . Če sta  $A$  in  $B$  pred tem že imela interakcije, potem  $A$  preveri oceno zaupanja do  $B$ , tj. vrednost  $\omega_{A,B}$ . Če je vrednost  $\omega_{A,B}$  nad določenim pragom  $th$ , potem agent  $A$  zahteva storitev od ponudnika  $B$ , sicer ne. Agent  $B$  po zahtevku ponudi storitev z zadovoljivo ali nezadovoljivo kakovostjo, odvisno od tega, ali je dober ali slab ponudnik storitve. Po opravljeni transakciji agent  $A$  vnese izid transakcije (tj. storitev z zadovoljivo kakovostjo ali storitev z nezadovoljivo kakovostjo) v zgodovino kakovosti transakcij z agentom  $B$ : v vektor  $H_{A,B}$ . Vektor  $H_{A,B}$  vključuje izide zadnjih štirih transakcij med agentoma  $A$  in  $B$ . Če sta imela agenta  $A$  in  $B$  skupno manj kot štiri transakcije, potem so v vektorju  $H_{A,B}$  shranjeni izidi *vseh* izvedenih transakcij. Glede na kakovost ponujenih

storitev agent  $A$  izračuna *neposredno oceno zaupanja* do agenta  $B$  na naslednji način:

$$\omega_{A,B}^- = \begin{cases} 2 & \text{če } |H_{A,B}(\bullet)| = 0 \\ 1 & \text{če } |H_{A,B}(\bullet)| = 1 \\ 0 & \text{če } |H_{A,B}(\bullet)| = 2 \\ -1 & \text{če } |H_{A,B}(\bullet)| = 3 \\ -2 & \text{če } |H_{A,B}(\bullet)| = 4 \end{cases} \quad (3.1)$$

Na primer: če so zadnje štiri izkušnje agenta  $A$  z agentom  $B$  take, da je  $B$  priskrbel eno storitev z nezadovoljivo kakovostjo (tj.  $|H_{A,B}(\bullet)| = 1$ ), potem je neposredna ocena zaupanja agenta  $A$  do agenta  $B$  ocenjena kot *delno zaupanja vreden odnos*, tj.  $\omega_{A,B}^- = 1$ .

V modelu zaupanja QAD je vrednost zaupanja med dvema agentom odvisna od neposredne ocene zaupanja in ocen drugih agentov do ocenjevanega agenta (glej definicijo 5). Torej, po izračunu  $\omega_{A,B}^-$  agent  $A$  svojo oceno zaupanja do  $B$  združi še z ocenami ostalih agentov v skupnosti, ki so že imeli interakcijo z  $B$ . Način združevanja vseh ocen zaupanja je definiran z operatorjem QAD (po definiciji 5), ki je pripisan agentu  $A$ .

Vrednosti odnosov zaupanja med ostalimi agenti, ki ne sodelujejo v transakciji, v danem simulacijskem koraku niso posodobljene.

Potek transakcije v posameznem koraku za scenarij Y je s psevdokodo prikazan v algoritmu 2.

### 3.4.2 Splošna analiza lastnosti modela zaupanja QAD

Namen prvega niza poskusov je splošna analiza lastnosti modela zaupanja QAD v simuliranem e-okolju. Primerjali bomo, koliko storitev z nezadovoljivo kakovostjo sprejmejo agenti z odločanjem, ki temelji na zaupanju (scenarij Y), v primerjavi z naključnim odločanjem (scenarij Z).

#### Parametri simulacije

V simulacijskem okolju je skupno  $n = 30$  agentov, od tega je 70 % dobrih agentov in 30 % slabih agentov. Slabi agenti priskrbijo storitev z nezadovoljivo kakovostjo s frekvenco  $r_{mal} = 0,8$ . Prag za sprejem ponujene odločitve na podlagi odločanja, ki temelji na zaupanju (scenarij Y), je  $th = 1$ . Torej, iskalec storitve  $A$  mora imeti vsaj *delno zaupanja vreden odnos* do ponudnika storitve  $B$ , da sprejme njegovo storitev.

*Algoritem 2*

Korak simulacije po scenariju Y

---

*Vhod:* naključno izbran agent iskalec storitve  $a_R$ , naključno izbran agent ponudnik storitve  $a_P$ , matrika zaupanja  $\mathcal{M}$

$st\_transakcij(a_R, a_P) \leftarrow zgod\_transakcij(a_R, a_P).prestej$

*if* ( $st\_transakcij = 0$ ) *then*

$izid\_transakcije \leftarrow naredi\_transakcijo(a_R, a_P)$

$zgod\_transakcij(a_R, a_P).dodaj(izid\_transakcije)$

$\omega_{R,P}^- \leftarrow izracunaj\_neposredno\_oceno(zgod\_transakcij(a_R, a_P))$

$\omega_{R,P}^+ \leftarrow ocena\_zaupanja\_QAD(a_R, a_P, operator(a_R), \mathcal{M})$

*else*

*if* ( $\omega_{R,P} \geq th$ ) *then*

$izid\_transakcije \leftarrow naredi\_transakcijo(a_R, a_P)$

$zgod\_transakcij(a_R, a_P).dodaj(izid\_transakcije)$

$\omega_{R,P}^- \leftarrow izracunaj\_neposredno\_oceno(zgod\_transakcij(a_R, a_P))$

$\omega_{R,P}^+ \leftarrow ocena\_zaupanja\_QAD(a_R, a_P, operator(a_R), \mathcal{M})$

*end if*

*end if*

---

Scenarij Y smo izvedli z različnimi porazdelitvami agentov s pripisanimi različnimi operatorji QAD:

- $Y_{mix}$ : 16,7 % agentov z  $\uparrow$ , 16,7 % agentov z  $\downarrow$ , 16,7 % agentov z  $\uparrow$ , 16,7 % agentov z  $\downarrow$ , 16,6 % agentov z  $\rightsquigarrow$ , 16,6 % agentov z  $\leftrightarrow$ ;
- $Y_{ex-o}$ : 100 % agentov z  $\uparrow$ ;
- $Y_{ex-p}$ : 100 % agentov z  $\downarrow$ ;
- $Y_{mod-o}$ : 100 % agentov z  $\uparrow$ ;
- $Y_{mod-p}$ : 100 % agentov z  $\downarrow$ ;
- $Y_{cent}$ : 100 % agentov z  $\rightsquigarrow$ ;
- $Y_{oppo}$ : 100 % agentov z  $\leftrightarrow$ .

Tabela 3.1

Količina sprejetih storitev z nezadovoljivo kakovostjo v scenarijih Y in Z.

SCENARIJ	POVPREČJE	ST. ODKLON	MIN	MAKS
Z	23,69 %	0,62 %	22,46 %	25,71 %
$Y_{mix}$	16,50 %	0,62 %	14,69 %	18,10 %
$Y_{ex-o}$	24,04 %	0,60 %	22,94 %	25,38 %
$Y_{ex-p}$	7,16 %	0,30 %	6,58 %	7,76 %
$Y_{mod-o}$	16,92 %	0,32 %	16,20 %	17,57 %
$Y_{mod-p}$	8,21 %	0,20 %	7,78 %	8,65 %
$Y_{cent}$	7,53 %	0,35 %	6,80 %	8,33 %
$Y_{oppo}$	24,09 %	0,66 %	22,86 %	25,70 %

V scenariju Z se agenti ne odločajo na podlagi zaupanja, zato nimajo pripisanih operatorjev zaupanja QAD.

Izvedli smo 5000 simulacijskih korakov. V tem času vsak agent naredi v povprečju vsaj pet transakcij z vsakim drugim agentom,<sup>1</sup> s čimer presežemo zgodovino pomnjenja izida transakcij med agentoma. Vsak tek simulacije smo ponovili 50-krat, pri čemer smo vsakič uporabili drugo vrednost semena za generator psevdonaključnih števil.

### Rezultati in analiza simulacij

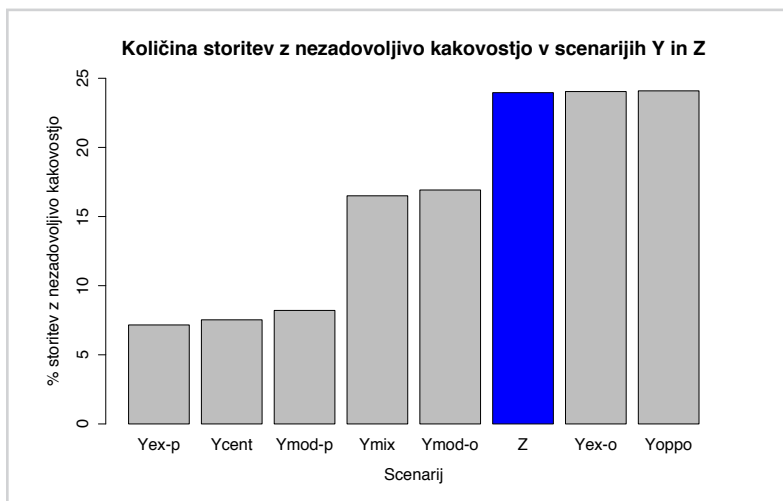
Vsako opisano konfiguracijo smo ponovili 50-krat in spremljali kakovost izbranih storitev v vsakem teku simulacije. Po koncu vsake ponovitve simulacij smo prešteli, koliko storitev z nezadovoljivo kakovostjo je bilo sprejetih s strani iskalcev storitev. Na podlagi rezultatov 50 ponovitev posamezne simulacijske konfiguracije smo izračunali, koliko storitev z nezadovoljivo kakovostjo je bilo v povprečju sprejetih v posameznem scenariju. Rezultati so prikazani v tabeli 3.1 in vključujejo tudi ostale statistične podatke: standardni odklon, minimalno vrednost in maksimalno vrednost. Grafičen prikaz rezultatov je podan na sliki 3.2.

V scenarijih Y smo za potrebe analize lastnosti modela zaupanja QAD vpeljali preprosto odločitvenimi model odločanja na podlagi zaupanja med agenti: iskalec storitve A sprejme storitev ponudnika storitev B, če je odnos agenta A do agenta B vsaj delno

<sup>1</sup>V simulaciji s 30 agenti vsak agent z vsakim drugim naredi *povprečno* 5 transakcij v  $30 \cdot 29 \cdot 5 = 4350$  korakih.

Slika 3.2

Primerjava količine sprejetih storitev z nezadovoljivo kakovostjo s strani agentov iskalcev storitev v scenarijih Y in Z.



zaupanja vreden, tj.  $\omega_{A,B} \geq 1$ . Na podlagi tega preprostega odločitvenega modela smo glede na "dovzetnost" operatorjev QAD za sprejem storitev z nezadovoljivo kakovostjo primerjali lastnosti operatorjev.

Če predpostavljamo, da so v skupnosti agenti z različnimi operatorji, kot je modelirano v scenariju  $Y_{mix}$ , potem že s preprostim odločitvenim modelom na podlagi zaupanja zmanjšamo količino sprejetih storitev z nezadovoljivo kakovostjo v primerjavi z naključnim odločanjem.

V splošnem je količina sprejetih storitev z nezadovoljivo kakovostjo manjša v skupnostih, kjer je malo agentov z ekstremno optimističnim ( $\uparrow$ ) in zmerno optimističnim ( $\uparrow$ ) operatorjem zaupanja QAD. Iz tega lahko izpeljemo, da operatorja  $\uparrow$  in  $\uparrow$  modelirata zaupanje (kot pove že ime) na optimističen način. Kljub slabim lastnim izkušnjam in slabim izkušnjam, ki so jih posredovali drugi agenti, so agenti z  $\uparrow$  in  $\uparrow$  nagnjeni k zaupanju in s tem dovzetni za sprejem slabih storitev.

Največ storitev z nezadovoljivo kakovostjo je bilo sprejetih v scenariju  $Y_{oppo}$ , kjer imajo vsi agenti v skupnosti pripisan oportunistični operator ( $\leftrightarrow$ ) za vrednotenje zaupanja. V splošnem obnašanje agentov z operatorjem  $\leftrightarrow$  ni tako naivno, kot izhaja iz rezultatov scenarija  $Y_{oppo}$ , kjer so v simuliranem okolju prisotni samo agenti z operatorjem  $\leftrightarrow$ . Za primer predpostavljamo, da so v skupnosti prisotni agenti z različnimi

operatorji, vključujoč agente s pripisanim operatorjem  $\leftrightarrow$ . Če bi bila večina ostalih agentov v skupnosti nagnjenih k ocenjevanju vrednosti zaupanja z oceno (delno) zaupanja nevreden odnos (to so tipično agenti s pripisanim operatorjem  $\Downarrow$  ali  $\Downarrow$ ), potem bi agenti z operatorjem  $\leftrightarrow$ , ki vrednotijo zaupanje tako, da povprečno oceno drugih agentov iz skupnosti zaokrožijo na vrednost, ki je bolj stran od vrednosti 0, izračunane vrednosti zaokrožili proti ocenama zaupanja nevreden odnos in delno zaupanja nevreden odnos (saj bi v tem primeru take ocene prevladovale). V opisanem primeru bi pri odločanju na podlagi zaupanja sprejeli manj storitev z nezadovoljivo kakovostjo, kot so jih v prikazanem scenariju  $Y_{oppo}$ .

Najmanj storitev z nezadovoljivo kakovostjo je bilo sprejetih v scenariju  $Y_{ex-p}$ . Operator  $\Downarrow$  modelira zaupanje agenta  $A$  tako, da agent  $A$  zavrne vse prihodnje transakcije s ponudnikom storitve  $B$ , takoj ko je zaupanje do agenta  $B$  nekdo ocenil s slabo oceno. Kljub temu, da je bilo v scenariju  $Y_{ex-p}$  sprejetih najmanj storitev z nezadovoljivo kakovostjo, je treba poudariti, da je bilo veliko ponudb zavrženih. Med njimi tudi take, ki bi imele za posledico storitev z zadovoljivo kakovostjo.<sup>2</sup>

Rezultati simulacij jasno nakazujejo, kako je modeliranje zaupanja v modelu QAD odvisno tudi od družbenih dejavnikov in vplivov okolja. Naj še enkrat poudarimo, da model zaupanja QAD matematično formulira proces vrednotenja zaupanja med agenti v e-okolju in *ne* predstavlja odločitvenega modela [57]. Preprost odločitveni model smo v simulacije vpeljali z namenom, da bolje ponazorimo delovanje in lastnosti operatorjev zaupanja QAD. Modeli zaupanja predstavljajo izhodišče za razvoj sistemov za obvladovanje zaupanja in ugleda, ti pa lahko pomenijo vhod odločitvenim modelom, ki (delno) temeljijo na zaupanju. Z analizo smo predstavili splošne lastnosti formalizacije zaupanja v modelu QAD. Hkrati pa podana analiza že predstavlja smernice za razvoj sistemov za obvladovanje zaupanja in ugleda ter različnih odločitvenih modelov, s katerimi lahko izboljšamo proces odločanja in posledično zmanjšamo količino sprejetih storitev z nezadovoljivo kakovostjo za vsakega agenta posamično.

### 3.4.3 Poglobljena analiza lastnosti operatorjev QAD

Namen drugega niza eksperimentov je poglobljena analiza lastnosti operatorjev QAD in zajem obnašanja agentov, ki ga operatorji določajo. Primerjali bomo, kakšen odnos zaupanja do slabih agentov imajo agenti z različnimi operatorji. Poleg tega bomo analizirali odločitve, ki jo jih sprejeli agenti z različnimi operatorji QAD in s tem analizirali

<sup>2</sup>Natančnejša analiza lastnosti operatorjev QAD je opisana v poglavju 3.4.3, kjer je opisan tudi ta aspekt.

modelirano dovezetnost k zaupanju definiranih operatorjev.

### *Parametri simulacije*

Za analizo lastnosti operatorjev QAD smo uporabili predlagani scenarij Y, ki je opisan v poglavju 3.4.1 (algoritem 2). Simulirali smo okolje s skupno  $n = 12$  agenti. Name-noma bomo predstavili rezultate simulacij z majhnim številom prisotnih agentov, saj sta interpretacija rezultatov in opis lastnosti operatorjev QAD s tem bolj nazorna. Ena-ke lastnosti veljajo tudi za simulacijska okolja z večjim številom agentov, vendar je za ponazoritev lastnosti potrebnih več simulacijskih korakov in več izvedenih interakcij, s čimer se jasnost in preglednost rezultatov bistveno zmanjšata. V simulacijskem okolju je eden izmed prisotnih agentov slab agent, ki v 80 % ponudi storitev z nezadovoljivo kakovostjo, tj.  $r_{mal} = 0,8$ . Slab agent je naključno določen na začetku simulacije in v času simulacije ostane slab. Ostali agenti so t. i. dobri agenti in vedno ponudijo storitev z zadovoljivo kakovostjo. Iskalec storitve  $A$  sprejme storitev naključno izbra-nega ponudnika  $B$ , če je stopnja zaupanja med njima vsaj delno zaupanja vreden odnos ( $\omega_{A,B} \geq 1$ ) oz. če gre za prvo interakcijo med njima. Agenti imajo naključno pripisan operator, pri čemer uporabimo porazdelitev operatorjev, kot jo definira scenarij  $Y_{mix}$ : 16,7 % agentov z  $\uparrow$ , 16,7 % agentov z  $\Downarrow$ , 16,7 % agentov z  $\uparrow$ , 16,7 % agentov z  $\Downarrow$ , 16,6 % agentov z  $\rightsquigarrow$  in 16,6 % agentov z  $\leftrightarrow$ . Izvedli smo 2000 korakov simulacije. V tem času se vrednost odnosa zaupanja agenta  $X$  do slabega agenta  $U$  stabilizira. Iz gra-fov, ki so podani v naslednjih poglavjih, je razvidno, da v tem času vrednost zaupanja preide na določeno vrednost, ki je neodvisna od na novo prihajajočih vplivov iz okolja. Vsak tek simulacije smo ponovili 50-krat ter pri tem vsakič uporabili drugo vrednost semena za generator psevdonaključnih števil. V 50 ponovitvah simulacij so se situacije vrednotenja zaupanja s posameznim operatorjem QAD stabilizirale in tako smo zajeli tipično obnašanje agenta, modelirano s specifičnim operatorjem.

### *Rezultati*

Na podlagi rezultatov simulacij smo analizirali obnašanje agenta, kot ga določa pripisa-ni operator zaupanja QAD. V vsaki simulaciji smo se osredotočili na izbranega agenta  $X$ , ki mu je pripisan določen operator, in analizirali odnos agenta  $X$  do slabega agenta  $U$ .

Rezultati analize so dodatno opremljeni z grafi. Rdeči simboli na grafu  $\bullet$ ,  $*$  ali  $\nabla$  označujejo transakcije, ki jih je AGENT  $X$  naredil s slabim agentom  $U$ . Sivi simboli  $\circ$ ,

\* ali  $\nabla$  označujejo transakcije, ki so jih s slabim agentom  $U$  naredili DRUGI AGENTI v skupnosti. Simbol  $\bullet$  označuje, da je agent  $U$  ponudil storitev z nezadovoljivo kakovostjo, simbol  $*$  označuje, da je agent  $U$  ponudil storitev z zadovoljivo kakovostjo in simbol  $\nabla$  označuje, da je iskalec storitve zavrnil transakcijo z agentom  $U$ . Simboli  $\bullet$ ,  $*$  in  $\nabla$  imajo različne vrednosti na ordinatni osi, ki odražajo neposredno oceno zaupanja agenta  $X$  do agenta  $U$  pred agregacijo z ocenami zaupanja ostalih agentov do agenta  $U$ , kot je definirano v enačbi 3.1. Če povzamemo: odražajo vrednost  $\omega_{X,U}^-$ . Modra črta na grafu predstavlja končno oceno zaupanja agenta  $X$  do agenta  $U$  po postopku vrednotenja zaupanja, ki ga definira model zaupanja in ugleda QAD. Torej, modra črta odraža končno oceno zaupanja  $X$  do  $U$  po združitvi ocen zaupanja iz podvektorja ugleda na način, ki ga definira operator QAD:  $\omega_{X,U}^+$ .

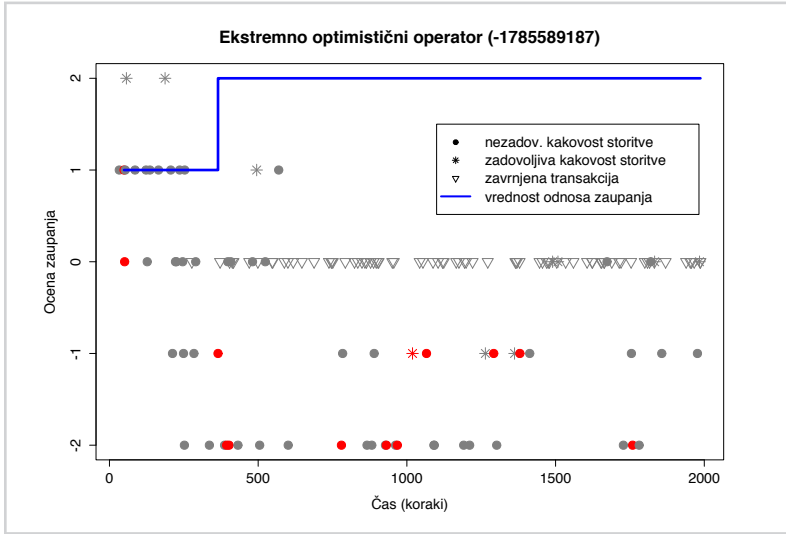
### Ekstremno optimistični operator ( $\Uparrow$ )

Obnašanje vrednosti odnosa zaupanja, ki je modeliran z ekstremnim optimističnim operatorjem ( $\Uparrow$ ), bomo analizirali na primeru -1785589187<sup>3</sup>, ki je prikazan na sliki 3.3. Operator  $\Uparrow$  modelira zelo optimistično obnašanje, kar pomeni, da agent  $A$  "verjame", da je agent  $B$  (delno) zaupanja vreden agent, takoj ko vsaj en agent iz skupnosti oceni odnos do agenta  $B$  kot (delno) zaupanja vreden odnos.

Na začetku so vsi odnosi med agenti nedefinirani. Po prvi transakciji med  $X$  in  $U$ , ki se v prikazanem primeru s slike 3.3 zgodi v 49. koraku simulacije, agent  $X$  oceni zaupanje do agenta  $U$  na podlagi neposredne izkušnje. Ker je agent  $U$  ponudil storitev z nezadovoljivo kakovostjo, je vrednost  $|H_{X,U}(\bullet)| = 1$ . Glede na vrednost  $|H_{X,U}(\bullet)|$  je neposredna ocena zaupanja enaka  $\omega_{X,U}^- = 1$ , kar je na grafu prikazano z rdečim simbolom  $\bullet$ . Agent  $X$  končno vrednost zaupanja do agenta  $U$  izračuna z obdelavo ocen zaupanja do agenta  $U$ , ki so jih podali ostali agenti iz skupnosti. Glede na predpis za operator  $\Uparrow$  (definicija 5) je končna vrednost zaupanja največja vrednost zaupanja izmed vseh ocen zaupanja do agenta  $U$ . Agent  $U$  je do 49. koraka ocenil en agent (agent  $Y$ ), in sicer z vrednostjo delno zaupanja vreden odnos, tj.  $\omega_{Y,U} = 1$ . Iz tega sledi, da je končna ocena zaupanja agenta  $X$  do agenta  $U$  delno zaupanja vreden odnos, tj.  $\omega_{X,U}^+ = 1$ , kar je prikazano z modro črto na grafu.

Sprememba vrednosti ocene zaupanja agenta  $X$  do agenta  $U$  se odraža v matriki za-

<sup>3</sup>To je vrednost semena za generator psevdonaključnih števil, ki je bila uporabljena pri danem teku simulacije. Vrednost semena za vsak tek simulacije je podana v naslovu grafa, ki prikazuje rezultate za dani tek simulacije.



Slika 3.3

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem  $\uparrow$ .

upanja  $\mathcal{M}$ . V nadaljevanju so izpisane vrednosti elementov matrike zaupanja v koraku št. 48 (pred transakcijo med  $X$  in  $U$ ) in v 49. koraku (po transakciji). Rdeče osenčen stolpec v matriki vsebuje ocene zaupanja do agenta  $U$ . Sivo osenčena vrstica matrike vsebuje ocene zaupanja agenta  $X$  do ostalih agentov v skupnosti.

$$\mathcal{M}_{48} = \begin{array}{l}
 (\Downarrow) \\
 (\rightsquigarrow) \\
 (\leftrightarrow) \\
 (\Uparrow) \\
 (\leftrightarrow) \\
 (\Uparrow) \\
 (\Downarrow) \\
 (\Downarrow) \\
 (\Uparrow) \\
 (\Downarrow) \\
 (\Uparrow) \\
 (\rightsquigarrow)
 \end{array} \left[ \begin{array}{cccccccccccc}
 / & 2 & / & / & / & / & 2 & 2 & 2 & / & / & / \\
 2 & / & / & 1 & 2 & / & / & / & 2 & 2 & / & / \\
 / & / & / & / & / & / & / & / & 2 & 2 & / & 2 & / \\
 / & 2 & / & / & / & / & / & / & / & 2 & / & / & / \\
 / & / & 2 & / & / & 2 & 2 & 2 & / & 2 & / & / & / \\
 / & / & 2 & / & / & / & / & / & / & 2 & / & / & / \\
 2 & 2 & / & / & / & 2 & / & / & / & / & / & / & 2 \\
 / & 2 & / & / & 2 & / & / & / & 2 & / & / & / & / \\
 2 & / & / & / & / & / & 2 & 2 & / & / & / & / & / \\
 / & / & / & / & 2 & / & / & 2 & / & / & 2 & / & / \\
 2 & / & 2 & / & / & / & / & / & / & / & / & / & / \\
 / & / & 2 & / & / & 2 & / & 2 & / & / & / & / & /
 \end{array} \right]_{t=48}$$

$\xrightarrow[49. \text{ korak}]{\omega_{X,U} = 1}$

$$\mathcal{M}_{49} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{matrix} \begin{bmatrix} / & 2 & / & / & / & / & 2 & 2 & 2 & / & / & / \\ 2 & / & / & 1 & 2 & / & / & / & 2 & 2 & / & / \\ / & / & / & / & / & / & / & 2 & 2 & / & 2 & / \\ / & 2 & / & / & / & / & / & / & / & 2 & / & / \\ / & / & 2 & / & / & 2 & 2 & 2 & / & 2 & / & / \\ / & / & 2 & / & / & / & / & / & / & 2 & / & / \\ 2 & / & / & 1 & / & / & 2 & 2 & / & / & / & / \\ / & / & / & / & 2 & / & / & 2 & / & / & 2 & / \\ 2 & / & 2 & / & / & / & / & / & / & / & / & / \\ / & / & 2 & / & / & 2 & / & 2 & / & / & / & / \end{bmatrix}_{t=49}$$

Pred in po drugi transakciji med agentoma  $X$  in  $U$ , ki se se zgodi v 51. koraku teka simulacije, so vrednosti v matriki zaupanja naslednje:

$$\mathcal{M}_{50} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{matrix} \begin{bmatrix} / & 2 & / & / & / & / & 2 & 2 & 2 & / & / & / \\ 2 & / & / & 1 & 2 & / & / & / & 2 & 2 & / & / \\ / & / & / & / & / & / & / & 2 & 2 & / & 2 & / \\ / & 2 & / & / & / & / & / & / & / & 2 & / & / \\ / & / & 2 & / & / & 2 & 2 & 2 & / & 2 & / & / \\ / & / & 2 & / & / & / & / & / & / & 2 & / & / \\ 2 & / & / & 1 & / & / & 2 & 2 & / & / & / & / \\ / & / & / & / & 2 & / & / & 2 & / & / & 2 & / \\ 2 & / & 2 & / & / & / & / & / & / & / & / & / \\ / & / & 2 & / & / & 2 & / & 2 & 2 & / & / & / \end{bmatrix}_{t=50} \xrightarrow[51. \text{ korak}]{\omega_{X,U}^- = 0}$$

$$\mathcal{M}_{51} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\rightsquigarrow) \end{matrix} \begin{bmatrix} / & 2 & / & / & / & / & 2 & 2 & 2 & / & / & / \\ 2 & / & / & 1 & 2 & / & / & / & 2 & 2 & / & / \\ / & / & / & / & / & / & / & / & 2 & 2 & / & 2 & / \\ / & 2 & / & / & / & / & / & / & / & 2 & / & / & / \\ / & / & 2 & / & / & 2 & 2 & 2 & / & 2 & / & / & / \\ / & / & 2 & / & / & / & / & / & / & 2 & / & / & / \\ 2 & 2 & / & / & / & 2 & / & / & / & / & / & / & 2 \\ / & 2 & / & / & 2 & / & / & / & / & 2 & / & / & / \\ 2 & / & / & 1 & / & / & 2 & 2 & / & / & / & / & / \\ / & / & / & / & 2 & / & / & / & 2 & / & / & 2 & / \\ 2 & / & 2 & / & / & / & / & / & / & / & / & / & / \\ / & / & 2 & / & / & 2 & / & 2 & 2 & / & / & / & / \end{bmatrix}_{t=51}$$

Po drugi transakciji med agentoma  $X$  in  $U$  (v 51. koraku teka simulacije, glej sliko 3.3) je neposredna ocena zaupanja agenta  $X$  do agenta  $U$  enaka  $\omega_{X,U}^- = 0$ , ker je  $U$  ponovno ponudil storitev z nezadovoljivo kakovostjo in je  $|H_{X,U}(\bullet)| = 2$ . V tem času drugi agenti niso imeli transakcije z agentom  $U$  in niso ocenili zaupanja do agenta  $U$ . Najvišja ocena zaupanja do agenta  $U$  je v tem času delno zaupanja vreden odnos, zato je končna ocena zaupanja agenta  $X$  do agenta  $U$  po 51. koraku delno zaupanja vreden odnos, tj.  $\omega_{X,U}^+ = 1$ .

Do tretje transakcije agenta  $X$  s slabim agentom  $U$  pride v koraku št. 365. Agent  $U$  agentu  $X$  zopet ponudi storitev z nezadovoljivo kakovostjo, tj.  $|H_{X,U}(\bullet)| = 3$ , zato je neposredna ocena zaupanja agenta  $X$  do agenta  $U$  enaka  $\omega_{X,U}^- = -1$ . Ker so v tem času drugi agenti v skupnosti ocenili agenta  $U$  z oceno *zaupanja vreden odnos* (bodisi na podlagi pozitivnih izkušenj z  $U$  bodisi na podlagi pozitivno usmerjenega operatorja zaupanja, npr.  $\Uparrow$  ali  $\Uparrow$ ), posledično tudi agent  $X$  "verjame", da je agent  $U$  zaupanja vreden. Vrednost zaupanja agenta  $X$  do agenta  $U$ , modelirana z operatorjem  $\Uparrow$ , je tako enaka 2 (tj. zaupanja vreden odnos), kar je na grafu na sliki 3.3 ponazorjeno z modro črto. Stanje matrice zaupanja v simulacijskih korakih št. 364 in 365 je, kot sledi:

$$\mathcal{M}_{364} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & / & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & / & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & / & 2 & 2 & 2 & / & 2 & / & 2 \\ 2 & 2 & 2 & 2 & / & / & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ \hline 2 & / & 2 & 1 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & / & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & / & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=364} \xrightarrow[\omega_{X,U} = -1]{365. \text{ korak}}$$

$$\mathcal{M}_{365} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & / & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & / & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & / & 2 & 2 & 2 & / & 2 & / & 2 \\ 2 & 2 & 2 & 2 & / & / & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ \hline 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & / & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & / & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=365}$$

Ocena zaupanja  $\omega_{X,U}$  do konca simulacije ostane nad določenim pragom za sprejem storitev ( $th = 1$ ), zato  $X$  še naprej sprejema storitve, ki jih ponudi agent  $U$ . Posledično agent  $X$  sprejme veliko storitev z nezadovoljivo kakovostjo, kar je tudi razvidno iz grafa na sliki 3.3.

Na sliki 3.3 so prav tako jasno prikazane lastnosti neposredne ocene zaupanja, ki temelji na zgodovini osebnih izkušenj s ponudnikom storitev  $U$ . V koraku št. 1019 agent  $X$  sprejme transakcijo z  $U$ , ki mu dostavi storitev z zadovoljivo kakovostjo, kar je na grafu označeno z rdečim simbolom  $*$ . Agent  $X$  na podlagi ene dobre in treh slabih izkušenj z  $U$  (upoštevajoč zgodovino zadnjih štirih izidov transakcij) ovrednoti neposredno oceno zaupanja z  $\omega_{X,U}^- = -1$ , saj je  $|H_{X,U}(\bullet)| = 3$ . V naslednjih treh transakcijah  $U$  ponudi storitve z nezadovoljivo kakovostjo. Neposredna ocena zaupanja, ki temelji na zgodovini izkušenj z agentom ponudnikom storitev, upošteva zadnje štiri izkušnje. Ker je imel  $X$  eno pozitivno izkušnjo in tri negativne izkušnje z agentom  $U$  in velja  $|H_{X,U}(\bullet)| = 3$ , neposredna ocena zaupanja  $X$  do  $U$  ostane ocenjena z vrednostjo  $-1$ . Ko agent  $X$  naredi naslednjo transakcijo z agentom  $U$  v koraku 1759 in agent  $U$  zopet ponudi storitev z nezadovoljivo kakovostjo, spomin na storitev z zadovoljivo kakovostjo zbledi in ne vpliva več na vrednost neposredne ocene zaupanja. Neposredna ocena zaupanja  $\omega_{X,U}^-$  se zmanjša na vrednost  $-2$ , tj. zaupanja nevreden odnos, kar je na grafu označeno z rdečim simbolom  $\bullet$ . Končna ocena zaupanja – združitve ocen iz podvektorja ugleda agenta  $U$  po pravilih, ki jih določa operator agenta  $X$ , ves ta čas zaradi zunanjih dejavnikov (tj. vsaj ene pozitivne ocene zaupanja drugih agentov do  $U$ ) ostaja nespremenjena in ima vrednost  $\omega_{X,U}^+ = 2$ , ki je na grafu prikazana z modro črto.

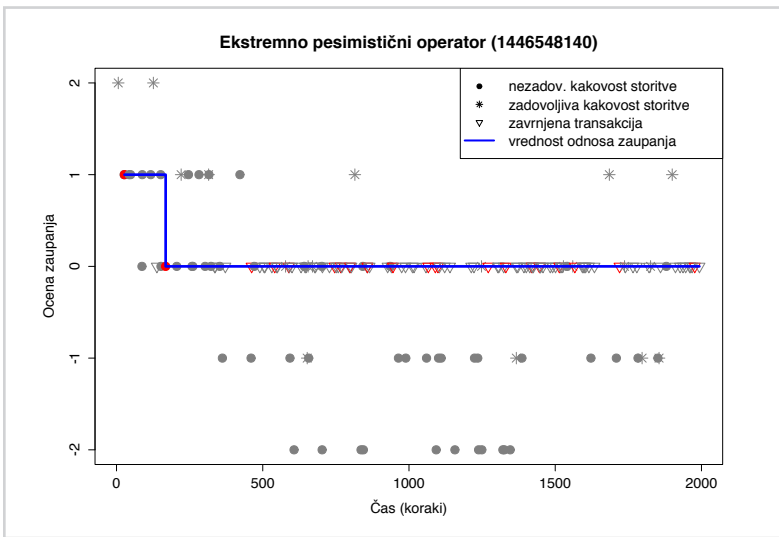
Vrednost matrike zaupanja po 2000 korakih je naslednja:

$$\mathcal{M}_{2000} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{matrix} \begin{bmatrix} / & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & / & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & / & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 \end{bmatrix} \quad t = 2000$$

Grafična analiza obnašanja vrednosti odnosa zaupanja, modeliranega z ekstremno optimističnim operatorjem, je za vseh 50 ponovitev simulacije prikazana na slikah A.1, A.2, A.3, A.4, A.5 v dodatku A.

### Ekstremno pesimistični operator ( $\Downarrow$ )

Operator  $\Downarrow$  modelira zaupanje agentov ravno nasprotno, kot je modeliran odnos zaupanja z operatorjem  $\Uparrow$ . Agent s pripisanim operatorjem  $\Downarrow$  je izjemno dovzeten za zaupanja nevrudne odnose drugih agentov v skupnosti. Pri vrednotenju zaupanja do določenega agenta upošteva najbolj negativno oceno zaupanja, ki so jo za ocenjevanega agenta podali drugi agenti v skupnosti.



Slika 3.4

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem  $\Downarrow$ .

Na sliki 3.4 je prikazan primer obnašanja vrednosti odnosa zaupanja, ki ga modelira ekstremno pesimistični operator  $\Downarrow$ .

Agent  $X$  oceni odnos do  $U$  kot neodločen odnos ( $\omega_{X,U}^+ = 0$ ), takoj ko tako nizko oceno do agenta  $U$  poda katerikoli agent v skupnosti.<sup>4</sup> V analiziranem primeru s

<sup>4</sup>V opisanem scenariju je meja sprejem storitev  $th = 1$ . Če je ocena zaupanja med agentom  $X$  in  $U$  nižja od te vrednosti, potem agent  $X$  ne naredi transakcije z agentom  $U$  ter posledično ne oceni znova vrednosti zaupanja do agenta  $U$ . Končna ocena zaupanja v opisanih scenarijih zato ne pade pod vrednost  $\omega_{X,U} = 0$ .

slike 3.4 se prva transakcija med agentoma  $X$  in  $U$  zgodi v 27. koraku. Vrednosti v matriki zaupanja v 26. in 27. koraku so naslednje:

$$\mathcal{M}_{26} = \begin{array}{l} (\Downarrow) \\ (\sim) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\sim) \end{array} \begin{array}{c} \left[ \begin{array}{cccccccccccc} / & / & / & / & / & 2 & / & / & 2 & / & / & / \\ / & / & / & / & / & / & / & / & 2 & / & / & / \\ 2 & / & / & / & / & / & / & / & / & / & 2 & / \\ / & / & / & / & 2 & 2 & / & / & / & / & / & / \\ / & / & / & / & / & / & / & / & 2 & / & / & 2 \\ / & / & / & / & / & / & / & / & / & / & / & / \\ / & / & / & / & 2 & / & / & / & 2 & / & / & / \\ / & 2 & / & / & / & / & / & / & / & / & / & / \\ / & / & / & / & / & 2 & / & / & / & / & / & / \\ / & / & / & / & / & / & 2 & / & / & / & / & 2 \\ / & 2 & 2 & / & / & / & / & 2 & / & / & / & / \end{array} \right]_{t=26}$$

$$\mathcal{M}_{27} = \begin{array}{l} (\Downarrow) \\ (\sim) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\sim) \end{array} \begin{array}{c} \left[ \begin{array}{cccccccccccc} / & / & / & / & / & 2 & / & 1 & 2 & / & / & / \\ / & / & / & / & / & / & / & / & / & 2 & / & / \\ 2 & / & / & / & / & / & / & / & / & / & / & 2 \\ / & / & / & / & 2 & 2 & / & / & / & / & / & / \\ / & / & / & / & / & / & / & / & / & 2 & / & 2 \\ / & / & / & / & / & / & / & / & / & / & / & / \\ / & / & / & / & 2 & / & / & 1 & / & / & / & / \\ / & / & / & 2 & / & 2 & / & / & 2 & / & / & / \\ / & 2 & / & / & / & / & / & / & / & / & / & / \\ / & / & / & / & / & 2 & / & / & / & / & / & / \\ / & / & / & / & / & / & 2 & / & / & / & / & 2 \\ / & 2 & 2 & / & / & / & / & 2 & / & / & / & / \end{array} \right]_{t=27}$$

V prvi transakciji agent  $U$  agentu  $X$  ponudi storitev z nezadovoljivo kakovostjo. Vrednost neposredne ocene zaupanja je  $\omega_{X,U}^- = 1$ , saj gre za prvo storitev z nezadovoljivo kakovostjo s strani agenta  $U$  do agenta  $X$ , tj.  $|H_{X,U}(\bullet)| = 1$ . V 27. koraku

simulacije je najslabša izmed ocen zaupanja, s katerimi so ostali agenti ocenili agenta  $U$ , ocena 1 (delno zaupanja vreden odnos). Končna vrednost zaupanja, izračunana z operatorjem  $\uparrow$ , je zato  $\omega_{X,U}^+ = 1$ .

Druga transakcija med agentoma  $X$  in  $U$  se zgodi v koraku št. 168. Vrednosti v matriki zaupanja pred in po opravljeni prvi transakciji med agentoma  $X$  in  $U$  so naslednje:

$$\mathcal{M}_{167} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \begin{array}{c} \left[ \begin{array}{cccccccccccc} / & / & 2 & / & 2 & 2 & 2 & 1 & 2 & / & 2 & 2 \\ 2 & / & / & / & / & 2 & / & 1 & 2 & / & / & / \\ 2 & / & / & / & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ / & / & / & / & 2 & 2 & / & / & 2 & 2 & / & 2 \\ / & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & / \\ / & / & 2 & / & 2 & 2 & / & 0 & 2 & 2 & / & / \\ 2 & 2 & / & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / \\ 2 & 2 & 2 & 2 & 2 & 2 & / & / & / & 2 & 2 & 2 \\ / & / & 2 & 2 & 2 & 2 & 2 & 1 & / & / & 2 & 2 \\ / & 2 & 2 & 2 & 2 & / & 2 & / & / & / & / & 2 \\ / & 2 & 2 & / & 2 & 2 & 2 & 1 & 2 & 2 & 2 & / \end{array} \right]_{t=167}

$\xrightarrow[168. \text{ korak}]{\omega_{X,U}^- = 0}$$$

$$\mathcal{M}_{168} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \begin{array}{c} \left[ \begin{array}{cccccccccccc} / & / & 2 & / & 2 & 2 & 2 & 0 & 2 & / & 2 & 2 \\ 2 & / & / & / & / & 2 & / & 1 & 2 & / & / & / \\ 2 & / & / & / & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ / & / & / & / & 2 & 2 & / & / & 2 & 2 & / & 2 \\ / & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & / \\ / & / & 2 & / & 2 & 2 & / & 0 & 2 & 2 & / & / \\ 2 & 2 & / & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / \\ 2 & 2 & 2 & 2 & 2 & 2 & / & / & / & 2 & 2 & 2 \\ / & / & 2 & 2 & 2 & 2 & 2 & 1 & / & / & 2 & 2 \\ / & 2 & 2 & 2 & 2 & / & 2 & / & / & / & / & 2 \\ / & 2 & 2 & / & 2 & 2 & 2 & 1 & 2 & 2 & 2 & / \end{array} \right]_{t=168}$$

Zaradi druge storitve z nezadovoljivo kakovostjo s strani agenta  $U$  se neposredna ocena zaupanja agenta  $X$  do agenta  $U$  zmanjša na vrednost neodločen odnos ( $\omega_{X,U}^- = 0$ ). Končna ocena zaupanja, izračunana z operatorjem  $\uparrow$ , je  $\omega_{X,U}^+ = 0$ , kot je tudi razvidno iz vrednosti matrice zaupanja pred in po drugi transakciji med agentoma  $X$  in  $U$ .

Glede na frekvenco zlonamernega obnašanja agenta  $U$  imata tipično prvi dve storitvi, ki ju bo agent  $X$  sprejel od agenta  $U$ , nezadovoljivo kakovost, zaradi česar agent  $X$  oceni agenta  $U$  z oceno  $\omega_{X,U} = 0$ . Če do prve interakcije med agentoma  $X$  in  $U$  pride kasneje, je pred tem zaradi slabih storitev agenta  $U$  z nizko oceno zaupanja ocenil kateri od ostalih agentov v simulirani e-skupnosti, zaradi česar agent  $X$  prav tako oceni agenta  $U$  z oceno  $\omega_{X,U} = 0$ . Oba dejavnika vplivata na to, da vrednost zaupanja  $X$  do  $U$  pade pod določeno vrednost ( $th = 1$ ) za sprejem storitev od ponudnika  $U$ . Agent  $X$  posledično zavrne vse nadaljnje transakcije z agentom  $U$ . Zavrjene transakcije so na grafu označene s simbolom  $\nabla$ .

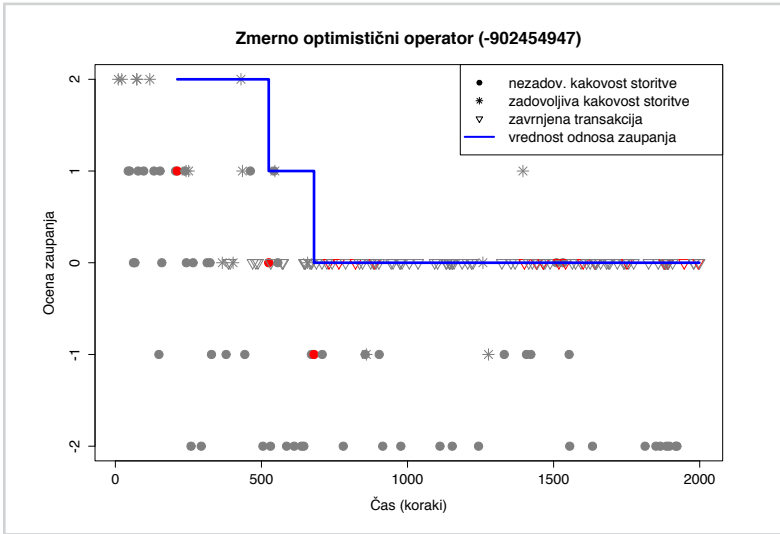
Obnašanje vrednosti odnosa zaupanja, kot ga definira operator  $\Downarrow$ , je za vseh 50 povnitev simulacije prikazano na slikah [A.6](#), [A.7](#), [A.8](#), [A.9](#), [A.10](#) v dodatku [A](#).

#### *Zmerno optimistični operator ( $\uparrow$ )*

V splošnem je obnašanje agenta  $A$  z operatorjem  $\uparrow$  modelirano tako, da na vrednost odnosa zaupanja vplivajo pozitivne ocene zaupanja iz okolja. Agent  $A$  za eno stopnjo zviša oceno zaupanja do agenta  $B$ , če imajo ostali agenti v skupnosti v povprečju višjo oceno zaupanja do  $B$ , kot jo ima  $A$ . Če imajo ostali agenti v okolju v povprečju nižjo oceno zaupanja do  $B$ , potem to ne vpliva na spremembo vrednosti odnosa zaupanja agenta  $A$  do agenta  $B$ .

Na sliki [3.5](#) je prikazano značilno obnašanje vrednosti odnosa zaupanja, modeliranega z operatorjem  $\uparrow$ .

Do prve transakcije med agentoma  $X$  in  $U$  pride v koraku št. 212. Neposredna ocena zaupanja ima vrednost  $\omega_{X,U}^- = 1$ . V tem času je povprečna ocena zaupanja ostalih agentov v skupnosti do agenta  $U$  več kot 1, zato ocene iz okolja pozitivno vplivajo na končno oceno zaupanja. Končna ocena zaupanja je glede na definicijo operatorja  $\uparrow$  za eno stopnjo višja in ima vrednost  $\omega_{X,U}^+ = 2$ . Vrednosti v matriki zaupanja v koraku št. 211 in koraku št. 212 so naslednje:



Slika 3.5

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem ↑.

$$\mathcal{M}_{211} = \begin{matrix}
 (\Downarrow) \\
 (\sim) \\
 (\leftrightarrow) \\
 (\Uparrow) \\
 (\leftrightarrow) \\
 (\Uparrow) \\
 (\Downarrow) \\
 (\Downarrow) \\
 (\Uparrow) \\
 (\Downarrow) \\
 (\Uparrow) \\
 (\sim)
 \end{matrix} \begin{bmatrix}
 / & 2 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\
 2 & / & 2 & 2 & 1 & 2 & / & 2 & 2 & 2 & 2 & 2 \\
 / & / & / & 2 & 1 & 2 & 2 & 2 & / & 2 & 2 & / \\
 2 & 2 & 2 & / & / & 2 & 2 & 2 & 2 & 2 & / & 2 \\
 2 & 2 & / & 2 & / & 2 & / & 2 & 2 & 2 & 2 & 2 \\
 2 & / & / & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & / \\
 / & 2 & 2 & 2 & 1 & 2 & / & 2 & / & 2 & 2 & 2 \\
 2 & / & 2 & 2 & 0 & 2 & 2 & / & 2 & / & 2 & 2 \\
 / & / & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & / \\
 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & / & 2 & 2 \\
 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & / & / \\
 2 & 2 & / & 2 & 1 & / & 2 & 2 & 2 & 2 & 2 & /
 \end{bmatrix} \xrightarrow[\omega_{X,U}^- = 1]{212. \text{ korak}}$$

$t = 211$

$$\mathcal{M}_{212} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{matrix} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & / & 2 & 2 & 1 & 2 & / & 2 & 2 & 2 & 2 & 2 \\ / & / & / & 2 & 1 & 2 & 2 & 2 & / & 2 & 2 & / \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 \\ 2 & 2 & / & 2 & / & 2 & / & 2 & 2 & 2 & 2 & 2 \\ 2 & / & / & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & / \\ / & 2 & 2 & 2 & 1 & 2 & / & 2 & / & 2 & 2 & 2 \\ 2 & / & 2 & 2 & 0 & 2 & 2 & / & 2 & / & 2 & 2 \\ / & / & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & / \\ 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & / & / \\ 2 & 2 & / & 2 & 1 & / & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=212}$$

Agent  $X$  po drugi opravljeni transakciji z agentom  $U$  oceni odnos do  $U$  kot neodločen, tj.  $\omega_{X,U}^- = 0$ . V tem času je povprečna ocena zaupanja ostalih agentov v skupnosti do agenta  $U$  več kot 0, zato ocene iz okolja tudi tokrat vplivajo na končno vrednost zaupanja. Vrednost te se zato (kljub ponovni slabi izkušnji) poveča za eno stopnjo in je enaka  $\omega_{X,U}^+ = 1$ , kar je na grafu nakazano z modro črto. Spremembe v matriki zaupanja pred in po drugi transakciji med agentoma  $X$  in  $U$ , ki se zgodi v 525. simulacijskem koraku, so, kot sledi:

$$\mathcal{M}_{524} = \begin{matrix} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{matrix} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & / & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & / & 2 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 1 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 0 & 2 & / & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 0 & 2 & 2 & / & 2 & / & 2 & 2 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 \\ 2 & 2 & 2 & 2 & 0 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / & 2 \\ 2 & 2 & 2 & 2 & 0 & / & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=524}$$

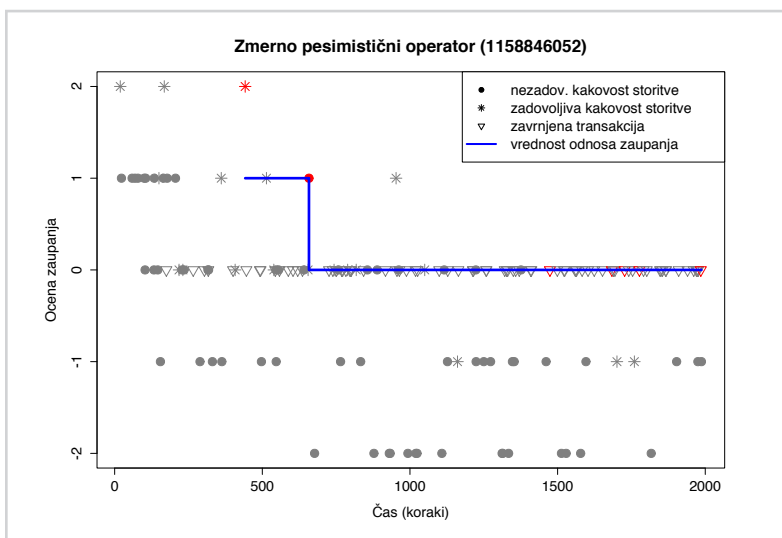
525. korak  
 $\xrightarrow{\omega_{X,U}^- = 0}$



Grafi, ki prikazujejo obnašanje vrednosti odnosa zaupanja, modeliranega z zmernim optimističnim operatorjem, so za vseh 50 ponovitev simulacije podani na slikah A.11, A.12, A.13, A.14, A.15 v dodatku A.

### Zmerno pesimistični operator ( $\Downarrow$ )

V nasprotju z operatorjem  $\Uparrow$  zmerno pesimistični operator  $\Downarrow$  modelira obnašanje agenta, na katerega vplivajo samo bolj negativne ocene zaupanja drugih agentov od njegove lastne. V splošnem agent  $A$  za eno stopnjo zniža oceno zaupanja do agenta  $B$ , če je povprečna ocena ostalih agentov iz skupnosti do agenta  $B$  nižja od *neposredne* ocene zaupanja agenta  $A$  do agenta  $B$ .



Slika 3.6

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem  $\Downarrow$ .

Na sliki 3.6 je prikazan primer obnašanja vrednosti odnosa zaupanja, ki je modeliran z zmernim pesimističnim operatorjem. Agent  $X$  z agentom  $U$  prvič naredi transakcijo v 442. koraku. Agent  $U$  ponudi storitev z zadovoljivo kakovostjo, zato je neposredna ocena zaupanja agenta  $X$  do agenta  $U$  ovrednotena z  $\omega_{X,U}^- = 2$ , kar je na grafu prikazano s simbolom \*. Ostali agenti v skupnosti so agenta  $U$  ocenili z vrednostmi, ki so v povprečju nižje od 2. Končna ocena zaupanja, kot jo definira operator  $\Downarrow$ , je torej za

eno stopnjo nižja ocena:  $\omega_{X,U}^+ = 1$ . Vrednosti v matriki zaupanja pred in po izvedeni transakciji med agentoma  $X$  in  $U$  so naslednje:

$$\mathcal{M}_{441} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc|c} / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & 0 \\ \hline 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & / \\ / & 2 & 2 & 2 & / & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=441}$$

$\xrightarrow[442. \text{ korak}]{\omega_{X,U}^- = 2}$

$$\mathcal{M}_{442} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc|c} / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & 0 \\ \hline 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 1 \\ / & 2 & 2 & 2 & / & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=442}$$

Neposredna ocena zaupanja agenta  $X$  do agenta  $U$  po drugi transakciji v koraku št. 658 je  $\omega_{X,U}^- = 1$ , saj je agent  $U$  ponudil storitev z nezadovoljivo kakovostjo. Ker je povprečna vrednost ocene zaupanja do agenta  $U$  manj kot 1, je končna vrednost

zaupanja enaka  $\omega_{X,U}^+ = 0$ , kot je tudi razvidno iz vrednosti matrik zaupanja v 657. in 658. simulacijskem koraku:

$$\mathcal{M}_{657} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc|c} / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & 0 \\ \hline 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=657}$$

$\xrightarrow{\text{658. korak}} \omega_{X,U}^- = 1$

$$\mathcal{M}_{658} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc|c} / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & / & 2 & 2 & 2 & 0 \\ \hline 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 0 \\ 2 & / & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & / \end{array} \right]_{t=658}$$

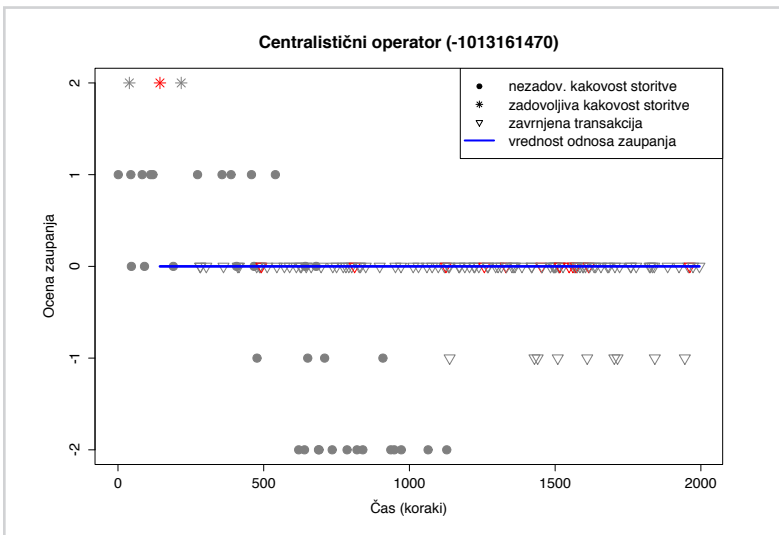
Agent  $X$  je zaradi pripisanega operatorja  $\Downarrow$  zelo dovzeten za pesimistične ocene zaupanja ostalih agentov v danem e-okolju, zato se ocena zaupanja agenta  $X$  do agenta  $U$

hitro niža in pade pod predpisano mejo za opravljanje transakcij. Posledično agent  $X$  zavrta transakcije z agentom  $U$ , kar je na grafih označeno s simbolom  $\nabla$ .

Na slikah A.16, A.17, A.18, A.19, A.20 v dodatku A so prikazani primeri obnašanja vrednosti zaupanja, kot ga definira operator  $\downarrow$ , za vseh 50 ponovitev simulacije.

### Centralistični operator ( $\rightsquigarrow$ )

Centralistični operator  $\rightsquigarrow$  modelira "nesamozavestno" obnašanje agenta. Agent  $A$  sledi sredinskemu mnenju in ocenam zaupanja, ki jih imajo ostali agenti v skupnosti do ocenjevanega agenta. Končna vrednost zaupanja je namreč povprečna<sup>5</sup> vrednost podanih ocen zaupanja, zaokrožena proti vrednosti 0.



Slika 3.7

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem  $\rightsquigarrow$ .

Značilno obnašanje vrednosti ocene zaupanja agenta  $X$  do slabega agenta  $U$  je prikazano na sliki 3.7. Prva transakcija med agentoma  $X$  in  $U$  se zgodi v koraku št. 144. Agent  $U$  ponudi storitev z zadovoljivo kakovostjo, zato je neposredna ocena zaupanja

<sup>5</sup>Zaupanje je v modelu QAD definirano s kvalitativnimi ocenami, ki simbolizirajo *zaupanja neveden* (-2), *delno zaupanja neveden* (-1), *neodločen* (0), *delno zaupanja vreden* (1) in *zaupanja vreden odnos* (2). Za potrebe izračuna povprečne vrednosti se ocene pretvorijo v pripadajoče numerične vrednosti.

$\omega_{X,U}^- = 2$ . Agenti  $U$  so v tem času ocenili tudi drugi agenti iz skupnosti, in sicer z ocenami, ki so razvidne iz matrike zaupanja:

$$\mathcal{M}_{143} = \begin{array}{l} (\Downarrow) \\ (\sim) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\sim) \end{array} \left[ \begin{array}{cccc|cccc|cccc} / & 2 & 2 & 0 & / & 2 & 2 & 2 & / & / & 2 & 2 \\ 2 & / & 2 & / & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ / & 2 & / & / & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ / & / & 2 & / & / & 2 & 2 & 2 & / & / & / & 2 \\ 2 & 2 & 2 & / & / & 2 & / & / & 2 & 2 & 2 & 2 \\ / & 2 & 2 & / & 2 & / & / & 2 & 2 & 2 & 2 & / \\ / & 2 & / & 0 & 2 & 2 & / & 2 & 2 & 2 & 2 & / \\ 2 & 2 & 2 & 0 & / & 2 & / & / & 2 & / & 2 & / \\ / & 2 & 2 & / & 2 & 2 & 2 & 2 & / & / & 2 & 2 \\ 2 & 2 & / & 0 & / & / & 2 & 2 & 2 & / & / & / \\ 2 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & / & 2 & / & 2 \\ 2 & 2 & 2 & / & 2 & / & 2 & / & 2 & / & / & / \end{array} \right] \xrightarrow[144. \text{ korak}]{\omega_{X,U}^- = 2} t = 143$$

$$\mathcal{M}_{144} = \begin{array}{l} (\Downarrow) \\ (\sim) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\sim) \end{array} \left[ \begin{array}{cccc|cccc|cccc} / & 2 & 2 & 0 & / & 2 & 2 & 2 & / & / & 2 & 2 \\ 2 & / & 2 & 0 & 2 & 2 & 2 & 2 & 2 & / & 2 & 2 \\ / & 2 & / & / & 2 & / & 2 & 2 & 2 & 2 & 2 & 2 \\ / & / & 2 & / & / & 2 & 2 & 2 & / & / & / & 2 \\ 2 & 2 & 2 & / & / & 2 & / & / & 2 & 2 & 2 & 2 \\ / & 2 & 2 & / & 2 & / & / & 2 & 2 & 2 & 2 & / \\ / & 2 & / & 0 & 2 & 2 & / & 2 & 2 & 2 & 2 & / \\ 2 & 2 & 2 & 0 & / & 2 & / & / & 2 & / & 2 & / \\ / & 2 & 2 & / & 2 & 2 & 2 & 2 & / & / & 2 & 2 \\ 2 & 2 & / & 0 & / & / & 2 & 2 & 2 & / & / & / \\ 2 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & / & 2 & / & 2 \\ 2 & 2 & 2 & / & 2 & / & 2 & / & 2 & / & / & / \end{array} \right] t = 144$$

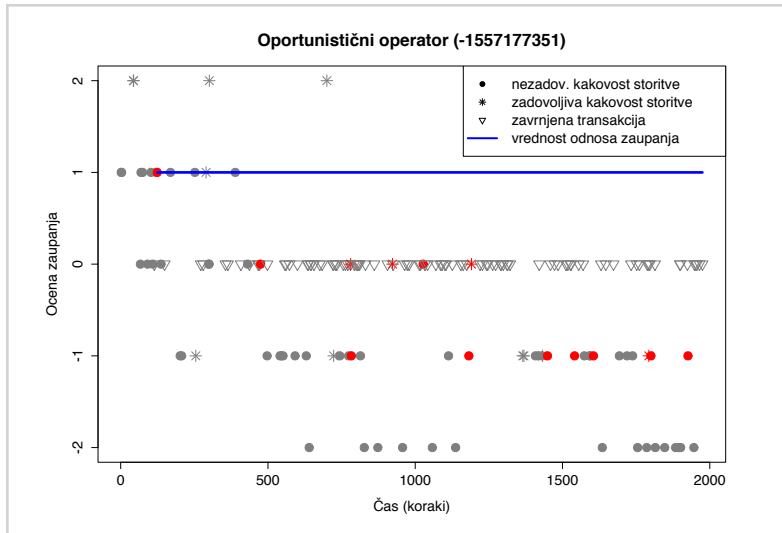
Povprečna vrednost ocen zaupanja do agenta  $U$ , zaokrožena proti vrednosti 0, je neodločen odnos zaupanja. Torej, končna ocena zaupanja agenta  $X$  do agenta  $U$  je po definiciji preslikave za operator  $\sim$  enaka  $\omega_{X,U}^+ = 0$ .

Zaradi zaokroževanja proti vrednosti 0 ocena zaupanja agenta  $X$  do agenta  $U$  relativno hitro pade na vrednost 0 (neodločen odnos zaupanja), in to kljub morebitnim dobrim izkušnjam z agentom  $U$ . Glede na vpeljani preprosti model odločanja na podlagi zaupanja to pomeni, da pade vrednost zaupanja na vrednost, ki je pod določenim pragom ( $th = 1$ ) za opravljanje transakciji. Agent  $X$  zato nadaljnje transakcije z agentom  $U$  zavrne.

Slike A.21, A.22, A.23, A.24, A.25 v dodatku A prikazujejo primere obnašanja vrednosti zaupanja, kot ga definira operator  $\rightsquigarrow$ , za vseh 50 ponovitev simulacije.

### Oportunistični operator ( $\leftrightarrow$ )

Podobno kot centralistični operator  $\rightsquigarrow$  tudi oportunistični operator  $\leftrightarrow$  modelira obnašanje odnosa zaupanja agenta, ki je v veliki meri odvisno od mnenj ostalih agentov v skupnosti. Agent ocenjevalec  $A$  oceni vrednost zaupanja do ocenjevanega agenta  $B$ , tako da izračuna povprečno vrednost ocen zaupanja ostalih agentov iz skupnosti do agenta  $B$  in dobljeno vrednost zaokroži na celo število, ki je dlje od vrednosti 0.



Slika 3.8

Primer obnašanja vrednosti odnosa zaupanja, modeliranega z operatorjem  $\leftrightarrow$ .

Slika 3.8 prikazuje značilno vrednotenje odnosa zaupanja, ki je modelirano z operatorjem  $\leftrightarrow$ . Po prvi transakciji med agentoma  $X$  in  $U$ , ki se zgodi v 124. koraku

simulacije, agent  $X$  izračuna neposredno oceno zaupanja do agenta  $U$  kot delno zaupanja vreden odnos, tj.  $\omega_{X,U}^- = 1$ , kar je na grafu prikazano z rdečim simbolom  $\bullet$ . Nato agent  $X$  izračuna povprečno oceno ostalih agentov v skupnosti do agenta  $U$ . Ta vrednost je enaka 0,875, kot je tudi razvidno iz vrednosti v matriki zaupanja:

$$\mathcal{M}_{123} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 2 & / & / & 2 & 2 & / & / & 2 & 0 \\ 2 & / & / & 2 & / & / & / & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & / & 2 & / & 2 & 2 & / & / & 2 & 2 & / \\ / & 2 & / & / & / & / & / & / & / & 2 & / & 1 \\ / & / & 2 & 2 & / & / & 2 & / & / & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & / & 2 & / & / & / & 2 & 1 \\ 2 & / & 2 & / & / & / & / & / & / & 2 & 2 & / \\ / & / & 2 & 2 & 2 & 2 & 2 & / & / & 2 & 2 & 0 \\ / & 2 & / & / & 2 & 2 & / & / & / & / & / & / \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / & / & 0 \\ / & 2 & 2 & 2 & 2 & / & / & / & 2 & / & / & 2 \\ 2 & / & / & 2 & 2 & 2 & 2 & 2 & 2 & / & / & / \end{array} \right]_{t=123}$$

$\xrightarrow[124. \text{ korak}]{\omega_{X,U}^- = 1}$

$$\mathcal{M}_{124} = \begin{array}{l} (\Downarrow) \\ (\rightsquigarrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\leftrightarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\rightsquigarrow) \end{array} \left[ \begin{array}{cccccccccccc} / & 2 & 2 & 2 & / & / & 2 & 2 & / & / & 2 & 0 \\ 2 & / & / & 2 & / & / & / & 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & / & 2 & / & 2 & 2 & / & / & 2 & 2 & 1 \\ / & 2 & / & / & / & / & / & / & / & 2 & / & 1 \\ / & / & 2 & 2 & / & / & 2 & / & / & 2 & 2 & 2 \\ 2 & 2 & 2 & / & 2 & / & 2 & / & / & / & 2 & 1 \\ 2 & / & 2 & / & / & / & / & / & / & 2 & 2 & / \\ / & / & 2 & 2 & 2 & 2 & 2 & / & / & 2 & 2 & 0 \\ / & 2 & / & / & 2 & 2 & / & / & / & / & / & / \\ 2 & 2 & 2 & 2 & 2 & / & 2 & 2 & 2 & / & / & 0 \\ / & 2 & 2 & 2 & 2 & / & / & / & 2 & / & / & 2 \\ 2 & / & / & 2 & 2 & 2 & 2 & 2 & 2 & / & / & / \end{array} \right]_{t=124}$$



Po drugi transakciji z agentom  $U$  je neposredna ocena zaupanja med agentoma ovrednotena z  $\omega_{X,U}^- = 0$ , saj je agent  $U$  zopet ponudil storitev z nezadovoljivo kakovostjo. Povprečna ocena agentov v skupnosti do agenta  $U$  je s tem v danem času enaka 0,4, kot je tudi razvidno iz vrednosti v matriki zaupanja v danem simulacijskem koraku.

Izračunana povprečna vrednost se glede na definicijo operatorja  $\leftrightarrow$  zaokroži na vrednost 1, kar tudi predstavlja končno oceno zaupanja agenta  $X$  do agenta  $U$  po drugi transakciji ( $\omega_{X,U}^+ = 1$ ).

V danem simulacijskem scenariju je vrednost večine ocen zaupanja do agenta  $U$  najmanj 0. Če je vrednost ocene zaupanja enaka 0, potem agent iskalec storitve namreč zavrne storitev, ki jo ponuja agent  $U$ , saj je vrednost zaupanja do  $U$  pod definiranim pragom za sprejem storitve. Zavrnitev storitve pomeni tudi, da se vrednost zaupanja do  $U$  ne izračuna več na novo, saj ponovno vrednotenje zaupanja poteka po opravljeni transakciji, tj. po sprejeti storitvi. Torej, ko vrednost  $\omega_{X,U}$  pade pod prag za opravljanje transakcij  $th = 1$ , vrednost ostane enaka do konca simulacije. Iz tega razloga je končna ocena zaupanja, kot jo modelira operator  $\leftrightarrow$ , višja od vrednosti 1 (delno zaupanja vreden odnos). Vrednost  $\omega_{X,U} = 1$  zadošča pogoju za opravljanje transakcij, tj.  $\omega_{X,U} = 1$  je nad definiranim pragom, zato agent  $X$  še naprej sprejema storitve od slabega agenta  $U$ . V nadaljevanju agent  $X$  od agenta  $U$  zato sprejme veliko storitev z nezadovoljivo kakovostjo, kar še dodatno pojasni rezultate iz prejšnjega poglavja, ki so prikazani na sliki 3.2.

Grafična analiza obnašanja vrednosti odnosa zaupanja, modeliranega z oportunističnim operatorjem, je za vseh 50 ponovitev simulacije prikazana na slikah A.26, A.27, A.28, A.29, A.30 v dodatku A.

*Metoda za obvladovanje  
zlaganih ocen zaupanja*

### 4.1 *Predstavitev problema*

Eden izmed najpomembnejših odprtih problemov na področju obvladovanja zaupanja in ugleda v e-okoljih je prisotnost neresničnih<sup>1</sup> ocen zaupanja. Če agent nima lastnih izkušenj z ocenjevanim agentom, so ocene zaupanja, ki so jih podali ostali agenti v danem sistemu, pogosto edini viri informacij za ovrednotenje zaupanja do ocenjevanega agenta. Če agent ocenjevalec izračuna zaupanje z združevanjem posredovanih ocen, bo tak izračun napačen, če so posredovane ocene zlagane.

O problemu zlaganih ocen govorimo, kadar agent v dani skupnosti posreduje oceno zaupanja, ki ne odraža njegove resnične izkušnje. Napade, ki temeljijo na posredovanju zlaganih ocen, v osnovi razvrstimo v dve skupini [50, 74–77]:

- napadi z očrnitvijo (angl. bad-mouthing attacks) in
- napadi s hvaljenjem (angl. ballot-stuffing attacks).

Napad s hvaljenjem pomeni, da agenti posredujejo pretirano visoke ocene zaupanja do ciljnega/napadenega agenta – ne glede na njihovo dejansko izkušnjo. S tem se splošen ugled napadenega agenta na nepošten način poveča. Pri napadu z očrnitvijo agenti posredujejo neresnične nizke vrednosti zaupanja do napadenih agentov, s čimer zmanjšajo njihov ugled.

Obe vrsti napadov dalje delimo v dve podkategoriji:

- samostojni napadi (angl. individual user attacks) in
- skupinski napadi (angl. collaborative users attacks).

Pri samostojnem napadu gre za napad s strani agenta posameznika, ki širi neresnično visoke ali nizke ocene zaupanja do drugih agentov. V skupinskem napadu neresnične ocene zaupanja širi skupina med seboj povezanih agentov s skupnim ciljem, da poneverijo ugled ciljne množice agentov v danem e-okolju.

Problematika zlaganih ocen zaupanja je ključnega pomena, saj agenti v e-okolju nimajo nadzora nad tem, kateri agenti so pošteni in kateri krivični ter posledično, katere ocene zaupanja so resnične in katere potvorbene. Agenti se za izbiro partnerjev odločajo (tudi) na podlagi ocenjenega zaupanja. Zaradi neresničnih ocen zaupanja lahko agent

---

<sup>1</sup>Neresnične ocene zaupanja v doktorski disertaciji imenujemo tudi zlagane ali potvorbene ocene zaupanja. Vse izraze (neresnične, zlagane in potvorbene) uporabljamo izmenljivo, pri čemer pa pomenijo enako.

napačno ovrednoti zaupanje do drugega agenta in posledično sprejme napačne (slabe) odločitve, kar dodatno kaže na pomembnost problematike zlaganih ocen.

Za rešitev problema kot prvo predlagamo razširitev modela zaupanja QAD s konstrukti, ki so potrebni za modeliranje zlaganih ocen zaupanja. Nato predlagamo izvirno metodo za identifikacijo in izločanje neresničnih ocen zaupanja, imenovano metoda QADE-fILTER.

## 4.2 Predlagana rešitev

### 4.2.1 Model zaupanja in ugleda QADE

Osnovni model zaupanja QAD predpostavlja, da so ocene zaupanja, ki so shranjene v matriki zaupanja  $\mathcal{M}$  resnične (agenti ocene iz matrike  $\mathcal{M}$  uporabijo pri vrednotenju zaupanja do drugih, nepoznanih agentov). V resnici to ne drži vedno in dober matematični model zaupanja mora predpostavljati, da lahko agenti lažjejo o vrednostih ocen zaupanja. S tem namenom smo osnovni model zaupanja razširili in ga poimenovali QADE (QAD Extended).

Najprej uvedemo agentov *zasebni* vektor ocen zaupanja.

*Definicija 6:* Zasebni vektor ocen zaupanja vsebuje resnične ocene zaupanja agenta  $a_i$  do ostalih agentov v skupnosti. Označimo ga z  $\mathcal{Z}_i = (\zeta_{i,1}, \zeta_{i,2}, \dots, \zeta_{i,n})$ , kjer element  $\zeta_{i,j} \in \Omega$  predstavlja resnično oceno zaupanja agenta  $a_i$  do agenta  $a_j$ . Dalje,  $\underline{\mathcal{Z}}_i$  je zasebni podvektor ocen zaupanja agenta  $a_i$ , v katerem so izpuščene vrednosti "p".

Za razliko od vrednosti  $\omega_{ij}$ , ki so podane v matriki zaupanja  $\mathcal{M}$  in so vidne vsem agentom v skupnosti, so vrednosti iz zasebnega vektorja ocen zaupanja  $\mathcal{Z}_i$  vidne in na razpolago samo agentu  $a_i$ , ki je lastnik vektorja.

Ocene zaupanja, ki so vsebovane v matriki zaupanja  $\mathcal{M}$ , predstavljajo aktualne (nazadnje izračunane in podane) ocene zaupanja. Model zaupanja QADE pri modeliranju zaupanja upošteva tudi pretekle ocene zaupanja. V ta namen definiramo zgodovinske matrike zaupanja.

*Definicija 7:* Pretekle ocene zaupanja med agenti so podane v zgodovinskih matrikah zaupanja  $\mathcal{M}^{-l}$ , kjer je  $\{l : l \in \mathbb{N} \cup \{0\}\}$ . Element matrike  $\omega_{ij}^{-l}$  definira  $l$ -to

preteklo oceno zaupanja agenta  $a_i$  do agenta  $a_j$ .

Nazadnje podane ocene zaupanja med agenti, tj. ocene zaupanja, ki so bile podane po zadnji evalvaciji zaupanja, so shranjene v zgodovinski matriki zaupanja  $\mathcal{M}^0$ . Matrika zaupanja  $\mathcal{M}^0$  predstavlja trenutne vrednosti ocen zaupanja med agenti. Predzadnje podane ocene zaupanja so shranjene v zgodovinski matriki zaupanja  $\mathcal{M}^{-1}$ . Ocene zaupanja, ki so bile podane pri predpredzadnjem vrednotenju zaupanja, so shranjene v zgodovinski matriki  $\mathcal{M}^{-2}$  itd.

Poleg zgodovinskih matrik zaupanja model QADE shranjuje tudi zgodovino zasebnih vektorjev ocen zaupanja.

*Definicija 8:* Pretekle resnične ocene zaupanja agenta  $a_i$  do drugih agentov so shranjene v zgodovinskih zasebnih vektorjih ocen  $\mathcal{Z}^{-l}$ , pri čemer je  $\{l : l \in \mathbb{N} \cup \{0\}\}$ . Element  $\zeta_{i,j}^{-l}$  označuje  $l$ -to preteklo resnično oceno zaupanja agenta  $a_i$  do agenta  $a_j$ .

Zgodovinske matrike zaupanja in zgodovina zasebnih vektorjev ocen zaupanja agentu ocenjevalcu med drugim omogočajo spremljanje posredovanih in resničnih vrednosti zaupanja skozi čas ter s tem zaznavanje morebitnih sprememb v obnašanju ocenjenih agentov. Agenti lahko s strateškim spreminjanjem obnašanja škodujejo drugim agentom. Tipičen primer predstavlja vzorec obnašanja, ko se agent obnaša dobro, dokler si ne pridobi visoke vrednosti zaupanja drugih agentov, nato pa pridobljeno zaupanje izkoristi in se začne obnašati zlonamerno. Ker tovrstno spreminjanje obnašanja agentov predstavlja način ogrožanja mehke varnosti agentov, je prepoznavanje namernega (strateškega) spreminjanja obnašanja agentov odprt problem na področju sistemov za obvladovanje zaupanja in ugleda [78–80]. Čeprav je prepoznavanje strateških sprememb obnašanja agentov izven obsega doktorske disertacije, predstavljajo zgodovinski zasebni vektorji ocen in zgodovinske matrike zaupanja podlago za reševanje opisanega problema, tj. zaznavanja in obvladovanja sprememb v obnašanju agentov.

Vse pretekle ocene zaupanja med agenti so kategorizirano shranjene v naslednjih večkratnih množicah:

*Definicija 9:*  $Z \omega_{i,j}^{0:-\{i,j\}} = \{\omega_{i,j}^0, \omega_{i,j}^{-1}, \dots, \omega_{i,j}^{-\{i,j\}}\}$  označimo večkratno množico vseh ocen zaupanja agenta  $a_i$  do agenta  $a_j$ , vključujoč zadnjo in pretekle ocene. Tej ustre-

zno večkratno množico resničnih ocen zaupanja agenta  $a_i$  do agenta  $a_j$ , vključujoč zadnje in pretekle resnične ocene, označimo z  $\zeta_{ij}^{0:-t[i,j]} = \{\zeta_{ij}^0, \zeta_{ij}^{-1}, \dots, \zeta_{ij}^{-t[i,j]}\}$ .

*Definicija 10:* Večkratna množica  $\mathcal{M}_{k,n}^{0:-t} = \{\omega_{k,1}^{0:-t[k,1]}, \omega_{k,2}^{0:-t[k,2]}, \dots, \omega_{k,n}^{0:-t[k,n]}\}$ , kjer je  $-t = \min\{-t[k,1], -t[k,2], \dots, -t[k,n]\}$ , vsebuje trenutno in pretekle ocene zaupanja agenta  $a_k$  do drugih agentov. Dalje, večkratna množica  $\mathcal{M}_{k,n_k}^{0:-t} = \{\omega_{k,1}^{0:-t[k,1]}, \omega_{k,2}^{0:-t[k,2]}, \dots, \omega_{k,n_k}^{0:-t[k,n_k]}\}$ , kjer je  $\underline{-t} = \min\{-t[k,1], -t[k,2], \dots, -t[k,n_k]\}$ , definira večkratno množico trenutnih in preteklih ocen zaupanja agenta  $a_k$  do drugih agentov, pri čemer so vrednosti "f" izpuščene.

*Definicija 11:* V večkratni množici  $\mathcal{M}_{n,k}^{0:-t} = \{\omega_{1,k}^{0:-t[1,k]}, \omega_{2,k}^{0:-t[2,k]}, \dots, \omega_{n,k}^{0:-t[n,k]}\}$ , pri čemer je  $-t = \min\{-t[1,k], -t[2,k], \dots, -t[n,k]\}$ , so vsebovane trenutna in pretekle ocene zaupanja iz vektorja ugleda agenta  $a_k$ . Dalje,  $\mathcal{M}_{n_k,k}^{0:-t} = \{\omega_{1,k}^{0:-t[1,k]}, \omega_{2,k}^{0:-t[2,k]}, \dots, \omega_{n_k,k}^{0:-t[n_k,k]}\}$ , kjer je  $\underline{-t} = \min\{\underline{-t}[1,k], \underline{-t}[2,k], \dots, \underline{-t}[n_k,k]\}$ , definira večkratno množico ocen zaupanja iz podvektorja ugleda agenta  $a_k$ .

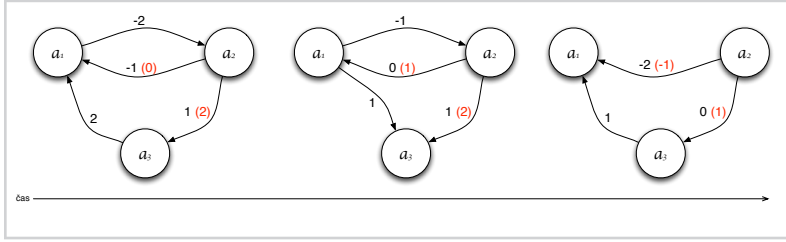
Za ponazoritev definicij 9, 10 in 11 bomo podali naslednji primer. Slika 4.1 prikazuje interakcije med agenti  $a_1$ ,  $a_2$  in  $a_3$  v treh časovnih intervalih. Ocena zaupanja med agentoma je vrednost na usmerjeni povezavi med vozliščema grafa, ki predstavljata dana agenta. Agent  $a_2$  v matriko zaupanja  $\mathcal{M}$  sporoča vrednosti, ki se razlikujejo od resničnih vrednosti zaupanja. Sporočene vrednosti so označene rdeče. Resnične ocene zaupanja so označene s črno. Agenti  $a_1$  in  $a_3$  v matriko zaupanja  $\mathcal{M}$  sporočata resnične vrednosti zaupanja, zato sporočene vrednosti niso dodatno navedene.

V danem primeru so vrednosti večkratnih množic naslednje:

- $\omega_{1,1}^{0:0} = \{ "f" \}, \zeta_{1,1}^{0:0} = \{ "f" \}$
- $\omega_{1,2}^{0:-1} = \{-1, -2\}, \zeta_{1,2}^{0:-1} = \{-1, -2\}$
- $\omega_{1,3}^{0:0} = \{1\}, \zeta_{1,3}^{0:0} = \{1\}$
- $\omega_{2,1}^{0:-2} = \{-1, 1, 0\}, \zeta_{2,1}^{0:-2} = \{-2, 0, -1\}$

Slika 4.1

Primer: Ocene zaupanja med agenti  $a_1, a_2$  in  $a_3$  v treh časovnih intervalih.



- $\omega_{2,2}^{0:0} = \{“/”\}, \zeta_{2,2}^{0:0} = \{“/”\}$
- $\omega_{2,3}^{0:-2} = \{1, 2, 2\}, \zeta_{1,3}^{0:-2} = \{0, 1, 1\}$
- $\omega_{3,1}^{0:-1} = \{1, 2\}, \zeta_{3,1}^{0:-1} = \{1, 2\}$
- $\omega_{3,2}^{0:0} = \{“/”\}, \zeta_{3,2}^{0:0} = \{“/”\}$
- $\omega_{3,3}^{0:0} = \{“/”\}, \zeta_{3,3}^{0:0} = \{“/”\}$
- $\mathcal{M}_{1,n=3}^{0:-1} = \{“/”, -1, -2, 1\}, \underline{\mathcal{M}}_{1,n_k=2}^{0:-1} = \{-1, -2, 1\}$
- $\mathcal{M}_{2,n=3}^{0:-2} = \{-1, 1, 0, “/”, 1, 2, 2\}, \underline{\mathcal{M}}_{2,n_k=2}^{0:-2} = \{-1, 1, 0, 1, 2, 2\}$
- $\mathcal{M}_{3,n=3}^{0:-1} = \{1, 2, “/”, “/”\}, \underline{\mathcal{M}}_{3,n_k=1}^{0:-1} = \{1, 2\}$
- $\mathcal{M}_{n=3,1}^{0:-2} = \{“/”, -1, 1, 0, 1, 2\}, \underline{\mathcal{M}}_{n_k=2,1}^{0:-2} = \{-1, 1, 0, 1, 2\}$
- $\mathcal{M}_{n=3,2}^{0:-1} = \{-1, -2, “/”, “/”\}, \underline{\mathcal{M}}_{n_k=1,2}^{0:-1} = \{-1, -2\}$
- $\mathcal{M}_{n=3,3}^{0:-2} = \{1, 1, 2, 2, “/”\}, \underline{\mathcal{M}}_{n_k=2,3}^{0:-2} = \{1, 1, 2, 2\}$

V nadaljevanju bomo  $\mathcal{M}^0, \omega_{i,j}^0, \mathcal{Z}_i^0, \zeta_{i,j}^0, \mathcal{M}_{k,n}^0, \mathcal{M}_{n,k}^0, \underline{\mathcal{M}}_{k,n_k}^0$  in  $\underline{\mathcal{M}}_{n_k,k}^0$  označevali z  $\mathcal{M}, \omega_{i,j}, \mathcal{Z}_i, \zeta_{i,j}, \mathcal{M}_{k,n}, \mathcal{M}_{n,k}, \underline{\mathcal{M}}_{k,n_k}$  in  $\underline{\mathcal{M}}_{n_k,k}$ , razen če ne bo eksplicitno navedeno drugače.

V osnovnem modelu zaupanja QAD agenti ovrednotijo zaupanje do drugih agentov, kot je določeno z operatorji zaupanja QAD (definicija 5). V modelu zaupanja QADE postopek ocenjevanja zaupanja med agenti modificiramo na naslednji način. Končna ocena zaupanja med agentoma  $a_i$  in  $a_j$  se izračuna na podlagi:

- zadnje ter preteklih *resničnih* ocen zaupanja agenta  $a_i$  do agenta  $a_j$ ;
- zadnje ter preteklih (javnih) ocen zaupanja drugih agentov do ocenjevanega agenta  $a_j$ .

Zadnja in pretekle resnične ocene zaupanja agenta  $a_i$  so podane v zgodovinskih zasebnih vektorjih ocen  $\mathcal{Z}^{-1}$ , medtem ko so zadnje in pretekle ocene zaupanja drugih agentov do agenta  $a_j$  shranjene v zgodovinskih matrikah zaupanja  $\mathcal{M}^{-1}$ .

*Definicija 12:* Operatorji QADE so funkcije  $f_i \in \Psi$ , tako da  $f_i : \widetilde{\mathcal{M}}_{n_a, j} = [\underline{\mathcal{M}}_{n_i, j}^{0:-t} \setminus \omega_{i,j}^{0:-t[i,j]}] \cup \zeta_{i,j}^{0:-t[i,j]} \rightarrow \zeta_{i,j}^+$ . Preslikave za operatorje so definirane v definiciji 5, pri čemer je  $n_a = \left| [\underline{\mathcal{M}}_{n_k, k}^{0:-t}(\omega_{i,k}^{0:-t[i,j]})] \cup \zeta_{i,j}^{0:-t[i,j]} \right|$ .

Za ponazoritev definicije 12 bomo dodali primer izračuna zaupanja agenta  $a_2$  do agenta  $a_1$  glede na ocene zaupanja, kot so prikazane na sliki 4.1. Vrednosti (večkratnih) množic, ki so vsebovane v definiciji, so naslednje:  $\underline{\mathcal{M}}_{n_k=2,1}^{0:-2} = \{-1, 1, 0, 1, 2\}$ ,  $\omega_{2,1}^{0:-2} = \{-1, 1, 0\}$  in  $\zeta_{2,1}^{0:-2} = \{-2, 0, -1\}$ . Iz danih vrednosti izpeljemo vrednosti  $n_a = |\{1, 2\} \cup \{-2, 0, -1\}| = 5$  in  $\underline{\mathcal{M}}_{n_a=5,2} = [1, 2, -2, 0, -1]$ . Naj ima agent  $a_2$  pripisan centralistični operator  $\rightsquigarrow$ . Nova vrednost zaupanja, izračunana po definiciji 12, je  $\zeta_{i,j}^+ = 0$ .

Zasebni vektorji ocen zaupanja agentov vsebujejo njihove resnične ocene zaupanja do drugih agentov. Agent lahko ocene iz zasebnega vektorja ocen zaupanja deli z ostalimi agenti v skupnosti, tj. jih posreduje v vsem vidno matriko zaupanja  $\mathcal{M}$ , v spremenjeni obliki. Agent, ki v matriko zaupanja  $\mathcal{M}$  posreduje spremenjene ocene zaupanja, imenujemo *napadalec*.

*Definicija 13:* Agent  $a_i$  je napadalec, če  $\mathcal{Z}_i \neq \mathcal{M}_{i,n}$ , tj. vrednosti v zasebnem vektorju ocen zaupanja agenta  $a_i$  se razlikujejo od vrednosti v (javnem) vektorju ocen zaupanja agenta  $a_i$ .

Operator QADE je funkcija, ki pri evalvaciji zaupanja upošteva subjektivno naravo posameznega agenta in vplive iz okolja kot tudi zgodovino preteklih ocen zaupanja. Operatorji lahko za enake vhodne podatke izračunajo različne končne ocene zaupanja zaradi drugačnega dojemanja vplivov iz okolja, tj. ocen zaupanja agentov svetovalcev

do ocenjevanega agenta. Če agenta  $a_i$  in  $a_j$  v matriko zaupanja  $\mathcal{M}$  posredujeta različni oceni zaupanja do agenta  $a_k$ , to ne pomeni nujno, da je eden izmed njiju napadalec. Pomeni lahko, da agenta zaupanje vrednotita na različne načine, tj. z različnima operatorjema QADE. Agent  $a_i$  mora ocene zaupanja drugih agentov obravnavati na različne načine, kot je običajno tudi v realnem svetu. Na primer, ljudje obravnavajo mnenja tistih, ki se pritožujejo nad vsem, drugače od mnenj tistih ljudi, ki so (pretirano) navdušeni nad vsem.

V modelu zaupanja QADE definiramo funkcijo podobnosti, s katero agent določi, kako podoben je ostalim agentom v skupnosti. Agent  $a_i$  dojame agenta  $a_j$  kot podobnega, če oba vrednotita zaupanje na podoben oziroma enak način. V formalizaciji zaupanja to pomeni, da za vrednotenje zaupanja uporabljata isti operator QADE. Z uporabo istega operatorja QADE so ocene zaupanja agenta  $a_i$  do drugih agentov v enakih okoliščinah podobne ali enake ocenam zaupanja, ki jih je podal agent  $a_j$ . Pravimo, da imata agenta  $a_i$  in  $a_j$  podobno *splošno mnenje* do drugih agentov v skupnosti.

Izračun podobnosti med agentoma  $a_i$  in  $a_j$  v modelu zaupanja in ugleda QADE je dvodelen:

- ocene zaupanja obeh agentov primerjamo paroma (angl. pairwise similarity) in
- upoštevamo primerjavo splošne miselnosti (angl. general mindset) obeh agentov, ki se odraža preko operatorjev QADE, s katerimi je modeliran odnos zaupanja primerjanih agentov.

*Definicija 14:* Splošno miselnost agenta  $a_i$  definiramo z normaliziranim histogramom porazdelitve ocen zaupanja iz zasebnega podvektorja ocen zaupanja  $\underline{\mathcal{Z}}_i$ :

$$htg(\underline{\mathcal{Z}}_i) = (\gamma_i^{[-2]}, \gamma_i^{[-1]}, \gamma_i^{[0]}, \gamma_i^{[1]}, \gamma_i^{[2]}),$$

kjer elementi  $\gamma_i^{[-2]}, \gamma_i^{[-1]}, \gamma_i^{[0]}, \gamma_i^{[1]}$  in  $\gamma_i^{[2]}$  označujejo relativno frekvenco zaupanja nevrednih, delno zaupanja nevrednih, neodločenih, delno zaupanja vrednih in zaupanja vrednih resničnih ocen zaupanja do drugih agentov v skupnosti. Dalje, izpeljano splošno miselnost agenta  $a_i$  definiramo z normaliziranim histogramom po-

razdelitve ocen zaupanja iz podvektorja ocen zaupanja  $\underline{\mathcal{M}}_{i,n}$ :

$$htg(\underline{\mathcal{M}}_{i,n}) = (\gamma_i^{[-2]}, \gamma_i^{[-1]}, \gamma_i^{[0]}, \gamma_i^{[1]}, \gamma_i^{[2]}),$$

kjer elementi  $\gamma_i^{[-2]}, \gamma_i^{[-1]}, \gamma_i^{[0]}, \gamma_i^{[1]}$  in  $\gamma_i^{[2]}$  označujejo relativno frekvenco zaupanja nevrednih, delno zaupanja nevrednih, neodločenih, delno zaupanja vrednih in zaupanja vrednih ocen zaupanja do drugih agentov v skupnosti, ki so vsebovane v podvektorju ocen zaupanja agenta  $a_i$ .

Agent  $a_i$  lahko izračuna splošno miselnost le za agenta  $a_i$ , za vsakega drugega agenta pa lahko izračuna le izpeljano splošno miselnost. Zato bomo v nadaljevanju za splošno miselnost agenta in izpeljano splošno miselnost agenta uporabljali isti izraz, in sicer splošna miselnost. V kolikor ne bo iz konteksta razvidno, ali gre za splošno miselnost ali izpeljano splošno miselnost, bomo to eksplicitno navedli.

*Definicija 15:* Funkcija podobnosti  $sim : \underline{\mathcal{Z}}_i \times \underline{\mathcal{M}}_{j,n_j} \rightarrow [0, 1]$  je definirana kot:

$$sim(\underline{\mathcal{Z}}_i, \underline{\mathcal{M}}_{j,n_j}) = s_{ij} =$$

$$1 - \sqrt{\sum_{\substack{\zeta_{i,k} \in \underline{\mathcal{Z}}_i \\ \omega_{j,k} \in \underline{\mathcal{M}}_{j,n_j}}} \frac{\left( (\zeta_{i,k} - \omega_{j,k}) * (1 - \gamma_i^{[\omega_{j,k}]}) + (|\bar{\zeta}_i - \bar{\omega}_j|) / \maxDst \right) / 2}{\min(|\underline{\mathcal{Z}}_i|, |\underline{\mathcal{M}}_{j,n_j}|) * \maxDst^2}},$$

kjer je  $\bar{\zeta}_i = \frac{1}{|\underline{\mathcal{Z}}_i|} \sum_{\zeta_{i,k} \in \underline{\mathcal{Z}}_i} \zeta_{i,k}$ ,  $\bar{\omega}_j = \frac{1}{|\underline{\mathcal{M}}_{j,n_j}|} \sum_{\omega_{j,k} \in \underline{\mathcal{M}}_{j,n_j}} \omega_{j,k}$  in  $\maxDst = \max(\Omega) - \min(\Omega)$ .<sup>a</sup>

<sup>a</sup>Za model QAD velja, da je  $\max(\Omega) = 2$  in  $\min(\Omega) = -2$ .

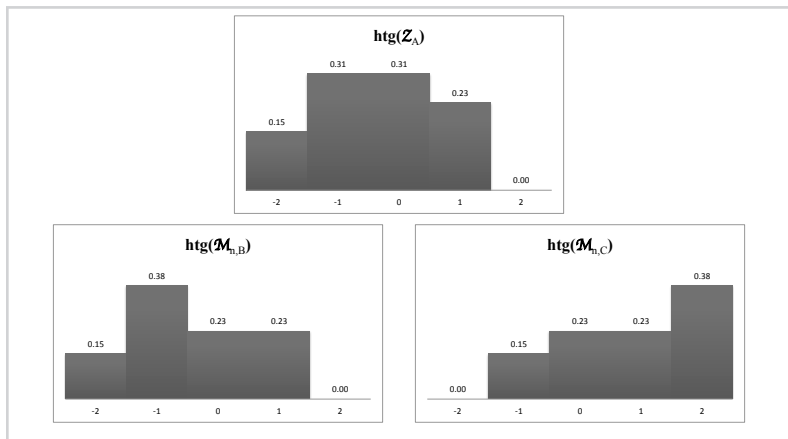
Podobnost med agentoma  $a_i$  in  $a_j$  je izračunana kot obtežena evklidska razdalja med elementi iz njunih podvektorjev ocen zaupanja. Uteži so definirane tako, da ustrezajo splošni miselnosti agenta  $a_i$ , ki ga primerjamo z drugimi agenti. Primer: izračunajmo podobnost agenta  $a_A$  z agentoma  $a_B$  in  $a_C$ . Lastnosti agentov  $a_A$ ,  $a_B$  in  $a_C$  so opisane v tabeli 4.1. Histogrami, ki opisujejo splošne miselnosti agentov

$a_A$ ,  $a_B$  in  $a_C$ , so prikazani na sliki 4.2. Agent  $a_A$  izračuna naslednje vrednosti splošnih miselnosti agentov:  $htg(\underline{\mathcal{Z}}_A) = (0, 33, 0, 17, 0, 39, 0, 11, 0, 00)$ ,  $htg(\underline{\mathcal{M}}_{B,n}) = (0, 33, 0, 17, 0, 22, 0, 28, 0, 00)$  in  $htg(\underline{\mathcal{M}}_{C,n}) = (0, 00, 0, 50, 0, 17, 0, 22, 0, 11)$ .

Tabela 4.1

Primer: Lastnosti agentov  $a_A$ ,  $a_B$  in  $a_C$ .

AGENT	OPERATOR	VEKTOR OCEN ZAUPANJA
$a_A$	$\downarrow$	$\underline{\mathcal{Z}}_A = [0, -1, -2, -2, 0, -2, 0, 1, 0, 1, 0, -1, -1, 0, -2, 0, -2, -2]$
$a_B$	$\downarrow$	$\underline{\mathcal{M}}_{B,n} = [1, -2, -2, -1, 1, -1, 1, 0, 1, 0, 1, -2, -2, 0, -2, 0, -2, -1]$
$a_C$	$\rightsquigarrow$	$\underline{\mathcal{M}}_{C,n} = [1, -1, -1, -1, 0, -1, 1, 2, 1, 1, 0, -1, -1, 0, -1, 1, -1, -1]$



Slika 4.2

Primer: Splošna miselnost agentov  $a_A$ ,  $a_B$  in  $a_C$ .

Najprej izračunajmo normalizirano evklidsko razdaljo (*NED*) (angl. normalized euclidean distance) med elementi podvektorjev ocen zaupanja. Vrednosti rezultatov  $1 - NED(A, B) = 0,79$  in  $1 - NED(A, C) = 0,80$  kažeta na to, da je agent  $a_C$  bolj podoben agentu  $a_A$  kot agent  $a_B$ . Poleg primerjave parov elementov podvektorjev ocen zaupanja sedaj upoštevajmo še splošne miselnosti agentov  $a_A$ ,  $a_B$  in  $a_C$ . Agenti  $a_A$  in  $a_B$  zaupanje ocenjujeta z zmernim pesimističnim operatorjem ( $\downarrow$ ), kar pomeni, da sta nagnjena k bolj negativnemu vrednotenju zaupanja. Ostale agente večinoma ocenjujeta kot zaupanja nevredne in delno zaupanja nevredne. Agent  $a_C$  zaupanje vrednoti s cen-

tralističnim operatorjem ( $\rightsquigarrow$ ), kar pomeni, da pri ocenjevanju zaupanja sledi povprečni oceni zaupanja, s katero so ostali agenti ocenili ocenjevanega agenta. V skupnostih, kjer so prisotni agenti s pripisanimi različnimi operatorji, bo agent s centralističnim operatorjem ( $\rightsquigarrow$ ) večinoma ocenil zaupanje do ostalih agentov z vrednostjo 0. Vidimo, da zgolj primerjava parov elementov iz vektorjev ocen zaupanja ne zadostuje. Primerjajmo oceno zaupanja  $\zeta_{A,j} = 1$  z ocenama zaupanja  $\omega_{B,j} = 0$  in  $\omega_{C,j} = 2$  (z rdečo označene vrednosti v tabeli 4.1). Oba elementa,  $\omega_{B,j} = 0$  in  $\omega_{C,j} = 2$ , sta od vrednosti  $\zeta_{A,j} = 1$  oddaljena za eno stopnjo, vendar namigujeta na različen način vrednotenja zaupanja agentov  $a_B$  in  $a_C$ . Vrednost  $\omega_{C,j} = 2$  nakazuje, da nabor ocen zaupanja agenta  $a_C$  do drugih agentov vključuje tudi ocene, ki opisujejo zaupanja vreden odnos. Ocenjevanje z visokimi vrednostmi zaupanja ne ustreza splošni miselnosti agenta  $a_A$  (glej sliko 4.2). Predlagana funkcija podobnosti *sim* zato zmanjša razdaljo med elementi, če element iz vektorja ocen zaupanja agenta, s katerim primerjamo podobnost (v našem primeru je to agent  $a_B$  ali  $a_C$ ), sovпада s splošno miselnostjo primerjanega agenta (v našem primeru je to agent  $a_A$ ). Pri izračunu podobnosti med agenti  $a_A$ ,  $a_B$  in  $a_C$  s predlagano funkcijo podobnosti *sim* dobimo naslednje rezultate:  $s_{A,B} = 0,85$  in  $s_{A,C} = 0,77$ . Glede na zadnji izračun je tako agentu  $a_A$  agent  $a_B$  bolj podoben kot agent  $a_C$ .

*Definicija 16:* Faktorji podobnosti med agenti so podani z matriko podobnosti  $\mathcal{S}$ . Element matrike  $s_{i,j}$  določa podobnost med agentoma  $a_i$  in  $a_j$ , izmerjeno s strani agenta  $a_i$ .

Splošna oblika matrike podobnosti v skupnosti z  $n$  agenti je naslednja:

$$\mathcal{S} = \begin{bmatrix} 1 & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & 1 & \dots & s_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \dots & 1 \end{bmatrix}.$$

Vrednosti  $s_{i,j}$  in  $s_{j,i}$  sta lahko različni, saj  $s_{i,j}$  določa podobnost med agentoma  $a_i$  in  $a_j$ , izmerjeno s strani agenta  $a_i$ , vrednost  $s_{j,i}$  pa podobnost med agentoma  $a_i$  in  $a_j$ , izmerjeno s strani agenta  $a_j$ . Če imajo vsi agenti v skupnosti enak operator in v dani skupnosti ni napadalcev, ki poročajo zlagane vrednosti, potem je matrika podobnosti simetrična. Sicer matrika  $\mathcal{S}$  ni nujno simetrična.

### Modeliranje napadov

O problemu zlaganih ocen zaupanja govorimo, če agenti v matriko zaupanja  $\mathcal{M}$  posredujejo vrednosti, ki ne odražajo resničnih ocen zaupanja do drugih agentov, ampak so posredovane vrednosti bodisi prenizke bodisi previsoke. Napade s širitvijo zlaganih ocen zaupanja delimo v skupini napadov z očrnitvijo in napadov s hvaljenjem, znotraj osnovnih skupin pa dalje na samostojne napade in skupinske napade. Tipi napadov so opisani v poglavju 4.1.

Razširjeni model zaupanja QADE omogoča modeliranje napadov s širitvijo zlaganih vrednosti, kar bomo opisali v nadaljevanju poglavja. Formalizirali bomo *samostojni napad z očrnitvijo*, *skupinski napad z očrnitvijo*, *samostojni napad s hvaljenjem* in *skupinski napad s hvaljenjem*. V nadaljevanju z  $a_A$  označimo AGENTA NAPADALCA (angl. attacker agent), z  $a_T$  pa NAPADENEGA AGENTA (angl. targeted agent).

Pri napadu s hvaljenjem agenti napadalci v matriko  $\mathcal{M}$  sporočajo *visoke* ocene zaupanja, ne glede na njihove dejanske izkušnje z ocenjevanimi agenti. Zaradi (pre)visokih ocen zaupanja se ugled napadenih agentov nepravilno poveča.

*Definicija 17:* Naj bo  $a_A, a_T \in \mathcal{A}$ ,  $\zeta_{A,T} \in \mathcal{Z}_A$  in  $\omega_{A,T} \in \mathcal{M}$ . Za predikat “ $BS(a_A) \equiv a_A$  je samostojni agent napadalec s hvaljenjem” velja:

$$\forall a_A : BS(a_A) \iff \exists a_T : (\omega_{A,T} > \zeta_{A,T})$$

Za potrebe formalizacije skupinskih napadalcev naj  $A \subseteq \mathcal{A}$  označuje skupino agentov napadalcev, ki sodelujejo pri posredovanju zlaganih vrednosti ocen zaupanja do skupine napadenih agentov  $T \subseteq \mathcal{A} \setminus A$ .

*Definicija 18:* Naj bo  $a_A \in A \subseteq \mathcal{A}$ ,  $a_T \in T \subseteq \mathcal{A} \setminus A$ ,  $\zeta_{A,T} \in \mathcal{Z}_A$  in  $\omega_{A,T} \in \mathcal{M}$ . Za predikat “ $BSG(a_A) \equiv a_A$  je skupinski agent napadalec s hvaljenjem” velja:

$$\forall a_A : BSG(a_A) \iff \left( (a_A \in A) \wedge \left( \exists a_T : \left( (a_T \in T) \wedge (\omega_{A,T} > \zeta_{A,T}) \right) \right) \right)$$

Pri napadu z očrnitvijo agenti napadalci v matriko  $\mathcal{M}$  posredujejo nepošteno *nizke* ocene zaupanja, kar posledično zmanjša ugled napadenih agentov.

*Definicija 19:* Naj bo  $a_A, a_T \in \mathcal{A}$ ,  $\zeta_{A,T} \in \mathcal{Z}_A$  in  $\omega_{A,T} \in \mathcal{M}$ . Za predikat “ $BM(a_A) \equiv a_A$  je samostojni agent napadalec z očrnitvijo” velja:

$$\forall a_A : BM(a_A) \iff \exists a_T : (\omega_{A,T} < \zeta_{A,T})$$

Pri skupinskem napadu z očrnitvijo skupina agentov napadalcev v matriko zaupanja  $\mathcal{M}$  posreduje nepravilno nizke ocene zaupanja.

*Definicija 20:* Naj bo  $a_A \in A \subseteq \mathcal{A}$ ,  $a_T \in T \subseteq \mathcal{A} \setminus A$ ,  $\zeta_{A,T} \in \mathcal{Z}_A$  in  $\omega_{A,T} \in \mathcal{M}$ . Za predikat “ $BMG(a_A) \equiv a_A$  je skupinski agent napadalec z očrnitvijo” velja:

$$\forall a_A : BMG(a_A) \iff \left( (a_A \in A) \wedge \left( \exists a_T : ((a_T \in T) \wedge (\omega_{A,T} < \zeta_{A,T})) \right) \right)$$

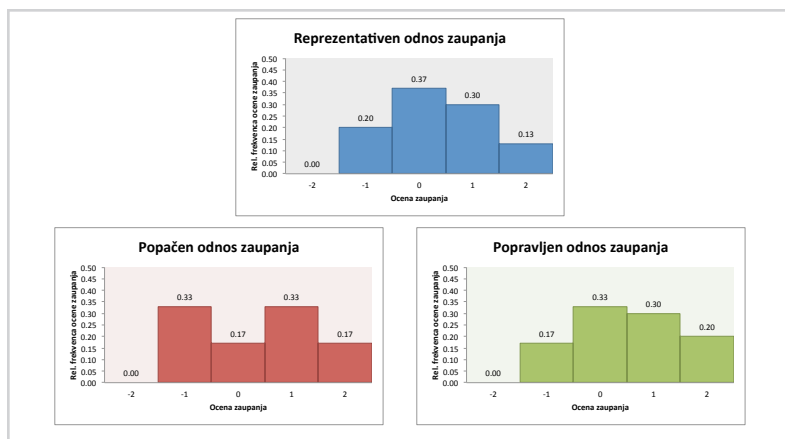
Formalizacija napadov je definirana s konstrukti, ki smo jih definirali v razširjenem modelu zaupanja in ugleda QADE. Opisana formalizacija predstavlja temelj za predlog metode za izračun zaupanja z obvladovanjem neresničnih ocen zaupanja, ki bo opisana v naslednjem razdelku.

#### 4.2.2 Metoda za izračun zaupanja z obrambo pred napadi QADE-filter

Agent oceni zaupanje do drugih agentov, tako da združi lastne ocene zaupanja iz  $\mathcal{Z}^{-1}$  in ocene, ki so jih posredovali drugi agenti v skupnosti in so shranjene v zgodovinskih matrikah zaupanja  $\mathcal{M}^{-1}$ . Relativne frekvence posameznih ocen zaupanja v agentovem podvektorju ocen zaupanja predstavljajo agentov *represntativen* odnos zaupanja. Potvorjene ocene zaupanja, na podlagi katerih agent ocenjevalec ovrednoti zaupanje do ocenjevanega agenta, odnos zaupanja agenta ocenjevalca *popačijo*. Za pravilno vrednotenje zaupanja mora biti agent ocenjevalec zmožen prepoznati, katere ocene so neresnične, in jih pri postopku vrednotenja zaupanja izločiti ali jih ustrezno modificirati.<sup>2</sup> S tem ko (domnevno) zlagane ocene zaupanja izloči ali jih ustrezno modificira, *popravi*

<sup>2</sup>Agent zlagane ocene prepozna z uporabo učinkovitega sistema za obvladovanje zaupanja in ugleda. Ta mora temeljiti na ustreznem matematičnem modelu, ki definira in omogoča zaznavanje potvorjenih vrednosti ocen zaupanja [7].

svoj odnos zaupanja. Primer reprezentativnega, popačenega in popravljenega odnosa zaupanja agenta ocenjevalca je prikazan na sliki 4.3.



Slika 4.3

Primer reprezentativnega, popačenega in popravljenega odnosa zaupanja agenta.

V splošnem obstajata dva osnovna pristopa za obvladovanje zlaganih ocen zaupanja: *endogene* in *eksogene* metode [6]. Endogene metode izključijo domnevno zlagane vrednosti glede na statistično primerjavo posredovanih ocen zaupanja. Pri eksogenih metodah se domnevno zlagane vrednosti izločijo glede na ugled agenta svetovalca, ki je vrednosti posredoval ostalim agentom, tj. v matriko zaupanja  $\mathcal{M}$ .

Naša rešitev je endogena in eksogena. V predlagani metodi agent ocenjevalec poišče agente s podobno splošno miselnostjo in njihove posredovane ocene zaupanja obravnava kot zanesljiv vir informacij. V tem pogledu je naša metoda eksogena. Agent ocenjevalec zatem lastne ocene zaupanja primerja z ocenami agentov, ki so bili prepoznani kot zanesljivi, v čemer je predlagana metoda endogena.

Predlagamo metodo za izključevanje zlaganih ocen zaupanja pri izračunu nove ocene zaupanja do ocenjevanega agenta, kot je opisano v algoritmih 3, 4 in 5.

Agent ocenjevalec  $a_i$  najprej izračuna, kateri ostali agenti v skupnosti imajo podobno splošno miselnost. S podobnimi agenti primerja oceno zaupanja do ocenjevanega agenta. Če se oceni razlikujeta za več kot eno stopnjo, agent ocenjevalec obravnava oceno zaupanja kot zlagano in jo označi s t. i. *črno piko*,<sup>3</sup> kot je opisano v algoritmu 4.

<sup>3</sup>Črna pika predstavlja prepričanje, da je ocena zaupanja neresnična, velikost črne pike pa stopnjo prepričanja

---

*Algoritem 3*Metoda *QADE-filter*

---

*Vhod:* agent ocenjevalec  $a_i$ , ocenjevani agent  $a_j$ , prag za mero podobnosti  $simT$ , prag za dopustno velikost črne pike  $bmT$

 $Ocene \leftarrow \{\}$ *for*  $a_k \in \mathcal{A}$  *do* $s_{i,k} \leftarrow \text{Izracunaj\_Podobnost}(a_i, a_k)$ *if*  $s_{i,k} > simT$  *then**if*  $|\zeta_{i,j} - \omega_{k,j}| > 1$  *then* $\omega_{k,j} \cdot \text{Oznaci\_S\_Crno\_Piko}(s_{i,k}, a_i, a_j, a_k)$ *else* $\omega_{k,j} \cdot \text{Zmanjsaj\_Crno\_Piko}(a_i, a_j, a_k)$ *end if**if*  $\text{Crna\_Pika}(a_k, a_i) < bmT$  *then* $Ocene \leftarrow Ocene \cup \omega_{k,j}^{0:-t[k,j]}$ *end if**end if**end for**return*  $Ocene$ 

---

Črna pika ima vrednost  $bm \in [0, 1]$ . Različne vrednosti črne pike nakazujejo, kako prepričan je agent ocenjevalec v zlaganost ocene. Velikost črne pike je odvisna od tega, kako podobna sta si agent svetovalec in agent ocenjevalec. Bolj kot sta si podobna, večja je velikost črne pike, s katero se označi domnevno zlagana ocena zaupanja. Velikost črne pike se poveča, če več agentov vrednost označi za zlagano, torej jo označi s črno piko. Največja vrednost črne pike  $bm = 1$  pomeni, da so vsi agenti iz skupnosti dano oceno zaupanja označili kot zlagano vrednost, pri čemer so imeli enako splošno miselnost. Če so v naslednjih evalvacijah zaupanja ocene zaupanja agenta ocenjevalca in agenta svetovalca bolj podobne, se velikost črne pike zmanjša (algoritem 5). Velikost črne pike  $bm = 0$  pomeni, da nobeden od agentov dane ocene zaupanja ni prepoznal kot zlagane ocene. Agent ocenjevalec pri izračunu zaupanja do ocenjevanega agenta  $a_j$

---

v neresničnost ocene. Zaradi večje izrazne moči pojmovanja "črna pika", bomo uporabljali to besedno zvezo.



zaupanja so naslednje:

$$\mathcal{M}_{712} = \begin{array}{l} (\leadsto) \\ (\Downarrow) \\ (\Downarrow) \\ (\Uparrow) \\ (\leadsto) \\ (\leadsto) \end{array} \left[ \begin{array}{cccccc} 0 & 0 & 0 & -1 & 0 & 0 \\ -2 & 0 & -2 & -2 & 0 & 0 \\ -2 & -2 & -2 & -2 & 0 & 0 \\ 0 & -2 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{array} \right]_{t=712}$$

Agent  $a_6$  v matriko zaupanja  $\mathcal{M}$  sporoča ocene zaupanja v nespremenjeni obliki, zato so vrednosti iz  $\mathcal{M}$  enake vrednostim v zasebnem vektorju zaupanja, ki ga iz tega razloga ne bomo eksplicitno zapisali. Podobnost agenta  $a_6$  z ostalimi je, kot sledi:  $s_{6,1} = 1,0$ ,  $s_{6,2} = 0,815$ ,  $s_{6,3} = 0,762$ ,  $s_{6,4} = 0,801$  in  $s_{6,5} = 0,989$ . Agent  $a_6$  iz izračuna zaupanja izloči ocene agenta  $a_3$ , saj je podobnost z njim manjša od določenega praga  $simT = 0,8$ . Z ostalimi agenti primerja ocene, ki so jih podali za  $a_2$  z lastno oceno za  $a_2$ . Agenti  $a_1$ ,  $a_2$  in  $a_4$  so podali enako oceno zaupanja do agenta  $a_2$ , zato se za te agente zmanjša vrednost črne pike:

- star delež črne pike( $a_1, a_2, a_6$ ) = 0,981; stara vrednost črne pike( $a_1, a_2$ ) = 0,467; nova vrednost črne pike( $a_1, a_2$ ) = 0,304;
- star delež črne pike( $a_2, a_2, a_6$ ) = 0,806; stara vrednost črne pike( $a_2, a_2$ ) = 0,134; nova vrednost črne pike( $a_2, a_2$ ) = 0,0;
- star delež črne pike( $a_5, a_2, a_6$ ) = 0,976; stara vrednost črne pike( $a_5, a_2$ ) = 0,461; nova vrednost črne pike( $a_5, a_2$ ) = 0,298.

Agent  $a_4$  je podal oceno zaupanja do agenta  $a_2$ , ki se razlikuje od ocene, ki jo je podal agent ocenjevalec  $a_6$ , zato mu ta dodeli črno piko z naslednjimi vrednostmi:

- star delež črne pike( $a_4, a_2, a_6$ ) = 0,801; nov delež črne pike( $a_4, a_2, a_6$ ) =  $s_{6,4} = 0,801$ ; stara vrednost črne pike( $a_4, a_2$ ) = 0,671; nova vrednost črne pike( $a_4, a_2$ ) = 0,671.

Ker je agent  $a_6$  pred tem z enako vrednostjo že prispeval k velikosti črne pike agenta  $a_4$ , je nova vrednost črne pike enaka stari. Ker je vrednost črne pike agenta  $a_4$  za podajanje ocen zaupanja do agenta  $a_2$  nad določenim pragom  $bmT = 0,35$ , ocene

zaupanja agenta  $a_4$  niso upoštevane pri evalvaciji zaupanja agenta  $a_6$  do agenta  $a_2$ . V izračune nove ocene zaupanja so poleg ocen iz množice  $\zeta_{6,2}^{0-19}$  upoštevane tudi trenutne in pretekle ocene agentov  $a_1, a_2$  in  $a_5$ , ki so v predstavljenem primeru naslednje:

$$\omega_{1,2}^{0-12} = \{0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 2\},$$

$$\omega_{2,2}^{0-20} = \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\} \text{ in}$$

$$\omega_{5,2}^{0-18} = \{0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1\}.$$

Kot vidimo, smo s predlagano metodo QADE-filter pravilno izločili ocene zaupanja agenta  $a_4$  do agenta  $a_2$ , ki so neresnične. Agentu ocenjevalcu je pripisan centralistični ( $\rightsquigarrow$ ) operator zaupanja. Glede na predpis operatorjev QADE (definicija 12) je končna vrednost zaupanja agenta  $a_6$  do agenta  $a_2$  enaka  $\zeta_{6,2} = 0$ .

### 4.3 Ovrednotenje metode QADE-filter

Predlagano metodo za izračun vrednosti zaupanja z izločanjem neresničnih ocen zaupanja smo ovrednotili s simulacijami, ki smo jih izvedli v simulacijskem orodju Zaupnik. Uspešnost naše metode QADE-filter smo primerjali z najbolj prepoznavno<sup>4</sup> eksogeno metodo, ki jo predlaga model TRAVOS [31], in z najbolj prepoznavno endogeno metodo statističnega filtriranja, ki so jo predlagali Whitby in drugi [46]. Metodi v nadaljevanju krajše poimenujemo TRAVOS-filter in Whitby filter.

Izračun zaupanja pri vseh treh metodah poteka, kot je opisano v algoritmu 6. Agent ocenjevalec oceni zaupanje do ocenjevanega agenta  $a_j$  glede na definicijo operatorja zaupanja QADE, ki mu je pripisan. Pri tem vzame v izračun lastne ocene zaupanja in *nekatero* ocene, ki so jih posredovali drugi agenti iz skupnosti. Katere ocene bodo izključene iz izračuna, določa izbrana metoda za izločanje neresničnih ocen, tj. metoda QADE-filter, metoda TRAVOS-filter ali metoda Whitby filter. Agent ocenjevalec po ovrednotenju zaupanja resnično oceno zaupanja  $\zeta_{i,j}$  shrani v zasebni vektor ocen zaupanja  $\mathcal{Z}_i$  in v matriko zaupanja  $\mathcal{M}$  posreduje oceno zaupanja  $\omega_{i,j}$ . Če je agent ocenjevalec  $a_i$  napadalec, bo ocena zaupanja  $\omega_{i,j}$ , ki jo posreduje v matriko zaupanja, drugačna od ocene zaupanja  $\zeta_{i,j}$ , ki jo shrani v zasebni vektor ocen zaupanja. V kolikor  $a_i$  ni napadalec, bo sporočena vrednost  $\omega_{i,j}$  enaka resnični oceni zaupanja  $\zeta_{i,j}$ .

Metoda TRAVOS-filter pri izločanju neresničnih ocen zaupanja najprej oceni verjetnost, da je agent svetovalec posredoval točne, tj. resnične, ocene. Agent ocenjevalec točnost posredovanih ocen izračuna tako, da primerja pretekle ocene agenta svetoval-

<sup>4</sup>Glede na število citatov.

---

*Algoritem 6*Ovrednotenje zaupanja z izključitvijo zlaganih ocen zaupanja - splošno

---

*Vhod:* agent ocenjevalec  $a_i$ , ocenjevani agent  $a_j$   
 $Ocene \leftarrow \{\}$   
 $LastneOcene \leftarrow Ocene \cup \zeta_{i,j}^{0:-t[i,j]}$   
 $SporoceneOcene \leftarrow FILTER(a_i, imeMetode, \mathcal{M}^{-1})$   
 $Ocene \leftarrow LastneOcene \cup SporoceneOcene$   
 $\zeta_{i,j}^+ \leftarrow QADE\_Operator(a_i, a_j, op_i, Ocene)$   
 $Posreduj\_V\_Matriko\_Zaupanja(\zeta_{i,j}^+, \check{a}_i, \omega_{i,j}^+)$

---

ca s svojimi (resničnimi) ocenami zaupanja do drugih agentov. Točnost se oceni z vrednostjo  $\rho \in [0, 1]$ , kjer  $\rho = 1$  pomeni, da se pretekle ocene agenta svetovalca popolnoma skladajo z ocenami agenta ocenjevalca, in  $\rho = 0$  pomeni, da je agent svetovalec v preteklosti posredoval netočne, tj. zlagane, ocene zaupanja. Glede na točnost v preteklosti posredovanih ocen se uravna vpliv ocen zaupanja, ki jih agent svetovalec posreduje. Posredovana mnenja agentov, ki so v preteklosti podali netočne ocene, imajo tako pri izračunu zaupanja manjši vpliv od posredovanih ocen agentov, ki so v preteklosti podali točne ocene zaupanja. Na primer, ocene zaupanja agentov z ocenjeno točnostjo  $\rho = 0$  nimajo nikakršnega vpliva in niso vključene v izračun nove vrednosti ocene zaupanja agenta ocenjevalca do ocenjevanega agenta. Pseudokoda algoritma za izločanje neresničnih ocen zaupanja z metodo TRAVOS-filter je opisana v algoritmu 7.

---

*Algoritem 7*Metoda TRAVOS-filter

---

*Vhod:* agent ocenjevalec  $a_i$ , ocenjevani agent  $a_j$   
 $Ocene \leftarrow \{\}$   
*for*  $a_k \in \mathcal{A}$  *do*  
    $\rho_k \leftarrow Izracunaj\_Tocnost\_Ocen\_Porocevalca(\mathcal{Z}^{-1}, \mathcal{M}^{-1})$   
    $OcenePorocevalca \leftarrow Zmanjsaj\_Ucinek\_Zlaganih\_Ocen(\omega_{k,j}^{0:-t[k,j]}, \rho_k)$   
    $Ocene \leftarrow Ocene \cup OcenePorocevalca$   
*end for*  
*return*  $Ocene$

---

Metoda Whitby filter izloči ocene zaupanja tistih agentov, ki so sporočili ocene zaupanja, ki ne sovpadajo z večino vseh ocen. Agent ocenjevalec preveri vrednost ocen, ki so jih agenti svetovalci podali za ocenjevanega agenta. Če se ocena zaupanja agenta svetovalca do ocenjevanega agenta bistveno razlikuje od povprečja ocen, ki so bile podane za ocenjevanega agenta, potem so ocene tega agenta svetovalca izvzete pri vrednotenju zaupanja. Postopek izločanja potvrjenih ocen zaupanja z metodo Whitby filter podaja psevdokoda v algoritmu 8.

---

#### Algoritem 8

Metoda *Whitby filter*

---

*Vhod:* agent ocenjevalec  $a_i$ , ocenjevani agent  $a_j$

$Ocene \leftarrow \{\}$

$l \leftarrow \text{Spodnji\_Kvantil}(\mathcal{M}_{n,j}^{0:-t}, q)$

$u \leftarrow \text{Zgornji\_Kvantil}(\mathcal{M}_{n,j}^{0:-t}, q)$

for  $a_k \in \mathcal{A}$  do

  if  $\omega_{k,j} > l$  AND  $\omega_{k,j} < u$  then

$Ocene \leftarrow Ocene \cup \omega_{k,j}^{0:-t\{k,j\}}$

  end if

end for

return  $Ocene$

---

S simulacijami smo primerjali učinkovitost izločanja zlaganih ocen zaupanja z metodami QADE-filter, TRAVOS-filter in Whitby filter. Zaupanje med agenti je modelirano z modelom zaupanja in ugleda QADE.

#### 4.3.1 Opis simulacij

V simulacijskem okolju sodeluje  $n$  agentov. Vsak agent ima določen operator QADE, ki določa njegovo obnašanje in način vrednotenja zaupanja. Vrednosti zaupanja med agenti so na začetku simulacije nedefinirane, tj.  $\omega_{i,j} = \text{"/"}$ ,  $\forall i, j$ . Nekateri agenti so *dobri*, drugi so *slabi*. Razlika med dobrimi in slabimi agenti bo opisana v nadaljevanju odstavka. Vsak agent je lahko bodisi *iskalec storitve* (*angl. service requester*) ali *ponudnik storitve* (*angl. service provider*). V vsako transakcijo sta vključena dva agenta: prvi agent v vlogi iskalca storitve in drugi v vlogi ponudnika storitve. Po opravljeni transakciji iskaletc storitve ovrednoti zaupanje do ponudnika storitve. Iskaletc storitve je agent

ocenjevalec in ponudnik storitve je ocenjevani agent. Simulacija poteka v diskretnih korakih, pri čemer je v vsakem koraku lahko narejena ena transakcija, tj. transakcija med iskalcem storitve in ponudnikom storitve. Med ostalimi agenti se v danem koraku ne izvrši nobena transakcija. Rezultat opravljene transakcije, ki se odraža s kakovostjo ponujene storitve, definira vrednost neposredne ocene zaupanja agenta iskalca storitve  $a_i$  do agenta ponudnika storitve  $a_j$ . Če je ponudnik storitve  $a_j$  dober agent, bo nova neposredna ocena zaupanja iskalca storitve do ponudnika storitve  $\omega_{i,j} = 0$ ,  $\omega_{i,j} = 1$  ali  $\omega_{i,j} = 2$ , tj. neodločen, delno zaupanja vreden ali zaupanja vreden odnos. V nasprotnem primeru, če je  $a_j$  slab agent, bo nova vrednost neposredne ocene zaupanja  $\omega_{i,j} = -2$ ,  $\omega_{i,j} = -1$  ali  $\omega_{i,j} = 0$ , tj. zaupanja nevreden, delno zaupanja nevreden ali neodločen odnos. V modelu zaupanja QADE na končno vrednost ocene zaupanja poleg neposredne (lastne) ocene zaupanja vplivajo tudi ocene drugih agentov iz skupnosti do ocenjevanega agenta, saj je vpliv okolja eden izmed dejavnikov, ki vplivajo na vrednost zaupanja med agenti. Aktualna in zgodovinske vrednosti neposrednih ocen zaupanja agenta ocenjevalca ter trenutne in zgodovinske ocene drugih agentov svetovalcev se združijo glede na definicijo operatorja QADE, ki je pripisan agentu ocenjevalcu. Agent  $a_i$  izračunano vrednost ocene zaupanja shrani v zasebni vektor ocen zaupanja  $\mathcal{Z}_i$  in oceno zaupanja posreduje v matriko zaupanja  $\mathcal{M}$ . Nekateri agenti v skupnosti so napadalci (*att*) in ti v matriko zaupanja posredujejo potvorjene vrednosti zaupanja do napadenih agentov (*tar*). Ocena zaupanja do napadenega agenta, ki jo v matriko zaupanja posreduje agent napadalec, ima drugačno vrednost od ocene, ki jo agent napadalec shrani v zasebni vektor ocen zaupanja. Da se agent ocenjevalec  $a_i$  izogne neresničnim ocenam zaupanja v postopku vrednotenja zaupanja, uporabi izbrano metodo za izločanje neresničnih ocen zaupanja: QADE-filter, TRAVOS-filter ali Whitby filter.

Simulacije smo izvedli s skupno  $n = 30$  agentov in simulacijo za vsako konfiguracijo ponovili 30-krat. V vsaki konfiguraciji smo agentom določili operatorje QADE, tako da je v skupnosti naslednji odstotek agentov z določenim operatorjem: 16,7 % z  $\uparrow$ , 16,7 % z  $\downarrow$ , 16,7 % z  $\uparrow$ , 16,7 % z  $\downarrow$ , 16,6 % z  $\rightsquigarrow$  in 16,6 % z  $\leftrightarrow$ . V simulacijskem okolju je 50 % dobrih agentov in 50 % slabih agentov. Prag za mero podobnosti agentov je nastavljen na vrednost  $simT = 0,8$  in še dovoljena vrednost črne pike je  $bmT = 0,35$ . V vsakem koraku naključno izberemo agenta  $a_i$ , ki je iskalec storitve, in agenta  $a_j$ , ki je ponudnik storitve. Izbrana agenta v danem koraku naredita transakcijo, po opravljeni transakciji pa agent  $a_i$  oceni vrednost zaupanja do agenta

$a_j$ . Ostali agenti v tem koraku ne ocenjujejo zaupanja do drugih agentov. Definirali smo simulacijske konfiguracije z različnim številom agentov napadalcev in napadenih agentov v skupnosti, kot je opisano v tabeli 4.2. Vsaka konfiguracija je bila izvedena z 10000 simulacijskimi koraki. V 10000 korakih vsak agent naredi v povprečju 333 transakcij, kar pomeni, da v povprečju 333-krat ovrednoti zaupanje do drugega agenta. Vsak agent vsakega drugega v povprečju oceni 11-krat, kar je dovolj, da se ocene zaupanja stabilizirajo. S tem zagotovimo tudi učinkovitost metode TRAVOS-filter, ki za izračun točnosti ocen agenta svetovalca predvideva zgodovino interakcij med agentom svetovalcem in ocenjevanim agentom. Pri vsaki ponovitvi zagona simulacije smo za inicializacijo generatorja psevdonaključnih števil uporabili drugačno vrednost semena. S tem smo zagotovili naključnost populacije v simulacijskem okolju (v smislu porazdelitve operatorjev QADE) kot tudi naključnost izbire iskalcev in ponudnikov storitev v simuliranih transakcijah (v vsakem koraku simulacije). Začetna porazdelitev ali vrstni red transakcij tako ne vplivata na končno porazdelitev ocen zaupanja, s čimer dobimo neodvisne simulacijske rezultate.

Na podlagi rezultatov simulacij smo ovrednotili našo predlagano metodo za izračun zaupanja z izločanjem zlaganih ocen. Po izvedenih ponovitvah simulacij z "Reprezentativno konfiguracijo" smo zajeli reprezentativno obnašanje agentov z različnimi operatorji QADE. Po 10000 simulacijskih korakih smo izračunali statistične lastnosti (kot so povprečna vrednost, mediana, standardni odklon in varianca) porazdelitve ocen zaupanja v vektorjih ocen zaupanja posameznih agentov. Zaupanje v modelu QADE je definirano s kvalitativnimi ocenami, ki simbolizirajo zaupanja nevreden (-2), delno zaupanja nevreden (-1), nedoločen (0), delno zaupanja vreden (1) in zaupanja vreden (2) odnos, zato smo ocene zaupanja za potrebe izračuna statističnih lastnosti pretvorili v pripadajoče numerične vrednosti in jih tako obravnavali. Ker v "Reprezentativni konfiguraciji" ni agentov napadalcev, ki bi poročali zlagane ocene, zajeto obnašanje agentov predstavlja njihov reprezentativni odnos zaupanja. Nato smo v simulacijsko okolje vključili tudi agente napadalce, ki v matriko zaupanja poročajo neresnične ocene zaupanja. Izvedli smo simulacije z različnim številom agentov napadalcev in različnim številom napadenih agentov, pri čemer so agenti izračunavali zaupanje brez izločanja ocen zaupanja in z uporabo metode QADE-filter, metode TRAVOS-filter ali metode Whitby filter za izločevanje domnevno zlaganih vrednosti. Izračunali smo statistične lastnosti porazdelitve ocen zaupanja pri vrednotenju zaupanja brez in z izločanjem zlaganih mnenj. Na podlagi dobljenih relativnih frekvenc ocen zaupanja smo za agen-

Tabela 4.2

Odstotek agentov napadalcev in napadenih agentov v simulacijskih konfiguracijah.

KONFIGURACIJA	ATT	TAR	IZRAČUN ZAUPANJA
Reprezentativna	0 %	0 %	QADE (brez napadov)
Napad z očrnitvijo (samostojni)	5 %–50 %	20 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
	20 %	5 %–50 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
Napad z očrnitvijo (skupinski)	5 %–50 %	20 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
	20 %	5 %–50 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
Napad s hvaljenjem (samostojni)	5 %–50 %	20 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
	20 %	5 %–50 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
Napad s hvaljenjem (skupinski)	5 %–50 %	20 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter
	20 %	5 %–50 %	QADE (z napadi), QADE-filter TRAVOS-filter, Whitby filter

te izpeljali njihove popačene odnose zaupanja (pri izračunu zaupanja *brez* izločanja zlaganih ocen zaupanja) in popravljene odnose zaupanja (pri vrednotenju zaupanja z izločanjem domnevno zlaganih ocen zaupanja). Za ocenitev učinkovitosti posamezne filtrirne metode smo statistično primerjali reprezentativne, popačene in popravljene odnose zaupanja posameznih agentov.

V poglavjih 4.3.2 in 4.3.3 bomo podrobneje opisali simulacijske parametre za vsak niz poskusov in predstavili rezultate.

#### 4.3.2 Ovrednotenje metod za izločanje zlaganih vrednosti pri napadih z očrnitvijo

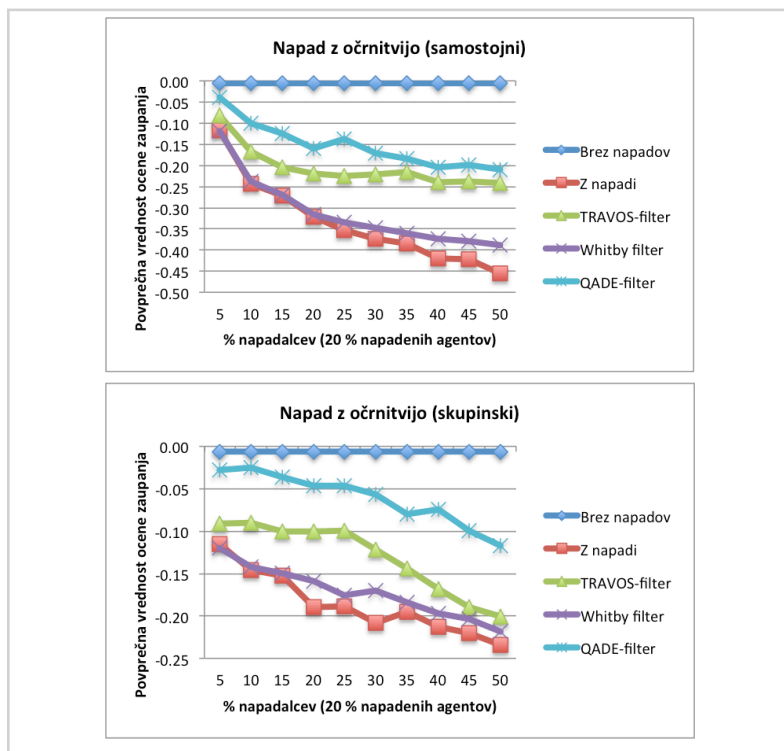
Za analizo uspešnosti filtrirnih metod smo simulirali e-okolje z različnim številom agentov napadalcev in napadenih agentov. V tem poglavju bomo ovrednotili učinkovitost metode QADE-filter pri napadih z očrnitvijo. Za potrebe ovrednotenja me-

tod za izločanje zlaganih vrednosti smo zbrali porazdelitve ocen zaupanja 900 agentov. Iz dobljenih vzorcev ocen zaupanja smo izračunali razliko med reprezentativnim in popačenim odnosom zaupanja ter razliko med reprezentativnim in popravljenim odnosom zaupanja za rezultate simulacij, kjer so bile uporabljene različne metode za izračun zaupanja z izločanjem zlaganih ocen. Razliko med reprezentativnim in popačenim/popravljenim odnosom smo izračunali za vsak operator QADE ter vse agente iz skupnosti.

#### *Rezultati ovrednotenja metod glede na različno število agentov napadalcev*

Slika 4.4 prikazuje povprečno vrednost ocen zaupanja v skupnostih z različnim številom agentov napadalcev in konstantnim številom napadenih agentov (20 %) po 10000 korakih simulacije. Ker smo simulacije izvedli v okolju, v katerem je 50 % dobrih agentov in 50 % slabih agentov, je povprečna vrednost ocene zaupanja v konfiguracijah brez napadalcev okrog 0. V analiziranih ponovitvah simulacij je bila vrednost ocen zaupanja v povprečju enaka  $-0,01$ . Če v skupnost vpeljemo napadalce z očrnitvijo, se povprečna vrednost ocen zaupanja zmanjša. Napadalci z očrnitvijo namreč v matriko zaupanja poročajo neresnične nizke ocene zaupanja. Povprečna vrednost ocen zaupanja je med  $-0,12$  v skupnosti s 5 % samostojnimi napadalci in  $-0,46$  v skupnosti s 50 % samostojnimi napadalci; ter med  $-0,11$  v skupnosti s 5 % skupinskimi napadalci in  $-0,23$  v skupnosti s 50 % skupinskimi napadalci. V vseh primerih je v simuliranem e-okolju 20 % napadenih agentov. Razlika v povprečju ocen zaupanja je manjša v konfiguracijah s skupinskimi napadalci, saj je v tem primeru omejeno število vrednosti, ki imajo lahko nepoštene vrednosti. Zlaganih je lahko namreč največ 20 % posredovanih ocen.

Namen filtrirnih metod za izločanje zlaganih ocen zaupanja je zmanjšati učinek takih ocen pri vrednotenju zaupanja. Boljše filtriranje pomeni, da je popravljen odnos zaupanja agenta bolj podoben reprezentativnemu. V konfiguracijah s samostojnimi agenti napadalci z očrnitvijo je razlika med popravljenimi in reprezentativnimi odnosi zaupanja agentov najmanjša pri uporabi metode QADE-filter in največja pri uporabi metode Whitby filter, kar je razvidno tudi s slike 4.4. Numerični izračuni rezultatov simulacij so podani na slikah B.1 in B.2. Metoda QADE-filter prekaša metodi Whitby filter in TRAVOS-filter tudi v konfiguracijah, v katerih so prisotni skupinski napadalci z očrnitvijo.



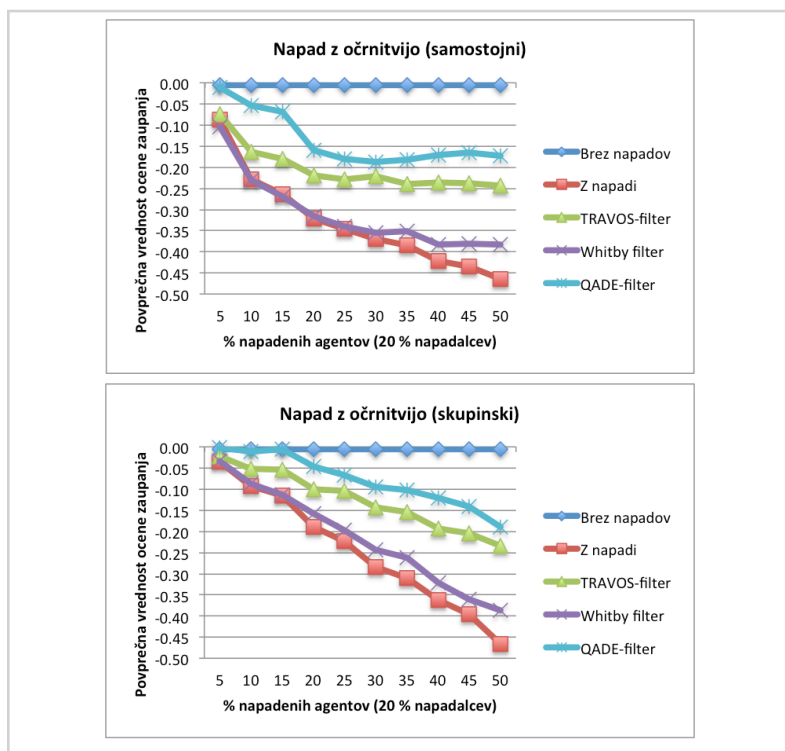
Slika 4.4

Povprečna ocena zaupanja v konfiguraciji brez napadov in v konfiguracijah z napadi z očrnitvijo pri različnem številu agentov napadalcev.

### Rezultati ovrednotenja metod glede na različno število napadenih agentov

Na sliki 4.5 so prikazane povprečne vrednosti ocen zaupanja po 10000 izvedenih korakih v skupnosti z 20 % agentov napadalcev in različnim odstotkom (med 5 % in 50 %) napadenih agentov. Samostojni agenti napadalci z očrnitvijo zmanjšajo povprečno vrednost ocen zaupanja, tako da je vrednost te med -0,09 (v skupnosti s 5 % napadenih agentov) in -0,46 (v skupnosti s 50 % napadenih agentov). Z uvedbo skupinskih agentov napadalcev z očrnitvijo se povprečna vrednost zmanjša na vrednosti med -0,04 (v skupnosti s 5 % napadenih agentov) in -0,47 (v skupnosti s 50 % napadenih agentov). Kot je razvidno s slike 4.5, je razlika med reprezentativnim in popravljenim odnosom zaupanja agentov najmanjša pri uporabi metode QADE-filter za vse definirane kon-

figuracije. Numerični izračuni za konfiguracije “Napad z očitvijo (samostojni)” in “Napad z očitvijo (skupinski)” so podani na slikah B.1 in B.2. Na primer, v skupnosti s 50 % napadenih agentov in 20 % skupinskih napadalcev je povprečna ocena zaupanja  $-0,19$  pri uporabi metode QADE-filter,  $-0,23$  pri uporabi metode TRAVOS-filter in  $-0,39$  pri uporabi metode Whitby filter. Vidimo, da je metoda QADE-filter uspešnejša pri izločanju zlaganih vrednosti od preostalih dveh metod.



Slika 4.5

Povprečna ocena zaupanja v konfiguraciji brez napadov in v konfiguracijah z napadi z očitvijo pri različnem številu napadenih agentov.

### Interpretacija in povzetek rezultatov

Razliko med povprečno oceno zaupanja v reprezentativnem, popačenem in popravljenem odnosu zaupanja smo izmerili kot normirano razdaljo z vrednostjo na intervalu  $[0, 1]$ . Razdalja med reprezentativnim odnosom, ki je izid scenarija brez napa-

dov, in popačenim odnosom, ki je posledica vpeljave napadalcev v okolje, je enaka  $dist(repr, skew(x \% att, y \% tar, attType)) = 1$ . Določa razdaljo med reprezentativnim in popačenim odnosom zaupanja v vsaki posamezni konfiguraciji, tj. za vsako vrsto napada, za različno število napadalcev in različno število napadenih agentov (opis konfiguracij je v tabeli 4.2). Vrednost  $dist(repr, skew(x \% att, y \% tar, attType)) = 1$  predstavlja *popolni vpliv* zlaganih vrednosti na agentov odnos zaupanja. Izračunamo jo kot

$$dist(repr, skew(x \% att, y \% tar, attType)) = \frac{PVOC(QADE(z \text{ napadi}, x \% att, y \% tar, attType) - PVOC(QADE(\text{brez napadov}))}{PVOC(QADE(z \text{ napadi}, x \% att, y \% tar, attType) - PVOC(QADE(\text{brez napadov}))}$$

pri čemer PVOC pomeni "povprečna vrednost ocene zaupanja" v dani konfiguraciji.

Kot primer bomo izračunali vrednost  $dist(repr, skew(20 \% att, 40 \% tar, BM))$ , tj. za konfiguracijo, v kateri je vrsta napada BM = "Napad z očrnitvijo (samostojni)", imamo 20 % napadalcev in 40 % napadenih agentov. Rezultati za dano simulacijsko konfiguracijo so grafično prikazani na sliki 4.5 in numerično podani na sliki B.1 v dodatku B. Za dano konfiguracijo velja:

$$PVOC(QADE(\text{brez napadov})) = -0,01$$

in

$$PVOC(QADE(z \text{ napadi}, 20 \% att, 40 \% tar, BM)) = -0,42.$$

Iz tega sledi

$$dist(repr, skew(20 \% att, 40 \% tar, BM)) = \frac{-0,42 - (-0,01)}{-0,42 - (-0,01)} = 1,$$

kar predstavlja *popolni vpliv* (se pravi, vrednost 1) zlaganih vrednosti na agentov odnos zaupanja v dani konfiguraciji.

Povprečno razdaljo med povprečno oceno zaupanja v reprezentativnem in popačenim odnosu zaupanja za ves nabor konfiguracij z različnim odstotkom napadalcev in fiksnim odstotkom napadenih agentov ter različnim odstotkom napadenih agentov in fiksnim odstotkom napadalcev označimo in izračunamo kot povprečje razdalj za vse konfiguracije za določen tip napada:

$$dist(repr, skew(attType)) =$$

$$= \frac{1}{|X \times Y|} \sum_{(x,y) \in X \times Y} \text{dist}(\text{repr}, \text{skew}(x \% \text{att}, y \% \text{tar}, \text{attType})),$$

kjer je

$$\begin{aligned} X \times Y = \{ & (5, 20), (10, 20), (15, 20), (20, 20), (25, 20), (30, 20), (35, 20), \\ & (40, 20), (45, 20), (50, 20), (20, 5), (20, 10), (20, 15), (20, 25), (20, 30), (20, 35), \\ & (20, 40), (20, 45), (20, 50) \}. \end{aligned}$$

Namen filtrirnih metod je izničiti učinek zlaganih vrednosti na vrednotenje zaupanja – in posledično, na odnos zaupanja. Z metodami za obvladovanje zlaganih ocen zaupanja želimo *popraviti* odnos zaupanja, da bo ta čim bolj podoben reprezentativnemu. Razdaljo med reprezentativnim in popravljenim odnosom za konfiguracijo z določenim odstotkom agentov napadalcev in napadanih agentov ter določeno vrsto napada, pri čemer smo za filtriranje neresničnih ocen zaupanja uporabili metodo “Filter” (za vse konfiguracije, ki so opisane v tabeli 4.2), označimo in izračunamo kot:

$$\begin{aligned} \text{dist}(\text{repr}, \text{corr}(\text{Filter}, x \% \text{att}, y \% \text{tar}, \text{attType})) &= \\ &= \frac{\text{PVOC}(\text{Filter}, x \% \text{att}, y \% \text{tar}, \text{attType}) - \text{PVOC}(\text{QADE}(\text{brez napadov}))}{\text{PVOC}(\text{QADE}(z \text{ napadi}), x \% \text{att}, y \% \text{tar}, \text{attType}) - \text{PVOC}(\text{QADE}(\text{brez napadov}))}. \end{aligned}$$

Manjša vrednost  $\text{dist}(\text{repr}, \text{corr}(\text{Filter}, x \% \text{att}, y \% \text{tar}, \text{attType}))$  torej pomeni bolj uspešno izločanje zlaganih vrednosti. Vrednost razdalje  $\text{dist}(\text{repr}, \text{corr}(\text{Filter}, x \% \text{att}, y \% \text{tar}, \text{attType})) = 0$  pomeni popolno izločitev zlaganih vrednosti. Predstavlja stanje *brez vpliva* zlaganih vrednosti na odnos zaupanja agenta, saj je popravljen odnos zaupanja enak reprezentativnemu.

Izračunali bomo vrednosti  $\text{dist}(\text{repr}, \text{corr}(\text{Filter}, 20 \% \text{att}, 40 \% \text{tar}, \text{BM}))$  (kot nadaljevanje prejšnjega primera) za primerjane metode za obvladovanje zlaganih ocen zaupanja, tj. za Filter = “QADE-filter”, Filter = “TRAVOS-filter” in Filter = “Whitby filter”. Za dane konfiguracije smo s simulacijami dobili naslednje vrednosti:

- $\text{PVOC}(\text{QADE}(\text{brez napadov})) = -0,01$ ,
- $\text{PVOC}(\text{QADE}(z \text{ napadi}), 20 \% \text{att}, 40 \% \text{tar}, \text{BM}) = -0,42$ ,
- $\text{PVOC}(\text{QADE-filter}, 20 \% \text{att}, 40 \% \text{tar}, \text{BM}) = -0,17$ ,
- $\text{PVOC}(\text{TRAVOS-filter}, 20 \% \text{att}, 40 \% \text{tar}, \text{BM}) = -0,24$ ,

- $PVOC(\text{Whitby filter}, 20\% \text{ att}, 40\% \text{ tar}, BM) = -0,38$ .

Iz tega sledijo naslednji izračuni:

- $dist(repr, corr(\text{QADE-filter}, 20\% \text{ att}, 40\% \text{ tar}, BM)) = \frac{-0,17 - (-0,01)}{-0,42 - (-0,01)} = 0,39$ ,
- $dist(repr, corr(\text{TRAVOS-filter}, 20\% \text{ att}, 40\% \text{ tar}, BM)) = \frac{-0,24 - (-0,01)}{-0,42 - (-0,01)} = 0,56$   
in
- $dist(repr, corr(\text{Whitby filter}, 20\% \text{ att}, 40\% \text{ tar}, BM)) = \frac{-0,24 - (-0,01)}{-0,42 - (-0,01)} = 0,90$ .

Izračuni vrednosti  $dist(repr, corr(\text{Filter}, x\% \text{ att}, y\% \text{ tar}, attType))$  za vse konfiguracije so podani v dodatku B.

Povprečno razdaljo med povprečno oceno zaupanja v reprezentativnem in popravljenem odnosu zaupanja za ves nabor konfiguracij označimo in izračunamo, kot sledi:

$$\begin{aligned} & dist(repr, corr(\text{Filter}, attType)) = \\ & = \frac{1}{|X \times Y|} \sum_{(x,y) \in X \times Y} dist(repr, corr(\text{Filter}, x\% \text{ att}, y\% \text{ tar}, attType)), \end{aligned}$$

kjer je

$$\begin{aligned} X \times Y = \{ & (5, 20), (10, 20), (15, 20), (20, 20), (25, 20), (30, 20), (35, 20), \\ & (40, 20), (45, 20), (50, 20), (20, 5), (20, 10), (20, 15), (20, 25), (20, 30), (20, 35), \\ & (20, 40), (20, 45), (20, 50) \}. \end{aligned}$$

V konfiguracijah "Napad z ocrnitvijo (samostojni)" (krajše BM) je vrednost razdalje med reprezentativnim in popravljenim odnosom zaupanja v povprečju enaka  $dist(repr, corr(\text{QADE-filter}, BM)) = 0,39$  pri uporabi metode QADE-filter, enaka  $dist(repr, corr(\text{TRAVOS-filter}, BM)) = 0,62$  pri uporabi metode TRAVOS-filter in enaka  $dist(repr, corr(\text{Whitby filter}, BM)) = 0,95$  pri uporabi metode Whitby filter. To predstavlja 23 % izboljšanje metode QADE-filter metode v primerjavi z metodo TRAVOS-filter in 56 % izboljšanje v primerjavi z metodo Whitby filter.

Izračunane razdalje za konfiguracije "Napad z ocrnitvijo (skupinski)" (krajše collBM) so  $dist(repr, corr(\text{QADE-filter}, collBM)) = 0,23$  pri uporabi metode QADE-filter,  $dist(repr, corr(\text{TRAVOS-filter}, collBM)) = 0,57$  pri uporabi metode TRAVOS-filter in  $dist(repr, corr(\text{Whitby filter}, collBM)) = 0,90$  pri uporabi metode Whitby

filter. To pomeni povprečno 34 % in 67 % boljše izločanje zlaganih ocen z metodo QADE-filter v primerjavi z metodama TRAVOS-filter in Whitby filter.

Dodatni numerični rezultati simulacij so prikazani na sliki B.1 za konfiguracije “Napad z očrnitvijo (samostojni)” in na sliki B.2 za konfiguracije “Napad z očrnitvijo (skupinski)”, ki se nahajata v dodatku B. Glede na dobljene rezultate lahko potrdimo učinkovitost naše predlagane metode.

#### *4.3.3 Ovrednotenje metod za izločanje zlaganih vrednosti pri napadih s hvaljenjem*

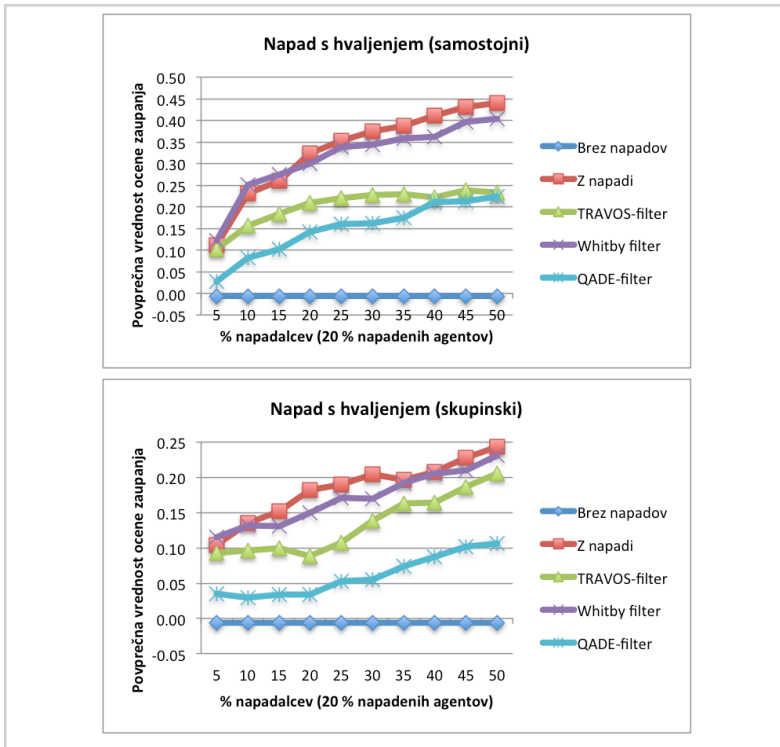
V drugem delu ovrednotenja metod za izločanje zlaganih vrednosti smo analizirani njihovo učinkovitost pri napadih s hvaljenjem, ko agenti svetovalci v matriko zaupanja posredujejo nepravilno visoke vrednosti ocen zaupanja. Izvedli smo simulacije z različnim številom (med 5 % in 50 %) napadalcev s hvaljenjem ob nespremenjenem številu (20 %) napadenih agentov. Prav tako smo izvedli simulacije z različnim številom (med 5 % in 50 %) napadenih agentov ob konstantnem številu agentov napadalcev (20 %). Simulacije smo izvedli tako za skupinske kot za samostojne napade. Kot je opisano v razdelku 4.3.1, smo vsako simulacijsko konfiguracijo ponovili 30-krat in primerjali razlike med dobljenimi reprezentativnimi, popačenimi in popravljeni odnosi zaupanja agentov.

#### *Rezultati ovrednotenja metod glede na različno število agentov napadalcev*

Po izvršenih 10000 simulacijskih korakov smo izračunali povprečne vrednosti ocen zaupanja za simulacijske konfiguracije z različnim številom agentov napadalcev in konstantnim številom napadenih agentov, ki so grafično prikazane na sliki 4.6. Povprečna vrednost ocen zaupanja v “Reprezentativni konfiguraciji” je enaka kot v prejšnjem nizu eksperimentov (poglavje 4.3.2), tj.  $-0,01$ . V konfiguracijah, kjer so v skupnosti prisotni agenti napadalci s hvaljenjem, se povprečna ocena zaupanja med agenti zviša, saj napadalci s hvaljenjem v matriko zaupanja poročajo ocene zaupanja, ki imajo nepošteno visoke vrednosti. V konfiguracijah “Napad s hvaljenjem (samostojni)” je povprečna ocena zaupanja med  $0,11$  (v simulacijskem okolju s 5 % napadalcev) in  $0,44$  (v simulacijskem okolju s 50 % napadalcev). V konfiguracijah “Napad s hvaljenjem (skupinski)” povprečna ocena zaupanja narašča od vrednosti  $0,10$  (v simulacijskem okolju s 5 % napadalcev) do  $0,24$  (v simulacijskem okolju s 50 % napadalcev).

Primerjava povprečnih ocen zaupanja pokaže, da smo popravljen odnos zaupanja, ki se najbolj ujema z reprezentativnim odnosom, dosegli z metodo QADE-filter. Po

učinkovitosti ji sledi metoda TRAVOS-filter, na zadnjem mestu je metoda Whitby filter, kar velja tako za konfiguracije s samostojnimi kot skupinskimi napadalci. Rezultati so grafično prikazani na sliki 4.6 in numerično podani na slikah B.3 in B.4.



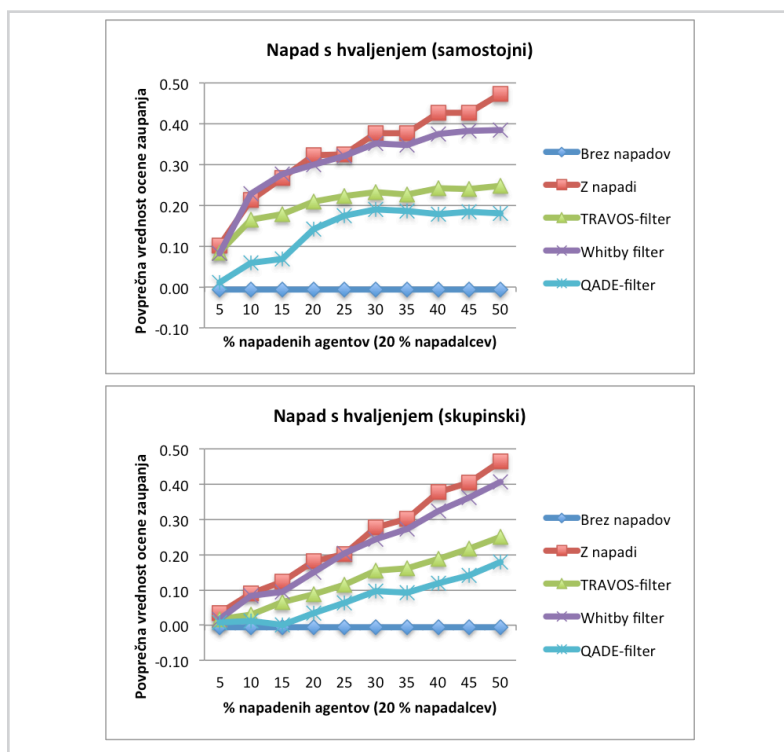
Slika 4.6

Povprečna ocena zaupanja v konfiguraciji brez napadov in v konfiguracijah z napadi s hvaljenjem pri različnem številu agentov napadalcev.

### Rezultati ovrednotenja metod glede na različno število napadenih agentov

Slika 4.7 prikazuje povprečne vrednosti ocen zaupanja v konfiguracijah s konstantnim številom agentov napadalcev in spremenljivim številom napadenih agentov. Povprečje ocen zaupanja v konfiguracijah s prisotnimi agenti napadalci narašča v primerjavi s povprečjem ocen zaupanja v "Reprezentativni konfiguraciji", kjer agenti napadalci niso prisotni. V konfiguracijah "Napad s hvaljenjem (samostojni)" je vrednost povprečja ocen zaupanja med 0,10 (v simulacijski skupnosti s 5 % napadenih agentov) in 0,47 (v

simulaciji s 50 % napadenih agentov). Podobno je v konfiguracijah skupinskega napada s hvaljenjem: povprečje ocen zaupanja je med 0,03 (simulacija s 5 % napadenih agentov) in 0,46 (v simulaciji s 50 % napadenih agentov). Cilj metod za filtriranje zlaganih ocen zaupanja je popravljen odnos zaupanja čim bolj približati reprezentativnemu odnosu zaupanja. Razlika med reprezentativnim in popravljenim odnosom zaupanja je najmanjša z uporabo metode QADE-filter. Sledita ji metodi TRAVOS-filter in Whitby filter. Numerični izračuni za konfiguracije “Napad s hvaljenjem (samostojni)” in “Napad s hvaljenjem (skupinski)” so podani na slikah B.3 in B.4.



Slika 4.7

Povprečna ocena zaupanja v konfiguraciji brez napadov in v konfiguracijah z napadi s hvaljenjem pri različnem številu napadenih agentov.

### Interpretacija in povzetek rezultatov

Enako kot pri ovrednotenju metod za izločanje potvorjenih ocen zaupanja pri napadih z očrnitvijo (glej poglavje 4.3.2) smo tudi v tem primeru izračunali normirano vrednost razdalje med reprezentativnimi in popravljenimi odnosi zaupanja agentov. V povprečju je vrednost normirane razdalje med reprezentativnim in popravljenim odnosom zaupanja agenta enaka  $dist(repr, corr(QADE\text{-}filter, BS)) = 0,43$  pri uporabi metode QADE-filter,  $dist(repr, corr(TRAVOS\text{-}filter, BS)) = 0,65$  pri uporabi metode TRAVOS-filter in  $dist(repr, corr(Whitby\ filter, BS)) = 0,95$  pri uporabi metode Whitby filter, kar veča za konfiguracije s samostojnimi napadalci s hvaljenjem (krajše BS). Dobljeni rezultati pomenijo 22 % in 52 % izboljšanje metode QADE-filter v primerjavi z metodama TRAVOS-filter in Whitby filter.

V konfiguracijah "Napad s hvaljenjem (skupinski)" (krajše collBS) smo v povprečju izmerili razdaljo  $dist(repr, corr(QADE\text{-}filter, collBS)) = 0,34$  pri uporabi metode QADE-filter, razdaljo  $dist(repr, corr(TRAVOS\text{-}filter, collBS)) = 0,65$  pri uporabi metode TRAVOS-filter in razdaljo  $dist(repr, corr(Whitby\ filter, collBS)) = 0,90$  pri uporabi metode Whitby filter. Rezultati predstavljajo za 31 % oziroma za 56 % boljše učinkovitost naše predlagane metode v primerjavi z metodo TRAVOS-filter oziroma metodo Whitby filter.

Rezultati simulacij za konfiguracije z različnim odstotkom napadenih agentov in različnim odstotkom agentov napadalcev so v numerični obliki prikazani na sliki B.3 za konfiguracije "Napad s hvaljenjem (samostojni)" in na sliki B.4 za konfiguracije "Napad s hvaljenjem (skupinski)". Sliki se nahajata v dodatku B.

Rezultati simulacij so pokazali, da je popravljen odnos zaupanja bližje reprezentativnemu odnosu zaupanja, če za obvladovanje zlaganih ocen zaupanja uporabimo metodo QADE-filter, kot pri uporabi eksogene metode TRAVOS-filter ali uporabi endogene metode Whitby filter. Metoda QADE-filter je uspešnejša v konfiguracijah, ki vključujejo samostojne ali skupinske napadalce z očrnitvijo kot tudi v konfiguracijah s samostojnimi ali skupinskimi napadalci s hvaljenjem. Predlagana metoda QADE-filter je učinkovitejša od metod Whitby filter in TRAVOS-filter v obeh primerih: v okolju z različnim številom agentov napadalcev in različnim številom napadenih agentov.

#### 4.4 *Kvalitativna primerjava metode QADE-filter z alternativnimi pristopi*

V tem poglavju bomo opisali metode za obvladovanje zlaganih ocen zaupanja in druge sorodne pristope ter jih kvalitativno primerjali s predlagano metodo QADE-filter za izločanje zlaganih vrednosti.

Predlagana metoda QADE-filter za prepoznavanje zlaganih ocen zaupanja v prvem koraku delovanja izračuna stopnjo podobnosti med agenti. Koncept podobnosti se uporablja pri metodah *kolektivnega filtriranja* (angl. *collaborative filtering*) in posledično v *sistemih za priporočila* (angl. *recommender systems*) [81–83]. Splošna ideja sistemov za priporočila je, da najdejo ostale uporabnike s podobnimi preferencami oz. zanimanji in te podatke uporabijo za nadaljnje priporočanje izdelkov. Naš pristop se bistveno razlikuje od sistemov za priporočanje, in sicer v tem, da poišče druge agente s podobnim *odnosom zaupanja*. To pomeni, da zaupanje dojemajo in vrednotijo na podoben način, pri čemer pa ni nujno, da imajo podobne preference in zanimanja v smislu izbire določenega tipa storitev.

Elemente kolektivnega filtriranja je v svojem pristopu uporabil Dellarocas [74]. Njegov model z metodo kolektivnega filtriranja poišče tiste agente, ki so najbližje agentu ocenjevalcu. Podobnost je izračunana glede na korelacijo preteklih ocen zaupanja do istih ocenjevanih agentov. Ocene zaupanja se nato razvrstijo v dve skupini: skupino visokih in skupino nizkih ocen. Visoke ocene zaupanja se obravnavajo kot zlagane, tj. posredovane s strani agentov napadalcev s hvaljenjem, in se izločijo iz nadaljnega vrednotenja zaupanja. V Dellarocasovem modelu sta agenta podobna, če sta imela v preteklosti transakcije z istimi agenti. Podobnost med agenti v modelu zaupanja QA-DE se nanaša na primerjavo splošnih miselnosti agentov do vrednotenja zaupanja, v čemer se naš model bistveno razlikuje od modela, ki ga je predlagal Dellarocas.

Pristope iz metode kolektivnega filtriranja sta uporabila tudi Chen in Singh [71]. Definirala sta model, ki izračuna zaupanje do agenta v treh korakih. Najprej se izračunata kakovost in verodostojnost ocen oz. mnenj, ki jih je posredoval ocenjevani agent. Kakovost in verodostojnost se izračunata za oceno vsakega ocenjenega objekta znotraj določene skupine; ti vrednosti sta poimenovani kot *lokalno ujemanje* in *lokalna verodostojnost*. V naslednjem koraku se ocene kakovosti in verodostojnosti mnenj za vsako skupino združijo in izračunata se t. i. *globalno ujemanje* in *globalna verodostojnost*. Zaupanje do agenta se nato izračuna glede na vrednosti globalnih ujemanj in globalnih

verodostojnosti vseh skupin. Ocene, ki jih posredujejo agenti z nižjo oceno zaupanja, se upoštevajo zadržano, z nižjo utežjo. V nasprotju z modelom QADE je njun model zaupanja in ugleda, vključujoč metodo za prepoznavanje zlaganih mnenj, popolnoma objektivni in ne upošteva človeških (subjektivnih) dejavnikov fenomena zaupanja.

Na podobnosti temelji tudi model zaupanja in ugleda, ki so ga predstavili Tavakoli-fard in drugi [84]. Podobnost je definirana na osnovi omrežja agentov, in sicer v tradicionalnem smislu “prijatelj” in “prijatelj-od-prijatelj” (angl. friend-of-a-friend, FOAF). Tradicionalno omrežje FOAF so razširili z definicijo, da sta agenta podobna, če bodisi zaupata istim agentom (tj. imata iste prijatelje) bodisi jima zaupajo isti agenti (tj. sta prijatelja istih agentov). Kljub uporabi koncepta podobnosti je njihov pristop modeliranja zaupanja bistveno drugačen od našega. V modelu zaupanja QADE se podobnost izračuna na podlagi primerjave parov ocen zaupanja in primerjave splošnih miselnosti agentov, medtem ko ocen zaupanja *do* agenta ocenjevalca ne upoštevamo pri izračunu podobnosti.

Whitby in soavtorji [46] so predlagali model zaupanja in ugleda, ki temelji na sistemu ugleda BRS (angl. Beta Reputation System, tudi Bayesian Reputation System). Model BRS so razširili s statistično filtrirno metodo za izločevanje nepoštenih ocen zaupanja. Metoda izloči tiste vrednosti zaupanja, ki so zunaj  $q$ -kvantila in zunaj  $(1 - q)$ -kvantila glede porazdelitev vseh ocen zaupanja do ocenjevanega agenta. Ocene zaupanja, ki jih posredujejo agenti v skupnosti, so formalizirane s porazdelitveno funkcijo beta (glej razlago v razdelku 2.2.3). Ključna pomanjkljivost Whitbyjevega modela je, da predlagani model ne domneva, da je bolj pozitivna ali bolj negativna ocena zaupanja lahko tudi odraz drugačne splošne miselnosti agenta ocenjevalca in ne pomeni nujno zlagane vrednosti. Predlagani model zaupanja QADE in metoda za izločanje zlaganih mnenj QADE-filter v nasprotju z modelom, ki so ga predlagali Whitby in soavtorji, upošteva družbeno-sociološke faktorje zaupanja in nima teh omejitev. Metodo za izločanje zlaganih mnenj Whitby filter smo kot reprezentativno endogeno metodo v prejšnjem poglavju tudi kvantitativno primerjali z metodo QADE-filter.

Na beta porazdelitveni funkciji temelji tudi model TRAVOS, ki so ga definirali Teacy in drugi [31, 85]. Model TRAVOS vključuje metodo za izločanje zlaganih ocen zaupanja, ki deluje v dveh korakih. Agent ocenjevalec najprej oceni verjetnost, da so posredovane ocene zaupanja točne. Verjetnost je ocenjena na primerjavi beta porazdelitvenih funkcij preteklih ocen zaupanja agenta ocenjevalca in agenta svetovalca, ki je ocene posredoval. V drugem koraku agent ocenjevalec modificira posredovane

vrednosti glede na ocenjeno verjetnost točnosti ocen. Čeprav je predlagana metoda TRAVOS-filter bolj dovršena in učinkovita v primerjavi z metodo Whitby filter, je glavna pomanjkljivost modela TRAVOS neupoštevanje družbeno-kognitivnih procesov, ki vplivajo na vrednotenje zaupanja. V nasprotju z njihovim pristopom, predlagani model zaupanja QADE to upošteva in temu primerno obravnava (izloča) posredovane ocene zaupanja.

Omenjene družbene, psihološke in sociološke dejavnike, ki niso dovolj dobro naslovljeni v opisanih modelih zaupanja in upoštevani pri predlaganih metodah za obvladovanje zlaganih ocen, sta pri modeliranju zaupanja v veliki meri upoštevala Castelfranchi in Falcone [86]. Predstavila sta kognitivni model zaupanja, v katerem zaupanje opredelita kot stanje miselnosti (angl. mental state) in kot družbeno relacijo med agenti. Kot osnovne elemente zaupanja definirata specifična prepričanja in cilje agentov. Dalje, zaupanje definirata kot povod za prenos in dodeljevanje nalog agentom. Predstavljeni model ima dobre psihološke temelje, podprte z matematično formalizacijo. Kakorkoli, avtorja v predstavljenem kognitivnem modelu zaupanja nista predpostavila in formalizirala obnašanja nepoštenih agentov in njihovega posredovanja potvorjenih ocen zaupanja drugim agentom.

Družbeno usmerjen model zaupanja in ugleda je definiral tudi Mezzetti [87]. V predlaganem modelu uvede kontekste pristojnosti. To pomeni, da ima agent pristojnost v določenem kontekstu (oziroma situaciji), znotraj katerega je zaupanja vreden posredovalec informacij in priporočil. V nasprotju s predlaganim modelom zaupanja QADE Mezzetti [87] ne modelira človeškega faktorja in ne obravnava možnosti in metod za obvladovanje ocen zaupanja, posredovanih s strani nepoštenih agentov.

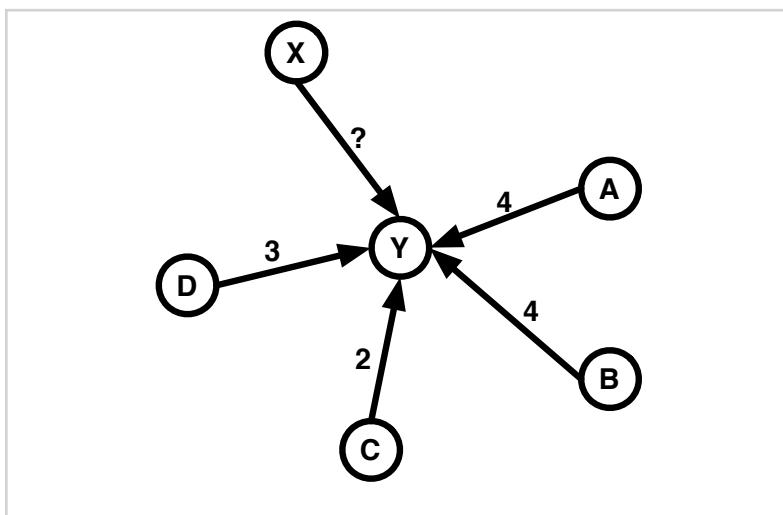
Naš pristop se razlikuje od drugih pristopov v tem, da pri zaznavanju zlaganih ocen upošteva psihološke in sociološke dejavnike zaupanja. Predlagana metoda QADE-filter metoda upošteva dejstvo, da sta različni oceni zaupanja, ki ju imata agenta ocenjevalca do ocenjevanega agenta, lahko odraz različnih procesov razumevanja oziroma vrednotenja zaupanja in ugleda drugih agentov. Metoda QADE-filter uvede primerjavo podobnosti med agenti, ki se nanaša na podobnost njihovega dojemanja odnosa zaupanja. Glede na dane lastnosti predstavlja metoda QADE-filter izviren pristop za obvladovanje zlaganih ocen zaupanja.

*Metoda za prilagoditev ocen  
zaupanja z izločitvijo  
dejavnika subjektivnosti*

5

### 5.1 Predstavitev problema

Naslednji nerešen problem na področju modeliranja zaupanja in ugleda v e-okoljih, ki ga bomo obravnavali v sklopu doktorske disertacije, je subjektivnost ocen zaupanja. Ocene zaupanja, ki si jih med seboj izmenjujejo člani e-skupnosti, lahko izgubijo svoj pravi pomen, saj je zaupanje subjektivni pojem [15, 88–91]. Zaradi različnih nagnjenj uporabnikov, ki so ocene podali, so te lahko v neskladju, čeprav naj bi pomenile isto. Na primer: Agent  $X$  se odloča, ali bo naredil transakcijo z agentom  $Y$ . V ta namen uporabi sistem za obvladovanje zaupanja in ugleda, prek katerega preveri, kakšne ocene zaupanja do  $Y$  so podali drugi agenti v danem e-okolju. Agenti  $A$ ,  $B$ ,  $C$  in  $D$  so podali ocene zaupanja, kot je prikazano na sliki 5.1. Primer bomo v nadaljevanju nadgradili z uporabo informacij iz sistema za obvladovanje zaupanja in ugleda, ki ga podpira spletišče za recenzije Yelp [62], zato imajo ocene zaupanja diskretne vrednosti med 1 (zaupanja nevreden agent) in 5 (zaupanja vreden agent).



Slika 5.1

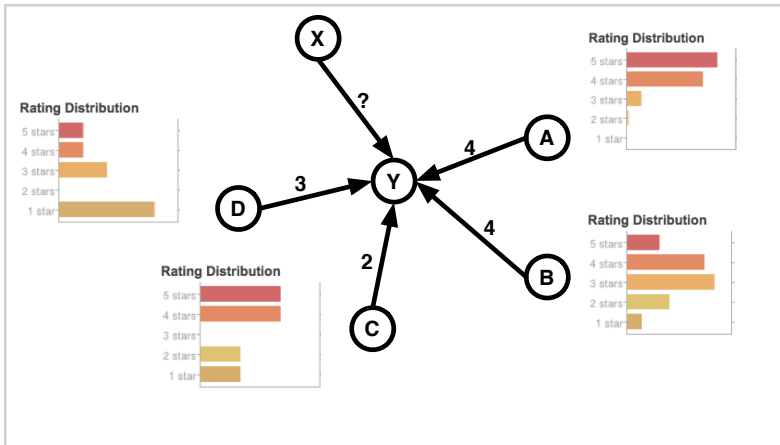
Primer: Ocene zaupanja do agenta  $Y$ , ki so jih podali agenti  $A$ ,  $B$ ,  $C$  in  $D$ .

Na prvi pogled se zdi, da:

- agenta  $A$  in  $B$  razmeroma zaupata agentu  $Y$ ;
- je agent  $C$  izrazil izkušnjo z  $Y$  kot precej zaupanja nevredno;

- je izkušnja agenta *D* z agentom *Y* “nekje v sredini”.

Vendar pa so ocene zaupanja subjektivne, kar je treba upoštevati pri združevanju teh [88, 89]. Zato naj še nadalje analiziramo podane ocene zaupanja glede na to, kako posamezen agent ocenjuje zaupanje *v splošnem*. Slika 5.2 prikazuje porazdelitev ocen zaupanja agentov *A, B, C* in *D*.<sup>1</sup>



Slika 5.2

Primer: Ocene zaupanja do uporabnika *Y*, ki so jih podali uporabniki *A, B, C* in *D*, skupaj s porazdelitvijo ocen zaupanja.

Sklepamo lahko, da je agent *A* nagnjen k ocenjevanju zaupanja z visokimi vrednostmi. Zato ocena zaupanja z vrednostjo 4 nakazuje, da je *A* ocenil *Y* kot *povprečno* zaupanja vrednega. Najpogostejša ocena agenta *B* je 3, kar pomeni, da izkušnje z drugimi agenti večinoma ocenjuje kot povprečne – ne dobre ne slabe. Vrednost 4, s katero je ocenil zaupanje do agenta *Y*, pomeni, da *B* dojema agenta *Y* kot (*rahlo*) *nadpovprečno* zaupanja vrednega v njegovem pogledu. Njegova ocena zaupanja do *Y* z vrednostjo 4 ima drugačen pomen kot ocena zaupanja, ki jo je podal uporabnik *A* in ki ima prav tako vrednost 4. Nagnjenje k zaupanju agenta *C* je skrajno v pozitivno in negativno smer. Agent *C* namreč ocenjuje zaupanje drugih agentov kot pozitivno ali negativno in nikoli kot povprečno. Ocena zaupanja z vrednostjo 2 nakazuje, da je agent *C* ocenil agenta *Y* slabo, vendar spet *ne tako zelo slabo*. Značaj agenta *D* je precej pesimističen

<sup>1</sup>Porazdelitve ocen zaupanja na sliki 5.2 so porazdelitve ocen zaupanja naključnih uporabnikov Yelp spletne skupnosti. Pridobljeno na <http://www.yelp.com>. V pričujočem primeru agente ocenjevalce predstavljajo človeški uporabniki sistema, ocenjevani agenti pa so storitve, za katere so uporabniki podali oceno in recenzijo.

in zato ocena zaupanja 3 pomeni, da agent  $D$  ocenjuje agenta  $Y$  kot *v splošnem zaupanja vrednega*. Predlagana razlaga ocen zaupanja seveda ni edina možna. Služi kot ponazoritev problema in poudarja potrebo po nadaljnji analizi in personalizaciji ocen, ki bodo ustrezale vsakemu agentu posebej.

Do sedaj problem subjektivnosti ocen zaupanja še ni bil ustrezno obravnavan. Kot rešitev problema predlagamo inovativno metodo za prilagoditev ocen zaupanja, imenovano HOMRA (angl. Human-Oriented Method for Ratings Adaptation), in v ta namen osnovni model zaupanja QAD razširimo z novimi konstrukti. Prepričljiv model zaupanja in ugleda mora namreč upoštevati tudi aspekt subjektivnosti ocen zaupanja in ocene ustrezno prilagoditi.

## 5.2 Predlagana rešitev

### 5.2.1 Model zaupanja in ugleda QADES

Model zaupanja QADES (QAD Extended with Subjectivity) predstavlja formalizacijo zaupanja in ugleda, ki temelji na modelu kvalitativne dinamike ocenjevanja QAD (glej opis v poglavju 3.1) in njegovih razširitvah (glej opis v poglavju 4.2.1). V sklopu formalizacije zaupanja in ugleda QADES bomo podali še nekatere dopolnilne elemente, ki se sicer ne nanašajo neposredno na modeliranje zaupanja in ugleda, a jih bomo uporabili pri ovrednotenju modela in ovrednotenju predlagane rešitve za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti. Taki elementi so dodatne lastnosti agenta (npr. dobrohotnost in lastništvo denarja), formalizacija transakcij med agenti, ki implementirajo igro investicij ipd.

Lastnosti agenta  $a_i$  opišemo s parom  $a_i = (m, b)$ . Element  $m \in \mathbb{R}$  predstavlja agentov znesek denarja. Agenti v e-okolju opravljajo različne transakcije in interakcije drug z drugim. Znesek denarja se po vsaki transakciji, ki jo agenti opravi, lahko poveča ali zmanjša. Element  $b \in [0, 1]$  predstavlja dobrohotnost agenta in se skozi čas ne spreminja. Vrednost  $b = 1$  pomeni, da je dobrohotnost agenta na najvišji možni ravni. Tak agent zagotavlja dobre storitve (tj. transakcije z zadovoljivim izidom), s katerimi bodo partnerji zadovoljni. Nasprotno, agent s stopnjo dobrohotnosti  $b = 0$  priskrbi slabe storitve, nad katerimi bo agent, ki je dano storitev zahteval, najverjetneje razočaran.

Model QAD kot enega izmed ključnih dejavnikov zaupanja definira tudi *Razlikovanje zaupanja* (glej opis v poglavju 3.1). Razlikovanje zaupanja se nanaša na različno

izražanje in vrednotenje odnosa zaupanja. Poleg različnih sposobnosti izražanja agentov je eden izmed povodov za razlikovanje zaupanja tudi subjektivnost pojmovanja zaupanja. V modelu zaupanja in ugleda QADES bomo poudarili lastno subjektivno dojetje zaupanja agentov in s tem namenom iz procesa ocenjevanja zaupanja izključili vse zunanje dejavnike. *Zunanje dejavnike*, ki predstavljajo vpliv mnenja družbe in se odražajo prek posredovanih ocen zaupanja agentov svetovalcev, pa bomo upoštevali pri definiciji ugleda in v predlagani metodi za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti.

V modelu zaupanja in ugleda QADES je zaupanje “odnos med agentom ocenjevalcem in ocenjevanim agentom” (definicija 1). Poudariti velja, da agent ocenjevalec *subjektivno* oceni stopnjo zaupanja do ocenjevanega agenta. Ocena zaupanja temelji na njegovem lastnem dojetanju izkušenj. Agenti ne poznajo dejavnikov, meril ali povodov, na podlagi katerih ostali agenti v danem e-okolju ocenjujejo vrednosti zaupanja. Različna agenta tako lahko ocenita zaupanje do istega agenta različno, kljub temu da se ta do obeh obnaša popolnoma enako in obema nudi identične storitve. Predpostavlja se namreč, da lahko agenti zaupanje dojemajo na različne načine. Na primer, agent  $a_x$  oceni agenta  $a_z$  kot delno zaupanja vrednega ( $\omega_{x,z} = 1$ ), medtem ko agent  $a_y$  istega agenta za popolnoma enako storitev oceni kot delno zaupanja nevrednega ( $\omega_{y,z} = -1$ ). Oceni zaupanja do agenta  $a_z$  se razlikujeta zaradi različnega dojetanja zaupanja agentov  $a_x$  in  $a_y$ , ki sta oceni podala.

Ocene zaupanja agentov v danem e-okolju so predstavljene v matriki zaupanja  $\mathcal{M}$  (definicija 2). Prek matrike zaupanja se odraža *splošno mnenje skupnosti* o zaupanju posameznim agentom – ugled agentov. Ugled agenta je izpeljan z združevanjem ocen zaupanja agentov, ki so prisotni v danem e-okolju, do izbranega agenta. Ocene zaupanja do agenta  $a_k$ , ki so jih posredovali v matriko zaupanja ostali agenti v skupnosti, so shranjene v podvektorju ugleda agenta  $\underline{\mathcal{M}}_{n_k,k}$  (definicija 4). Izhajajoč iz tega, model QADES definira *ugled*, kot sledi:

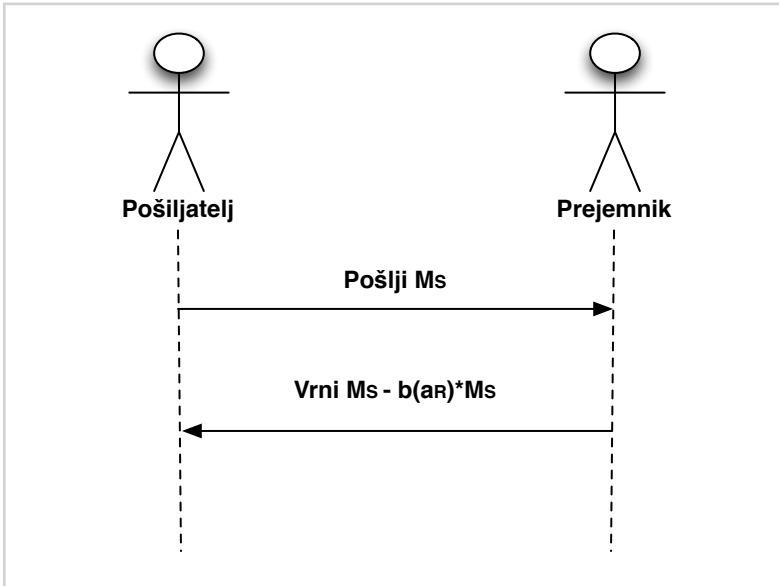
*Definicija 21:* Ugled agenta  $a_k$  je poenoteno mnenje agentov iz skupnosti o vrednosti zaupanja do agenta  $a_k$ . Ugled agenta je definiran kot povprečna vrednost elementov podvektorja ugleda:  $\rho(a_k) = \text{avg}(\underline{\mathcal{M}}_{n_k,k}) \in [-2, 2]$ .

V podvektorju ugleda agenta  $a_k$  (definicija 4) so shranjene ocene zaupanja, ki so jih podali agenti, ki so naredili transakcijo oziroma so na podlagi drugih informacij ocenili

vrednost zaupanja do agenta  $a_k$ . V odprtih spletnih okoljih je mnenje agentov v obliki ocene zaupanja pogosto edini vir informacij, na podlagi katerega lahko ostali agenti v sistemu ocenijo vrednost zaupanja drug do drugega. Sistemi za obvladovanje zaupanja in ugleda najpogosteje agentu vrnejo samo končne ocene zaupanja, ki so jih posredovali agenti svetovalci, tj. brez kriterijev za ocenjevanje, kot so na primer čas dostave storitve, cena, zanesljivost ipd. Glavna naloga sistemov za obvladovanje zaupanja in ugleda v e-okoljih je, da združijo ocene iz vektorja ugleda in s tem izračunajo ugled agentov. Za združevanje ocen so bili predlagani različni pristopi – nekateri med njimi so opisani v poglavju 2.2. Najpogostejši pristop v komercialnih sistemih za obvladovanje zaupanja in ugleda, katerih tipični predstavniki so spletni sistemi za ocenjevanje (angl. online rating systems), je izračun ugleda kot povprečje vseh ocen [1, 39, 92]. Iz tega razloga smo tudi v našem modelu ugled agentov definirali s povprečjem ocen iz podvektorja ugleda. S tem je metoda za prilagoditev ocen zaupanja, ki je opisana v nadaljevanju poglavja, splošno uporabna in obenem primerna za izboljšavo komercialnih sistemov za obvladovanje zaupanja in ugleda.

Transakcije med agenti modeliramo s pomočjo pravil igre investicij (angl. investment game) [93], ki jo v literaturi pogosto imenujejo tudi igra zaupanja (angl. trust game). Igro investicij se pogosto uporablja za preučevanje razvoja zaupanja in ugleda v različnih skupnostih, vključujoč skupnosti v e-okoljih [94–96]. Igra investicij je zasnovana tako, da omogoča nadaljnje razširitve in modifikacije igre. Zato je primerna za poskuse, povezane s preučevanjem zaupanja in ugleda, zlasti v kombinaciji s koristnostjo, dobičkom, vzorcem obnašanja in drugimi zanimivimi dejavniki e-okolja.

Igro investicij igrata agenta  $a_S$  in  $a_R$ . Agent  $a_S$  ima vlogo pošiljatelja (angl. sender), agent  $a_R$  pa vlogo prejemnika (angl. receiver). V prvem delu igre pošiljatelj  $a_S$  prejemniku  $a_R$  pošlje  $M_S$  denarnih enot. Po pravilih igre investicij [93] se poslani znesek potroji in prejemnik prejeme trikratni znesek poslanega denarja:  $3M_S$ . Pravila igre investicij bomo za namene modeliranja transakcij med agenti v našem primeru modificirali. Namesto potrojenega zneska, tj.  $3M_S$ , v modificirani različici igre investicij prejemnik prejeme poslani znesek denarja:  $M_S$ . S tem smo dosegli, da imamo na voljo omejeno količino denarja, ki je razporejena med agenti. Tak pristop nam omogoča, da lahko razlikujemo uspešne in manj uspešne agente. V drugem delu igre se prejemnik odloči, kolikšen znesek (med 0 in  $M_S$ ) bo poslal nazaj pošiljatelju. V naši različici igre investicij je to odvisno od dobrohotnosti prejemnika: prejemnik pošlje pošiljatelju znesek v višini  $b(a_R) * (M_S)$ , pri čemer  $b(a_R)$  označuje stopnjo dobrohotnosti agenta



Slika 5.3

Grafični prikaz poteka igre investicij.

prejemnika.

Izkupiček (angl. payoff) agenta pošiljatelja  $a_S$  je:

$$PO_S(M_S, b(a_R)) = -M_S + b(a_R) * (M_S);$$

izkupiček agenta prejemnika  $a_R$  pa:

$$PO_R(M_S, b(a_R)) = M_S - b(a_R) * (M_S).$$

Potek igre investicij je grafično prikazan na sliki 5.3.

Po opravljeni transakciji pošiljatelj oceni stopnjo zaupanja prejemnika. V danem kontekstu je pošiljatelj agent ocenjevalec, prejemnik pa ocenjevani agent. Agent ocenjevalec ovrednoti zaupanje do ocenjevanega agenta glede na izkupiček v dani transakciji. Poleg tega je ocena zaupanja odvisna tudi od "tipa osebnosti" agenta ocenjevalca in njegovega nagnjenja k zaupanju.

V modelu zaupanja in ugleda QADES uvedemo tipe osebnosti, ki predstavljajo modifikacijo operatorjev zaupanja QAD (definicija 5). Lastnosti operatorjev QAD so

modificirane tako, da so iz procesa vrednotenja izvzeti zunanji dejavniki, ki se nanašajo na vpliv mnenja skupnosti pri evalvaciji zaupanja. Zunanje dejavnike smo izključili, saj želimo dodatno poudariti vpliv subjektivnosti v ocenah zaupanja.

*Definicija 22:* Naj bo  $QoI_{i,j} \in [0, 1]$  objektivna kakovost izida interakcije agenta  $a_i$  z agentom  $a_j$ . Tipe osebnosti agentov definiramo z množico  $\Psi = \{\uparrow, \uparrow, \leftrightarrow, \downarrow, \downarrow, \rightsquigarrow\}$ , pri čemer simboli predstavljajo optimističen, zmerno optimističen, centralističen, zmerno pesimističen, pesimističen in ekstremističen tip osebnosti. Tip osebnosti je funkcija  $\psi \in \Psi$ , ki definira vrednotenje zaupanja agenta  $a_i$  do agenta  $a_j$  glede na objektivno kakovost izida transakcije. Funkcije tipov osebnosti agentov so definirane, kot sledi:

$$\uparrow_i: \omega_{i,j}(QoI_{i,j}) = \begin{cases} 0; & QoI_{i,j} < 0,25 \\ 1; & 0,25 \leq QoI_{i,j} < 0,50 \\ 2; & 0,50 \leq QoI_{i,j} \end{cases}$$

$$\uparrow_i: \omega_{i,j}(QoI_{i,j}) = \begin{cases} -2; & QoI_{i,j} < 0,10 \\ -1; & 0,10 \leq QoI_{i,j} < 0,30 \\ 0; & 0,30 \leq QoI_{i,j} < 0,50 \\ 1; & 0,50 \leq QoI_{i,j} < 0,70 \\ 2; & 0,70 \leq QoI_{i,j} \end{cases}$$

$$\leftrightarrow_i: \omega_{i,j}(QoI_{i,j}) = \begin{cases} -2; & QoI_{i,j} < 0,20 \\ -1; & 0,20 \leq QoI_{i,j} < 0,40 \\ 0; & 0,40 \leq QoI_{i,j} < 0,60 \\ 1; & 0,60 \leq QoI_{i,j} < 0,80 \\ 2; & 0,80 \leq QoI_{i,j} \end{cases}$$

$$\downarrow_i: \omega_{i,j}(QoI_{i,j}) = \begin{cases} -2; & QoI_{i,j} < 0,30 \\ -1; & 0,30 \leq QoI_{i,j} < 0,50 \\ 0; & 0,50 \leq QoI_{i,j} < 0,70 \\ 1; & 0,70 \leq QoI_{i,j} < 0,90 \\ 2; & 0,90 \leq QoI_{i,j} \end{cases}$$

$$\Downarrow_i: \omega_{ij}(QoI_{ij}) = \begin{cases} -2; & QoI_{ij} < 0,50 \\ -1; & 0,50 \leq QoI_{ij} < 0,75 \\ 0; & 0,75 \leq QoI_{ij} \end{cases}$$

$$\rightsquigarrow_i: \omega_{ij}(QoI_{ij}) = \begin{cases} -2; & QoI_{ij} < 0,25 \\ -1; & 0,25 \leq QoI_{ij} < 0,50 \\ 1; & 0,50 \leq QoI_{ij} < 0,75 \\ 2; & 0,75 \leq QoI_{ij} \end{cases}$$

Iz definicij tipov osebnosti vidimo, da so ocene zaupanja, s katerimi agenti ocenjevalci ocenijo zaupanje do ocenjevanih agentov, subjektivne in odražajo različno dojemanje zaupanja agentov ocenjevalcev. Da bi vsi agenti enako razumeli ocene, je treba posredovane ocene zaupanja ustrezno prilagoditi vsakemu agentu posebej. V ta namen smo predlagali novo metodo za prilagoditev ocen zaupanja, imenovano HOMRA (angl. Human-Oriented Method for Ratings Adaptation), ki je opisana v naslednjem poglavju. Cilj predlagane metode je, da iz ocen zaupanja izloči faktor subjektivnosti in jih prilagodi, tako da ustrezajo nagnjenju k zaupanju agenta, ki te ocene prejme in jih uporabi za nadaljnje vrednotenje zaupanja in ugleda.

Naša rešitev temelji na primerjavi nagnjenj k zaupanju med agenti. V ta namen mora vsak agent najprej določiti, kakšno je njegovo nagnjenje oziroma *dojemanje zaupanja*, preden ga primerja s stališči ostalih agentov. V danem kontekstu agenti tipično predstavljajo človeške uporabnike sistema. Določanje agentovega dojemanja zaupanja poteka na različne načine [88]:

- Agent sam zase določi svoje dojemanje zaupanja. Tak pristop je enostaven, vendar zahteva, da agent eksplicitno sodeluje v procesu določanja dojemanja zaupanja. Poleg tega je vprašljivo, ali lahko agent sam zase pravilno oceni, kako (npr. pesimistično, optimistično, realistično) usmerjen je.
- Agent izpolni vprašalnik, ki je zasnovan za presojo oziroma ugotavljanje vedenjskih lastnosti, kot so test osebnosti Briggs-Myers [97], t. i. Interpersonal Trust Scale [98] in drugi testi [99]. Tudi ta pristop zahteva eksplicitno sodelovanje agenta v procesu določanja dojemanja zaupanja. Reševanje vprašalnikov je časovno potratno, zaradi česar je možen odpor agentov do sodelovanja.

- Agentovo dojetanje zaupanja izpeljemo iz vzorca njegovih ocen zaupanja do ostalih agentov v e-okolju. Zgodovinski podatki se pogosto uporabljajo za določanje vzorca obnašanja uporabnikov in napovedi njihovih prihodnjih dejanj, npr. v [100–103]. Eksplicitno sodelovanje agentov pri tem pristopu ni potrebno. Rezultat določitve agentovega dojetanja zaupanja, ki je pridobljen izključno na podlagi njegovih ocen zaupanja, je tako odvisen samo od metode, s katero so podatki obdelani.

Predlagana metoda za prilagoditev ocen zaupanja HOMRA izpelje agentove značilnosti dojetanja zaupanja na podlagi ocen, s katerimi je ocenil druge agente in so shranjene v njegovem vektorju ocen zaupanja (definicija 3). Izpeljava dojetanja zaupanja agenta temelji na porazdelitvi podanih ocen zaupanja. Dojetanje zaupanja je definirano, kot sledi:

*Definicija 23:* Dojetanje zaupanja agenta  $a_i$  označimo s  $F_{i,n_i}(\omega)$ . Predstavljeno je z empirično kumulativno porazdelitveno funkcijo elementov iz podvektorja ocen zaupanja  $\underline{\mathcal{M}}_{i,n_i}$  in je definirano kot  $F_{i,n_i}(\omega) = \frac{1}{n} \sum_{j=1}^{n_i} (I(\omega_{i,j}) \leq \omega)$ , kjer je  $(I(\omega_{i,j}) \leq \omega)$  indikatorska funkcija, ki vrne vrednost 1, če je  $(\omega_{i,j}) \leq \omega$ , in sicer vrednost 0.

Definicijo 23 bomo ponazorili z naslednjim primerom. V skupnosti imamo agente  $a_1, a_2, a_3$  in  $a_4$  s pripadajočimi vektorji ocen zaupanja:

- $\underline{\mathcal{M}}_{1,n_1} = [-2, -2, 1, 1]$ ,
- $\underline{\mathcal{M}}_{2,n_2} = [-2, 0, 1, 0]$ ,
- $\underline{\mathcal{M}}_{3,n_3} = [-2, 2, 2, 2]$ ,
- $\underline{\mathcal{M}}_{4,n_4} = [1, 0, 0, 1]$ .

Iz danih podatkov izračunamo empirične kumulativne porazdelitvene funkcije elementov iz podvektorjev ocen zaupanja, ki predstavljajo dojetanje zaupanja agentov:

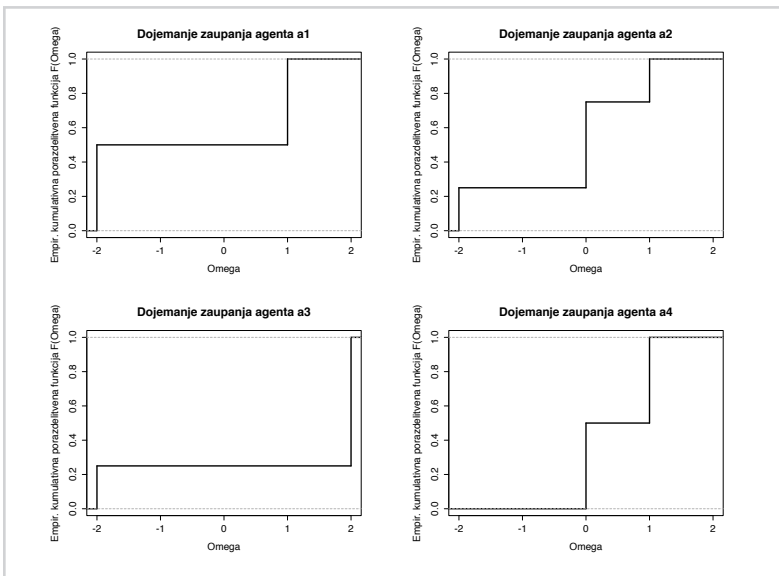
$$F_{1,n_1}(\omega) = \begin{cases} 0 & \text{za } x < -2 \\ 0,5 & \text{za } -2 \leq x < 1 \\ 1,0 & \text{za } x \geq 1 \end{cases}$$

$$F_{2,n_2}(\omega) = \begin{cases} 0 & \text{za } x < -2 \\ 0,25 & \text{za } -2 \leq x < 0 \\ 0,75 & \text{za } 0 \leq x < 1 \\ 1,0 & \text{za } x \geq 1 \end{cases}$$

$$F_{3,n_3}(\omega) = \begin{cases} 0 & \text{za } x < -2 \\ 0,25 & \text{za } -2 \leq x < 2 \\ 1,0 & \text{za } x \geq 2 \end{cases}$$

$$F_{4,n_4}(\omega) = \begin{cases} 0 & \text{za } x < 0 \\ 0,5 & \text{za } 0 \leq x < 1 \\ 1,0 & \text{za } x \geq 1 \end{cases}$$

Dojemanje zaupanja agentov  $a_1$ ,  $a_2$ ,  $a_3$  in  $a_4$  iz danega primera je grafično prikazano na sliki 5.4.



Slika 5.4

Primer: Dojemanje zaupanja agentov  $a_1$ ,  $a_2$ ,  $a_3$  in  $a_4$ .

V naši predlagani rešitvi agentov s podobnimi dojemANJI zaupanja ne razvrščamo v skupine (kot predlagajo nekateri pristopi, npr. [10, 104]), ampak vsakega od njih

obravnavamo posamično. Na ta način lahko dojezanje zaupanja posameznega agenta primerjamo z dojezanjem katerega koli drugega agenta povsem neodvisno, in ne samo glede na pripadnost agentov določeni skupini.

Primerjava dojezanj zaupanja dveh agentov je primerjava dveh neodvisnih porazdelitev ocen zaupanja. Za primerjavo porazdelitev podatkov obstaja mnogo statističnih testov. Ker agenti v e-okoljih tipično ne poznajo parametrov porazdelitev ocen zaupanja drugih agentov, za primerjavo dojezanj zaupanja agentov predlagamo uporabo neparametričnih statističnih testov. Pri testiranju z neparametričnimi statističnimi testi namreč ni predpostavk o vrsti porazdelitve ali velikosti vzorca.

Najbolj pogosto uporabljani neparametrični testi za določanje, če sta porazdelitvi podatkov enaki, so Wilcoxonov test z rangi, Mann-Whitneyev test, Kruskal-Wallisov test in Kolmogorov-Smirnov test za dva vzorca [105–109]. Wilcoxonov test z rangi se uporablja za primerjavo dveh odvisnih vzorcev podatkov in zato ni primeren za primerjavo dojezanj zaupanja dveh agentov, saj gre v tem primeru za *neodvisne* podatke. Kruskal-Wallisov test je sicer primeren za primerjavo neodvisnih vzorcev podatkov, vendar pa predpostavlja, da imata vzorca enako obliko porazdelitev – razen morebitnih razlik v mediani. Navadno se uporablja za primerjavo vsaj treh vzorcev. Ker ne moremo predpostavljati, da bosta porazdelitvi ocen zaupanja dveh agentov imeli enako obliko, tudi ta test ni primeren za primerjavo dojezanj zaupanja. Mann-Whitneyev test izračuna razliko med dvema statistično neodvisnima populacijama in temelji na razliki med rangi obeh populacij. Kot tak je test sicer primeren za primerjavo dojezanj zaupanja dveh agentov, vendar je Kolmogorov-Smirnov test za dva vzorca bolj zmožljiv. Kolmogorov-Smirnov test za dva vzorca testira, kakšne so splošne razlike med dvema porazdelitvama in ne temelji samo na razlikah v povprečjih. Kolmogorov-Smirnov test za dva vzorca tako prepozna tudi razlike v obliki obeh porazdelitev, kot so razpršenost, asimetrija ipd.

Statistiko Kolmogorov-Smirnovega testa za dva vzorca uporabimo za merjenje *podobnosti dojezanja zaupanja* (angl. *trust dispositions similarity*, *tds*) dveh agentov.

*Definicija 24:* Naj bosta  $F_{i,n_i}(\omega)$  in  $F_{j,n_j}(\omega)$  dojezanja zaupanja agentov  $a_i$  in  $a_j$ . Podobnost dojezanja zaupanja agentov  $a_i$  in  $a_j$  definiramo s funkcijo  $tds(a_i, a_j) = 1 - \sup_{\omega} |F_{i,n_i}(\omega) - F_{j,n_j}(\omega)|$ , kjer je  $tds(a_i, a_j) = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ .

Višja vrednost funkcije  $tds(a_i, a_j)$  pomeni večjo podobnost med dvema agentoma

glede na primerjavo njunih dojemanj zaupanja, nižja vrednost pa manjšo podobnost dveh agentov. Kolmogorov-Smirnov test za dva vzorca izračuna največjo vertikalno diferenco med primerjanima empiričnima kumulativnima porazdelitvama. Kot primer bomo izračunali podobnost dojemanja agenta  $a_1$  z agenti  $a_2$ ,  $a_3$  in  $a_4$  (iz prejšnjega primera):

- $tds(a_1, a_2) = 1 - 0,25 = 0,75$ ;
- $tds(a_1, a_3) = 1 - 0,75 = 0,25$ ;
- $tds(a_1, a_4) = 1 - 0,50 = 0,50$ .

Rezultati funkcij

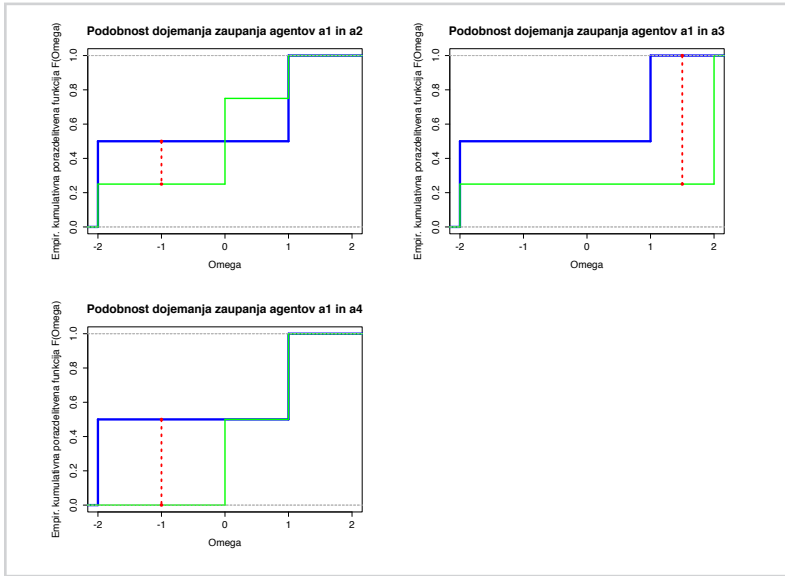
$$\begin{aligned} & \sup_{\omega} |F_{1,n_1}(\omega) - F_{2,n_2}(\omega)|, \\ & \sup_{\omega} |F_{1,n_1}(\omega) - F_{3,n_3}(\omega)| \text{ in} \\ & \sup_{\omega} |F_{1,n_1}(\omega) - F_{4,n_4}(\omega)| \end{aligned}$$

so grafično prikazani na sliki 5.5.

Agenti s podobnimi dojemANJI zaupanja podobno vrednotijo zaupanje do drugih agentov. *Razlikovanje zaupanja* – eden izmed glavnih dejavnikov, povezanih s pojmom zaupanja – je tako manjše, če primerjamo ocene zaupanja, ki so jih podali agenti s podobnimi dojemANJI zaupanja. Podobnost dojemanja zaupanja uporabimo za prilagoditev vektorja ugleda posameznim agentom:

*Definicija 25:*  $Z \overleftarrow{\mathcal{M}}_{[n_k],k}^i = [\omega_{[1],k}^i, \omega_{[2],k}^i, \dots, \omega_{[n_k],k}^i]$  označimo urejen vektor ugleda, tako da velja  $\forall [p] < [r] : tds(a_i, a_{[p]}) \geq tds(a_i, a_{[r]})$ . Vektor  $\overleftarrow{\mathcal{M}}_{[n_s],k}^i = [\omega_{[1],k}^i, \omega_{[2],k}^i, \dots, \omega_{[n_s],k}^i], \forall \omega_{[j],k}^i : [j] \geq [tdsTh]$  označuje agentu  $a_i$  prilagojen vektor ugleda agenta  $a_k$ .

Agentu  $a_i$  prilagojen vektor ugleda agenta  $a_k$  vsebuje samo tiste ocene zaupanja do agenta  $a_k$ , ki so jih podali agenti, ki zaupanje dojemajo podobno kot agent  $a_i$ . Prilagojen vektor ugleda na ta način vsebuje  $tdsTh$  ocen zaupanja do agenta  $a_k$ . Vrednost  $tdsTh$  lahko predstavlja absolutno število najbolj podobnih agentov (npr. prilagojen



Slika 5.5

Primer: Podobnost dojenja zaupanja agenta  $a_1$  z agenti  $a_2$ ,  $a_3$  in  $a_4$ .

vektor ugleda vsebuje ocene zaupanja od 50 agentov, ki imajo najbolj podobno dojenje zaupanja) ali relativno število (npr. prilagojen vektor ugleda vsebuje ocene zaupanja od 20 % najbolj podobnih agentov).

Na primer, agent  $a_1$  (iz prejšnjega primera) ima najbolj podobno dojenje zaupanja z agentom  $a_2$  ( $t ds(a_1, a_2) = 0,75$ ), sledita mu  $a_4$  ( $t ds(a_1, a_4) = 0,50$ ) in  $a_3$  ( $t ds(a_1, a_3) = 0,25$ ). Iz definicije funkcije podobnosti dojenja zaupanja sledi, da je  $t ds(a_1, a_1) = 1$ . Izračunali bomo agentu  $a_1$  prilagojen vektor ugleda agenta  $a_2$ . Ocene zaupanja, ki so jih podali agenti o  $a_2$ , so:  $\omega_{1,2} = -2$ ,  $\omega_{2,2} = 0$ ,  $\omega_{3,2} = 2$  in  $\omega_{4,2} = 0$ . Urejen vektor ugleda ima vrednosti  $\overleftarrow{\mathcal{M}}_{[n_2],2}^1 = [\omega_{1,2}^1, \omega_{2,2}^1, \omega_{4,2}^1, \omega_{3,2}^1] = [-2, 0, 0, 2]$ . Naj bo  $t dsTh = 2$ , kar pomeni, da prilagojen vektor ugleda vsebuje 2 vrednosti. Agentu  $a_1$  prilagojen vektor ugleda agenta  $a_2$  je tako  $\overleftarrow{\mathcal{M}}_{[n_s],2}^i = [-2, 0]$ .

Ocene zaupanja, ki so vsebovane v *prilagojenem* vektorju ugleda, se uporabijo pri izračunu prilagojene vrednosti ugleda agenta.

*Definicija 26:* Povprečna vrednost elementov vektorja  $\rho^i(a_k) = \text{avg}(\overleftarrow{\mathcal{M}}_{[n_s],k}^i)$  definiira agentu  $a_i$  prilagojeno vrednost ugleda agenta  $a_k$ .

Iz definicije 26 sledi, da je agentu  $a_1$  prilagojena vrednost ugleda agenta  $a_2$  v našem primeru enaka  $\rho^1(a_2) = \text{avg}([-2, 0]) = -1$ .

Prilagojene vrednosti ugleda agentov imajo lahko različne vrednosti, čeprav podajajo ugled za istega agenta. Vrednosti so prilagojene posameznemu agentu, ki te prejme ter uporabi glede na njegove vedenjske značilnosti in nagnjenje k zaupanju. Z uporabo vrednosti ugleda agentov, ki so prilagojene posameznim agentom v danem e-okolju, ti posredovane vrednosti ugleda razumejo na enak (ali vsaj podoben) način in s tem dobijo enakovredne možnosti pri nadaljnjem odločanju.

### 5.2.2 Metoda za prilagoditev ocen zaupanja HOMRA

Na podlagi konstruktov *prilagojenega vektorja ugleda agenta* in *prilagojene vrednosti ugleda agenta* predlagamo k človeku usmerjeno metodo za prilagoditev ocen zaupanja HOMRA (angl. Human-Oriented Method for Ratings Adaptation), kot je prikazano v algoritmu 9. Cilj metode HOMRA je izračun agentu ocenjevalcu prilagojene vrednosti ugleda ocenjevanega agenta.

Predlagana metoda najprej izračuna dojemanje zaupanja agenta ocenjevalca  $a_i$ . Dojemanje zaupanja je ocena teoretične kumulativne porazdelitvene funkcije ocen zaupanja agenta  $a_i$ , ki jo dobimo z empirično kumulativno porazdelitveno funkcijo (angl. empirical cumulative distribution function, ecdf). Empirična kumulativna porazdelitvena funkcija izračuna delež ocen zaupanja v danem podvektorju ocen zaupanja agenta  $a_i$ , ki so manjše ali enake od določene vrednosti  $\omega$ .

V naslednjem koraku metoda HOMRA izpelje dojemanje zaupanja ostalih agentov v skupnosti – agentov svetovalcev. Tudi za agente svetovalce se dojemanje zaupanja izpelje na podlagi empiričnih kumulativnih porazdelitvenih funkcij ocen zaupanja iz njihovih podvektorjev.

Po izpeljavi vrednosti dojemanj zaupanja agentov svetovalcev metoda HOMRA izračuna podobnost izpeljanih dojemanj z dojemanjem zaupanja agenta ocenjevalca  $a_i$ . Podobnost dojemanj zaupanja se izračuna kot obratna vrednost največje razlike med dojemanjem zaupanja agenta ocenjevalca in agenta svetovalca, tj. med empiričnima kumulativnima porazdelitvenima funkcijama njunih ocen zaupanja. Če sta porazdeli-

*Algoritem 9*

Izračun prilagojene vrednosti ugleda agenta (HOMRA)

---

*Vhod:* agent ocenjevalec  $a_i$ , ocenjevani agent  $a_k$

$$F_{i,n_i}(\omega) \leftarrow \text{izpelji\_dojemanje\_zaupanja}(\underline{\mathcal{M}}_{n_i,i})$$

for  $a_j \in A$  do

$$F_{j,n_j}(\omega) \leftarrow \text{izpelji\_dojemanje\_zaupanja}(\underline{\mathcal{M}}_{n_j,j})$$

$$tds(a_i, a_j) \leftarrow \text{primerjaj\_dojemanji}(F_{i,n_i}, F_{j,n_j})$$

end for

$$\underline{\mathcal{M}}_{[n_k],k}^i \leftarrow \text{razvrsti\_glede\_na\_}(tds(a_i, a_j))$$

$$\underline{\mathcal{M}}_{[n_s],k}^i \leftarrow \{\}$$

for all  $\omega_{[j],k}^i \in \underline{\mathcal{M}}_{[n_k],k}^i$  AND  $[j] \geq [tdsTh]$  do

$$\underline{\mathcal{M}}_{[n_s],k}^i \leftarrow \underline{\mathcal{M}}_{[n_s],k}^i \cup \omega_{[j],k}^i$$

end for

$$\rho^i(a_k) \leftarrow \text{izracunaj\_prilagojen\_ugled}(\underline{\mathcal{M}}_{[n_s],k}^i)$$


---

tvi ocen zaupanja agenta ocenjevalca in agenta svetovalca enaki, potem faktor podobnosti stališč zaupanja  $tds(a_i, a_j)$  konvergira proti 1. Če sta njuni stališči *dovolj* podobni, kar je določeno s pragom  $tdsTh$ , potem se ocena zaupanja agenta  $a_j$  do ocenjevanega agenta  $a_k$  shrani v agentu  $a_i$  prilagojen vektor ugleda agenta  $a_k$ :  $\underline{\mathcal{M}}_{n_k,k}^i$ .

Nazadnje metoda HOMRA izračuna agentu ocenjevalcu  $a_i$  prilagojeno vrednost ugleda ocenjevanega agenta  $a_k$ , in sicer kot povprečje vrednosti iz agentu  $a_i$  prilagojenega vektorja ugleda agenta  $a_k$ . Agent ocenjevalec s tem dobi vrednost ugleda drugega agenta, ki ustreza njegovemu osebnostnemu tipu in mu tako v nadaljnjem postopku odločanja omogoča nepristransko primerjavo agentov.

S predlagano metodo za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti HOMRA se zmanjša možnost morebitne napačne interpretacije ocen zaupanja, ki so jih podali ostali (neznani) agenti v danem e-okolju. Metoda od agentov (uporabnikov) ne zahteva, da bi ti morali svoje ocene zaupanja na kakršenkoli način uravnati ali kako drugače spremeniti. Poleg tega tudi sam proces (subjektivnega) ocenjevanja zaupanja ostane nespremenjen.

### 5.3 Ovrednotenje metode HOMRA

V prejšnjem poglavju smo opisali metodo za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti, poimenovano HOMRA. V nadaljevanju bomo predlagano metodo empirično ovrednotili. Vrednotenje bomo izvedli na podlagi večagentnih simulacij s simulacijskih orodjem Zaupnik.

#### 5.3.1 Opis simulacij

V simulacijskem okolju je  $n$  sodelujočih agentov, ki predstavljajo člane virtualne skupnosti. Na začetku je odnos zaupanja med agenti nedefiniran, tj.  $\omega_{ij} = "I"$ ,  $\forall i, j$ . Na začetku simulacije vsakemu agentu  $a_i$  določimo začetno stopnjo dobrohotnosti  $b(a_i)$ , katere vrednost ostane v času simulacije nespremenjena. Poleg tega vsakemu agentu določimo začetni znesek denarja, ki pa ga lahko agent z opravljenimi transakcijami investicijske igre v času simulacije poveča ali pomanjša. Agenti nimajo denarnih omejitev, ker v simulacijah ne bomo uporabili nobenega odločitvenega modela. Denar potrebujejo za opravljanje transakcij in (posledično) za ocenjevanje zaupanja do drugih agentov, ki temelji na izidu izvršene transakcije igre investicij.

V vsakem koraku simulacije naključno izberemo dva agenta (pošiljatelja in prejemnika), ki bosta opravila transakcijo. Vsak agent v simulaciji je lahko bodisi pošiljatelj bodisi prejemnik. Ostali agenti iz skupnosti v danem koraku ne opravijo transakcije. Po zaključeni transakciji investicijske igre agent pošiljatelj oceni zaupanje do agenta prejemnika in oceno zaupanja prek matrike zaupanja  $\mathcal{M}$  posreduje ostalim agentom. Vsak agent zaupanje vrednoti na edinstven (subjektiven) način, kot smo podali v formalizaciji zaupanja in ugleda QADES.

V sklopu formalizacije QADES smo definirali šest skupin tipov osebnosti: optimiste, zmerne optimiste, centraliste, zmerne pesimiste, pesimiste in ekstremiste. V transakcijah igre investicij izid interakcije med agentom pošiljateljem in agentom prejemnikom ( $QOI_{S,R}$ ) predstavlja izkupiček pošiljatelja v dani transakciji. Vrednost izkupička za agenta pošiljatelja je  $-M_S \leq PO_S(M_S, b(a_R)) \leq M_S$ , s čimer je določena objektivna kakovost izida interakcije. Agenti pošiljatelji z različnimi tipi osebnosti bodo zaupanje do agenta prejemnika ocenili glede na definicije tipov osebnosti (definicija 22), kot sledi:

$$\uparrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} 0 & \text{če } PO_S(M_S, b(a_R)) < -0,75M_S \\ 1 & \text{če } -0,75M_S \leq PO_S(M_S, b(a_R)) < -0,5M_S \\ 2 & \text{če } -0,5M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

$$\uparrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} -2 & \text{če } PO_S(M_S, b(a_R)) < -0,9M_S \\ -1 & \text{če } -0,9M_S \leq PO_S(M_S, b(a_R)) < -0,7M_S \\ 0 & \text{če } -0,7M_S \leq PO_S(M_S, b(a_R)) < -0,5M_S \\ 1 & \text{če } -0,5M_S \leq PO_S(M_S, b(a_R)) < -0,3M_S \\ 2 & \text{če } -0,3M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

$$\leftrightarrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} -2 & \text{če } PO_S(M_S, b(a_R)) < -0,8M_S \\ -1 & \text{če } -0,8M_S \leq PO_S(M_S, b(a_R)) < -0,6M_S \\ 0 & \text{če } -0,6M_S \leq PO_S(M_S, b(a_R)) < -0,4M_S \\ 1 & \text{če } -0,4M_S \leq PO_S(M_S, b(a_R)) < -0,2M_S \\ 2 & \text{če } -0,2M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

$$\downarrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} -2 & \text{če } PO_S(M_S, b(a_R)) < -0,7M_S \\ -1 & \text{če } -0,7M_S \leq PO_S(M_S, b(a_R)) < -0,5M_S \\ 0 & \text{če } -0,5M_S \leq PO_S(M_S, b(a_R)) < -0,3M_S \\ 1 & \text{če } -0,3M_S \leq PO_S(M_S, b(a_R)) < -0,1M_S \\ 2 & \text{če } -0,1M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

$$\downarrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} -2 & \text{če } PO_S(M_S, b(a_R)) < -0,5M_S \\ -1 & \text{če } -0,5M_S \leq PO_S(M_S, b(a_R)) < -0,25M_S \\ 0 & \text{če } -0,25M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

$$\rightsquigarrow_S: \omega_{S,R}(PO_S(M_S, b(a_R))) = \begin{cases} -2 & \text{če } PO_S(M_S, b(a_R)) < -0,75M_S \\ -1 & \text{če } -0,75M_S \leq PO_S(M_S, b(a_R)) < -0,5M_S \\ 1 & \text{če } -0,5M_S \leq PO_S(M_S, b(a_R)) < -0,25M_S \\ 2 & \text{če } -0,25M_S \leq PO_S(M_S, b(a_R)) \end{cases}$$

Ponazorimo ocenjevanje zaupanja agentov z različnimi tipi osebnosti še na primeru: pošiljatelj  $a_{S1}$  in  $a_{S2}$  opravita transakcijo z istim agentom prejemnikom  $a_R$ . Oba agenta pošiljatelja investirata enak znesek denarja  $M_{S1} = M_{S2} = 10$ . Prejemnik obojima agentoma vrne znesek v vrednosti 7,5 enot (interpretirano tudi kot: oba agenta sta prejela storitev s popolnoma enako kakovostjo, kar pomeni, da je bila objektivna kakovost izida interakcije v obeh primerih enaka). Recimo, da je agent  $a_{S1}$  centralist in agent  $a_{S2}$  zmerni pesimist. Agent  $a_{S1}$  prejemnika  $a_R$  na podlagi izkupička transakcije oceni kot delno zaupanja vrednega, tj.  $\omega_{S1,R} = 1$ , medtem ko agent  $a_{S2}$  prejemnika  $a_R$  oceni boljše – kot (popolnoma) zaupanja vrednega, tj.  $\omega_{S2,R} = 2$ .

Agent pošiljatelj  $a_S$ , ki je naključno izbran v danem koraku simulacije, *pred* izvršitvijo transakcije z naključno izbranim pošiljateljem  $a_R$  izračuna agentu  $a_S$  prilagojeno vrednost ugleda agenta  $a_R$ . *Po* izvršeni transakciji igre investicij agent  $a_S$  na podlagi izkupička transakcije oceni zaupanje do agenta  $a_R$ . Ocena zaupanja je odvisna od izkupička transakcije in tipa osebnosti agenta  $a_S$ . Naključno izbrani agent pošiljatelj  $a_S$  v vsakem primeru naredi transakcijo z naključno izbranim agentom prejemnikom  $a_R$ , četudi je ocenjena vrednost zaupanja do agenta  $a_R$  nizka. V simulacijah primerjamo izračunano vrednost ugleda, ki je izračunana pred opravljeno transakcijo, in resnično vrednost zaupanja, ki se ovrednoti po opravljeni transakciji. Cilj metod za prilagoditev ocen zaupanja je izločitev dejavnika subjektivnosti, s čimer se sporočene ocene zaupanja prilagodijo posameznemu agentu. Razlika med izračunano in resnično oceno zaupanja je s tem majhna, če uporabimo učinkovito metodo za prilagoditev ocen zaupanja. V nasprotnem primeru je razlika med izračunano in resnično oceno zaupanja velika. Učinkovitost metode za prilagoditev ocen je čim boljša, tem manjša je razlika med izračunano in resnično oceno zaupanja.

Učinkovitost predlagane metode HOMRA smo primerjali z metodo, ki so jo predlagali Hasan in drugi [88], in jo bomo v nadaljevanju poimenovali Hasanova metoda; in z metodo, ki sta jo predlagala Abdul-Rahman in Hailes [28], in jo bomo poimenovali Abdul-Rahmanova metoda. Hasanova in Abdul-Rahmanova metoda sta (po našem najboljšem prepričanju) edini metodi, predlagani v znanstveni literaturi, ki obravnavata problem subjektivnosti v ocenah zaupanja v enakem smislu kot subjektivnost obravnavamo v pričujoči doktorski disertaciji.

Ocena zaupanja, ki jo agent ocenjevalec dobi od agenta svetovalca v Hasanovi metodi, predstavlja relativno vrednost glede na ocene zaupanja agenta svetovalca. Izražena je s percentilom. Agent ocenjevalec posredovano relativno oceno zaupanja pretvori v

absolutno, tako da ustreza njegovemu dojetanju zaupanja.

V Abdul-Rahmanovi metodi agenti ocenjevalci ocene zaupanja, ki jih dobijo od agentov svetovalcev, naravnajo glede na to, kakšne so pretekle izkušnje agenta ocenjevalca z agentom svetovalcem. Agent ocenjevalec posredovano oceno zmanjša ali poveča glede na to, kako (in za koliko) se njegove pretekle ocene do ocenjevanih agentov razlikujejo od ocen zaupanja, ki jih je posredoval agent ocenjevalec.

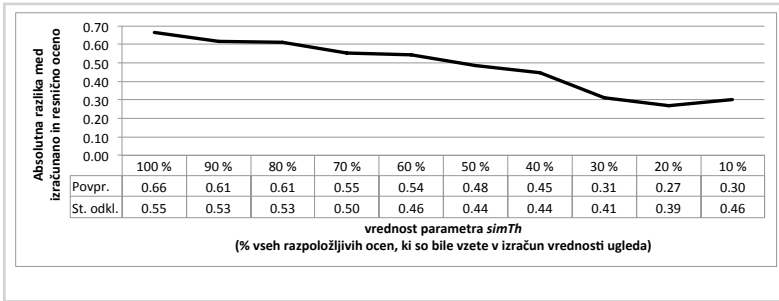
Učinkovitost metod za prilagoditev ocen zaupanja HOMRA, Hasanovo in Abdul-Rahmanovo smo testirali s štirimi nizi poskusov. Poleg tega smo izvedli simulacije brez prilagajanja ocen zaupanja, v katerih agenti izračunajo vrednost ugleda (definicija 2.1), namesto prilagojene vrednosti ugleda. V prvem nizu smo ovrednotili učinkovitost metode HOMRA glede na različne vrednosti parametra  $tdsTh$ . V drugem nizu eksperimentov smo analizirali učinkovitost metod glede na število agentov v skupnosti in v tretjem glede na število izvedenih transakcij. V zadnjem nizu smo primerjali učinkovitost metode HOMRA s Hasanovo in Abdul-Rahmanovo metodo glede na porazdelitev agentov z različnimi tipi osebnosti. V naslednjih poglavjih bomo opisali simulacijske parametre, ki so specifični za posamezen niz eksperimentov, in prikazali rezultate.

### 5.3.2 Analiza učinkovitosti metode HOMRA glede na vrednost parametra $tdsTh$

V prvem nizu eksperimentov smo torej ovrednotili učinkovitost predlagane metode za prilagoditev ocen zaupanja HOMRA glede na vrednost parametra  $tdsTh$ .

V simulacijski konfiguraciji je  $n = 240$  agentov z enakomerno porazdelitvijo različnih tipov osebnosti med agenti. Tako je v skupnosti 16,7 % optimistov, 16,7 % zmernih optimistov, 16,6 % centralistov, 16,7 % zmernih pesimistov, 16,7 % pesimistov in 16,6 % ekstremistov. Agenti imajo različne stopnje dobrohotnosti, ki je enakomerno porazdeljena, tako da velja:  $b(a_1) = 1/n$ ,  $b(a_2) = 2/n$ ,  $b(a_3) = 3/n$  itn. Določili smo simulacijske konfiguracije, v katerih smo uporabili različne vrednosti parametra za mero vključenih ocen zaupanja:  $tdsTh = 0,1 * n$ ,  $0,2 * n$ ,  $0,3 * n$ ,  $0,4 * n$ ,  $0,5 * n$ ,  $0,6 * n$ ,  $0,7 * n$ ,  $0,8 * n$ ,  $0,9 * n$ ,  $1,0 * n$ . V konfiguraciji z vrednostjo  $tdsTh = 1,0 * n$  simuliramo izračun ugleda brez prilagoditve ocen zaupanja, saj v izračunu upoštevamo ocene 100 % najbolj podobnih agentov, tj. vse ocene. Vsak tek simulacije smo ponovili 100-krat, pri čemer smo vsakič uporabili drugo vrednost semena za generator psevdonaključnih števil.

Učinkovitost metode HOMRA smo ovrednotili v simuliranem okolju, kjer je v povprečju na voljo 5 % vseh ocen zaupanja. To pomeni, da je v matriki zaupanja v povpre-



Slika 5.6

Absolutna razlika med izračunano in resnično oceno zaupanja glede na različne vrednosti parametra *tdsTh*. Manjša razlika pomeni boljše učinkovitost metode za prilagoditev ocen zaupanja.

čju 5 % ocen zaupanja definiranih, ostalih 95 % ocen pa ima vrednost “/”. Z drugimi besedami: vsakega agenta je v povprečju ocenilo 5 % drugih agentov iz skupnosti. Poleg tega morata biti v vektorju ocen zaupanja agenta pošiljatelja (ki je ocenjevalec zaupanja) vsebovani vsaj dve oceni zaupanja, da lahko izračunamo podobnost dojemanja zaupanja agentov (definicija 2.4). Če sta v vektorju ocen zaupanja manj kot dve oceni, potem agent pošiljatelj izračuna ugled agenta prejemnika (definicija 2.1) in ne prilagojene vrednosti ugleda (definicija 2.6). S tem dodatnim pogojem zberemo začetne vrednosti ocen zaupanja.

V vsakem koraku simulacije agent pošiljatelj z metodo HOMRA izračuna prilagojeno vrednost ugleda agenta prejemnika. Ta vrednost se izračuna pred izvršitvijo transakcije. Po izvršeni transakciji agent pošiljatelj ovrednoti *resnično* vrednost zaupanja do agenta prejemnika. Po vsaki transakciji smo izračunali absolutno razliko med izračunano in resnično oceno zaupanja. Slika 5.6 prikazuje povprečno vrednost absolutne razlike med izračunano in resnično oceno zaupanja za 100 ponovitev simulacije za vsako izmed določenih vrednosti *tdsTh*. Standardni odkloni so podani v tabeli pod grafom na sliki 5.6.

Z eksperimenti smo pokazali, da je razlika med izračunano in resnično oceno zaupanja najmanjša, če uporabimo ocene 20 % najbolj podobnih agentov iz skupnosti, tj.  $tdsTh = 0,2 * n$ . Vrednost parametra  $tdsTh = 0,2 * n$  nam zagotavlja največjo učinkovitost metode HOMRA v okolju, v katerem je v povprečju definiranih 5 % ocen zaupanja. Za analizo učinkovitosti metode HOMRA glede na različne vrednosti parametra *tdsTh* smo zasnovali konfiguracijo z majhnim odstotkom razpoložljivih ocen zaupanja, zato da zagotovimo učinkovitost metode HOMRA že od začetka, tj. v okoljih z majhnim številom razpoložljivih ocen zaupanja. Vsak sistem je namreč na

začetku "prazen". Prva ocena zaupanja je na voljo, ko dva agenta naredita transakcijo in za tem ocenita vrednost zaupanja. S tem ko agenti naredijo čedalje več transakcij, je na voljo čedalje več ocen zaupanja. Ker želimo zagotoviti učinkovitost metode HOMRA tudi v okoljih z majhnim odstotkom razpoložljivih ocen zaupanja, bomo v vseh nadaljnjih simulacijah uporabili vrednost parametra  $tdsTh = 0,2 * n$ .

### 5.3.3 Analiza učinkovitosti metode HOMRA glede na število agentov

V drugem nizu eksperimentov smo ovrednotili učinkovitost predlagane metode za prilagoditev ocen zaupanja HOMRA glede na število agentov v dani skupnosti. Določili smo simulacijske konfiguracije z različnim številom agentov:  $n = 60, 120, 240, 480$  in  $960$ . Stopnja dobrohotnosti agentov je enakomerno porazdeljena, prav tako njihovi tipi osebnosti (glej poglavje 5.3.2).

V vsakem koraku simulacije sta naključno izbrana dva agenta: agent pošiljatelj  $a_S$  in agent prejemnik  $a_R$ . Agent  $a_S$  naredi transakcijo z agentom  $a_R$  ne glede na vnaprej izračunan ugled agenta  $a_R$ . Če agent  $a_S$  uporabi metodo brez prilagajanja, potem izračuna povprečno vrednost vseh ocen iz vektorja ugleda. Sicer agent  $a_S$  zbere ocene iz vektorja ugleda in jih lokalno obdelata, tj. izračuna prilagojeno vrednost ugleda z uporabo ene izmed metod za prilagoditev ocen zaupanja: metodo HOMRA, Hasonovo metodo ali Abdul-Rahmanovo metodo. Ocene zaupanja v matriki zaupanja na ta način ostanejo nespremenjene. Pri uporabi metode HOMRA je vrednost parametra  $tdsTh = 0,2 * n$ , kar pomeni, da agent  $a_S$  za izračun prilagojene vrednosti ugleda agenta  $a_R$  uporabi ocene 20 % njemu najbolj podobnih agentov. Po izvršeni transakciji igre investicij agent  $a_S$  izračuna resnično vrednost zaupanja do  $a_R$  in to posreduje v matriko zaupanja.

Z vsako simulacijsko konfiguracijo smo izvedli toliko korakov, da smo zbrali 30 % ocen zaupanja. Na primer, v konfiguraciji z  $n = 60$  agentov, polna matrika zaupanja vsebuje  $60 * 60 = 3600$  ocen zaupanja. V tem primeru je vsak agent ocenil zaupanje do vsakega drugega agenta v skupnosti. Z izvršitvijo  $60 * 60 * 0,3 = 1080$  transakcij zberemo 30 % ocen zaupanja. Ostalih 70 % vrednosti v matriki zaupanja ima vrednost "/>. V vsakem koraku simulacije se izvrši ena transakcija. Konfiguracijo z  $n = 60$  agentov smo zato izvedli s 1080 koraki. Konfiguracije z  $n = 120, 240, 468$  oziroma 960 agenti smo izvedli s 4320, 17280, 69120 oziroma 276480 koraki, da smo zbrali 30 % razpoložljivih ocen zaupanja.

Po določenem številu korakov smo izračunali absolutno razliko med vnaprej izračunano in resnično oceno zaupanja. Vsako simulacijsko konfiguracijo smo izvedli z

uporabo metode HOMRA, Hasanove metode, Abdul-Rahmanove metode kot tudi z uporabo metode brez prilagajanja ocen zaupanja. Vsako konfiguracijo smo ponovili 100-krat z uporabo različnih vrednosti semen pri generaciji psevdonaključnih števil. Iz rezultatov 100 ponovitev vsake izmed simulacijskih konfiguracij smo izračunali povprečno vrednost absolutne razlike med izračunanimi in resničnimi ocenami zaupanja ter standardni odklon.

Slika 5.7 prikazuje povprečno vrednost absolutnih razlik med izračunanimi in resničnimi ocenami zaupanja za skupnosti z različnim številom agentov. Pod grafom so navedene tudi vrednosti standardnih odklonov. Razlike med izračunanimi in resničnimi ocenami zaupanja so bile največje pri uporabi metode brez prilagajanja ocen zaupanja, sledijo ji Hasanova in Abdul-Rahmanova metoda. Najmanjšo povprečno vrednost absolutne razlike med izračunano in resnično oceno zaupanja in s tem največjo učinkovitost smo dosegli z uporabo metode HOMRA.

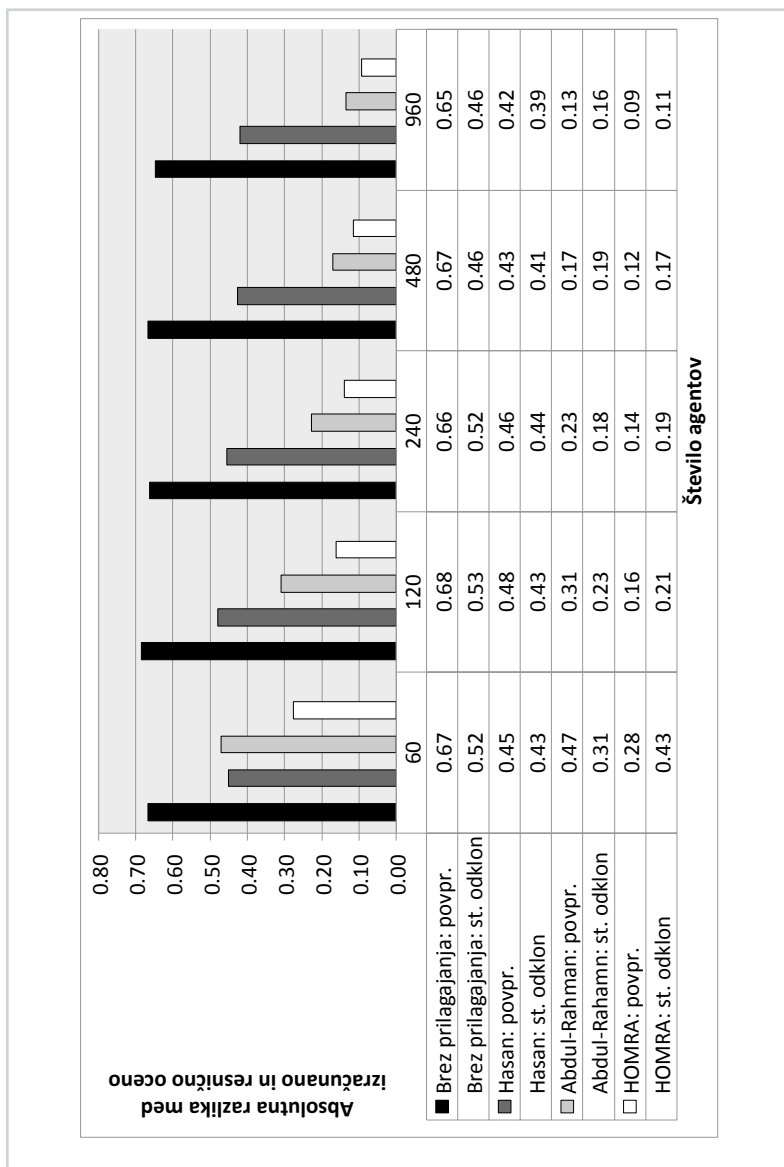
Povprečna razlika med izračunano in resnično oceno zaupanja je pri uporabi metode brez prilagajanja enaka 0,67. Razlika se ne zmanjša z večjim številom agentov, kot se to zgodi pri uporabi metode za prilagoditev ocen zaupanja, tj. Hasanove, Abdul-Rahmanove ali metode HOMRA. Število razpoložljivih ocen zaupanja, s katerimi agenti izračunajo prilagojene vrednosti ugleda drugih agentov, je večje v skupnosti z večjim številom agentov. Če je vsakega agenta v povprečju ocenilo 30 % agentov iz celotne skupnosti, to pomeni 18, 36, 72, 144 oziroma 288 razpoložljivih ocen zaupanja v skupnostih s 60, 120, 240, 480 oziroma 960 agenti. Metode za prilagoditev ocen so pod temi pogoji zato bolj učinkovite v skupnostih z večjim številom agentov.

Največjo učinkovitost smo dosegli z metodo HOMRA, kar velja za vse predlagane konfiguracije. Učinkovitost metode HOMRA smo primerjali z učinkovitostjo ostalih metod. V tabeli 5.1 so prikazani rezultati primerjave.

Metoda HOMRA je v konfiguracijah z različnim številom agentov bolj učinkovita od metode brez prilaganja ocen zaupanja za 58,2 % do 86,2 %. Dalje, od Hasanove metode je za 37,8 % do 78,6 % bolj učinkovita in od Abdul-Rahmanove metode za 29,4 % do 48,4 %.

#### 5.3.4 Analiza učinkovitosti metode HOMRA glede na število transakcij

V tretjem nizu eksperimentov smo ovrednotili učinkovitost metod za prilagoditev ocen zaupanja v odvisnosti od števila transakcij med agenti. Izvedli smo simulacije z  $n = 240$  agentov. Stopnje dobrohotnosti agentov in tipi osebnosti agentov so enakomerno



Slika 5.7

Absolutna razlika med izračunano in resnično oceno zaupanja, dosežena z uporabo različnih metod glede na število agentov v skupnosti. Manjša razlika pomeni boljšo učinkovitost metode za prilagoditev ocen zaupanja.

Tabela 5.1

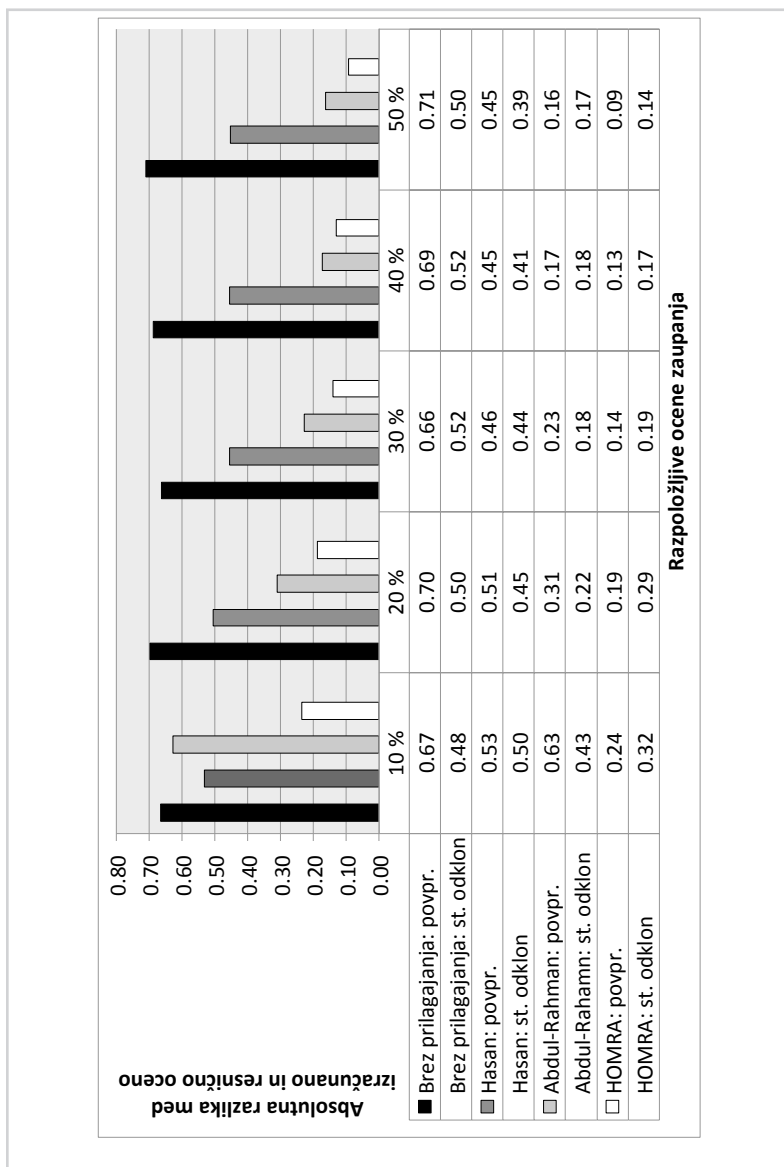
Primerjava učinkovitosti metode HOMRA z ostalimi metodami za konfiguracije z različnim številom agentov. Vrednost v tabeli predstavlja odstotek (%), koliko bolj učinkovita je metoda HOMRA v primerjavi z drugo metodo.

UČINKOVITOST METODE			
HOMRA V PRIMERJAVI			
Z UČINKOVITOSTJO			
DRUGE METODE /	<i>Brez prilagajanja</i>	<i>Hasanova</i>	<i>Abdul-Rahmanova</i>
ŠTEVILO AGENTOV			
60	58,2	37,8	40,4
120	76,5	66,7	48,4
240	78,8	69,6	39,1
480	82,1	72,1	29,4
960	86,2	78,6	30,8

porazdeljeni, kot je opisano v poglavju 5.3.2.

Vsaka izvedba simulacije je tekla določeno število korakov, ki so potrebni za nabor naslednjih odstotkov razpoložljivih ocen zaupanja: 10 %, 20 %, 30 %, 40 % in 50 %. V resničnih e-okoljih, ki uporabljajo sisteme za obvladovanje zaupanja in ugleda, je matrika zaupanja tipično zelo redka. Zaradi tega nismo načrtovali konfiguracij z več kot 50 % vrednosti v matriki zaupanja. Vsako konfiguracijo smo 100-krat ponovili, pri čemer smo za vsak tek simulacije uporabili drugo vrednost semena za generiranje psevdonaključnih števil. Nato smo izračunali povprečno vrednost in standardni odklon absolutnih razlik med izračunanimi in resničnimi ocenami zaupanja. Rezultati za konfiguracije z različnim odstotkom razpoložljivih ocen zaupanja so prikazani na sliki 5.8.

Učinkovitost Hasanove, Abdul-Rahmanove in metode HOMRA se z večjim odstotkom razpoložljivih vrednosti v matriki zaupanja poveča. Abdul-Rahmanova metoda prilagodi sporočene ocene zaupanja na podlagi preteklih izkušenj z agentom, ki je dane ocene posredoval v matriko zaupanja. Na primer: odnos zaupanja agenta svetovalca  $a_C$  do ocenjevanega agenta  $a_B$  je delno zaupanja vreden ( $\omega_{C,B} = 1$ ), medtem ko agent ocenjevalec  $a_A$  agenta  $a_B$  ocenjuje kot zaupanja vrednega ( $\omega_{A,B} = 2$ ). V tem primeru bo agent  $a_A$  vse nadaljnje ocene zaupanja, ki jih bo pridobil od agenta  $a_C$ , povečal za eno stopnjo in s tem zmanjšal t. i. *semantično razdaljo* (angl. *semantic distance*) med



Slika 5.8

Absolutna razlika med izračunano in resnično oceno zaupanja, dosežena z uporabo različnih metod glede na število razpoložljivih ocen zaupanja. Manjša razlika pomeni boljšo učinkovitost metode za prilagoditev ocen zaupanja.

agentoma. Slabost metode Abdul-Rahman je, da deluje le ob predpogoju, da sta imela agent ocenjevalec in agent svetovalec v preteklosti izkušnje z istimi agenti. Če množica agentov, s katerimi sta imela izkušnjo *oba* agenta, ocenjevalec in svetovalec, ne obstaja, potem agent ocenjevalec ocen od agenta svetovalca ne more prilagoditi. Zaradi opisane razloga je učinkovitost Abdul-Rahmanove metode slaba v konfiguracijah, kjer je na voljo manjši odstotek razpoložljivih ocen. Če je v matriki zaupanja samo 10 % vrednosti, potem je učinkovitost Abdul-Rahmanove metode skoraj enaka učinkovitosti metode, ki sploh ne prilagaja ocen zaupanja. Učinkovitost Abdul-Rahmanove metode z večjim odstotkom razpoložljivih ocen zaupanja narašča. V konfiguraciji, kjer je na voljo 50 % vseh ocen zaupanja, je tako povprečna vrednost razlike med izračunanimi in resničnimi ocenami zaupanja enaka 0,16.

Hasanova metoda in metoda HOMRA ne zahtevata, da imata agent ocenjevalec in agent svetovalec pretekle izkušnje z istimi agenti. Kljub temu je učinkovitost Hasanove metode in metode HOMRA višja v konfiguracijah z višjim odstotkom razpoložljivih ocen zaupanja. Metoda HOMRA izračuna podobnost dojemanja zaupanja med agenti in na podlagi tega zavrže ocene zaupanja agentov, ki zaupanje dojemajo drugače. Izračun podobnosti dojemanja agentov je natančnejši, če je v vektorjih ocen zaupanja večje število ocen zaupanja. Učinkovitost metode HOMRA se zato z večjim odstotkom razpoložljivih ocen zaupanja povečuje. Najbolj natančen izračun ocene zaupanja (glede na primerjavo z resnično oceno zaupanja) smo dosegli z uporabo metode HOMRA, in sicer v konfiguraciji, kjer je razpoložljivih 50 % možnih ocen zaupanja. V tem primeru je povprečna razlika med izračunano in resnično oceno zaupanja enaka 0,09.

Podobno velja za pretvorbe ocen zaupanja s Hasanovo metodo. Transformacija ocen zaupanja je natančnejša, če je v vektorjih ocen zaupanja na voljo večje število vrednosti. Učinkovitost Hasanove metode se zato s številom transakcij med agenti poveča, tj. poveča se z večjim odstotkom razpoložljivih ocen zaupanja. Povprečna razlika med izračunano in resnično oceno zaupanja je v konfiguraciji s 50 % razpoložljivih ocen zaupanja enaka 0,45.

Učinkovitost predlagane metode HOMRA smo primerjali z učinkovitostjo ostalih primerjanih metod. Rezultati primerjave so navedeni v tabeli 5.2.

Metoda HOMRA je v vseh konfiguracijah z različnim odstotkom razpoložljivih ocen zaupanja bolj učinkovita od Hasanove kot tudi Abdul-Rahmanove metode. V primerjavi z metodo, ki ne prilagaja ocen zaupanja, je metoda HOMRA za 64,2 % do 87,3 % bolj učinkovita. Od Hasanove je metode je učinkovitejša za 54,7 % do 80,0 %. Po-

Tabela 5.2

Primerjava učinkovitosti metode HOMRA z ostalimi metodami za konfiguracije z različnim odstotkom razpoložljivih ocen zaupanja. Vrednost v tabeli predstavlja odstotek (%), koliko bolj učinkovita je metoda HOMRA v primerjavi z drugo metodo.

UČINKOVITOST METODE			
HOMRA V PRIMERJAVI			
Z UČINKOVITOSTJO			
DRUGE METODE /	<i>Brez prilagajanja</i>	<i>Hasanova</i>	<i>Abdul-Rahmanova</i>
ODSTOTEK			
RAZPOLOŽLJIVIH OCEN			
10	64,2	54,7	61,9
20	72,9	62,7	38,7
30	78,8	69,6	39,1
40	81,2	71,1	23,5
50	87,3	80,0	43,8

leg tega je učinkovitost metode HOMRA za 23,5 % do 61,9 % boljša v primerjavi z učinkovitostjo Abdul-Rahmanove metode.

### 5.3.5 Analiza učinkovitosti metode HOMRA glede na porazdelitev agentov z različnimi tipi osebnosti

V zadnjem nizu eksperimentov smo ovrednotili učinkovitost metod za prilagoditev ocen zaupanja v skupnostih z različnimi porazdelitvami agentov, ki imajo različne tipe osebnosti. Izvedli smo simulacije z 240 agenti, ki so sodelovali v transakcijah igre investicij. Vsaka simulacija je trajala 17280 simulacijskih korakov. V tem času namreč agenti, ki v vsakem koraku v matriko zaupanja posredujejo eno oceno zaupanja, matriko zaupanja napolnijo s 30 % vrednosti. Ostalih 70 % ocen v matriki zaupanja ima nedefinirane vrednosti, tj. “/”.

Definirali smo šest vrst skupnosti z različnimi porazdelitvami agentov z različnimi tipi osebnosti:

- *Enotna*: V skupnosti je enako število agentov z optimističnim, zmerno optimističnim, centralističnim, zmerno pesimističnim, pesimističnim in ekstremnim tipom osebnosti.

Tabela 5.3

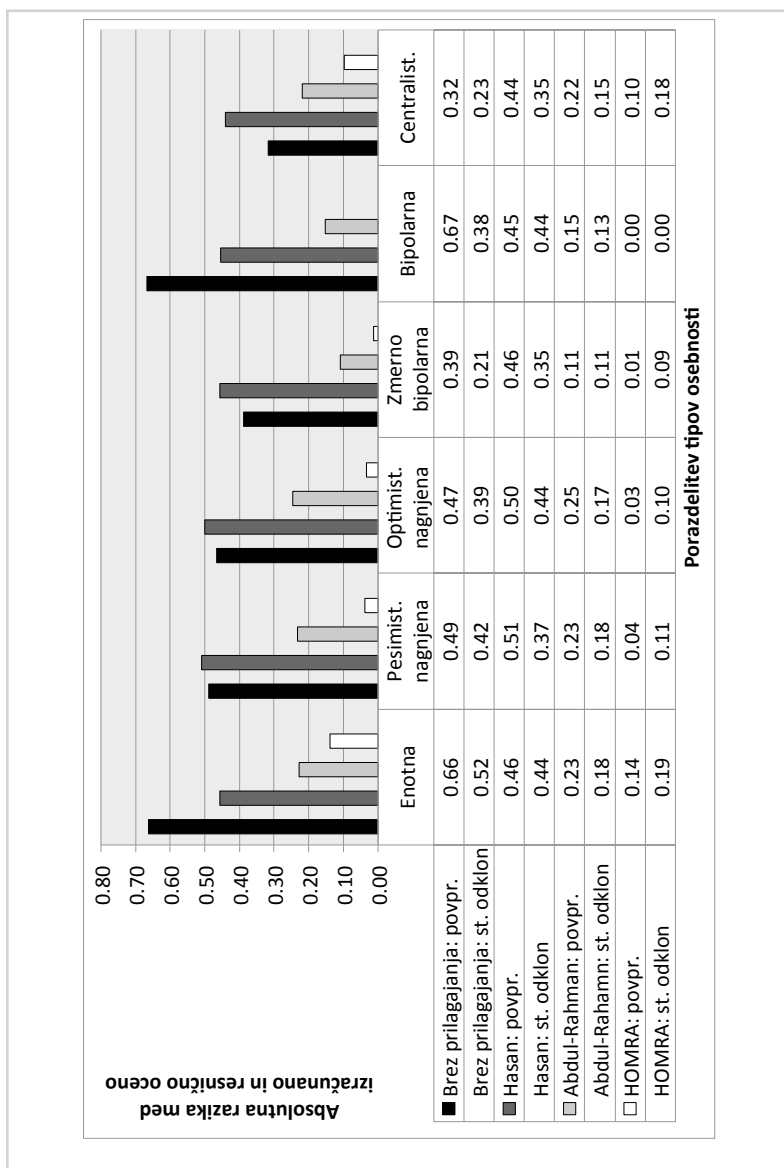
Porazdelitev agentov z različnimi tipi osebnosti (v %) v različnih simulacijskih skupnostih.

TIP OSEBNOSTI / SKUPNOST	OPT. (↑)	ZM. OPT. (↑)	CENTR. (↔)	ZM. PES. (↓)	PES. (↓)	EKSTR. (↘)
ENOTNA	16,7	16,7	16,6	16,7	16,7	16,6
PES. NAGNJ.	0,0	0,0	33,4	33,3	33,3	0,0
OPT. NAGNJ.	33,3	33,3	33,4	0,0	0,0	0,0
ZM. BIPOL.	0,0	50,0	0,0	50,0	0,0	0,0
BIPOLARNA	33,3	0,0	0,0	0,0	33,3	33,4
CENTRAL.	0,0	33,3	33,4	33,3	0,0	0,0

- *Pesimistično nagnjena*: Prisotni so agenti pesimisti, zmerni pesimisti in centralisti.
- *Optimistično nagnjena*: Skupnost sestavljajo agenti z optimističnim, zmerno optimističnim in centralističnim tipom osebnosti.
- *Zmerno bipolarna*: Prisotni so agenti z zmernim optimističnim in zmernim pesimističnim osebnostnim tipom.
- *Bipolarna*: V skupnosti so prisotni optimisti, pesimisti in agenti z ekstremnim osebnostnim tipom.
- *Centralistična*: Skupnost je sestavljena iz zmernih optimistov, centralistov in zmernih pesimistov. Agentov z bolj ekstremnimi tipi obnašanja v tej skupnosti ni.

Porazdelitve agentov z različnimi tipi osebnosti v posamezni skupnosti so številčno opisane v tabeli 5.3.

Enako kot v prejšnjih nizih eksperimentov smo vsako od simulacijskih konfiguracij ponovili 100-krat z uporabo različnih vrednosti semen pri inicializaciji generatorja psevdonaključnih števil. Nato smo izračunali povprečje in standardni odklon absolutnih razlik med izračunanimi in resničnimi ocenami zaupanja. Na sliki 5.9 so rezultati numerično in grafično podani za vse konfiguracije z različnimi porazdelitvami agentov z različnimi tipi osebnosti.



Slika 5.9

Absolutna razlika med izračunano in resnično oceno zaupanja, dosežena z uporabo različnih metod glede na porazdelitev agentov z različnimi tipi osebnosti. Manjša razlika pomeni boljše učinkovitost metode za prilagoditev ocen zaupanja.

Povprečna vrednost razlike med izračunanimi in resničnimi ocenami zaupanja pri uporabi metode brez prilagoditve ocen zaupanja se v različnih simulacijskih konfiguracijah razlikuje. Agenti z različnimi tipi osebnosti ovrednotijo zaupanje do drugih agentov in vrednosti posredujejo v matriko zaupanja. V "Enotni" skupnosti so prisotni agenti z vsemi osebnostnimi tipi, zaradi česar so vrednosti ocen zaupanja zelo razpršene. Na primer, agenti s pesimističnim tipom osebnosti ovrednotijo zaupanje do drugih agentov z nizkimi ocenami, tj. z  $-2$  ali  $-1$ , agenti s centralističnim tipom osebnosti ocenjujejo zaupanje do drugih agentov z vrednostmi okoli  $0$ , medtem ko agenti z optimističnim tipom osebnosti zaupanje do drugih agentov vrednotijo z visokimi ocenami, kot sta  $2$  ali  $1$ . Metoda brez prilagajanja ocen zaupanja pri izračunu ugleda izračuna povprečno vrednost vseh ocen, ki so jih o ocenjevanem agentu posredovali ostali agenti iz skupnosti (definicija 21). Zaradi tega so razlike med izračunanimi in resničnimi ocenami zaupanja velike: v povprečju je razlika  $0,66$ . Povprečna vrednost razlike med izračunanimi in resničnimi ocenami  $0,67$  v "Bipolarni" skupnosti je prav tako posledica razpršenih vrednosti ocen zaupanja, ki so posredovane v matriko zaupanja. V "Centralistični" skupnosti je situacija ravno obratna. Ker v tem scenariju ni agentov z ekstremnimi načini vrednotenja zaupanja (tj. ni agentov z ekstremističnim, pesimističnim ali optimističnim osebnostnim tipom), so vrednosti ocen v matriki zaupanja bolj zgoščene. Zaradi tega je povprečna razlika med izračunano in resnično oceno v "Centralistični" skupnosti enaka  $0,32$ . V "Optimistično nagnjeni", "Pesimistično nagnjeni" oziroma "Zmerno bipolarni" skupnosti so povprečne vrednosti razlik med izračunanimi in resničnimi ocenami enake  $0,47$ ,  $0,49$  oziroma  $0,39$ .

V Hasanovi metodi se ocene zaupanja transformirajo *neposredno* med agentom ocenjevalcem zaupanja in agentom svetovalcem, ki je podal ocene. Zaradi tega so razlike med izračunanimi in resničnimi ocenami zaupanja podobne v vseh simulacijskih konfiguracijah z različnimi porazdelitvami agentov z različnimi tipi osebnosti in so v povprečju enake  $0,47$ . Agent svetovalec agentu ocenjevalcu namreč posreduje percentil ocene zaupanja do ocenjevanega agenta glede na svetovalčeve ocene zaupanja do drugih agentov iz skupnosti. Agent ocenjevalec izpelje (absolutno) vrednost ocene zaupanja, tako da izračuna oceno zaupanja na danem percentilu glede na ocenjevalčev vektor ocen zaupanja do drugih agentov. Model zaupanja, ki smo ga uporabili v simulacijah, uporablja kvalitativne ocene zaupanja, zaradi česar je izpeljana ocena zaupanja pogosto za eno stopnjo previsoka ali prenizka glede na dejanski tip osebnosti agenta ocenjevalca in s tem povezanega načina za vrednotenje zaupanja. Poleg tega je slabša učinkovitost

Hasanove metode tudi posledica tega, da ima agent na voljo malo podatkov. Rezultati prikazujejo povprečno vrednost razlike med izračunanimi in resničnimi ocenami zaupanja, kjer ima agent pri evalvaciji zaupanja na voljo 30 % vrednosti iz vektorja ocen zaupanja. Učinkovitost Hasanove metode se poveča, če je na voljo več podatkov, kot je bilo pokazano v poglavju 5.3.4. Kakorkoli, v realnih e-okoljih, ki uporabljajo sisteme za obvladovanje zaupanja in ugleda, je tipično na voljo majhen odstotek vrednosti iz matrike zaupanja.

Povprečne razlike med izračunanimi in resničnimi ocenami pri uporabi Abdul-Rahmanove metode za prilagoditev ocen so 0,22, 0,23, 0,23 oziroma 0,25 v "Centralistični", "Enotni", "Pesimistično nagnjeni" oziroma "Optimistično nagnjeni" simulacijski skupnosti. Abdul-Rahmanova metoda je učinkovitejša v "Zmerno bipolarni" in "Bipolarni" skupnosti. V teh konfiguracijah namreč agenti ocenjujejo zaupanje zelo pozitivno ali zelo negativno, zato se ostali agenti lahko zelo hitro naučijo, kako drugi agenti vrednotijo zaupanje (prek izračunavanj in posodabljanj semantičnih razlik). Posledično je povprečna razlika med izračunanimi in resničnimi ocenami enaka 0,11 v "Zmerno bipolarni" skupnosti in 0,15 v "Bipolarni" skupnosti.

Naša predlagana metoda HOMRA je bolj učinkovita od Hasanove, Abdul-Rahmanove in metode brez prilagajanja ocen. Enako kot pri Abdul-Rahmanovi metodi tudi za metodo HOMRA velja, da je zelo učinkovita v konfiguracijah z agenti, ki imajo ekstremen način vrednotenja zaupanja; to so agenti s pesimističnim, optimističnim in ekstremnim tipom osebnosti. Metoda HOMRA primerja dojetanje zaupanja agentov. Dojetanje zaupanja agentov z ekstremnimi tipi osebnosti se bistveno razlikuje, zato z metodo HOMRA pravilno prepoznamo podobne agente in v izračun prilagojene vrednosti ugleda vzamemo primerne ocene zaupanja agentov svetovalcev. V "Bipolarni" skupnosti je učinkovitost metode HOMRA najvišja. Izračunane ocene zaupanja so namreč enake resničnim ocenam zaupanja. Povprečne razlike med izračunanimi in resničnimi ocenami zaupanja so sicer 0,01, 0,03, 0,04, 0,10 oziroma 0,14 v "Zmerno bipolarni", "Optimistično nagnjeni", "Pesimistično nagnjeni", "Centralistični" oziroma "Enotni" skupnosti.

V tabeli 5.4 smo primerjali učinkovitost metod za prilagoditev ocen zaupanja v konfiguracijah z različnimi porazdelitvami tipov osebnosti. Enako kot v prejšnjih nizih eksperimentov smo primerjali učinkovitost metode HOMRA v primerjavi z učinkovitostjo metode brez prilagajanja ocen, Hasanove in Abdul-Rahmanove metode.

Metoda HOMRA je bolj učinkovita od metode brez prilagajanja ocen, Abdul-Rah-

Tabela 5.4

Primerjava učinkovitosti metode HOMRA z ostalimi metodami za konfiguracije z različnimi porazdelitvami agentov z različnimi tipi osebnosti. Vrednost v tabeli predstavlja odstotek (%), koliko bolj učinkovita je metoda HOMRA v primerjavi z drugo metodo.

UČINKOVITOST METODE HOMRA V PRIMERJAVI Z UČINKOVITOSTJO DRUGE METODE /			
	<i>Brez prilagajanja</i>	<i>Hasanova</i>	<i>Abdul-Rahmanova</i>
SKUPNOST			
ENOTNA	78,8	69,6	39,1
PESIMISTIČNO NAGNJ.	91,8	92,2	82,6
OPTIMISTIČNO NAGNJ.	93,6	94,0	88,0
ZMERNO BIPOLARNA	97,4	97,8	90,9
BIPOLARNA	izračunane in resnične ocene so enake		
CENTRALISTIČNA	68,8	77,3	54,5

manove metode kot tudi Hasanove metode v vseh simulacijskih skupnostih. Je za 68,8 % do 97,4 % bolj učinkovita v primerjavi z metodo brez prilagajanja ocen. Metoda HOMRA je za 69,6 % do 97,8 % bolj učinkovita od Hasanove metode in za 39,1 % do 90,9 % učinkovitejša v primerjavi z Abdul-Rahmanovo metodo. Poleg tega z metodo HOMRA v "Bipolarni" skupnosti dosežemo cilj metod za prilagoditev ocen zaupanja. Pred transakcijo izračunane ocene so namreč enake resničnim ocenam zaupanja.

S simulacijami smo pokazali, da je predlagana metoda HOMRA uspešnejša od drugih dveh metod za prilagoditev ocen zaupanja, ki sta bili predlagani v literaturi. V skupnostih z različnim številom agentov je metoda HOMRA od Abdul-Rahmanove v povprečju boljša za 37,6 %, od metode Hasan pa za 65,0 %. Dalje, metoda HOMRA je v povprečju za 41,4 % boljša od Abdul-Rahmanove in za 67,6 % od Hasanove v skupnostih z različnim odstotkom razpoložljivih ocen zaupanja. V skupnostih z različnimi porazdelitvami agentov z optimističnim, zmerno optimističnim, centralističnim, zmerno pesimističnim, pesimističnim in ekstremističnim tipom osebnosti je metoda HOMRA od Abdul-Rahmanove metode v povprečju boljša za 71,0 % in za 86,2 % od Hasanove metode. V simulirani "Bipolarni" skupnosti z uporabo metode HOMRA dosežemo cilj: vnaprej izračunane in resnične ocene zaupanja so enake.

#### 5.4 *Kvalitativna primerjava metode HOMRA z alternativnimi pristopi*

Kljub temu, da je bilo predlaganih že veliko računskih modelov za obvladovanje zaupanja in ugleda v e-okoljih, problem subjektivnosti v ocenah zaupanja še ni bil zadovoljivo obravnavan in rešen. Nekateri avtorji so sicer prepoznali in opozorili na problematiko subjektivnosti pri modeliranju zaupanja, vendar niso predlagali ustreznih rešitev problema.

Hasan in drugi [88] so predlagali metodo za odstranitev dejavnika subjektivnosti iz ocen zaupanja z uporabo percentilov. Metoda deluje tako, da absolutne ocene, ki jih posreduje agent svetovalec, pretvori v relativne vrednostni glede na nagnjenje k zaupanju agenta svetovalca. Agent ocenjevalec, ki so mu posredovane ocene zaupanja, te pretvori nazaj v absolutno vrednost glede na njegovo nagnjenje k zaupanju. V primerjavi z našo metodo za prilagoditev ocen zaupanja so Hasan in soavtorji uporabili drugačen pristop, vendar je učinkovitost njihove metode v primerjavi z učinkovitostjo metode HOMRA manjša, kar smo pokazali z eksperimenti v poglavjih 5.3.3, 5.3.4 in 5.3.5.

Abdul-Rahman in Hailes [28] sta predlagala enega izmed prvih modelov zaupanja in ugleda v virtualnih skupnostih, v katerem sta predstavila koncept t. i. *semantične razlike* med agenti, ki si izmenjujejo ocene zaupanja. Semantična razlika implicira problem subjektivnosti v ocenah zaupanja. Agenti ocene zaupanja, ki jih posredujejo drugi agenti v skupnosti, povečajo ali pomanjšajo glede na pretekle razlike oziroma odstopanja ocen agenta ocenjevalca in agenta svetovalca. Glavna pomanjkljivost metode je pogoj, da sta imela agent ocenjevalec in agent svetovalec v preteklosti izkušnje z istimi agenti. Zaradi tega pogoja je Abdul-Rahmanova metoda neučinkovita v e-okoljih z velikim številom agentov in kjer ima vsak agent možnost interakcije z vsakim drugim agentom. Tipičen primer so spletna nakupovalna okolja C2C. V primerjavi z metodo HOMRA sta Abdul-Rahman in Hailes predlagala drugačen pristop, ki je (tudi zaradi prej opisanih slabosti) glede na eksperimentalno evalvacijo manj učinkovit od naše predlagane metode HOMRA.

V [38] je Marsh, ki velja za začetnika formalizacije zaupanja v računalništvu, definiral različna nagnjenja k zaupanju agentov. Predlagal je tri različne vrste obnašanja agentov (optimistično, pesimistično in realistično) in analiziral razlike med vrednotenjem zaupanja glede na vrsto obnašanja. Marsh je ustrezno obravnaval problem subjektiv-

nosti in predlagal formalizacijo zaupanja, ki vključuje tudi definicije različnih nagnjenj k zaupanju. Kljub ustreznosti formalnega modela nagnjenja k zaupanju agentov pa avtor ni predlagal metode za prilagoditev ocen zaupanja, s katero bi odstranili faktor subjektivnosti iz posredovanih ocen zaupanja.

Razširitev Marševe formalizacije zaupanja, ki naslavlja problem prilagoditve ocen zaupanja, so opisali Noorianova in drugi [110]. Predlagali so dvostopenjski filtrirni algoritem, ki združuje kognitivno in verjetnostno modeliranje zaupanja. Različna nagnjenja k zaupanju (optimizem, pesimizem in realizem) se odražajo v različnem vrednotenju zaupanja. V predlaganem modelu agent ocenjevalec od agenta svetovalca poleg ocen zaupanja prejme še kriterije za vrednotenje zaupanja. Na podlagi kriterijev agent ocenjevalec identificira vedenjske vzorce agenta svetovalca, ki je ocene podal. Agent ocenjevalec kriterije za vrednotenje uteži glede na lastne preference in na tak način izpelje prilagojene vrednosti ocen zaupanja. Pogoj za delovanje metode, ki jo predlaga Noorianova s soavtorji, je razkritje kriterijev za vrednotenje zaupanja. Poleg tega morata imeti agent ocenjevalec in agent svetovalec skupen nabor kriterijev, ki so pri vsakem od njiju lahko različno obteženi. Zahteva po razkritju predstavlja veliko slabost metode, če agenti niso pripravljene z ostalimi deliti kriterijev (razlogi za to so lahko povsem različni: kriteriji so zaupne narave, prevelik časovni napor, agent se kriterijev niti ne zaveda ipd.). Teh omejitev v naši predlagani metodi za prilagoditev ocen zaupanja HOMRA ni, saj metoda deluje na modifikaciji končnih ocen zaupanja in od agentov *ne* zahteva, da z ostalimi delijo tudi kriterije, cilje ali faktorje motivacije, na podlagi katerih so ovrednotili zaupanje.

Podoben pristop predlaga tudi Griffiths [89], ki zaupanje modelira z dimenzijami uspeha, cene, časa in kakovosti. Griffiths implicira subjektivnost zaupanja s tem, da različni agenti različno utežijo predlagane dimenzije. Naša predlagana metoda za prilagoditev ocen zaupanja HOMRA subjektivnost v ocenah zaupanja zajame širše, saj na podlagi ostalih ocen zaupanja zajame agentovo dožemanje zaupanja in temu primerno prilagodi posredovane ocene zaupanja.

Huynh [111] je predstavil ogrodje za avtomatizacijo ocenjevanja zaupanja. Defini-rano ogrodje vključuje tudi mehanizem za zajem vzorca obnašanja agenta, ki bo uporabnik ogrodja. Glede na vzorec obnašanja agenta je ustrezno izbran računski model zaupanja za nadaljnjo avtomatizacijo vrednotenja zaupanja. Huynh je pri zasnovi ogrodja poudaril problematiko subjektivne narave zaupanja in predlagal rešitev, ki določa, kako se različno predstavljene ocene zaupanja samodejno pretvorijo v obliko, ki

jo uporablja vsak posamezen agent. Predlagana rešitev dopolnjuje naše delo in je uporabna za predobdelavo ocen zaupanja. V našem modelu zaupanja in predlagani metodi za prilagoditev ocen je namreč definirano, da vsi agenti zaupanje vrednotijo z elementi iz množice  $\Omega = \{-2, -1, 0, 1, 2\}$ , ki simbolizirajo zaupanja nevreden, delno zaupanja nevreden, neodločen, delno zaupanja vreden in zaupanja vreden odnos. Kakorkoli, v modelu zaupanja QAD definirana množica ocen  $\Omega$  ne omejuje delovanja metode HOMRA, saj je ta zasnovana tako, da deluje tudi v sistemih, ki za ocenjevanje zaupanja uporabljajo drug nabor vrednosti. Huynh poleg prevedbe ocen v izbrano obliko ne predlaga drugih postopkov za odstranitev dejavnika subjektivnosti iz ocen zaupanja, do katere pride zaradi različnega dojemanja zaupanja posameznih agentov.

Capra [112] je predlagala k človeku usmerjeno ogrodje za obvladovanje zaupanja hTrust, ki obravnava subjektivnost evalvacije zaupanja. Ogrodje hTrust definira prilagoditvene funkcije, ki zaupanje ocenijo na različne načine – glede na različna nagnjenja k zaupanju uporabnikov ogrodja hTrust. Prilagoditvene funkcije na različne načine združijo lastne ocene zaupanja in ocene, posredovane s strani drugih agentov, iz razpona vrednosti zaupanja izračunajo oceno zaupanja, obravnavajo natančnost ocene zaupanja itd. Ogrodje predvideva, da uporabnik izbere t. i. *profil zaupanja*, tj. ustrezno kombinacijo prilagoditvenih funkcij. Capra je v ogrodju hTrust obravnavala problematiko različnega nagnjenja k zaupanju agentov in subjektivnosti procesa vrednotenja zaupanja, a ni razvila konkretnih metod in algoritmov za rešitev problema prilagoditve ocen zaupanja.

Koster in drugi [104, 113, 114] so predstavili argumentacijsko ogrodje, ki agentom omogoča, da izrazijo pomembnost posameznih kriterijev pri vrednotenju zaupanja. Predlagana rešitev zahteva, da agenti razkrijejo svoje modele zaupanja in drug drugemu sporočijo, kaj so njihovi cilji in kriteriji pri ocenjevanju zaupanja. Z analizo in medsebojno primerjavo faktorjev, ki vplivajo za vrednotenje zaupanja, se agent odloči, če in kako bo uporabil posredovane ocene zaupanja od drugih agentov. Naša predlagana metoda HOMRA od agentov ne zahteva, da razkrijejo kriterije in povode za dano oceno zaupanja. Še več, metoda HOMRA niti ne predpostavlja, da agenti sploh poznajo oz. se zavedajo teh kriterijev. Iz tega razloga je bolj splošno uporabna tudi za nadgradnje komercialnih sistemov za obvladovanje zaupanja in ugleda.

Problem subjektivnosti v ocenah zaupanja so obravnavali tudi Dondio in drugi [115] ter predlagali mehanizem za podobno razumevanje ocen zaupanja med agenti. Mehanizem predpostavlja, da agenti izmenjajo nekatere elemente modela zaupanja, kot so

metrike ocenjevanja, funkcije zadovoljstva ipd. Na podlagi primerjav elementov se določi, kako združljivi so modeli zaupanja različnih agentov. Podobno kot pri predlogu Kosterja [104, 113, 114], Noorianove [110] in Griffithsa [89] tudi ta rešitev zahteva, da agenti eksplicitno razkrijejo elemente modelov zaupanja, kot so kriteriji za ocenjevanje, osebnostni tipi, cilji ipd. Aktivno sodelovanje agentov v e-okolju je vprašljivo zaradi upora agentov do sodelovanja in tudi zmožnosti samoevalvacije [88]. Naša predlagana metoda HOMRA je zato zasnovana tako, da deluje brez eksplicitnega sodelovanja agentov in v tem prednjači pred predlogom, ki ga je podal Dondio.



# *Zaključek*

6

## 6.1 Sklepne ugotovitve

V doktorski disertaciji smo obravnavali področje računalniško podprtega obvladovanja zaupanja in ugleda v e-okoljih. Tako kot v fizičnem svetu je tudi v e-okoljih zaupanje med agenti ključnega pomena za učinkovito odločanje o medsebojnih interakcijah in transakcijah. Agenti lahko v navideznem okolju uspešno sodelujejo le, če je med njimi vzpostavljeno zadostno zaupanje [4, 78].

Zaradi vse večjega števila uporabnikov interneta in prenosa poslovanja v e-okolje je potreba po učinkovitem računalniško podprtem obvladovanju zaupanja in ugleda zelo velika. Ovrednotenje zaupanja v e-okoljih se razlikuje od pristopov ocenjevanja zaupanja v fizičnem okolju, saj so v navideznem okolju na voljo povsem drugačni viri informacij. Zato je treba smiselno identificirati dejavnike, ki vplivajo (in so hkrati dosegljivi) na vrednost zaupanja in ugleda v e-okoljih. V ta namen smo analizirali obstoječe matematične modele zaupanja in ugleda ter kot izhodišče za raziskovalno delo izbrali model kvalitativne dinamike ocenjevanja – model QAD [15, 16].

Lastnosti modela QAD smo analizirali s pomočjo računalniških simulacij. Za boljše ponazoritev procesa vrednotenja zaupanja, ki ga definira model QAD, smo v simulacije vpeljali preprost odločitveni model. Rezultati simulacij so pokazali, da formalizacija zaupanja in ugleda QAD lahko ustrezno upošteva družbeno in psihološko ozadje pojma zaupanja. Agenti z različnimi operatorji zaupanja QAD, ki definirajo proces vrednotenja zaupanja, so namreč v enakih pogojih vrednost zaupanja ocenili na različne načine. Agenti z ekstremno pesimističnim, zmerno pesimističnim in centralističnim operatorjem so manj dovzetni za zaupanje, medtem ko so agenti s preostalimi operatorji bolj dovzetni za zaupanje. Kljub temu, da modeli zaupanja in ugleda niso odločitveni modeli, podana analiza lastnosti modela QAD hkrati predstavlja smernice za razvoj odločitvenih modelov, ki (med drugim) temeljijo na vrednostih zaupanja in ugleda agentov v danem e-okolju.

Model QAD smo nato razširili in uvedli nove definicije in konstrukte. Najpomembnejši med njimi so zasebni vektor ocen zaupanja, zgodovinske matrike zaupanja, zgodovinski zasebni vektorji ocen, splošna miselnost agenta, operatorji zaupanja QADE, agent napadalec in funkcija podobnosti *sim*. Razširjeni model smo poimenovali model zaupanja in ugleda QADE in na podlagi razširitev obravnavali problematiko posredovanja zlaganih ocen zaupanja v sistemih za obvladovanje zaupanja in ugleda. Posredovanje napačnih ocen zaupanja je pogosta manipulacija sistemov za obvladovanje

zaupanja in ugleda v e-okoljih, s katero želijo zlonamerni agenti vplivati na vrednotenje zaupanja do napadenih agentov. Izračun ocene zaupanja bo napačen, če bodo v izračun vključene zlagane vrednosti, zato je obvladovanje nepoštenih ocen bistvenega pomena. Z upoštevanjem napačnih ocen zaupanja in posledično napačnim vrednotenjem zaupanja se odnos zaupanja agenta ocenjevalca *popači*<sup>1</sup>, saj pri vrednotenju zaupanja upošteva ocene drugih agentov, ki ne odražajo dejanskega stanja. V izogib temu predlagamo metodo za prepoznavanje in izločanje zlaganih ocen iz postopka vrednotenja zaupanja, metodo QADE-filter.

Predlagana metoda QADE-filter ponuja rešitev za identifikacijo nepoštenih ocen zaupanja, ki jih nato izključi iz izračuna zaupanja. Metoda temelji na primerjavi posredovanih ocen zaupanja s podobnimi agenti. Ocene zaupanja podobnih agentov, ki se razlikujejo od ocen agenta ocenjevalca, so označene s t. i. črno piko. Ocene zaupanja, ki so ovrednotene s preveliko vrednostjo črne pike, v nadaljnjem izračunu vrednosti zaupanja niso upoštevane. Podobnost med agenti je definirana s funkcijo podobnosti *sim*. Ta funkcija je osnovana kot obtežena evklidska razdalja med ocenami zaupanja primerjanih agentov, pri čemer so uteži konstruirane tako, da dodatno nagradijo ustrežanje splošni miselnosti agenta ocenjevalca, ki se primerja z drugimi agenti. S filtriranjem lažnih ocen se odnos zaupanja agenta ocenjevalca *popravi*; v najboljšem primeru do te mere, da ustreza oziroma se izenači z *reprezentativnim* odnosom zaupanja, kakršnega ima agent ocenjevalec, če so vse posredovane ocene zaupanja resnične.

Učinkovitost predlagane metode smo primerjali z reprezentativno eksogeno metodo, ki jo predlaga model TRAVOS [31] (metodo TRAVOS-filter), in z reprezentativno endogeno metodo, ki jo predlagajo Whibty in soavtorji [46] (metodo Whitby filter). Evalvacijo metod za prepoznavanje in izločanje napačnih vrednosti smo izvedli na podlagi večagentnih simulacij. Izmerili smo razliko med povprečno oceno zaupanja v reprezentativnem in popačenem odnosu zaupanja ( $dist(repr, skew(attType))$ ) ter razliko med reprezentativnim in popravljenim odnosom zaupanja agentov ( $dist(repr, corr(Filter, attType))$ ) za štiri vrste napadov *attType*. Razliko smo izmerili z normirano vrednostjo na intervalu [0, 1], pri čemer je  $dist(repr, skew(attType)) = 1$  in predstavlja popolni vpliv zlaganih vrednosti na agentov odnos zaupanja. Vrednosti razdalje  $dist(repr, corr(Filter, attType))$  predstavljajo učinkovitost filtrirne metode. Vrednost  $dist(repr, corr(Filter, attType)) = 0$  pomeni popolno izločitev zlaganih vrednosti.

<sup>1</sup>V poglavju 4.2.2 smo vpeljali termine *reprezentativen* odnos zaupanja, *popačen* odnos zaupanja in *popravljen* odnos zaupanja.

Predstavlja stanje brez vpliva zlaganih vrednosti na odnos zaupanja agenta, saj je popravljen odnos zaupanja enak reprezentativnemu. V povprečju so bile vrednosti razlik med reprezentativnim in popravljenim odnosom zaupanja agentov v simulacijskih okoljih s prisotnimi napadalci z očrnitvijo (tako skupinskimi kot samostojnimi) naslednje:  $dist(repr, corr(QADE\text{-}filter, BM + collBM)) = 0,31$  pri uporabi metode QADE-filter,  $dist(repr, corr(TRAVOS\text{-}filter, BM + collBM)) = 0,59$  pri uporabi metode TRAVOS-filter in  $dist(repr, corr(Whitby\ filter, BM + collBM)) = 0,92$  pri uporabi metode Whitby filter. Metoda QADE-filter je tako za 29 % boljša v primerjavi z metodo TRAVOS-filter in za 62 % v primerjavi z metodo Whitby filter. V simulacijskih okoljih s prisotnimi napadalci s hvaljenjem so bile razlike v povprečju enake  $dist(repr, corr(QADE\text{-}filter, BS + collBS)) = 0,38$  pri uporabi metode QADE-filter,  $dist(repr, corr(TRAVOS\text{-}filter, BS + collBS)) = 0,65$  pri uporabi metode TRAVOS-filter in  $dist(repr, corr(Whitby\ filter, BS + collBS)) = 0,92$  pri uporabi metode Whitby filter. V danih konfiguracijah je s tem naša predlagana metoda QADE-filter boljša za 27 % od metode TRAVOS-filter in za 54 % od metode Whitby filter. Z dobljenimi rezultati smo pokazali, da je predlagana metoda QADE-filter bolj učinkovita v primerjavi z najbolj značilno eksogeno metodo TRAVOS-filter in najbolj značilno endogeno metodo Whitby filter.

V nadaljevanju dela smo se dotaknili še enega nerešenega problema na področju obvladovanja zaupanja in ugleda v e-okoljih: subjektivnosti ocen zaupanja. Subjektivnost je v ocenah zaupanja prisotna zaradi različnega vrednotenja zaupanja posameznih agentov oziroma zaradi njihovih različnih tipov osebnosti. Ocene zaupanja, ki se po vrednostih razlikujejo zaradi različnega dojetanja zaupanja agentov, ki so posredovali ocene, v resnici pa pomenijo isto, je treba ustrezno prilagoditi agentu, ki bo posredovane ocene zaupanja uporabil.

Za rešitev problema smo ustrezno spremenili in razširili osnovni model zaupanja in ugleda QAD ter predlagali metodo HOMRA, ki prilagodi ocene zaupanja posameznim agentom glede na njihova nagnjenja k zaupanju. Prilagojene predstavitve vrednosti zaupanja in ugleda omogočajo agentom enake možnosti v nadaljnjih postopkih odločanja, neodvisno od njihovega nagnjenja k zaupanju. Metoda HOMRA iz preteklih ocen zaupanja, s katerimi so agenti ocenili druge agente v e-okolju, izpelje njihova nagnjenja do zaupanja in jih primerja z uporabo neparametričnih statističnih testov. Pri izračunu posameznemu agentu prilagojene vrednosti ugleda drugih agentov upoštevajo ocene zaupanja, ki so jih podali agenti s podobnimi nagnjenji k zaupanju. S tem pre-

dlagana metoda širše zajema vrednosti posredovanih ocen zaupanja in upošteva njihov kontekstno ozadje.

Metodo HOMRA smo primerjali z drugima dvema metodama za izločitev dejavnika subjektivnosti in prilagoditev ocen zaupanja: metodo, ki jo je predlagal Hasan s soavtorji [88], in metodo, ki sta jo predlagala Abdul-Rahman in Hailes [28]. Vse tri metode smo ovrednotili na podlagi simulacij. Učinkovitost metod smo izmerili s primerjavo izračunane ocene zaupanja, izračunane *pred* izvršeno transakcijo med dvema agentoma z uporabo metode HOMRA, Abdul-Rahmanove ali Hasanove metode, in resnične ocene zaupanja, ki jo agent ovrednoti *po* opravljeni transakciji glede na dejansko izkušnjo. Cilj metod za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti je, da predstavijo ocene zaupanja v posameznemu agentu prilagojeni obliki, tako da je ta vrednost enaka dejanski oceni zaupanja, s katero bo agent ovrednotil zaupanje. V primerjavi metod glede na različno število agentov v simuliranem e-okolju je bila metoda HOMRA v povprečju za 37,6 % bolj učinkovita v primerjavi z uporabo Abdul-Rahmanove metode in za 65,0 % bolj učinkovita od Hasanove metode. V simulacijah z različnim odstotkom razpoložljivih ocen zaupanja je metoda HOMRA v povprečju za 67,6 % bolj učinkovita od Hasanove metode in za 41,4 % od Abdul-Rahmanove metode. Učinkovitost metod smo primerjali tudi glede na različne distribucije agentov z različnimi tipi osebnosti, tj. z različnimi nagnjenji k zaupanju. V tem primeru je bila predlagana metoda HOMRA povprečno za 86,2 % bolj učinkovita od Hasanove metode in za 71,0 % od Abdul-Rahmanove metode. S statistično primerjavo rezultatov smo pokazali, da smo predlagali ustrezno in učinkovito metodo za prilagoditev ocen zaupanja in ugleda z izločitvijo dejavnika subjektivnosti.

## 6.2 Prispevki k znanosti

Izvirni prispevki doktorske disertacije k znanosti so:

- Zasnovali smo model zaupanja in ugleda v e-okoljih, v katerega smo vključili elemente, s katerimi smo formalizirali obstoj neresničnih ocen zaupanja. Predlagani model smo razširili s konstrukti za ustrezno obravnavo faktorja subjektivnosti v ocenah zaupanja.
- Razvili smo metodo za obvladovanje zlaganih ocen zaupanja v e-okoljih. Na podlagi rezultatov večagentnih simulacij smo pokazali njeno učinkovitost v pri-

merjavi z najbolj značilno endogeno metodo in z najbolj značilno eksogeno metodo za obvladovanje zlaganih ocen zaupanja.

- Razvili smo metodo za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti. S simulacijami smo pokazali izboljšave predlagane metode v primerjavi z obstoječimi metodami za prilagoditev ocen zaupanja v sistemih za obvladovanje zaupanja in ugleda v e-okoljih.
- Podrobno smo analizirali obstoječe formalizacije zaupanja in ugleda v e-okoljih, pri čemer smo se osredotočili na razumno modeliranje človeških dejavnikov, ki vplivajo na pojma zaupanja in ugleda. Na podlagi rezultatov analize smo izbrali temeljni model zaupanja in ugleda v e-okoljih: model QAD.
- Model QAD, ki predstavlja izhodišče na novo predlaganega modela zaupanja in ugleda, ki upošteva prisotnost zlaganih ocen zaupanja in obstoj dejavnika subjektivnosti v ocenah zaupanja, smo podrobno analizirali in ga na podlagi simulacij kvalitativno ovrednotili.

Našteti izvirni prispevki doktorske disertacije k znanosti pomenijo pomemben in trajen doprinos na področju dveh zelo pomembnih vidikov upravljanja zaupanja in ugleda v e-okoljih: (1) identifikacije in rokovanja z neresničnimi ocenami zaupanja, ki so posredovane s strani zlonamernih agentov; in (2) obvladovanja dejavnika subjektivnosti v posredovanih ocenah zaupanja. Oba vidika sta izjemno pomembna za razvoj sistemov za obvladovanje zaupanja in ugleda, ki bodo univerzalno uporabni v različnih e-okoljih in bodo v tovrstnih okoljih omogočali vzpostavitev, formalizacijo in obvladovanje konceptov zaupanja in ugleda, kot smo jih vajeni iz fizičnega sveta.

Rezultati doktorske disertacije pomembno izboljšujejo in nadgrajujejo obstoječe dosežke na omenjenih področjih. Na področju obvladovanja zlaganih ocen zaupanja predlagana metoda QAD-filter izboljšuje obstoječe reprezentativne metode za obvladovanje neresničnih ocen zaupanja. Prav tako predstavlja znatne izboljšave obstoječih metod za prilagoditev ocen zaupanja z izločitvijo dejavnika subjektivnosti v disertaciji predlagana metoda HOMRA.

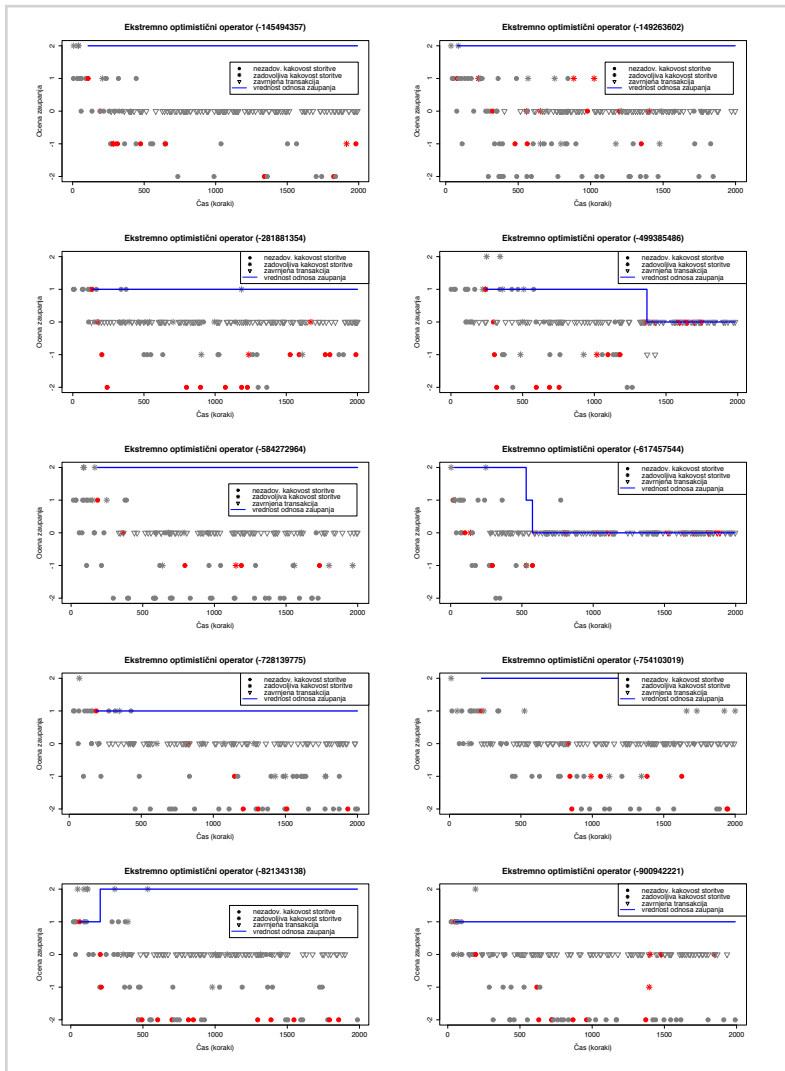
V tej doktorski disertaciji predstavljeni izvirni prispevki k znanosti imajo poleg teoretične tudi praktično uporabno vrednost in predstavljajo teoretično osnovo za implementacijo sistemov za obvladovanje zaupanja in ugleda, uporabnih v različnih e-okoljih, kot so npr. področja e-poslovanja z implementacijo B2C, B2B in ostalih in-

terakcij. Vodilni akterji na tem področju, kot so Amazon [55], eBay [56], Yelp [62] in ostali, imajo še vedno implementirane sisteme za obvladovanje zaupanja in ugleda, ki ne obravnavajo dejavnika subjektivnosti. Zato bo naše nadaljnje delo po doktoratu poleg dodatne izboljšave teoretičnih prispevkov tega dela usmerjeno tudi v praktično eksploatacijo rezultatov in realizacijo njihovega znanstvenega, raziskovalnega in družbeno-ekonomskega vpliva.



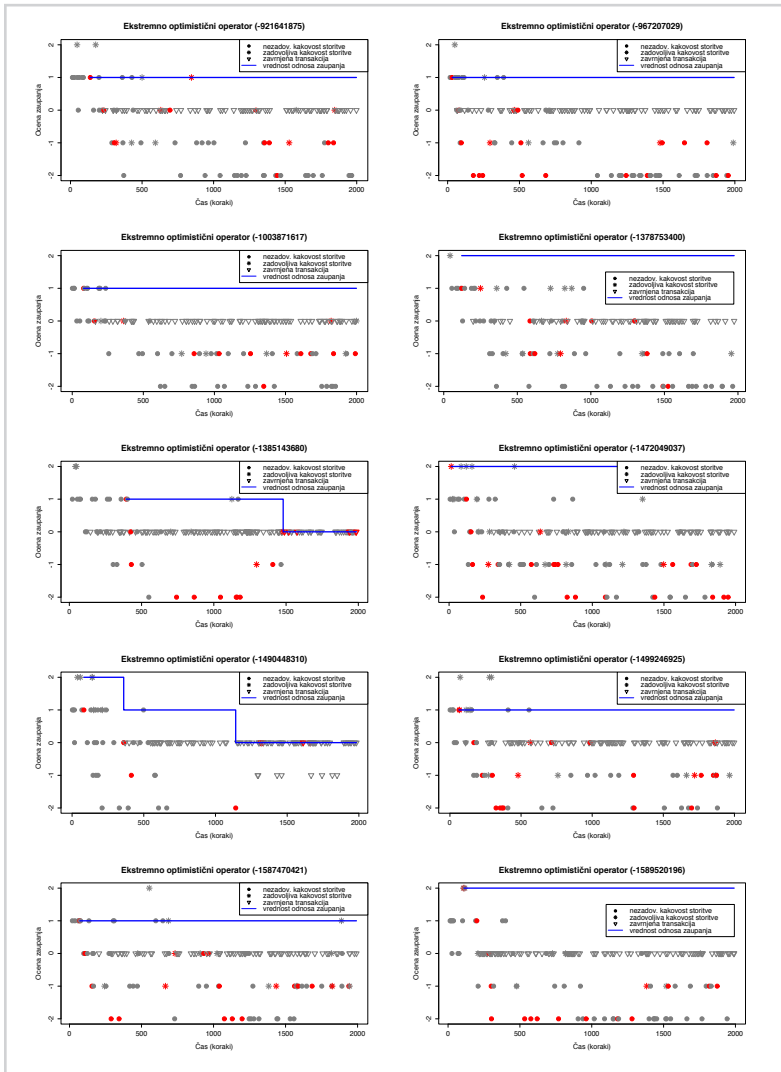
*Rezultati simulacij pri  
poglobljeni analizi lastnosti  
operatorjev QAD*

*A*



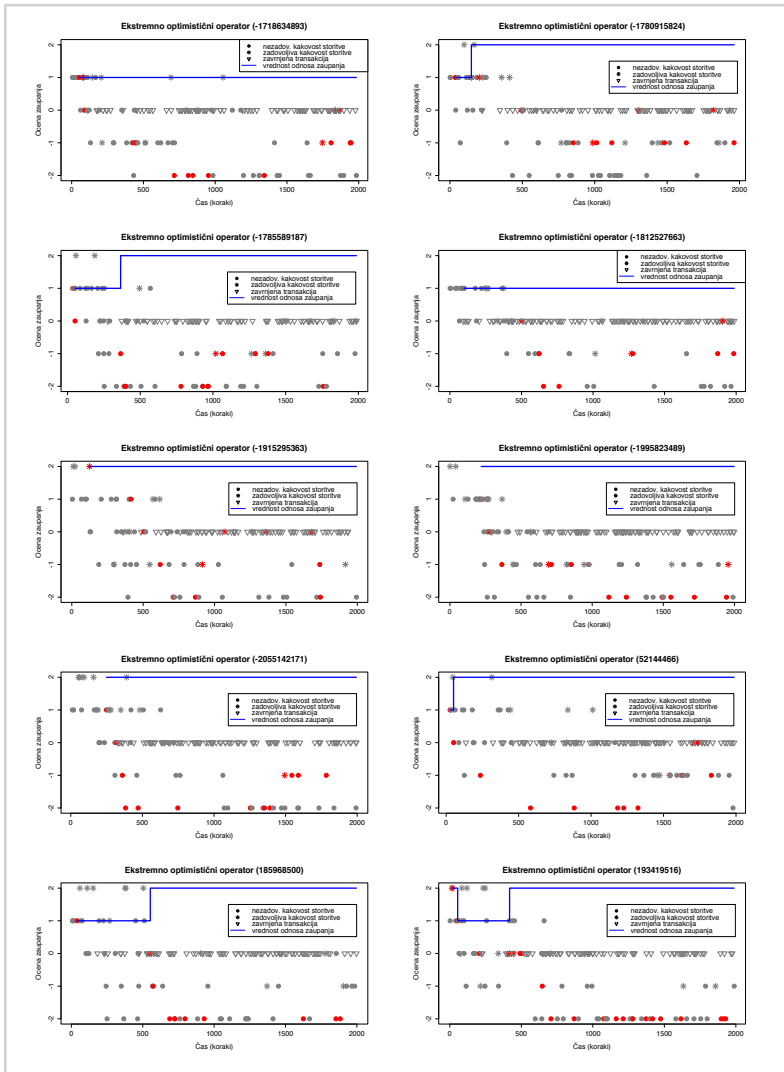
Slika A.1

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim optimističnim operatorjem  $\uparrow$  (1. del).



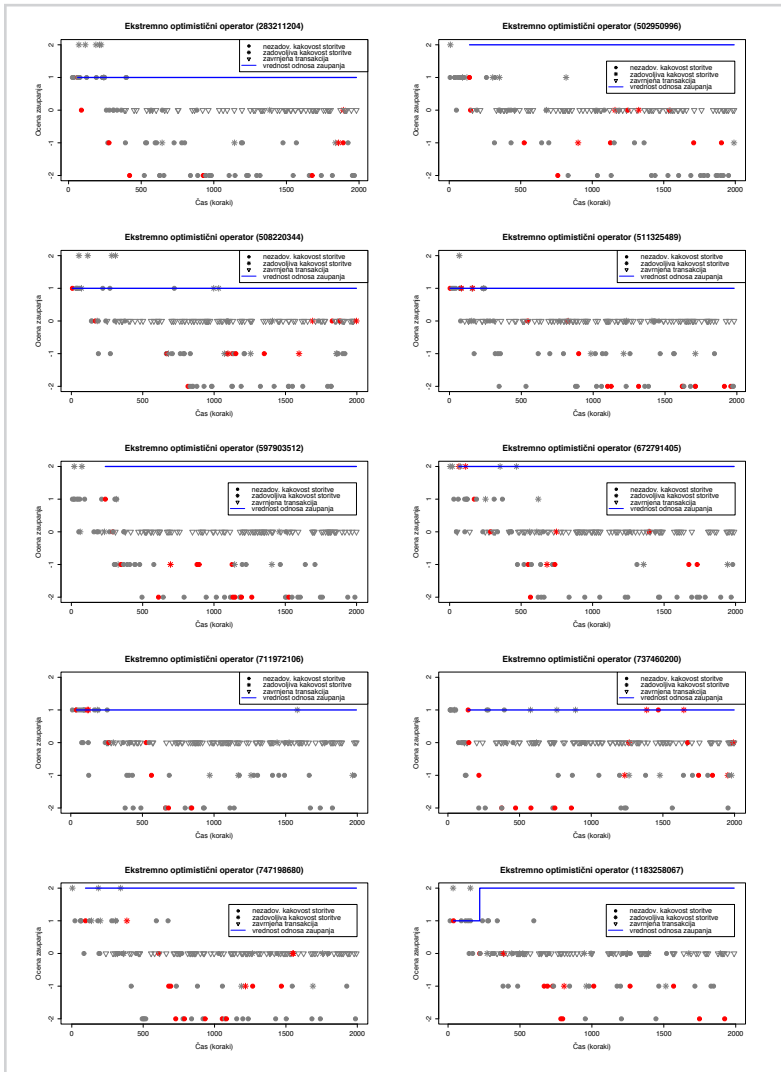
Slika A.2

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim optimističnim operatorjem  $\uparrow$  (2. del).



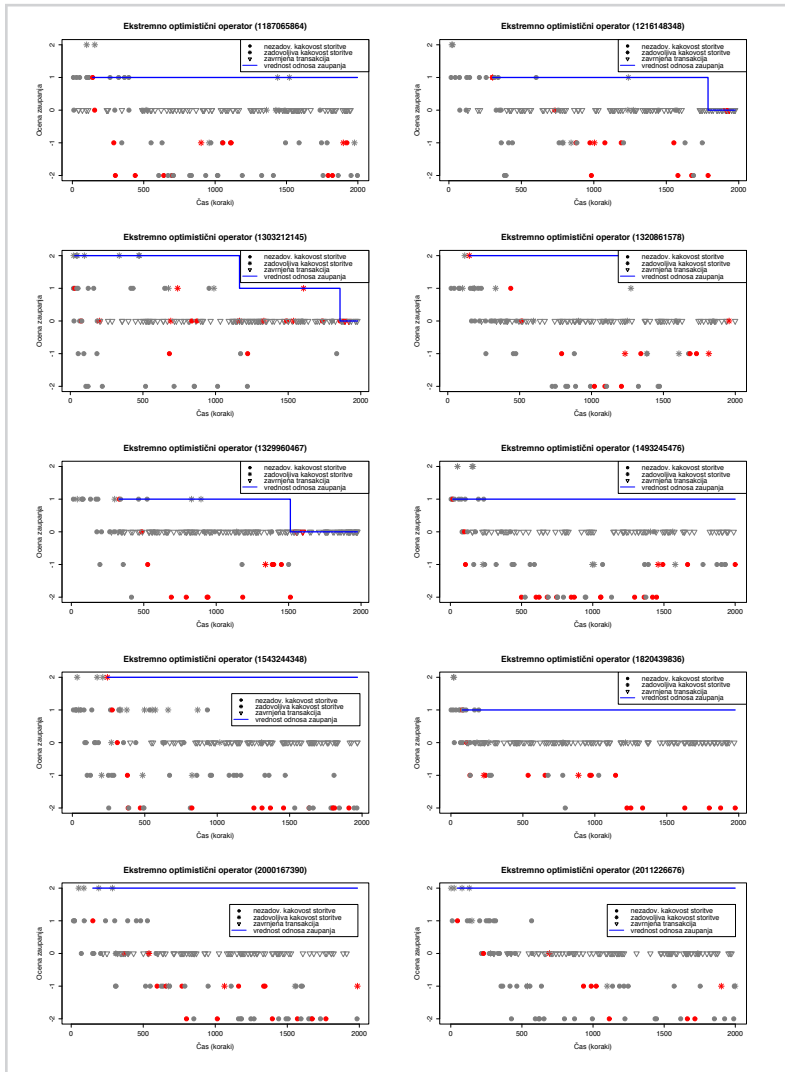
Slika A.3

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim optimističnim operatorjem  $\uparrow$  (3. del).



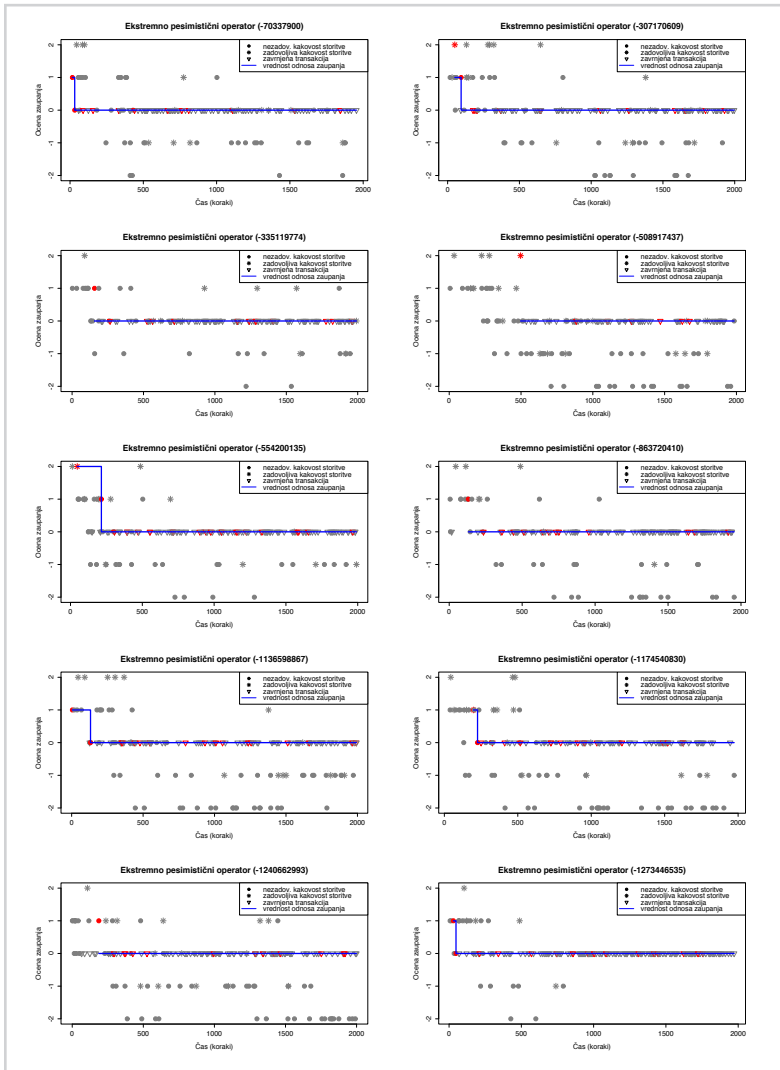
Slika A.4

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim optimističnim operatorjem  $\uparrow$  (4. del).



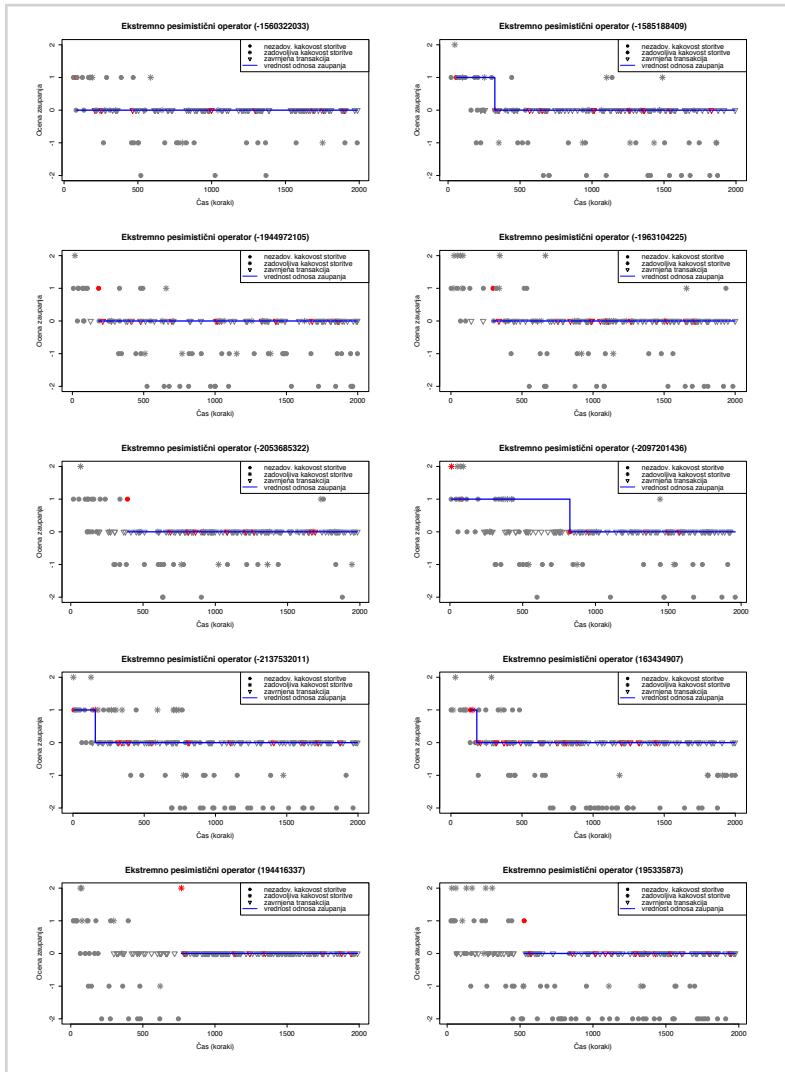
Slika A.5

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim optimističnim operatorjem  $\uparrow$  (5. del).



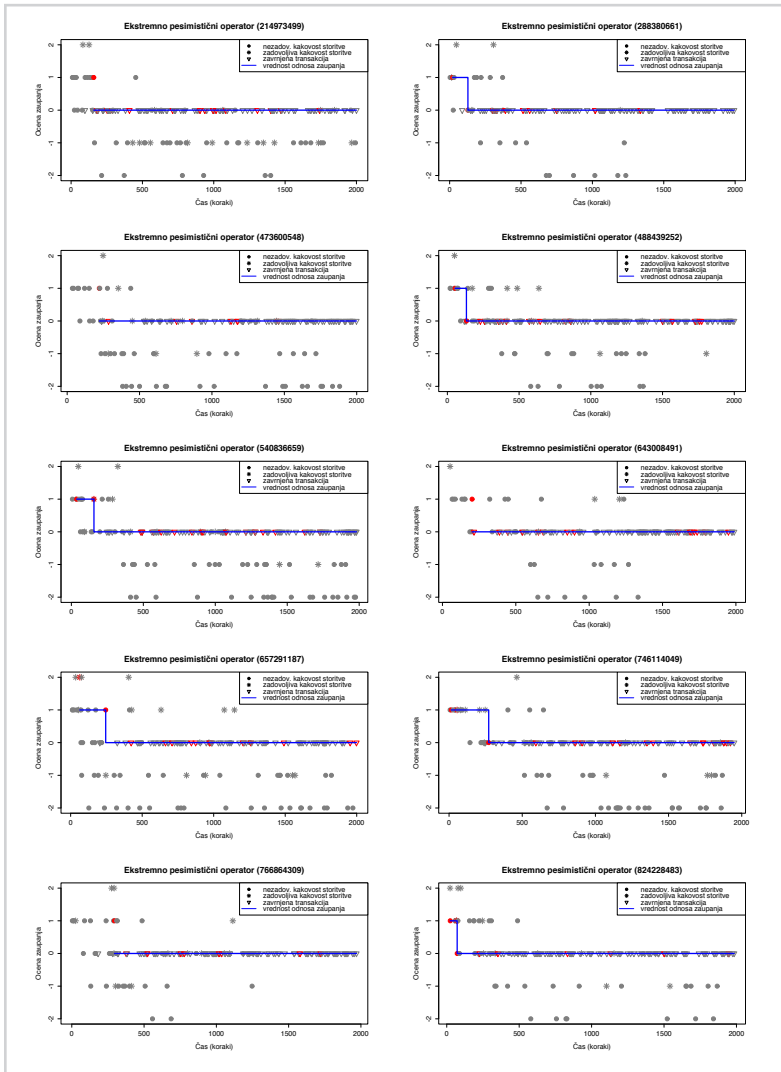
Slika A.6

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim pesimističnim operatorjem  $\Downarrow$  (1. del).



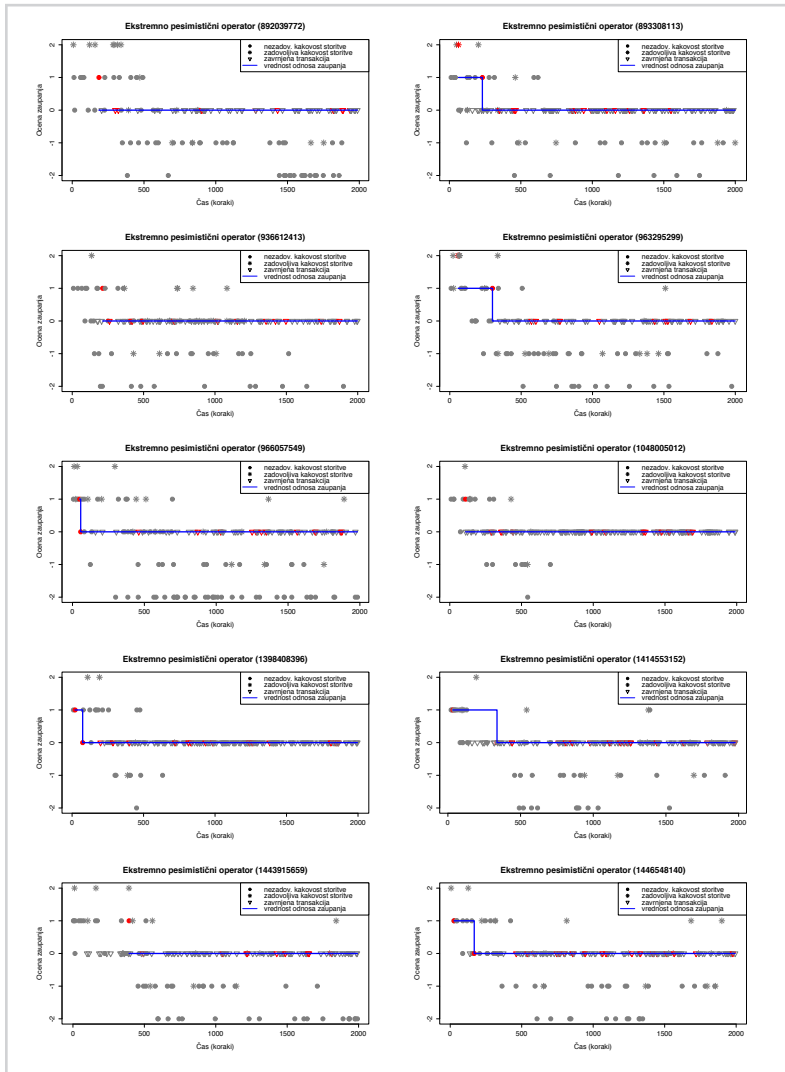
Slika A.7

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim pesimističnim operatorjem  $\downarrow$  (2. del).



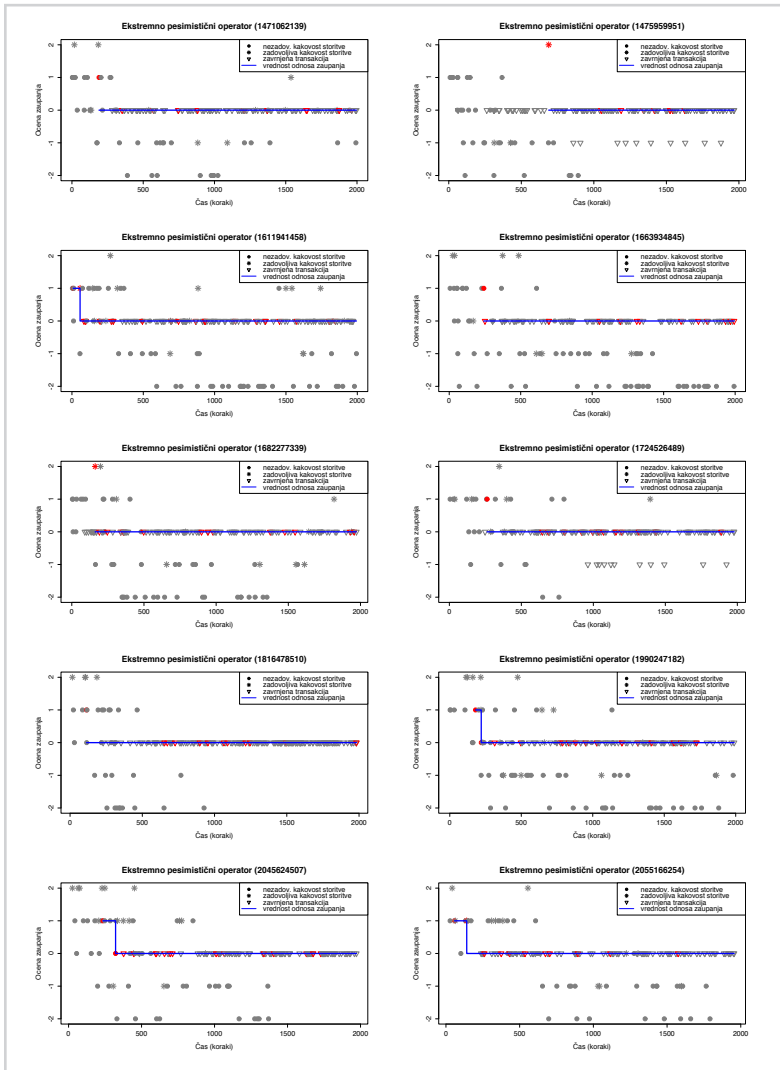
Slika A.8

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim pesimističnim operatorjem ↓ (3. del).



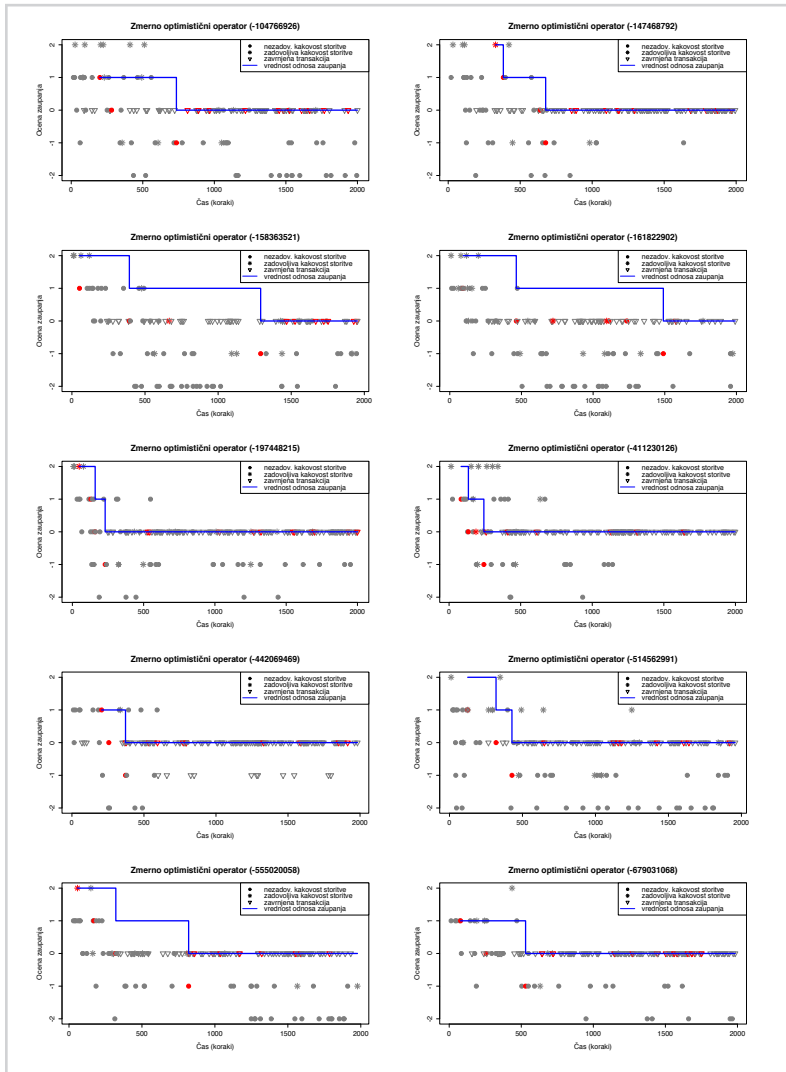
Slika A.9

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim pesimističnim operatorjem  $\downarrow$  (4. del).



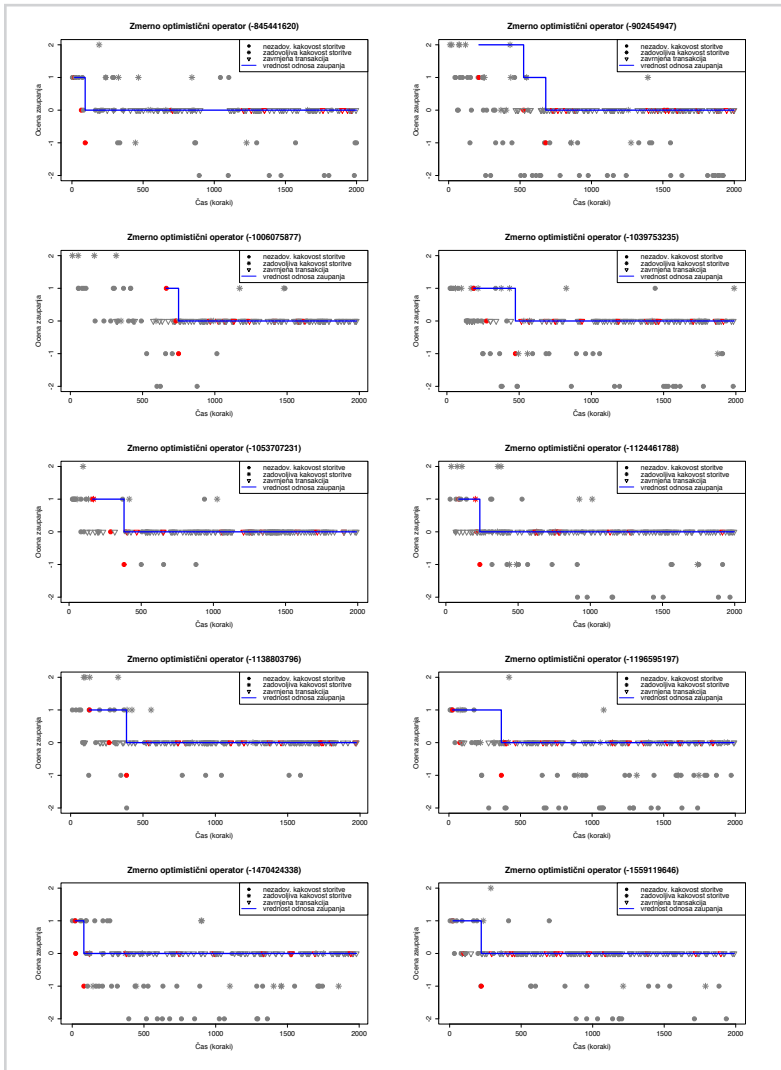
Slika A.10

Obnašanje vrednosti zaupanja, modeliranega z ekstremnim pesimističnim operatorjem  $\Downarrow$  (5. del).



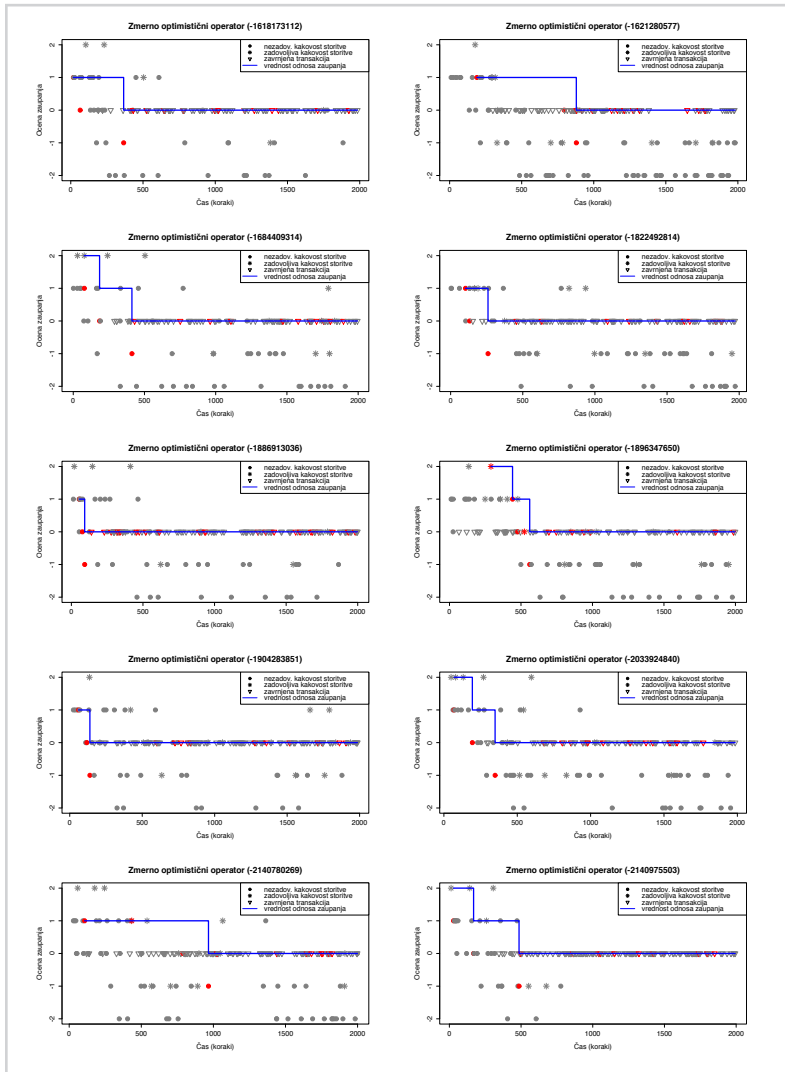
Slika A.11

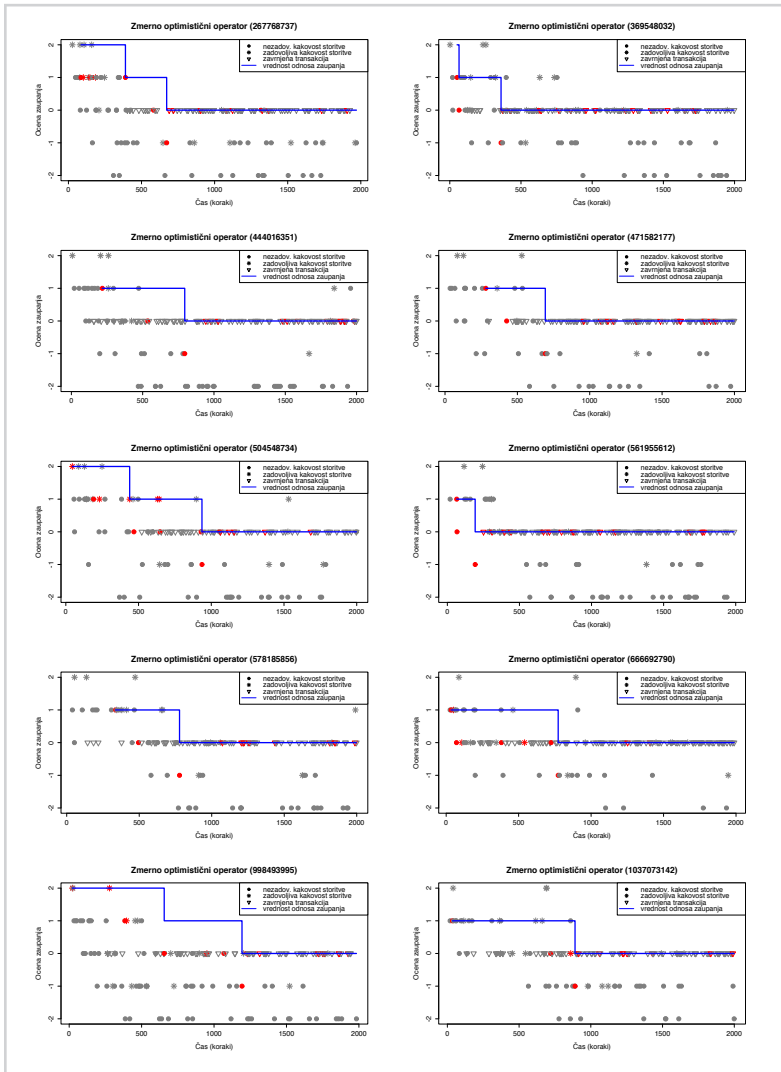
Obnašanje vrednosti zaupanja, modeliranega z zmernim optimističnim operatorjem  $\uparrow$  (1. del).



Slika A.12

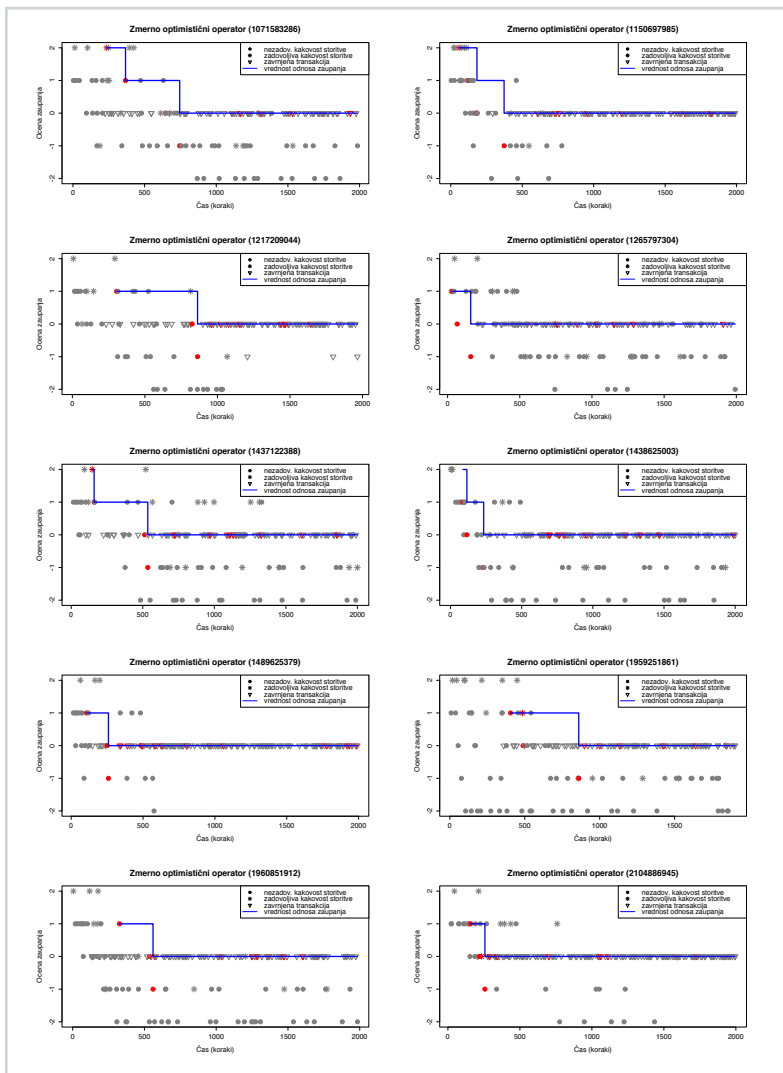
Obnašanje vrednosti zaupanja, modeliranega z zmernim optimističnim operatorjem  $\uparrow$  (2. del).





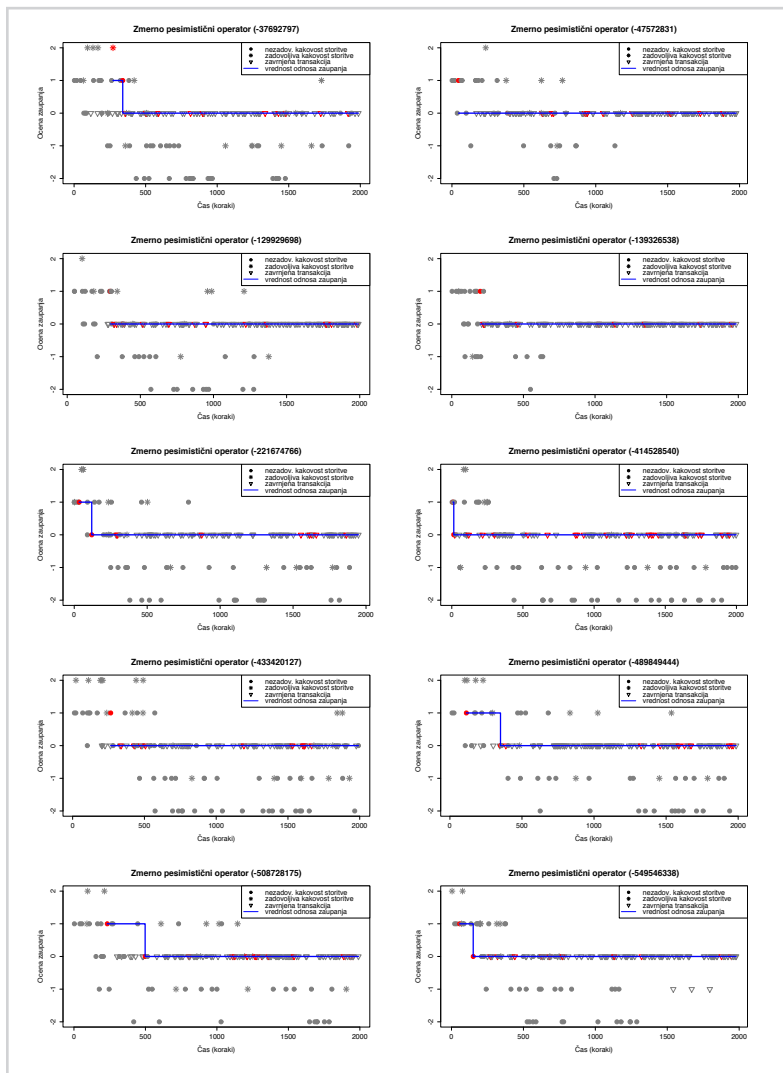
Slika A.14

Obnašanje vrednosti zaupanja, modeliranega z zmernim optimističnim operatorjem  $\uparrow$  (4. del).



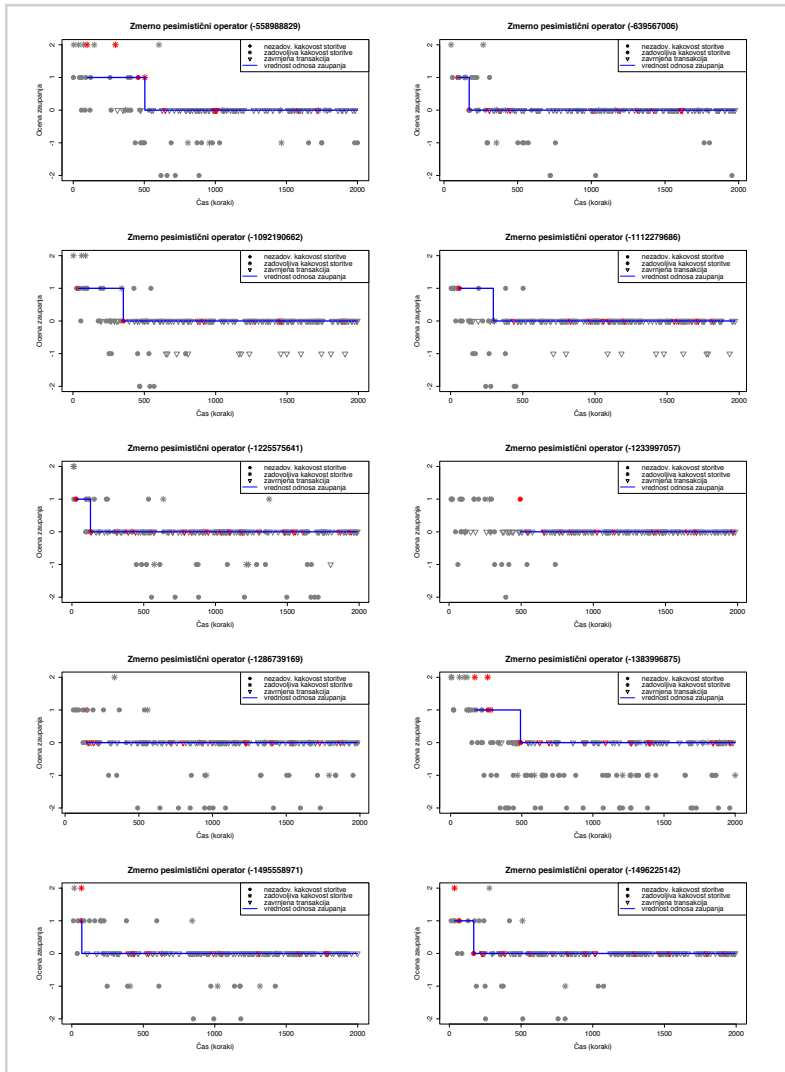
Slika A.15

Obnašanje vrednosti zaupanja, modeliranega z zmernim optimističnim operatorjem  $\uparrow$  (5. del).



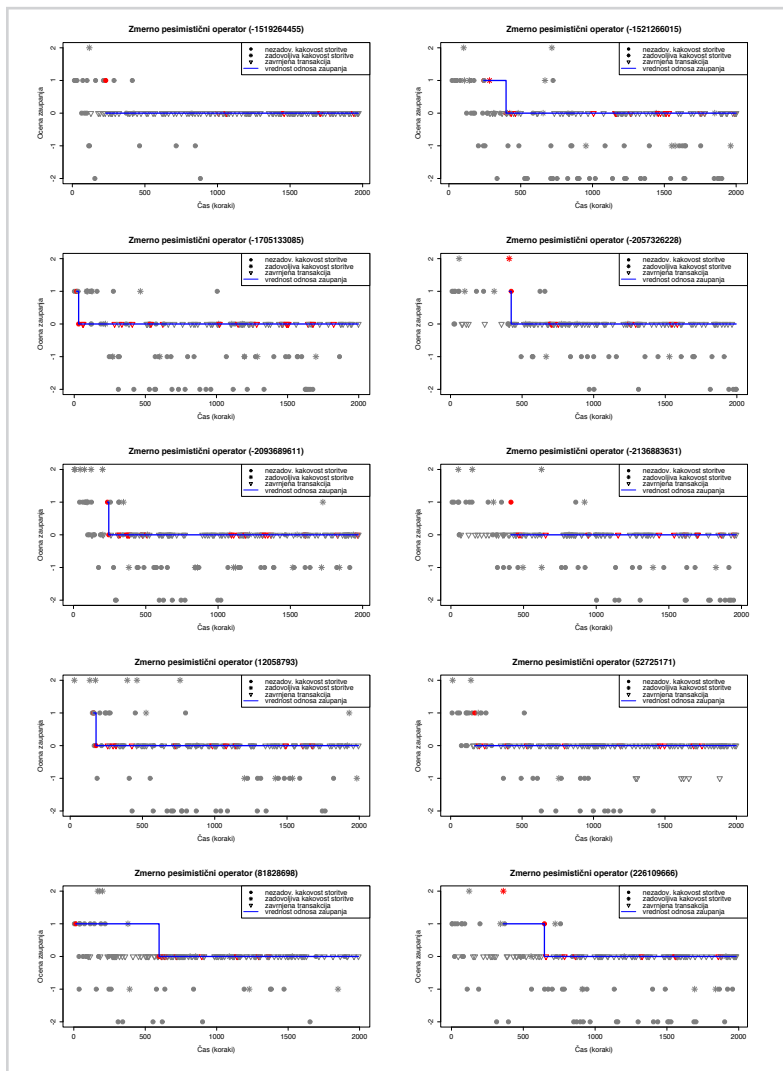
Slika A.16

Obnašanje vrednosti zaupanja, modeliranega z zmernim pesimističnim operatorjem  $\downarrow$  (t. del).



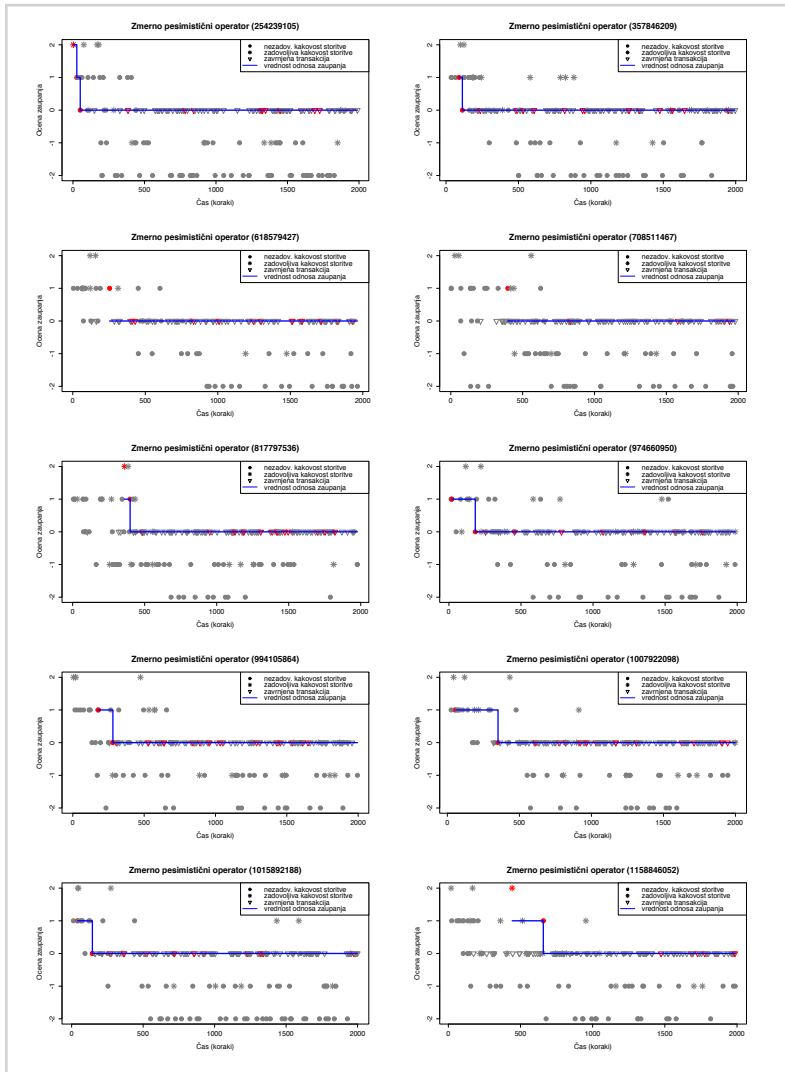
Slika A.17

Obnašanje vrednosti zaupanja, modeliranega z zmerim pesimističnim operatorjem ↓ (2. del).



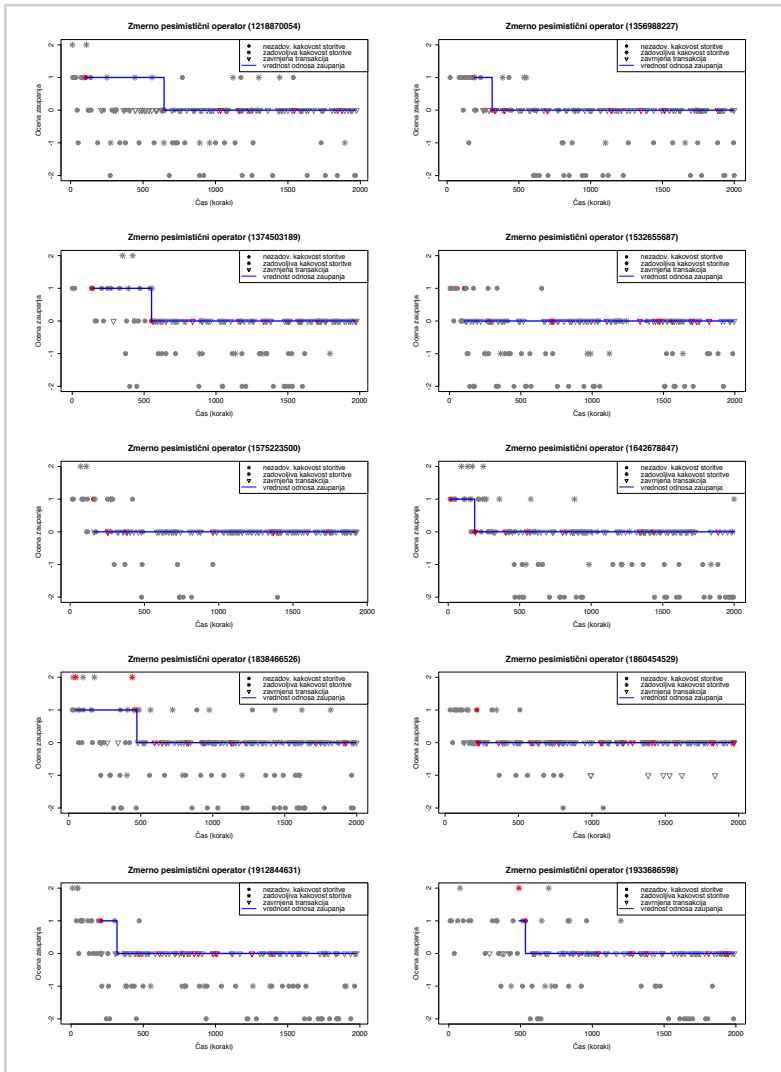
Slika A.18

Obnašanje vrednosti zaupanja, modeliranega z zmernim pesimističnim operatorjem  $\downarrow$  (3. del).



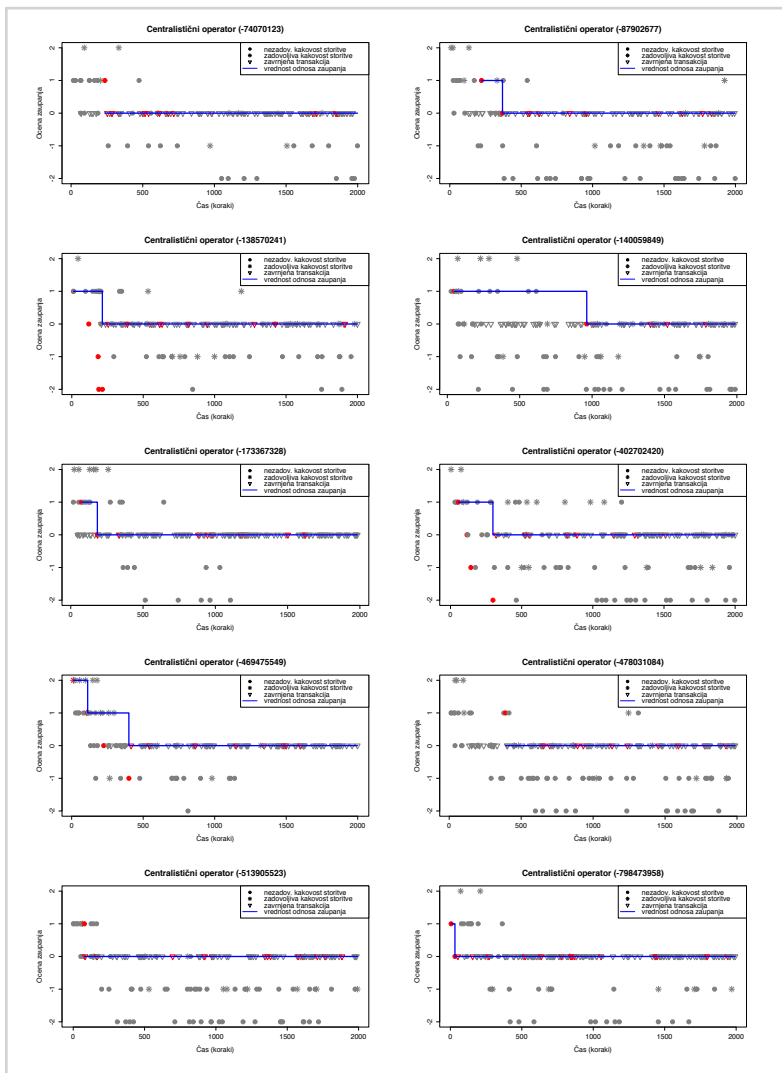
Slika A.19

Obnašanje vrednosti zaupanja, modeliranega z zmernim pesimističnim operatorjem  $\downarrow$  (4. del).



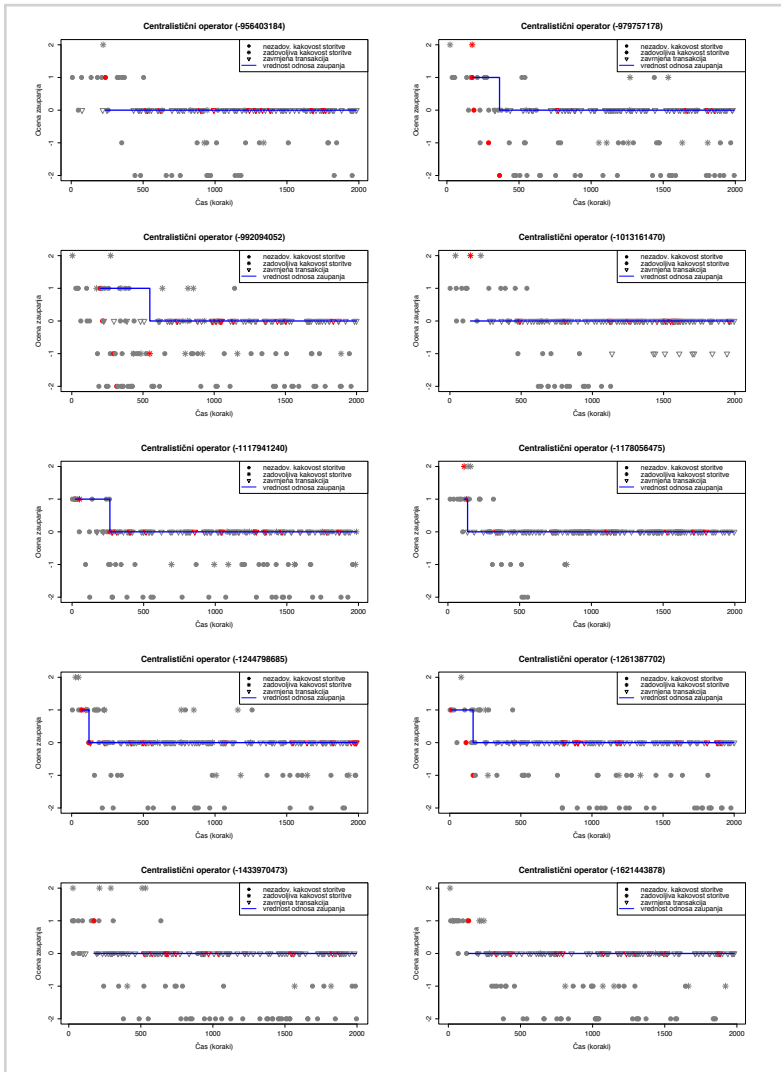
Slika A.20

Obnašanje vrednosti zaupanja, modeliranega z zmerim pesimističnim operatorjem  $\downarrow$  (5. del).



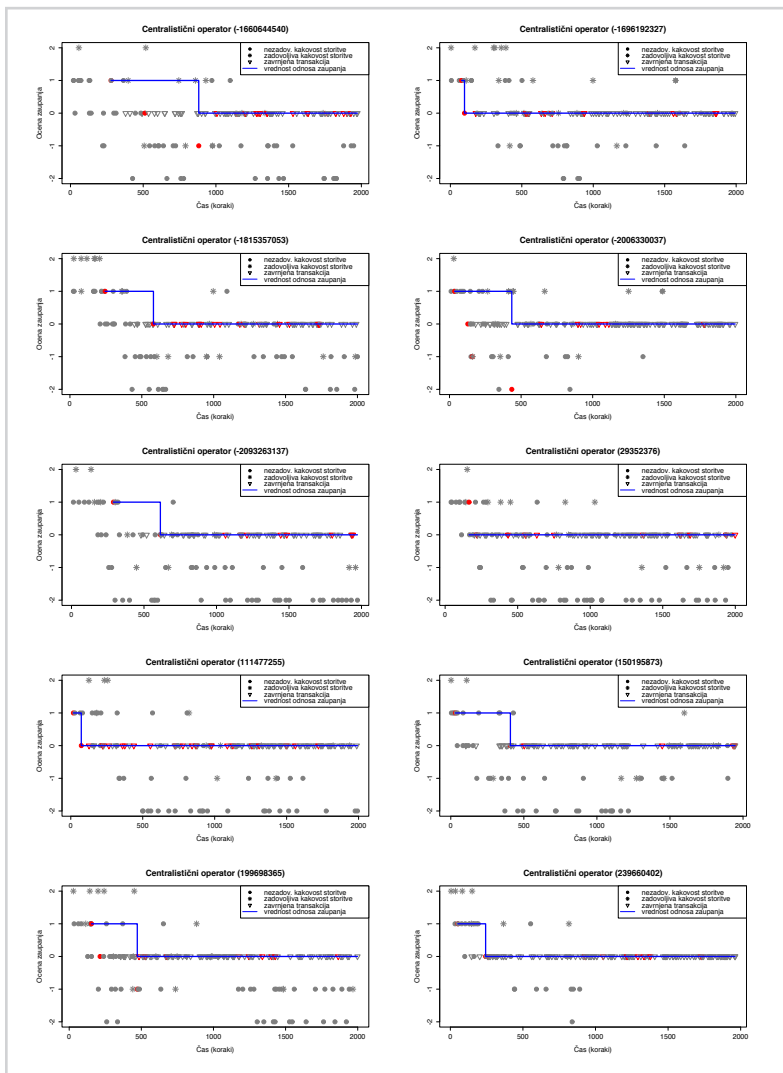
Slika A.21

Obnašanje vrednosti zaupanja, modeliranega s centralističnim operatorjem  $\rightarrow$  (1. del).



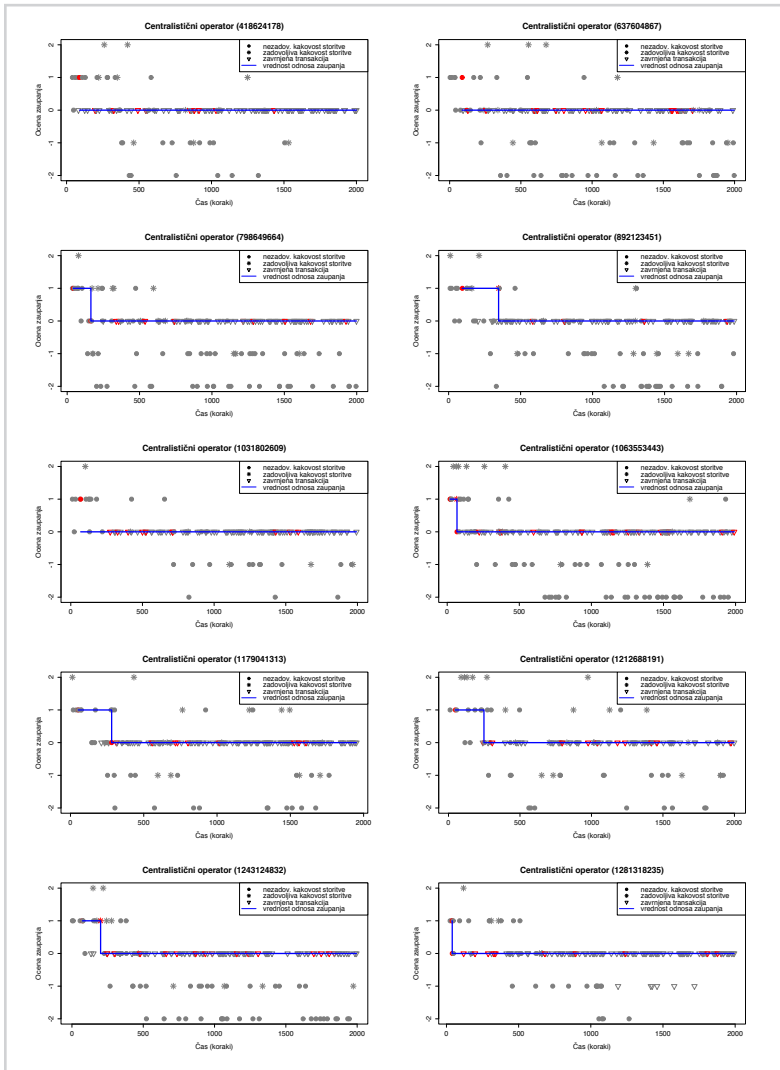
Slika A.22

Obnašanje vrednosti zaupanja, modeliranega s centralističnim operatorjem  $\rightarrow$  (2. del).



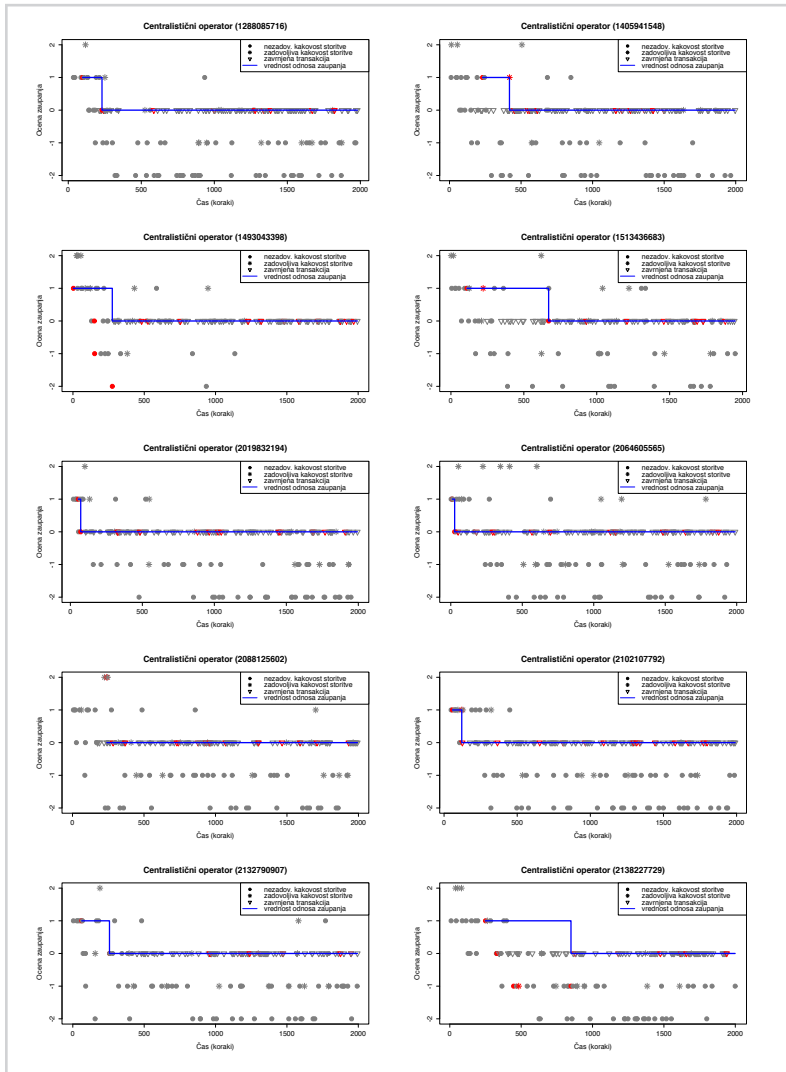
Slika A.23

Obnašanje vrednosti zaupanja, modeliranega s centralističnim operatorjem  $\rightarrow$  (3. del).



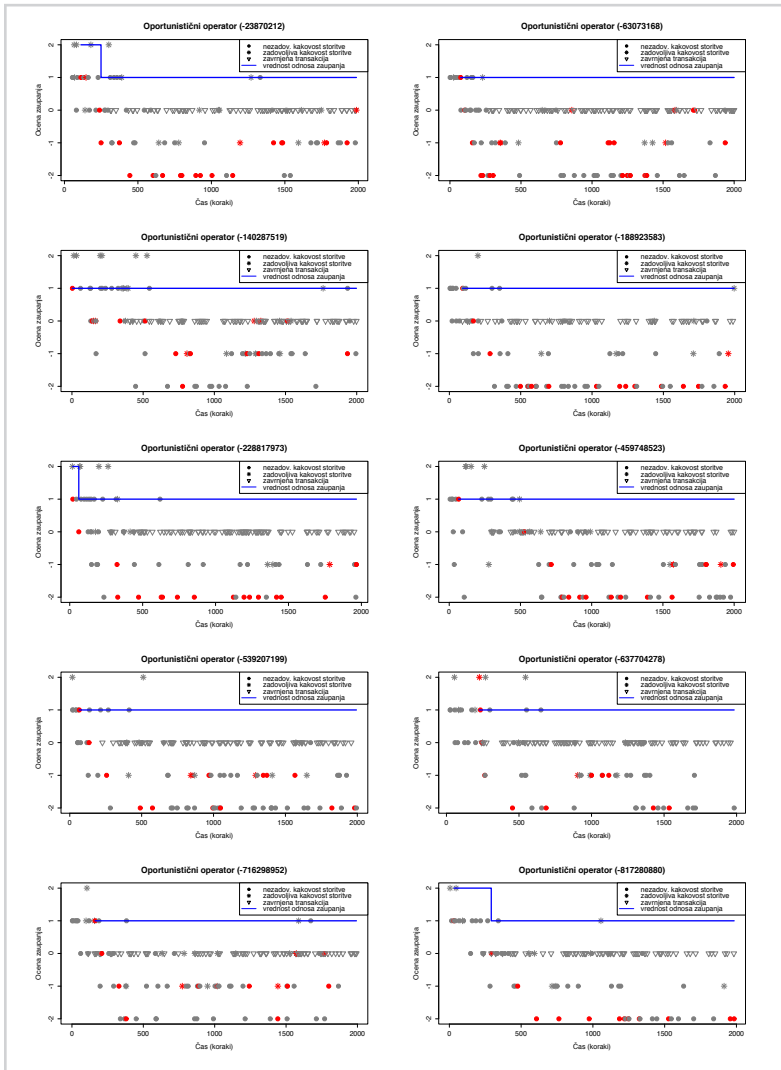
Slika A.24

Obnašanje vrednosti zaupanja, modeliranega s centralističnim operatorjem  $\rightarrow$  (4. del).



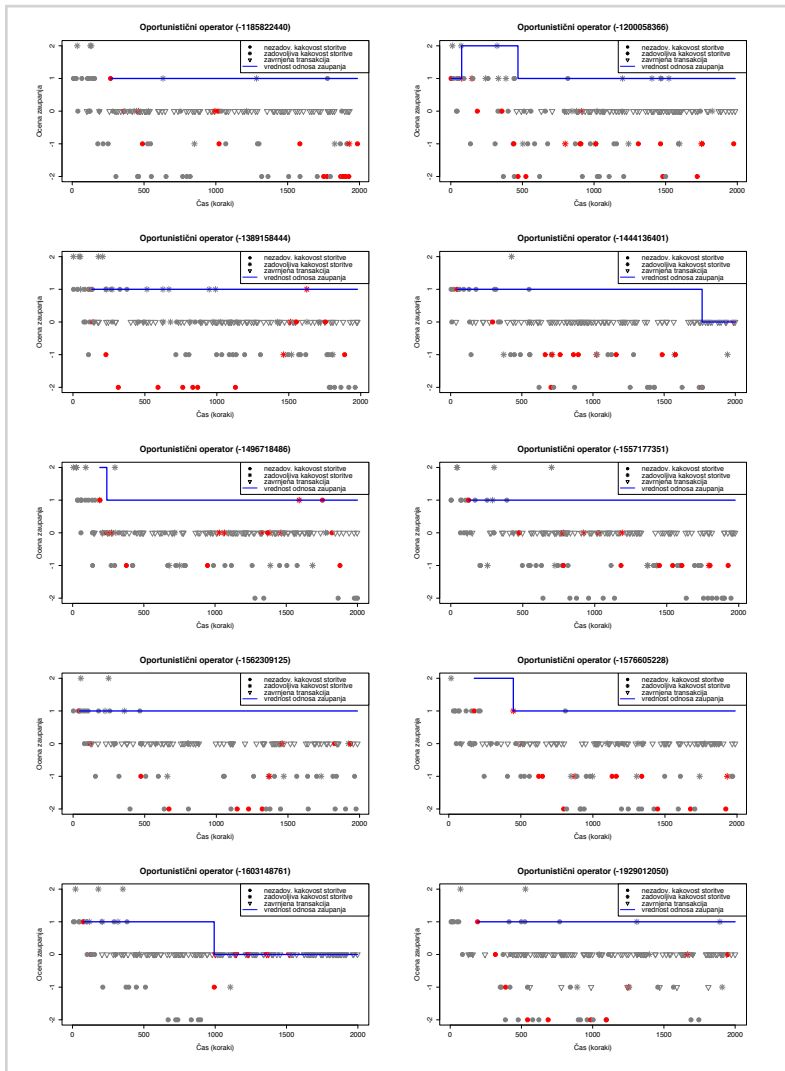
Slika A.25

Obnašanje vrednosti zaupanja, modeliranega s centralističnim operatorjem  $\rightarrow$  (5. del).



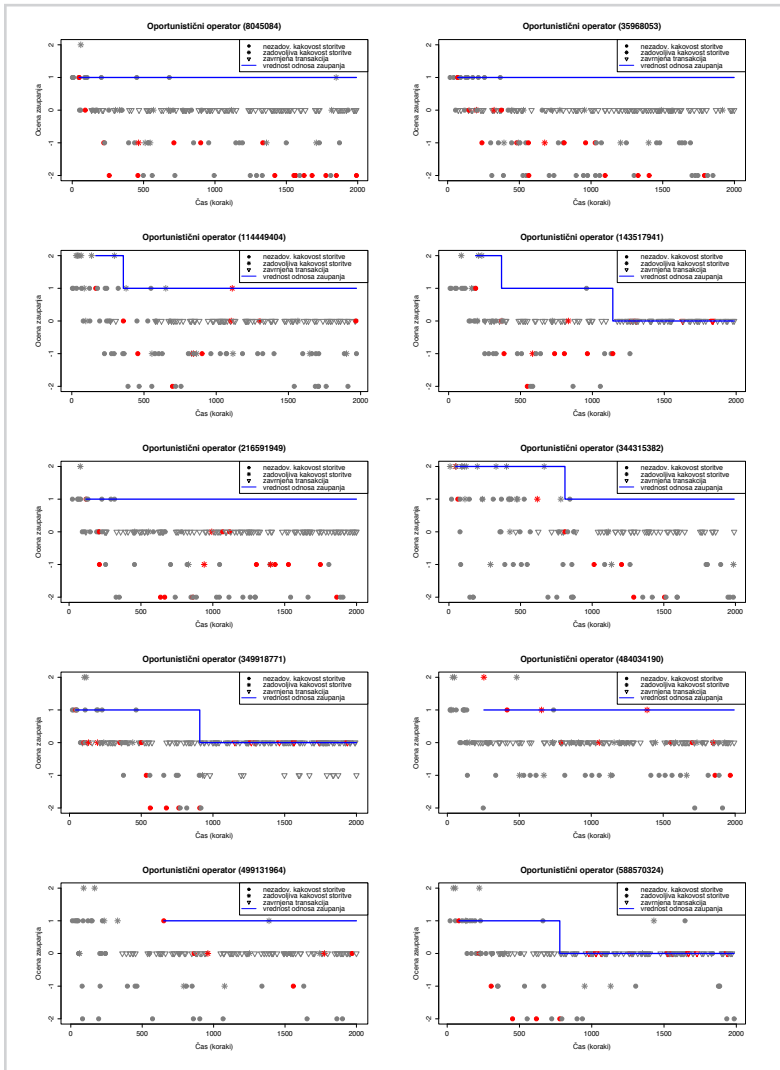
Slika A.26

Obnašanje vrednosti zaupanja, modeliranega z oportunističnim operatorjem  $\rightarrow$  (1. del).



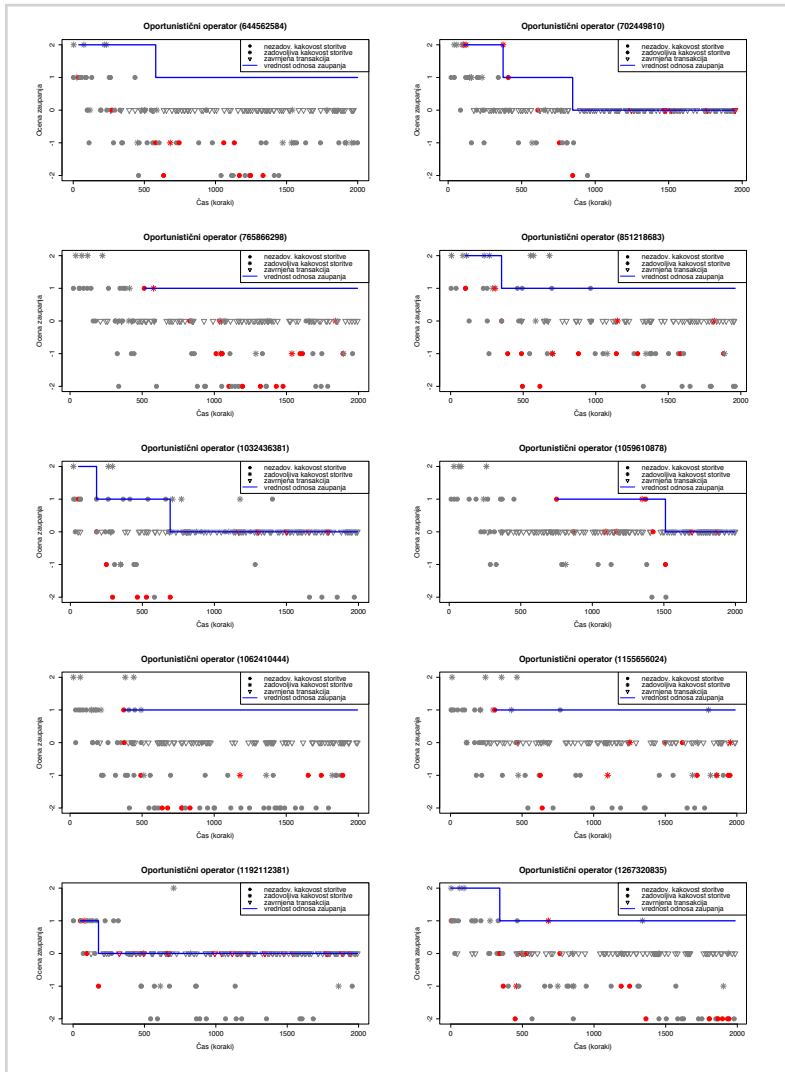
Slika A.27

Obnašanje vrednosti zaupanja, modelirana z oportunističnim operatorjem  $\rightarrow$  (2. del).



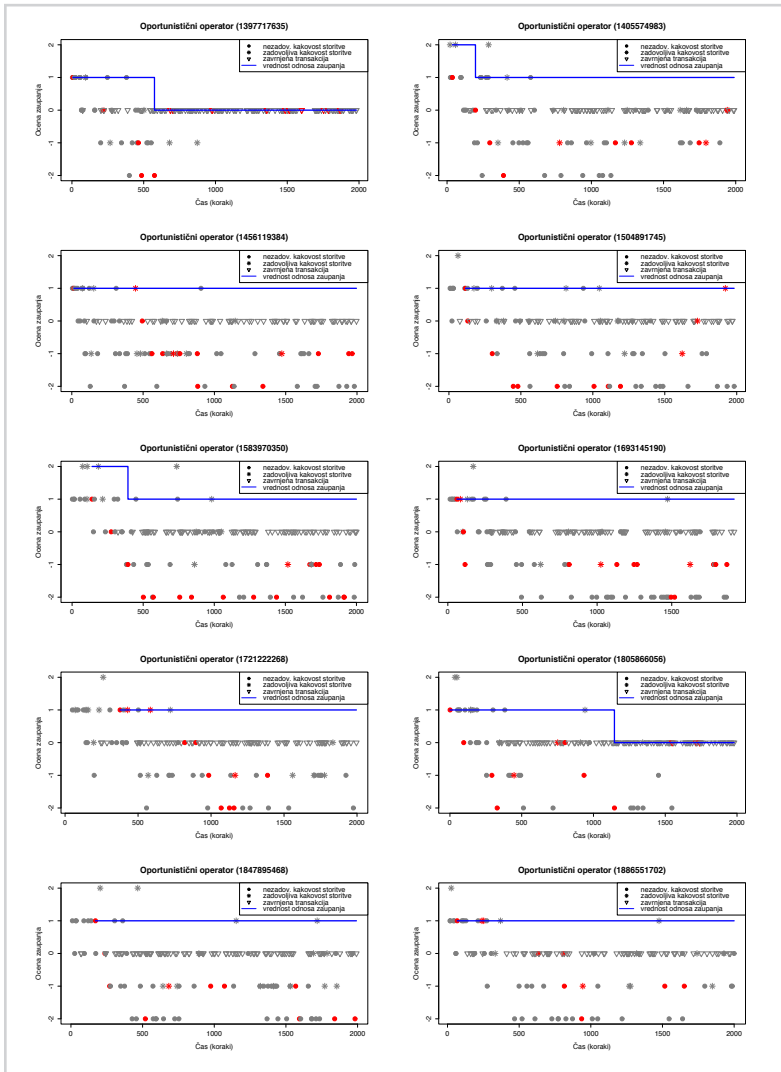
Slika A.28

Obnašanje vrednosti zaupanja, modeliranega z oportunističnim operatorjem  $\rightarrow$  (3. del).



Slika A.29

Obnašanje vrednosti zaupanja, modelirana z oportunističnim operatorjem  $\rightarrow$  (4. del).



Slika A.30

Obnašanje vrednosti zaupanja, modeliranega z oportunističnim operatorjem  $\rightarrow$  (5. del).



*Rezultati simulacij pri  
ovrednotenju metode  
QADE-filter*

*B*

Na slikah [B.1](#), [B.2](#), [B.3](#) in [B.4](#) so prikazani dodatni rezultati simulacij, s katerimi smo ovrednotili učinkovitost metod QADE-filter, TRAVOS-filter in Whitby filter. Potek evalvacij in rezultati so podrobno opisani v poglavju [4.3](#).

% agentov napadalcev	% napadenih agentov	Povprečna vrednost ocene zaupanja				
		Brez napadov	Samostojni BM napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	-0.01	-0.12	-0.04	-0.08	-0.12
10	20	-0.01	-0.24	-0.10	-0.17	-0.24
15	20	-0.01	-0.27	-0.12	-0.20	-0.27
20	20	-0.01	-0.32	-0.16	-0.22	-0.31
25	20	-0.01	-0.35	-0.14	-0.22	-0.33
30	20	-0.01	-0.37	-0.17	-0.22	-0.35
35	20	-0.01	-0.38	-0.18	-0.22	-0.36
40	20	-0.01	-0.42	-0.20	-0.24	-0.37
45	20	-0.01	-0.42	-0.20	-0.24	-0.38
50	20	-0.01	-0.46	-0.21	-0.24	-0.39
20	5	-0.01	-0.09	-0.01	-0.07	-0.10
20	10	-0.01	-0.23	-0.05	-0.16	-0.23
20	15	-0.01	-0.26	-0.07	-0.18	-0.27
20	20	-0.01	-0.32	-0.16	-0.22	-0.31
20	25	-0.01	-0.35	-0.18	-0.23	-0.34
20	30	-0.01	-0.37	-0.19	-0.22	-0.35
20	35	-0.01	-0.38	-0.18	-0.24	-0.35
20	40	-0.01	-0.42	-0.17	-0.24	-0.38
20	45	-0.01	-0.44	-0.17	-0.24	-0.38
20	50	-0.01	-0.46	-0.17	-0.24	-0.38

% agentov napadalcev	% napadenih agentov	Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja				
		Brez napadov	Samostojni BM napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	0.00	1.00	0.27	0.64	1.00
10	20	0.00	1.00	0.39	0.70	1.00
15	20	0.00	1.00	0.42	0.73	1.00
20	20	0.00	1.00	0.48	0.68	0.97
25	20	0.00	1.00	0.38	0.62	0.94
30	20	0.00	1.00	0.44	0.58	0.94
35	20	0.00	1.00	0.46	0.57	0.95
40	20	0.00	1.00	0.46	0.56	0.88
45	20	0.00	1.00	0.46	0.56	0.90
50	20	0.00	1.00	0.44	0.51	0.84
20	5	0.00	1.00	0.00	0.75	1.13
20	10	0.00	1.00	0.18	0.68	1.00
20	15	0.00	1.00	0.24	0.68	1.04
20	20	0.00	1.00	0.48	0.68	0.97
20	25	0.00	1.00	0.50	0.65	0.97
20	30	0.00	1.00	0.50	0.58	0.94
20	35	0.00	1.00	0.46	0.62	0.92
20	40	0.00	1.00	0.39	0.56	0.90
20	45	0.00	1.00	0.37	0.53	0.86
20	50	0.00	1.00	0.36	0.51	0.82
			<b>Povprečno:</b>	<b>0.39</b>	<b>0.62</b>	<b>0.95</b>
			Povprečje pri 20 % napadenih agentih:	0.42	0.61	0.94
			Povprečje pri 20 % agentih napadalcih:	0.35	0.62	0.96

Slika B.1

Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja pri samostojnem napadu z očitivijo.

% agentov napadalcev	% napadenih agentov	Povprečna vrednost ocene zaupanja				
		Brez napadov	Skupinski BM napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	-0.01	-0.11	-0.03	-0.09	-0.12
10	20	-0.01	-0.15	-0.02	-0.09	-0.14
15	20	-0.01	-0.15	-0.04	-0.10	-0.15
20	20	-0.01	-0.19	-0.05	-0.10	-0.16
25	20	-0.01	-0.19	-0.05	-0.10	-0.18
30	20	-0.01	-0.21	-0.06	-0.12	-0.17
35	20	-0.01	-0.19	-0.08	-0.14	-0.18
40	20	-0.01	-0.21	-0.07	-0.17	-0.20
45	20	-0.01	-0.22	-0.10	-0.19	-0.20
50	20	-0.01	-0.23	-0.12	-0.20	-0.22
20	5	-0.01	-0.04	0.00	-0.02	-0.03
20	10	-0.01	-0.09	-0.01	-0.05	-0.09
20	15	-0.01	-0.12	-0.01	-0.05	-0.11
20	20	-0.01	-0.19	-0.05	-0.10	-0.16
20	25	-0.01	-0.22	-0.07	-0.10	-0.20
20	30	-0.01	-0.28	-0.10	-0.14	-0.24
20	35	-0.01	-0.31	-0.10	-0.15	-0.26
20	40	-0.01	-0.36	-0.12	-0.19	-0.32
20	45	-0.01	-0.40	-0.14	-0.20	-0.36
20	50	-0.01	-0.47	-0.19	-0.23	-0.39

% agentov napadalcev	% napadenih agentov	Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja				
		Brez napadov	Skupinski BM napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	0.00	1.00	0.20	0.80	1.10
10	20	0.00	1.00	0.07	0.57	0.93
15	20	0.00	1.00	0.21	0.64	1.00
20	20	0.00	1.00	0.22	0.50	0.83
25	20	0.00	1.00	0.22	0.50	0.94
30	20	0.00	1.00	0.25	0.55	0.80
35	20	0.00	1.00	0.39	0.72	0.94
40	20	0.00	1.00	0.30	0.80	0.95
45	20	0.00	1.00	0.43	0.86	0.90
50	20	0.00	1.00	0.50	0.86	0.95
20	5	0.00	1.00	-0.33	0.33	0.67
20	10	0.00	1.00	0.00	0.50	1.00
20	15	0.00	1.00	0.00	0.36	0.91
20	20	0.00	1.00	0.22	0.50	0.83
20	25	0.00	1.00	0.29	0.43	0.90
20	30	0.00	1.00	0.33	0.48	0.85
20	35	0.00	1.00	0.30	0.47	0.83
20	40	0.00	1.00	0.31	0.51	0.89
20	45	0.00	1.00	0.33	0.49	0.90
20	50	0.00	1.00	0.39	0.48	0.83
		<b>Povprečno:</b>	<b>0.23</b>	<b>0.57</b>	<b>0.90</b>	
		Povprečje pri 20 % napadenih agentih:	0.28	0.68	0.94	
		Povprečje pri 20 % agentih napadalcih:	0.18	0.46	0.86	

Slika B.2

Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja pri skupinskem napadu z očitnitvijo.

% agentov napadalcev	% napadenih agentov	Povprečna vrednost ocene zaupanja				
		Brez napadov	Samostojni BS napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	-0.01	0.11	0.03	0.10	0.12
10	20	-0.01	0.23	0.08	0.16	0.25
15	20	-0.01	0.26	0.10	0.18	0.27
20	20	-0.01	0.32	0.14	0.21	0.30
25	20	-0.01	0.35	0.16	0.22	0.34
30	20	-0.01	0.38	0.16	0.23	0.34
35	20	-0.01	0.39	0.18	0.23	0.36
40	20	-0.01	0.41	0.21	0.22	0.36
45	20	-0.01	0.43	0.21	0.24	0.40
50	20	-0.01	0.44	0.22	0.23	0.40
20	5	-0.01	0.10	0.01	0.08	0.08
20	10	-0.01	0.21	0.06	0.16	0.23
20	15	-0.01	0.27	0.07	0.18	0.28
20	20	-0.01	0.32	0.14	0.21	0.30
20	25	-0.01	0.33	0.18	0.22	0.32
20	30	-0.01	0.38	0.19	0.23	0.35
20	35	-0.01	0.38	0.19	0.23	0.35
20	40	-0.01	0.43	0.18	0.24	0.37
20	45	-0.01	0.43	0.19	0.24	0.38
20	50	-0.01	0.47	0.18	0.25	0.38

% agentov napadalcev	% napadenih agentov	Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja				
		Brez napadov	Samostojni BS napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	0.00	1.00	0.33	0.92	1.08
10	20	0.00	1.00	0.38	0.71	1.08
15	20	0.00	1.00	0.41	0.70	1.04
20	20	0.00	1.00	0.45	0.67	0.94
25	20	0.00	1.00	0.47	0.64	0.97
30	20	0.00	1.00	0.44	0.62	0.90
35	20	0.00	1.00	0.48	0.60	0.93
40	20	0.00	1.00	0.52	0.55	0.88
45	20	0.00	1.00	0.50	0.57	0.93
50	20	0.00	1.00	0.51	0.53	0.91
20	5	0.00	1.00	0.18	0.82	0.82
20	10	0.00	1.00	0.32	0.77	1.09
20	15	0.00	1.00	0.29	0.68	1.04
20	20	0.00	1.00	0.45	0.67	0.94
20	25	0.00	1.00	0.56	0.68	0.97
20	30	0.00	1.00	0.51	0.62	0.92
20	35	0.00	1.00	0.51	0.62	0.92
20	40	0.00	1.00	0.43	0.57	0.86
20	45	0.00	1.00	0.45	0.57	0.89
20	50	0.00	1.00	0.40	0.54	0.81
			<b>Povprečno:</b>	<b>0.43</b>	<b>0.65</b>	<b>0.95</b>
			Povprečje pri 20 % napadenih agentih:	0.45	0.65	0.97
			Povprečje pri 20 % agentih napadalcih:	0.41	0.65	0.93

Slika B.3

Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja pri samostojnem napadu s hvaljenjem.

% agentov napadalcev	% napadenih agentov	Povprečna vrednost ocene zaupanja				
		Brez napadov	Skupinski BS napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	-0.01	0.10	0.04	0.09	0.12
10	20	-0.01	0.14	0.03	0.10	0.13
15	20	-0.01	0.15	0.03	0.10	0.13
20	20	-0.01	0.18	0.03	0.09	0.15
25	20	-0.01	0.19	0.05	0.11	0.17
30	20	-0.01	0.20	0.06	0.14	0.17
35	20	-0.01	0.20	0.07	0.16	0.19
40	20	-0.01	0.21	0.09	0.16	0.20
45	20	-0.01	0.23	0.10	0.19	0.21
50	20	-0.01	0.24	0.11	0.21	0.23
20	5	-0.01	0.03	0.01	0.02	0.02
20	10	-0.01	0.09	0.01	0.03	0.08
20	15	-0.01	0.12	0.00	0.06	0.09
20	20	-0.01	0.18	0.03	0.09	0.15
20	25	-0.01	0.20	0.06	0.11	0.20
20	30	-0.01	0.28	0.10	0.15	0.24
20	35	-0.01	0.30	0.09	0.16	0.27
20	40	-0.01	0.38	0.12	0.19	0.32
20	45	-0.01	0.40	0.14	0.22	0.36
20	50	-0.01	0.46	0.18	0.25	0.41

% agentov napadalcev	% napadenih agentov	Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja				
		Brez napadov	Skupinski BS napad	QADE-filter	TRAVOS-filter	Whitby filter
5	20	0.00	1.00	0.45	0.91	1.18
10	20	0.00	1.00	0.27	0.73	0.93
15	20	0.00	1.00	0.25	0.69	0.88
20	20	0.00	1.00	0.21	0.53	0.84
25	20	0.00	1.00	0.30	0.60	0.90
30	20	0.00	1.00	0.33	0.71	0.86
35	20	0.00	1.00	0.38	0.81	0.95
40	20	0.00	1.00	0.45	0.77	0.95
45	20	0.00	1.00	0.46	0.83	0.92
50	20	0.00	1.00	0.48	0.88	0.96
20	5	0.00	1.00	0.50	0.75	0.75
20	10	0.00	1.00	0.20	0.40	0.90
20	15	0.00	1.00	0.08	0.54	0.77
20	20	0.00	1.00	0.21	0.53	0.84
20	25	0.00	1.00	0.33	0.57	1.00
20	30	0.00	1.00	0.38	0.55	0.86
20	35	0.00	1.00	0.32	0.55	0.90
20	40	0.00	1.00	0.33	0.51	0.85
20	45	0.00	1.00	0.37	0.56	0.90
20	50	0.00	1.00	0.40	0.55	0.89
		<b>Povprečno:</b>		<b>0.34</b>	<b>0.65</b>	<b>0.90</b>
		Povprečje pri 20 % napadenih agentih:		0.36	0.75	0.94
		Povprečje pri 20 % agentih napadalcih:		0.31	0.55	0.87

#### Slika B.4

Normalizirana razlika v povprečni vrednosti ocene zaupanja glede na reprezentativni odnos zaupanja pri skupinskem napadu s hvaljenjem.

## LITERATURA

- [1] Paul Resnick and Richard Zeckhauser. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In Michael R. Baye, editor, *The Economics of the Internet and E-Commerce*, volume 11 of *Advances in Applied Microeconomics*, pages 127–157. Elsevier Science, 2002. doi: [10.1016/S0278-0984\(02\)11030-3](https://doi.org/10.1016/S0278-0984(02)11030-3).
- [2] Mark L. Knapp and Judith A. Hall. *Nonverbal Communication in Human Interaction*. Cengage Learning, 2009. ISBN 978-0-495-56869-8.
- [3] Andrew P. Bayliss and Steven P. Tipper. Predictive Gaze Cues and Personality Judgments: Should Eye Trust You? *Psychological Science*, 17(6):514–520, 2006. doi: [10.1111/j.1467-9280.2006.01737.x](https://doi.org/10.1111/j.1467-9280.2006.01737.x).
- [4] Audun Jøsang, Claudia Keser, and Theo Dimitrakos. Can we manage trust? In *Proceedings of the Third International Conference on Trust Management (iTrust), Versailles*, pages 93–107. Springer-Verlag, 2005.
- [5] Touhid Bhuiyan, Audun Jøsang, and Yue Xu. Trust and reputation management in web-based social network. In Zeeshan-Ul-Hassan Usmani, editor, *Web Intelligence and Intelligent Agents*, pages 207–232. InTech, 2010. ISBN 978-953-7619-85-5. doi: [10.5772/8375](https://doi.org/10.5772/8375).
- [6] Audun Jøsang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2):618–644, 2007. ISSN 0167-9236. doi: [10.1016/j.dss.2005.05.019](https://doi.org/10.1016/j.dss.2005.05.019).
- [7] Andreas Gutscher. A trust model for an open, decentralized reputation system. In Sandro Etalle and Stephen Marsh, editors, *Trust Management*, volume 238 of *IFIP International Federation for Information Processing*, pages 285–300. Springer US, 2007. ISBN 978-0-387-73654-9. doi: [10.1007/978-0-387-73655-6\\_19](https://doi.org/10.1007/978-0-387-73655-6_19).
- [8] Ilung Pranata, Geoff Skinner, and Rukshan Athauda. A Holistic Review on Trust and Reputation Management Systems for Digital Environments. *International Journal of Computer and Information Technology*, 01(01):44–53, 2012.
- [9] Isaac Pinyol and Jordi Sabater-Mir. Computational trust and reputation models for open multi-agent systems: a review. *Artificial Intelligence Review*, pages 1–25, 2011. ISSN 0269-2821. doi: [10.1007/s10462-011-9277-z](https://doi.org/10.1007/s10462-011-9277-z).
- [10] Zeinab Noorian and Mihaela Ulieru. The State of the Art in Trust and Reputation Systems: A Framework for Comparison. *Journal of theoretical and applied electronic commerce research*, 5(2):97–117, August 2010. ISSN 0718-1876. doi: [10.4067/S0718-18762010000200007](https://doi.org/10.4067/S0718-18762010000200007).
- [11] Babak Khosravifar, Jamal Bentahar, Maziar Gomrochi, and Rafiul Alam. CRM: An efficient trust and reputation model for agent computing. *Knowledge-Based Systems*, 30:1–16, June 2012. ISSN 0950-7051. doi: [10.1016/j.knsys.2011.01.004](https://doi.org/10.1016/j.knsys.2011.01.004).
- [12] Matthew S. Eastin. Diffusion of e-commerce: an analysis of the adoption of four e-commerce activities. *Telematics and Informatics*, 19(3):251–267, 2002. ISSN 0736-5853. doi: [10.1016/S0736-5853\(01\)00005-3](https://doi.org/10.1016/S0736-5853(01)00005-3).
- [13] Damjan Kovač. *Ovladovanje zaupanja v storitveno usmerjenih arhitekturah*. Phd, University of Ljubljana, Faculty of Computer and Information Science, 2009.
- [14] Chrysanthos Dellarocas. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10):1407–1424, October 2003. ISSN 0025-1909. doi: [10.1287/mnsc.49.10.1407.17308](https://doi.org/10.1287/mnsc.49.10.1407.17308).
- [15] Denis Trček. A formal apparatus for modeling trust in computing environments. *Mathematical and Computer Modelling*, 49(1-2):226–233, 2009. doi: [10.1016/j.mcm.2008.05.005](https://doi.org/10.1016/j.mcm.2008.05.005).
- [16] Denis Trček. Ergonomic Trust Management in Pervasive Computing Environments - Qualitative

- Assessment Dynamics. In *Pervasive Computing and Applications (ICPCA), 2010 5th International Conference on*, pages 7–13. IEEE, 2010.
- [17] ISO27001: Information Security Management System (ISMS) standard. October 2005. URL <http://www.27000.org/iso-27001.htm>.
- [18] Lars Rasmussen and Sverker Jansson. Simulated social control for secure Internet commerce. In *Proceedings of the 1996 workshop on New security paradigms*, pages 18–25, New York, New York, USA, 1996. ACM. doi: [10.1145/304851.304857](https://doi.org/10.1145/304851.304857).
- [19] Han Yu, Zhiqi Shen, Chunyan Miao, Cyril Leung, and Dusit Niyato. A Survey of Trust and Reputation Management Systems in Wireless Communications. *Proceedings of the IEEE*, 98(10):1755–1772, 2010. ISSN 0018-9219. doi: [10.1109/JPROC.2010.2059690](https://doi.org/10.1109/JPROC.2010.2059690).
- [20] *Slovar slovenskega knjižnega jezika*. Založba ZRC, ZRC SAZU, 2008. URL <http://bos.zrc-sazu.si/sskj.html>.
- [21] Huaizhi Li and Mukesh Singhal. Trust Management in Distributed Systems. *Computer*, 40(2):45–53, February 2007. ISSN 0018-9162. doi: [10.1109/MC.2007.76](https://doi.org/10.1109/MC.2007.76).
- [22] Diego Gambetta. Can we trust trust? In *Trust: Making and Breaking Cooperative Relations*, pages 213–237. Basil Blackwell, 1988.
- [23] Lik Mui, Mojdeh Mohtashemi, and Ari Halberstadt. A computational model of trust and reputation. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, volume 00, pages 2431–2439. IEEE, 2002. ISBN 0-7695-1435-9. doi: [10.1109/HICSS.2002.994181](https://doi.org/10.1109/HICSS.2002.994181).
- [24] Lik Mui, Mojdeh Mohtashemi, Cheewee Ang, Peter Szolovits, and Ari Halberstadt. Ratings in Distributed Systems: A Bayesian Approach. In *Workshop on Information Technologies and Systems (WITS)*, pages 1–7, 2001.
- [25] Jianshu Weng, Chunyan Miao, and Angela Goh. An entropy-based approach to protecting rating systems from unfair testimonies. *IEICE TRANSACTIONS on Information and Systems*, E89-D(9):2502–2511, September 2006. ISSN 0916-8532. doi: [10.1093/ietfisy/e89-d.9.2502](https://doi.org/10.1093/ietfisy/e89-d.9.2502).
- [26] Jordi Sabater-Mir and Carles Sierra. Reputation and social network analysis in multi-agent systems. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems (AAMAS '02)*, New York, New York, USA, 2002. ACM Press. ISBN 1-58113-480-0. doi: [10.1145/544741.544854](https://doi.org/10.1145/544741.544854).
- [27] Jordi Sabater-Mir and Carles Sierra. REGRET: reputation in gregarious societies. In *Proceedings of the fifth international conference on Autonomous agents, AGENTS '01*, pages 194–195, New York, NY, USA, 2001. ACM. ISBN 1-58113-326-X. doi: [10.1145/375735.376110](https://doi.org/10.1145/375735.376110).
- [28] Alfarez Abdul-Rahman and Stephen Hailes. Supporting trust in virtual communities. In *Proceedings of the 33rd Hawaii International Conference on System Sciences—Volume 6 - Volume 6, HICSS '00*, pages 6007–, Washington, DC, USA, 2000. IEEE Computer Society. ISBN 0-7695-0493-0.
- [29] Yao Wang and Julita Vassileva. Bayesian network trust model in peer-to-peer networks. In Gianluca Moro, Claudio Sartori, and Munindar P. Singh, editors, *Agents and Peer-to-Peer Computing*, volume 2872 of *Lecture Notes in Computer Science*, pages 23–34. Springer Berlin Heidelberg, 2005. ISBN 978-3-540-24053-2. doi: [10.1007/978-3-540-25840-7\\_3](https://doi.org/10.1007/978-3-540-25840-7_3).
- [30] Audun Jøsang and Roslan Ismail. The beta reputation system. In *Proceedings of the 15th Bled Electronic Commerce Conference*, volume 160, pages 41–55. Taylor & Francis, Palgrave Macmillan, 2002.
- [31] W. T. Luke Teacy, Jigar Parel, Nicholas R. Jennings, and Michael Luck. TRAVOS: Trust and Reputation in the Context of Inaccurate Information Sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, February 2006. ISSN 1387-2532. doi: [10.1007/s10458-006-5952-x](https://doi.org/10.1007/s10458-006-5952-x).
- [32] Bin Yu and Munindar P. Singh. An evidential model of distributed reputation management. In *Proceedings of the first international joint conference on Autonomous Agents and Multiagent Systems (AAMAS '02)*, page 294, New York, New York, USA, 2002. ACM Press. ISBN 1-58113-480-0. doi: [10.1145/544805.544809](https://doi.org/10.1145/544805.544809).
- [33] Audun Jøsang. A logic for uncertain probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(3):279–311, 2001. doi: [10.1142/S0218488501000311](https://doi.org/10.1142/S0218488501000311).
- [34] Moshe Tennenholtz. Game-theoretic recommendations: some progress in an uphill battle. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 1, AAMAS '08*, pages 10–16, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 978-0-9817381-0-9.
- [35] M. Harish, N. Anandavelu, N. Anbalagan, G.S. Mahalakshmi, and T.V. Geetha. Design and Analysis of a Game Theoretic Model for P2P Trust Management. In Tomasz Janowski and Hrushikesh Mohanty, editors, *Distributed Computing and Internet Technology*, volume 4882 of *Lecture Notes in Computer Science*, pages 110–115. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-77112-8. doi: [10.1007/978-3-540-77115-9\\_11](https://doi.org/10.1007/978-3-540-77115-9_11).

- [36] Patricia Victor, Chris Cornelis, Martine De Cock, and Enrique Herrera-Viedma. Practical aggregation operators for gradual trust and distrust. *Fuzzy Sets and Systems*, 184(1):126–147, December 2011. ISSN 0165-0114. doi: [10.1016/j.fss.2010.10.015](https://doi.org/10.1016/j.fss.2010.10.015).
- [37] Daniel W. Manchala. Trust metrics, models and protocols for electronic commerce transactions. In *Proceedings of the The 18th International Conference on Distributed Computing Systems*, ICDCS '98, pages 312–321, Washington, DC, USA, 1998. IEEE Computer Society. ISBN 0-8186-8292-2. doi: [10.1109/ICDCS.1998.679731](https://doi.org/10.1109/ICDCS.1998.679731).
- [38] Stephen Marsh. Optimism and pessimism in trust. In *Proceedings of the Ibero-American Conference on Artificial Intelligence (IBERAMIA '94)*, pages 1–12. McGraw-Hill Publishing, 1994.
- [39] Giorgos Zacharia and Pattie Maes. Trust management through reputation mechanisms. *Applied Artificial Intelligence*, 14(9):881–907, 2000. doi: [10.1080/08839510050144868](https://doi.org/10.1080/08839510050144868).
- [40] Denis Trček. Trust management methodologies for the web. In *Proceedings of the 7th international conference on Reasoning web: semantic technologies for the web of data*, RW'11, pages 445–459. Springer Berlin Heidelberg, 2011. ISBN 978-3-642-23031-8.
- [41] Jordi Sabater-Mir and Carles Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, September 2005. ISSN 0269-2821. doi: [10.1007/s10462-004-0041-5](https://doi.org/10.1007/s10462-004-0041-5).
- [42] Sarah N. Lim Choi Keung and Nathan Griffiths. Trust and Reputation. In Nathan Griffiths and Kuo-Ming Chao, editors, *Agent-Based Service-Oriented Computing*, pages 189–224. Springer London, London, 2010. ISBN 978-1-84996-040-3. doi: [10.1007/978-1-84996-041-0](https://doi.org/10.1007/978-1-84996-041-0).
- [43] Ling Liu and Malcolm Munro. Systematic analysis of centralized online reputation systems. *Decision Support Systems*, 52(2):438–449, January 2012. ISSN 0167-9236. doi: [10.1016/j.dss.2011.10.003](https://doi.org/10.1016/j.dss.2011.10.003).
- [44] Tyrone Grandison and Morris Sloman. A survey of trust in internet applications. *Communications Surveys Tutorials, IEEE*, 3(4):2–16, quarter 2000. ISSN 1553-877X. doi: [10.1109/COMST.2000.5340804](https://doi.org/10.1109/COMST.2000.5340804).
- [45] Stephen Paul Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, Department of Computing Science and Mathematics, University of Stirling, 1994.
- [46] Andrew Whitby, Audun Jøsang, and Jadwiga Indulska. Filtering Out Unfair Ratings in Bayesian Reputation Systems. In *Proc. 7th Int. Workshop on Trust in Agent Societies*, 2004. URL <http://svn.triabler.org/abc/branches/Leo/dataset/preferences/johan/johan-64.pdf>.
- [47] Jie Zhang and Robin Cohen. A personalized approach to address unfair ratings in multiagent reputation systems. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'06) Workshop on Trust in Agent Societies*, 2006.
- [48] Jie Zhang and Robin Cohen. Evaluating the Trustworthiness of Advice About Seller Agents in e-Marketplaces: A Personalized Approach. *Electronic Commerce Research and Applications*, 7(3): 330–340, November 2008. ISSN 1567-4223. doi: [10.1016/j.elerap.2008.03.001](https://doi.org/10.1016/j.elerap.2008.03.001).
- [49] Bin Yu and Munindar P. Singh. Distributed reputation management for electronic commerce. *Computational Intelligence*, 18(4):535–549, 2002. ISSN 1467-8640. doi: [10.1111/1467-8640.00202](https://doi.org/10.1111/1467-8640.00202).
- [50] Bin Yu and Munindar P. Singh. Detecting deception in reputation management. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems (AAMAS '03)*, pages 73–80, New York, New York, USA, 2003. ACM Press. ISBN 1-58113-683-8. doi: [10.1145/860586.860588](https://doi.org/10.1145/860586.860588).
- [51] Glenn Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.
- [52] Audun Jøsang. *Subjective Logic*. 2013. URL [http://folk.uio.no/josang/papers/subjective\\_logic.pdf](http://folk.uio.no/josang/papers/subjective_logic.pdf).
- [53] Jordi Sabater-Mir and Carles Sierra. Social regret, a reputation model based on social relations. *ACM SIGecom Exchanges - Chains of commitment*, 3(1):44–56, December 2001. ISSN 1551-9031. doi: [10.1145/844331.844337](https://doi.org/10.1145/844331.844337).
- [54] Gior Zacharia, Alexandros G Moukas, and Pattie Maes. Collaborative reputation mechanisms for electronic marketplaces. *Decision Support Systems*, 29(4): 371–388, 2000. doi: [10.1016/S0167-9236\(00\)00084-1](https://doi.org/10.1016/S0167-9236(00)00084-1).
- [55] Amazon. <http://www.amazon.com/>, jun 2014.
- [56] eBay. <http://www.ebay.com/>, jun 2014.
- [57] David Jelenc, Ramón Hermoso, Jordi Sabater-Mir, and Denis Trček. Decision making matters: A better way to evaluate trust models. *Knowledge-Based Systems*, 52:147–164, 2013. doi: [10.1016/j.knsys.2013.07.016](https://doi.org/10.1016/j.knsys.2013.07.016).
- [58] Baptiste Alcalde, Eric Dubois, Sjouke Mauw, Nicolas Mayer, and Saša Radomirović. Towards a decision model based on trust and security risk management. In Ljiljana Brankovic and Willy Susilo, editors, *Seventh Australasian Information Security Conference (AISC 2009)*, volume 98 of *CRPIT*, pages 61–69, Wellington, New Zealand, 2009. ACS.

- [59] Dan J. Kim, Donald L. Ferrin, and H. Raghav Rao. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, 44(2):544–564, January 2008. ISSN 01679236. doi: [10.1016/j.dss.2007.07.001](https://doi.org/10.1016/j.dss.2007.07.001).
- [60] Audun Jøsang. Trust-based decision making for electronic transactions. In *Proceedings of the 4th Nordic Workshop on Secure Computer Systems (NORDSEC'99)*, pages 1–21. Stockholm University, Sweden, 1999.
- [61] Epinions. <http://www.epinions.com/>, jun 2014.
- [62] Yelp. <http://www.yelp.com/>, jun 2014.
- [63] BizRate. <http://www.bizrate.com/>, jun 2014.
- [64] Runfang Zhou and Kai Hwang. PowerTrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing. *IEEE Transactions on Parallel and Distributed Systems*, 18(4):460–473, April 2007. ISSN 1045-9219. doi: [10.1109/TPDS.2007.1021](https://doi.org/10.1109/TPDS.2007.1021).
- [65] Ayad Barsoum and Anwar Hasan. Enabling dynamic data and indirect mutual trust for cloud computing storage systems. *IEEE Transactions on Parallel and Distributed Systems*, 24(12):2375–2385, Dec 2013. ISSN 1045-9219. doi: [10.1109/TPDS.2012.337](https://doi.org/10.1109/TPDS.2012.337).
- [66] Yu-Ting Caisy Hung, Alan R. Dennis, and Lionel Robert. Trust in virtual teams: Towards an integrative model of trust formation. In *Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 1 - Volume 1*, HICSS'04, pages 1–12, Washington, DC, USA, 2004. IEEE Computer Society. ISBN 0-7695-2056-1.
- [67] AllExperts. <http://www.allexperts.com/>, jun 2014.
- [68] Advogato. <http://www.advogato.org/>, jun 2014.
- [69] Audun Jøsang and Jennifer Golbeck. Challenges for Robust Trust and Reputation Systems. In *5th International Workshop on Security and Trust Management (STM 2009)*, number September, pages 1–12. Elsevier Science B. V., 2009.
- [70] Gayatri Swamynathan, Kevin C. Almeroth, and Ben Y. Zhao. The design of a reliable reputation system. *Electronic Commerce Research*, 10(3-4):239–270, August 2010. ISSN 1389-5753. doi: [10.1007/s10660-010-9064-y](https://doi.org/10.1007/s10660-010-9064-y).
- [71] Mao Chen and Jaswinder Pal Singh. Computing and using reputations for internet ratings. In *EC '01 Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 154–162, New York, New York, USA, 2001. ACM Press. ISBN 1-5811-3387-1. doi: [10.1145/501158.501175](https://doi.org/10.1145/501158.501175).
- [72] Jean Piaget. *Judgment and reasoning in the child*. Routledge, 1999.
- [73] Dorothy E. Denning. A new paradigm for trusted systems. In *Proceedings on the 1992-1993 workshop on New security paradigms*, NSPW '92-93, pages 36–41, New York, NY, USA, 1993. ACM. ISBN 0-8186-5430-9. doi: [10.1145/283751.283772](https://doi.org/10.1145/283751.283772).
- [74] Chrysanthos Dellarocas. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of the 2nd ACM conference on Electronic commerce (EC '00)*, pages 150–157, New York, New York, USA, 2000. ACM Press. ISBN 1-5811-3272-7. doi: [10.1145/352871.352889](https://doi.org/10.1145/352871.352889).
- [75] Bin Yu, Munindar P. Singh, and Katia Sycara. Developing trust in large-scale peer-to-peer systems. In *IEEE First Symposium on Multi-Agent Security and Survivability*, pages 1–10. IEEE, 2004. ISBN 0-7803-8799-6. doi: [10.1109/MASSUR.2004.1368412](https://doi.org/10.1109/MASSUR.2004.1368412).
- [76] Ya-Fei Yang, Qin-Yuan Feng, Yan (Lindsay) Sun, and Ya-Fei Dai. Dishonest behaviors in online rating systems: Cyber competition, attack models, and attack generator. *Journal of Computer Science and Technology*, 24(5):855–867, 2009. ISSN 1000-9000. doi: [10.1007/s11390-009-9277-5](https://doi.org/10.1007/s11390-009-9277-5).
- [77] Florian Kerschbaum, Jochen Haller, Yücel Karabulut, and Philip Robinson. PathTrust: A Trust-Based Reputation Service for Virtual Organization Formation. In Ketil Stolen, William H. Winsborough, Fabio Martinelli, and Fabio Massacci, editors, *Trust Management*, volume 3986 of *Lecture Notes in Computer Science*, pages 193–205. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-34295-3. doi: [10.1007/11755593\\_15](https://doi.org/10.1007/11755593_15).
- [78] Sini Ruohomaa, Lea Kutvonen, and Eleni Koutrouli. Reputation management survey. In *The Second International Conference on Availability, Reliability and Security (ARES 2007)*, pages 103–111. IEEE, 2007. doi: [10.1109/ARES.2007.123](https://doi.org/10.1109/ARES.2007.123).
- [79] Mudhakar Srivatsa and Ling Liu. Securing decentralized reputation management using TrustGuard. *Journal of Parallel and Distributed Computing*, 66(9):1217–1232, September 2006. ISSN 0743-7315. doi: [10.1016/j.jpdc.2006.04.003](https://doi.org/10.1016/j.jpdc.2006.04.003).
- [80] Mohammad Amin Morid and Mehdi Shajari. An enhanced e-commerce trust model for community based centralized systems. *Electronic Commerce Research*, 12(4):409–427, 2012. ISSN 1389-5753. doi: [10.1007/s10660-012-9099-3](https://doi.org/10.1007/s10660-012-9099-3).
- [81] John S. Breese, David Heckerman, and Carl Kadie. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In *UAI'98 Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, pages 43–52, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc. ISBN 1-55860-555-X.

- [82] Loren Terveen and David W. McDonald. Social matching: A framework and research agenda. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(3):401–434, 2005. doi: [10.1145/1096737.1096740](https://doi.org/10.1145/1096737.1096740).
- [83] Xiaoyuan Su and Taghi M. Khoshgoftaar. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence*, 2009(Section 3):1–19, 2009. ISSN 1687-7470. doi: [10.1155/2009/421425](https://doi.org/10.1155/2009/421425).
- [84] Mozghan Tavakolifard, Peter Herrmann, and Svein J. Knapskog. Inferring Trust Based on Similarity with TILLIT. In Elena Ferrari, Ninghui Li, Elisa Bertino, and Yücel Karabulut, editors, *Trust Management III*, volume 300 of *IFIP Advances in Information and Communication Technology*, pages 133–148. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-02055-1. doi: [10.1007/978-3-642-02056-8\\_9](https://doi.org/10.1007/978-3-642-02056-8_9).
- [85] W.T. Luke Teacy. *Agent-based trust and reputation in the context of inaccurate information sources*. PhD thesis, University of Southampton, 2006.
- [86] Rino Falcone and Cristiano Castelfranchi. Social trust: A cognitive approach. In Cristiano Castelfranchi and Yao-Hua Tan, editors, *Trust and Deception in Virtual Societies*, pages 55–90. Springer Netherlands, 2001. ISBN 978-90-481-5687-0. doi: [10.1007/978-94-017-3614-5\\_3](https://doi.org/10.1007/978-94-017-3614-5_3).
- [87] Nicola Mezzetti. A Socially Inspired Reputation Model. In *Proceedings of 1st European PKI Workshop*, volume 3093 of *Lecture Notes in Computer Science*, pages 191–204. Springer-Verlag, 2004. doi: [10.1007/978-3-540-25980-0\\_16](https://doi.org/10.1007/978-3-540-25980-0_16).
- [88] Omar Hasan, Lionel Brunie, Jean-Marc Pierson, and Elisa Bertino. Elimination of subjectivity from trust recommendation. In Elena Ferrari, Ninghui Li, Elisa Bertino, and Yücel Karabulut, editors, *Trust Management III*, volume 300 of *IFIP Advances in Information and Communication Technology*, pages 65–80. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-02055-1. doi: [10.1007/978-3-642-02056-8\\_5](https://doi.org/10.1007/978-3-642-02056-8_5).
- [89] Nathan Griffiths. Task Delegation using Experience-Based Multi-Dimensional Trust. In *International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '05)*, pages 489–496. ACM, 2005. doi: [10.1145/1082473.1082548](https://doi.org/10.1145/1082473.1082548).
- [90] Kui Meng, Xu Zhang, Xiao chun Xiao, and Geng du Zhang. A bi-rating based personalized trust management model for virtual communities. In *Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control (ICNSC '06)*, pages 260–265. IEEE, 2006. doi: [10.1109/ICNSC.2006.1673154](https://doi.org/10.1109/ICNSC.2006.1673154).
- [91] Ye Diana Wang and Henry H. Emurian. An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, 21(1):105–125, 2005. ISSN 0747-5632. doi: [10.1016/j.chb.2003.11.008](https://doi.org/10.1016/j.chb.2003.11.008).
- [92] Audun Jøsang. Trust and reputation systems. In Alessandro Aldini and Roberto Gorrieri, editors, *Foundations of Security Analysis and Design IV*, volume 4677 of *Lecture Notes in Computer Science*, pages 209–245. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-74809-0. doi: [10.1007/978-3-540-74810-6\\_8](https://doi.org/10.1007/978-3-540-74810-6_8).
- [93] Joyce Berg, John Dickhaut, and Kevin McCabe. Trust Reciprocity, and Social History. *Games and Economic Behavior*, 10(1):122–142, 1995. doi: [10.1006/game.1995.1027](https://doi.org/10.1006/game.1995.1027).
- [94] Giangiacomo Bravo, Flaminio Squazzoni, and Riccardo Boero. Trust and partner selection in social networks: An experimentally grounded model. *Social Networks*, 34(4):481–492, October 2012. ISSN 0378-8733. doi: [10.1016/j.socnet.2012.03.001](https://doi.org/10.1016/j.socnet.2012.03.001).
- [95] Noel D. Johnson and Alexandra A. Mislin. Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889, October 2011. ISSN 01674870. doi: [10.1016/j.joep.2011.05.007](https://doi.org/10.1016/j.joep.2011.05.007).
- [96] Claudia Keser. Experimental games for the design of reputation management systems. *IBM Systems Journal*, 42(3):498–506, 2003.
- [97] Isabel Briggs Myers. *Myers-Briggs Type Indicator*. Consulting Psychologists Press, 1962.
- [98] Julian B. Rotter. A new scale for the measurement of interpersonal trust. *Journal of personality*, 35(4):651–665, 1967. doi: [10.1111/j.1467-6494.1967.tb01454.x](https://doi.org/10.1111/j.1467-6494.1967.tb01454.x).
- [99] Anthony M. Evans and William Revelle. Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6):1585–1593, 2008. doi: [10.1016/j.jrp.2008.07.011](https://doi.org/10.1016/j.jrp.2008.07.011).
- [100] Ingrid Zukerman and David W. Albrecht. Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction*, 11(1-2):5–18, 2001. doi: [10.1023/A:1011175525451](https://doi.org/10.1023/A:1011175525451).
- [101] Vincenzo Moscato, Antonio Picariello, and Antonio M. Rinaldi. Towards a user based recommendation strategy for digital ecosystems. *Knowledge-Based Systems*, 37:165–175, January 2013. ISSN 0950-7051. doi: [10.1016/j.knsys.2012.07.021](https://doi.org/10.1016/j.knsys.2012.07.021).
- [102] Atsunori Minamikawa, Hamido Fujita, Jun Hakurab, and Masaki Kurematsu. Personality estimation approach for social media. In *New Trends in Software Methodologies, Tools and Techniques*, volume 246 of *Frontiers in Artificial Intelligence and Applications*, pages 327–335. IOS Press, 2012. doi: [10.3233/978-1-61499-125-0-327](https://doi.org/10.3233/978-1-61499-125-0-327).

- [103] Apostolos Georgakis and Haibo Li. User behavior modeling and content based speculative web page prefetching. *Data & Knowledge Engineering*, 59(3):770–788, 2006. ISSN 0169-023X. doi: [10.1016/j.datak.2005.11.005](https://doi.org/10.1016/j.datak.2005.11.005).
- [104] Andrew Koster, Marco Schorlemmer, and Jordi Sabater-Mir. Engineering trust alignment: Theory, method and experimentation. *International Journal of Human-Computer Studies*, 70(6):450–473, June 2012. ISSN 10715819. doi: [10.1016/j.ijhcs.2012.02.007](https://doi.org/10.1016/j.ijhcs.2012.02.007).
- [105] Peter Sprent and Nigel C. Smeeton. *Applied Non-parametric Statistical Methods*. Texts in Statistical Science Series. Taylor & Francis Group, 2007. ISBN 978-1584887010.
- [106] Elise Whitley and Jonathan Ball. Statistics review 6: Nonparametric methods. *Critical Care*, (6):509–513, 2002. doi: [10.1186/cc1820](https://doi.org/10.1186/cc1820).
- [107] Viv Bewick, Liz Cheek, and Jonathan Ball. Statistics review 10: Further nonparametric methods. *Critical Care*, 8(3):196–199, June 2004. ISSN 1466-609X. doi: [10.1186/cc2857](https://doi.org/10.1186/cc2857).
- [108] David J. Sheskin. *Handbook of Parametric and Nonparametric Statistical Procedures, Fifth Edition*. Taylor & Francis, 2011. ISBN 978-1439858011.
- [109] Erich Leo Lehmann and Howard JM D'Abbrera. *Non-parametrics: statistical methods based on ranks*. Springer New York, 2006.
- [110] Zeinab Noorian, Stephen Marsh, and Michael Fleming. Multi-layer cognitive filtering by behavioral modeling. In Tumer, Yolum, Sonenberg, and Stone, editors, *Proceedings of The 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pages 871–878, Taipei, Taiwan, 2011. International Foundation for Autonomous Agents and Multiagent Systems.
- [111] Trung Dong Huynh. A personalized framework for trust assessment. In *Proceedings of the 2009 ACM symposium on Applied Computing - SAC '09*, Honolulu, Hawaii, USA, 2009. ACM Press. ISBN 9781605581668. doi: [10.1145/1529282.1529574](https://doi.org/10.1145/1529282.1529574).
- [112] Licia Capra. Engineering human trust in mobile system collaborations. *SIGSOFT Softw. Eng. Notes*, 29(6):107–116, October 2004. ISSN 0163-5948. doi: [10.1145/1041685.1029912](https://doi.org/10.1145/1041685.1029912).
- [113] Andrew Koster, Jordi Sabater-Mir, and Marco Schorlemmer. Personalizing communication about trust. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, pages 517–524. International Foundation for Autonomous Agents and Multiagent Systems, 2012.
- [114] Andrew Koster, Jordi Sabater-Mir, and Marco Schorlemmer. Engineering trust alignment: a first approach. In *Proceedings of the Thirteenth International Workshop on Trust in Agent Societies at AAMAS 2010*, volume 108. International Foundation for Autonomous Agents and Multiagent Systems, 2010.
- [115] Pierpaolo Dondio, Luca Longo, and Stephen Barrett. A translation mechanism for recommendations. In Yücel Karabulut, John Mitchell, Peter Herrmann, and Christian Damsgaard Jensen, editors, *Trust Management II*, volume 263 of *IFIP – The International Federation for Information Processing*, pages 87–102. Springer US, 2008. ISBN 978-0-387-09427-4. doi: [10.1007/978-0-387-09428-1\\_6](https://doi.org/10.1007/978-0-387-09428-1_6).