VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INTELLIGENT SYSTEMS

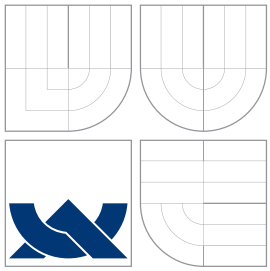
DŮVĚRA A REPUTACE V DISTRIBUOVANÝCH SYSTÉMECH

DISERTAČNÍ PRÁCE
PHD THESIS

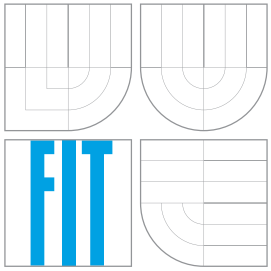
AUTOR PRÁCE
AUTHOR

Ing. JAN SAMEK

BRNO 2011



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INTELLIGENT SYSTEMS

DŮVĚRA A REPUTACE V DISTRIBUOVANÝCH SYSTÉMECH

TRUST AND REPUTATION IN DISTRIBUTED SYSTEMS

DISERTAČNÍ PRÁCE
PHD THESIS

AUTOR PRÁCE
AUTHOR

Ing. JAN SAMEK

VEDOUCÍ PRÁCE
SUPERVISOR

Doc. Dr. Ing. PETR HANÁČEK

BRNO 2011

Abstrakt

Tato disertační práce se zabývá problematikou modelování důvěry v distribuovaných systémech, konkrétně pak více-kontextové důvěry v distribuovaných multi-agentních systémech. V současné době existuje velké množství modelů důvěry či reputace, nicméně více-kontextová vlastnost důvěry či reputace v nich není často zohledněna. Z tohoto důvodu se práce zaměřuje především na analýzu více-kontextových modelů založených na důvěře a na jejím základě stanovuje předpoklady pro nový, vlastnost více-kontextovost plně podporující, model důvěry. Stěžejní částí práce je formální návrh nového více-kontextového modelu důvěry, který je schopen vytvářet, aktualizovat a uchovávat důvěru pro různé aspekty (kontexty) jedné entity multi-agentního systému. Důvěru lze v navrženém modelu budovat jak na základě přímých zkušeností, tak i na základě doporučení a reputace. Dalším aspektem navrženého modelu je schopnost odvozovat důvěru v různé kontexty na základě znalosti důvěry v kontexty jiné, což je zajištěno vytvořením hierarchické struktury jednotlivých kontextů jedné entity. Přínosem nového modelu je především zvýšení efektivity rozhodování agentů v rámci multi-agentního systému ve smyslu schopnosti výběru optimálního partnera pro provedení transakce. Návrh modelu byl ověřen implementací prototypu multi-agentního systému, ve kterém se agenti rozhodují a jednají na základě důvěry.

Abstract

This Ph.D. thesis deals with trust modelling for distributed systems especially to multi-context trust modelling for multi-agent distributed systems. There exists many trust and reputation models but most of them do not dealt with the multi-context property of trust or reputation. Therefore, the main focus of this thesis is on analysis of multi-context trust based models and provides main assumptions for new fully multi-contextual trust model on the bases of them. The main part of this thesis is in providing new formal multi-context trust model which are able to build, update and maintain trust value for different aspects (contexts) of the single entity in the multi-agent system. In our proposal, trust value can be built on the bases of direct interactions or on the bases on recommendations and reputation. Moreover we assume that some context of one agent is not fully independent and on the bases of trust about one of them we are able to infer trust to another's. Main contribution of this new model is increasing the efficiency in agent decision making in terms of optimal partner selection for interactions. Proposed model was verified by implementing prototype of multi-agent system when trust was used for agents' decision making and acting.

Klíčová slova

důvěra, více-kontextová důvěra, reputace, distribuované multi-agentní systémy, HMTC

Keywords

trust, multi-context trust, reputation, distributed multi-agent systems, HMTC

Citace

Jan Samek: Důvěra a reputace v distribuovaných systémech, disertační práce, Brno, FIT VUT v Brně, 2011

Důvěra a reputace v distribuovaných systémech

Prohlášení

Prohlašuji že jsem tuto práci vypracoval samostatně pod vedením Doc. Petra Hanáčka a Dr. Františka Zbořila. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Jan Samek
27. září 2011

Poděkování

Tato disertační práce vznikla na Ústavu inteligentních systémů Fakulty informačních technologií VUT v Brně. Výzkum v této práci vznikl částečně za podpory výzkumného záměru MŠMT MSM0021630528 a grantů GACR 102/09/H042 a BUT FIT-10-S-1.

Rád bych rovněž poděkoval všem, kteří jakýmkoliv způsobem přispěli k vytvoření této disertační práce. Jmenovitě pak především mým školitelům Doc. Petru Hanáčkovi a Dr. Františku Zbořilovi za odborné vedení a cenné rady během celého mého studia.

Velké díky patří také kolegům z doktorského studia, především pak i nejen kolegovi ale i dobrému příteli Ondrovi Malačkovi za *akademické pátky* a bouřlivé diskuze, které nezanedbatelnou měrou přispěly k této práci.

Poslední a největší poděkování patří mé rodině, především pak mým milujícím rodičům za jejich *důvěru*, morální a materiální podporu, bez které by tato práce nikdy nevznikla. Poslední díky patří Verunce, za všechno a za to, že to tak dlouho vydržela.

Obsah

Seznam obrázků

Seznam tabulek

Kapitola 1

Úvod

Žijeme ve světě, kde pojem důvěra, respektive jeho význam hraje zásadní roli. Tento pojem má své základy v sociologii, ale dnes je hojně využíván v celé řadě jiných vědních oborů jakými jsou např. psychologie, filosofie, ekonomie či informatika. Ač se zdá, že pojem důvěra je ve světě informačních technologií těžko použitelný (na rozdíl od ostatních jmenovaných), tak výzkum v této oblasti úspěšně probíhá již řadu let.

Důvěra tedy hraje důležitou roli v našem každodenním životě. Důvěřujeme našim nejbližším, naši rodině, našim rodičům že nás budou podporovat v našem jednání, v životě. Naopak, naše okolí důvěřuje nám, našim schopnostem v naši budoucnost. Stále více a více je také třeba budovat důvěru nejen v okruhu našich nejbližších, ale i v celé společnosti kde žijeme. Uvědomme si obzvláště v dnešní době, že bez důvěry bankovní instituce, jen velmi těžko získáme např. půjčku či hypotéku. Internet a globalizace, to jsou hlavní faktory, které snižují vzdálenosti mezi lidmi, skupinami a institucemi po celém světě. Vzhledem k tomu je třeba budovat důvěru napříč lidmi a institucemi nejen v rámci místa či regionu kde žijeme, ale i celosvětově. Každá naše sociální interakce, ať již přímá či nepřímá, způsobuje v našem okolí odezvu, která může mít za následek posílení důvěry či prohloubení nedůvěry.

Obrovský rozmach sociálních sítí typu Facebook v posledních letech jasně ukazuje směr vývoje v oblasti sociální interakce mezi lidmi. Důvěra mezi „virtuálními identitami“ vybudovaná na základě virtuálních vztahů a interakcí v prostředí internetu je velmi křehká a pochybná. Tisíce podvodných profilů a skupin, které bezhlavě – důvěřivě – následují tisíce „lidí“ je jasným příznakem toho, že není možné jednoduše přenést klasické koncepty důvěry z reálného světa do světa virtuálního.

V současné době probíhá mnoho výzkumů po celém světě v oblasti využití a aplikace sociologických aspektů důvěry ve virtuálním prostředí. Je možné, že v brzké době se role obrátí. Výsledky, které přinášíme na základě zkoumání lidského světa důvěry, aplikované do virtuálního prostředí pomohou člověku tak, aby byl v budoucnu schopen důvěřovat ve své virtuální okolí. Tato práce se snaží přispět k poznání fenoménu důvěry a jeho využití v oblasti distribuovaných systémů a umělé inteligence.

1.1 Motivace

Podíváme-li se na možnosti praktického využití důvěry a reputace v informačních technologiích a především pak v oblasti distribuovaných systémů zjistíme, že jsou nejčastěji využívány v oblastech jakými jsou *reputační systémy* a *kolaborativní filtrování* [JIB07, SK09]. V těchto oblastech je důvěra/reputace použita jako jedno z kritérií procesu rozhodovaní v momentě, kdy se pracuje s neúplnou nebo neurčitou informací [SS05, AR05]. V každém procesu rozhodování je vždy zahrnuta jistá míra rizika a nejistoty. Základní motivací využití důvěry/reputace je tedy to, že se snažíme minimalizovat rizika a nejistoty spojené s procesem rozhodovaní. Podíváme-li se na problematiku eliminace rizik, shledáváme, že existuje množství bezpečnostních funkcí, které se snaží riziko minimalizovat. Těmito bezpečnostními funkcemi máme na mysli především: řízení přístupu, identifikaci a autentizaci, kryptografické algoritmy a protokoly apod. Tyto bezpečnostní funkce jsou však většinou v distribuovaných systémech implementovány tak, že nepostihují problematiku práce s neúplnou nebo neurčitou informací. Šifrování, autentizace či identifikace je založena na přesnosti a úplnosti informace kterou šifrujeme, či ověřujeme a sebemenší změna této informace na vstupu generuje na výstupu diametrálně odlišné výsledky. Naproti tomu s využitím důvěry/reputace jsme schopni i na základě částečné informace na vstupu generovat relevantní výsledek na výstupu. Otázkou zůstává, do jaké míry je tento výsledek relevantní vzhledem k požadavkům na systém.

Je tedy zřejmé, že klasické bezpečnostní funkce generují na výstupu naprosto přesné údaje vzhledem ke vstupu. Naproti tomu, míra přesnosti výstupu bezpečnostní funkce založené na důvěře/reputaci je úměrná míře přesnosti informace na vstupu. Toto je také zásadní rozdíl mezi klasickým technickým „tvrdým“ přístupem a přístupem „měkkým“ založeným na sociálních aspektech: důvěře/reputaci.

Samozřejmě nelze říci, že pro všechny aplikace je *měkký* přístup vhodný, často je třeba naprosto striktně a neomylně provést rozhodnutí, kde není prostor pro neurčitost. Zde se zcela opodstatněně uplatní výhradně *tvrdý* přístup. V poslední době však vzniká velké množství aplikací, kde tvrdý přístup buď nelze použít vůbec anebo jeho použití negeneruje požadované výsledky [GBS08]. Jako příklad takové aplikace lze uvést využití distribuované agentní a multi-agentní systémy. Pomocí šifrování lze sice zajistit bezpečnou komunikaci mezi agenty, ale rozhodnutí o tom zda je protistrana důvěryhodná (ve smyslu schopnosti provedení nějaké interakce) lze pomocí tvrdých bezpečnostních mechanismů provést jen velmi obtížně anebo vůbec. A to i v případě, že jsme schopni zajistit identifikaci a autentizaci agentů, neboť v takovém případě stále není zaručeno, že komunikující jednoznačně identifikovaná protistrana není kompromitována útočníkem.

Naproti tomu, v případě využití měkkých dovedností v kombinaci s těmi tvrdými jsme schopni pro tento případ stanovit s určitou pravděpodobností to, že komunikující protistrana nebude nejvhodnější partner. To na základě sledovaní chování protistrany nebo na základě doporučení či reputace od ostatních agentů a nebo sociálních vazeb mezi nimi.

Zdá se tedy, že ideálním přístupem který je schopen odstranit riziko související s rozhodováním v distribuovaných systémech, je využití kombinace tvrdého a měkké přístupu. Nutno dodat, že vždy je třeba vyhodnotit vhodnost použití této kombinace pro konkrétní aplikace.

1.2 Cíle disertační práce

Předložena disertační práce se zabývá možnostmi modelování a využití důvěry a reputace v distribuovaných systémech. Konkrétně pak návrhem nového modelu důvěry, který by měl sloužit pro rozhodování inteligentních autonomních agentů v multi-agentním distribuovaném heterogenním prostředí.

Při studiu různých modelů důvěry a reputace v kontextu agentních systémech [KR03, Mui03, RHJ04, SS05, MP10] se ukazuje, že existuje celá řada modelů důvěry a reputace pro distribuované a agentní systémy. Tyto známé modely lze rozdělit do dvou kategorií: modely navržené pro konkrétní reálný problém [KSGM03, GS03, GBS08, WE05, AMDS06, MP08, MP09b] a modely navržené obecně bez vazby na konkrétní aplikaci [Mar94, ARH00, SS01, MMH02a, KR03, WS06]. V obou případech se však často ignoruje jedna ze základních přirozených vlastností důvěry a reputace, kterou je kontextová závislost nebo také více-kontextovost (z anglického *multi-context*). To znamená, že důvěra nebo reputace jakékoliv entity (člověka, organizace, služby, produktu, ...) může být posuzována z hlediska různých aspektů – kontextů. Dá se říci, že v současné době je jen velmi málo modelů, které zohledňují tuto přirozenou vlastnost důvěry a reputace, což dokládá i následující citace z významné vědecké publikace *Review on Computational Trust and Reputation Models* [SS05] autorů *Sabater & Sierra* (strana 39):

> *"... there are very few computation trust and reputation models that care about the multi-context nature of trust and reputation and event fewer that propose some kind of solution ..."*

Při zohlednění více-kontextového přístupu se tak otevírají nové možnosti v oblasti využití důvěry a reputace, jakožto kritéria při rozhodovaní a interakci autonomních entit v distribuovaných multi-agentních systémech. Hlavní cíle předložené práce tak můžeme shrnout do následujících bodů:

- Navrhnout model důvěry pro multi-agentní distribuovaný systém využívající vlastnost více-kontextovost s ohledem na možnost přenášení důvěry z jednoho kontextu na kontext druhý.
- Navrhnout způsob transformace přímých znalostí agenta (z přímých zkušeností) na odpovídající hodnotu důvěry, s níž bude schopen námi navržený více-kontextový model pracovat.
- Implementovat prototyp multi-agentního systému, který bude využívat navrženého modelu a způsobu transformace přímých zkušeností na hodnotu důvěry, na jejímž základě se agenti budou moci rozhodovat a jednat.

1.3 Struktura práce

Struktura předložené disertační práce je rozčleněna následovně: v první kapitole je představen úvod do problematiky, nastíněna motivace a stanoveny cíle disertační práce.

Druhá kapitola popisuje základní pojmy *důvěry* a *reputace* tak, jak jsou popisovány v různých současných publikacích různých autorů z různých vědních oborů. Dále jsou zde popsány různé vlastnosti důvěry, způsoby její reprezentace tak, jak jsou používány různými autory v relevantních vědeckých publikacích.

Třetí kapitola popisuje současný stav problematiky modelů důvěry, principy vytváření důvěry a reputace, doporučení pro tvorbu modelů založených na důvěře a klasifikaci těchto modelů. Poslední část této kapitoly popisuje reprezentativní formální modely důvěry a reputace, které využívají více-kontextového přístupu.

Čtvrtá kapitola se zabývá vlastním jádrem disertační práce, tedy návrhem samotného více-kontextového modelu důvěry pro agentní systémy, který je založen na definici vztahů mezi jednotlivými kontexty. Navržený model je zde formálně popsán včetně způsobu aktualizace důvěry pro jednotlivé kontexty a způsobu šíření důvěry mezi kontexty jedné entity.

Kapitola pátá se zabývá popisem modelu chování agenta v multi-agentním systému a způsobem, jak je jeho chování vyhodnocováno ve smyslu vytváření důvěry na základě přímých zkušeností (interakcí) mezi agenty. Tato kapitola tak úzce navazuje na jádro disertační práce a vytváří tak ucelený soubor metod pro robustní způsob odhadu důvěry na základě výsledku interakce mezi agenty v distribuovaném multi-agentním systému.

Následující kapitola šestá popisuje implementaci prototypu distribuovaného multi-agentního systému využívajícího navržený model důvěry. Na základě experimentů s tímto prototypem jsou zde diskutovány dosažené výsledky, výhody a nevýhody navrženého modelu.

Závěrečná kapitola shrnuje dosažené výsledky disertační práce a diskutuje možná rozšíření a směry případných pokračování založených na principu navrženého více-kontextového distribuovaného modelu důvěry pro multi-agentní systémy.

Kapitola 2

Pojem důvěra a reputace

Reputace, důvěra nebo zkušenost jsou pojmy, které známe z běžného života. Lze dokonce říci, že většina našich rozhodnutí je provedena právě na základě důvěry. Použití těchto pojmů ve světě informačních technologií není zcela obvyklé a občas i těžko představitelné. Oblast reputace a reputačních systémů, obecně tedy systémů založených na důvěře, začíná v poslední době nabývat důležitosti a významu v souvislosti s problematikou bezpečnosti počítačových systémů.

Problematika důvěry a reputace je však multi-disciplinární obor, který vychází ze sociologie a v současné době probíhá výzkum tohoto jevu v různých oborech lidského bádání, jako jsou: *psychologie*, *filozofie*, *ekonomie* (management a marketing) či *informační technologie* [WE05]. Právě vzhledem k této multi-disciplinární povaze a vzhledem ke komplexnosti problematiky důvěry a reputace se zaměříme výhradně na její využití v oboru informačních technologií a to konkrétně do oblasti distribuovaných agentních a multi-agentních systémů. Tato oblast je také často popisována jako *distribuovaná umělá inteligence* (zkratka DAI, z anglického *Distributed Artificial Intelligence*, jež používá ve své disertační práci věnované důvěře *Stephen P. Marsh* [Mar94]). Tato oblast zahrnuje pojetí konceptů distribuovanosti, inteligence, sociálnosti, nezávislosti (ve smyslu autonomie), odolnosti proti výpadkům (ve smyslu *fault-tolerant*) a lokálního rozhodování – jedná se tedy o ideální oblast a nástroje pro studium důvěry [Mar94].

2.1 Důvěra

V této kapitole si popíšeme význam pojmu *důvěra* tak, jak je popisován v různých vědních oborech a tak, jak je chápán v informatice a konkrétně pak v pracích zabývajících se modelováním důvěry či návrhem modelů založených na důvěře. Rovněž zde budou popsány vlastnosti důvěry, způsoby její reprezentace a principy jejího použití.

V následujících podkapitolách (*2.1 Důvěra* i *2.2 Reputace*) budeme používat dva pojmy, jejichž význam je nutné nejdříve vysvětlit. V první řadě to bude pojem *objekt*. Pod tímto pojmem budeme chápat nějaký prvek nebo entitu systému (obecně nějaké sociální skupiny) pro než je důvěra nebo reputace stanovena. Objekt lze označit jako *cíl důvěry*, vzhledem ke kterému je hodnota důvěry nebo reputace vztažena, spravována nebo vyhodnocována. Objekt nemusí být „inteligentní“ entita, může to být například software nebo kryptografický klíč. V anglické literatuře je tento pojem označen slovem *trustee*, volně přeloženo jako „svěřenec“.

Druhým pojmem, který s předchozím bezprostředně souvisí, je pojem *subjekt*. Tento

pojem bude označovat entitu, z jejíhož pohledu je reputace nebo důvěra vyhodnocována vzhledem k objektu. V tomto smyslu je třeba, aby subjekt byl nějakou formou „inteligentní“ entita tak, aby mohl s důvěrou či reputací vůči objektu smysluplně naložit. V některých modelech důvěry je explicitně definováno, zda je důvěra stanovena mezi subjektem a subjektem a nebo mezi subjektem a objektem. Konkrétně např. v práci [Net09] je důvěra mezi dvěma subjekty pojmenována jako *personální důvěra* a důvěra vzhledem k nějakému objektu pojmenována jako *fenomenální důvěra* (objekt je zde označován jako *produkt*, nebo *fenomén*). V anglické terminologii se používá pojem *trustor* (nebo také *truster* [Cof07]), což by se dalo volně přeložit jako „svěřitel“. Vztah mezi subjektem a objektem lze popsat například větou „důvěra subjektu v objekt“ a nebo „na základě reputace objektu je důvěra subjektu vzhledem k objektu vysoká …“.

2.1.1 Definice

Důvěra existuje od počátku lidské existence a sociální interakce. Protože koncept důvěry je velmi komplexní a ve své podstatě abstraktní, je velmi těžké ji přesně definovat a identifikovat části, které ji utváří. Původní význam slova důvěra však pochází ze sociologie, kde je sociologického slovníku [WP97] definována takto:

> *“Důvěra je typ postoje a zároveň mezilidského vztahu, který vyvolává pocit jistoty plynoucí z přesvědčení, že partner komunikace (osoba, instituce) splní určitá očekávání.”*

V dalších odborných sociologických pracích pak nacházíme velmi často používanou a citovanou definici, kterou ve svém článku *Can We Trust Trust?* [Gam00] používá profesor sociologie *Diego Gambetta*[1] a která zní následovně:

> *„… důvěra (nebo symetricky nedůvěra) je určitý stupeň subjektivní pravděpodobnosti se kterou agent odhadne, že jiný agent nebo skupina agentů provede určitou akci a to ještě předtím, než je schopen zpozorovat tuto akci (nezávisle na tom, zda je schopen takovou akci zpozorovat) a v závislosti na tom provede svoji vlastní akci …“*

Gambettovu definici pak přebírají i autoři, zabývající se důvěrou v oblasti informatiky. Zjednodušenou a upravenou formulaci tak používá např. autor *Lik Mui*, který ji definuje ve svém článku *A Computational Model of Trust and Reputation* [MMH02a] a následně ji používá ve své disertační práci [Mui03] takto:

> *“Důvěra je subjektivní víra v agenta, že se bude chovat za určitých okolností určitým, předem očekávaným způsobem, na základě jeho předchozího chování.”*

Mui tedy definuje důvěru jako jakousi míru pravděpodobnosti, která je určena nebo vypočtena na základě předchozího chování entity. Přitom pojem *pravděpodobnost*, který použil i *Gambetta*, zde má své jasné opodstatnění, neboť princip výpočtu důvěry/reputace v jeho modelu se zakládá na použití aplikace podmíněných pravděpodobností a Bayesova vzorce [Mui03]. Tuto upravenou definici používají následně i další autoři jako např. *Kinatede et al.* v článku *Architecture and Algorithms for a Distributed Reputation System* [KR03].

[1] *Diego Gambetta* je profesor sociologie na Oxfordské univerzitě, který svojí prací významně přispěl ke studiu fenoménu důvěry [Gam10].

Další významný článek *Can We Manage Trust?* [sKD05] od autorů *Jøsang et al.* z oblasti informatiky, která se zabývá otázkou „zvládnutí důvěry“ (ve smyslu vytváření důvěry, stanovování důvěryhodnosti a rozhodování na základě důvěryhodnosti) popsal důvěru jako *„orientovaný vztah mezi dvěma účastníky nazvanými subjekt a objekt“*. Pojem orientovaný je použit ve smyslu jasného rozlišení *zdroje* (subjekt) a *cíle* (objekt) vztahu. Autoři dále definují dva typy důvěry v závislosti na tom, zda je *kontextově nezávislá* (kterou také nazývají jako *spolehlivost* – reliability trust) či *kontextově závislá* (nebo také *rozhodovací důvěra* – decision trust). Kontextově nezávislou důvěru, tedy spolehlivost pak popisují jako „spolehlivost něčeho nebo někoho nezávisle na kontextu“ a následně pro její definici používají upravenou *Gambettovu* definici důvěry uvedenou výše.

Na kontextovou závislost důvěry (*kontextovost*), jakožto význačnou vlastnost důvěry upozorňují nejen *Jøsang et al.* ale i řada dalších autorů [FC01, ARH00, Gra06]. Přičemž kontextovost je v práci [sKD05] popsána ve smyslu situační závislosti důvěry a pojem *kontext* může nabývat mnoha významů. Definice *rozhodovací důvěry* (kontextově závislá důvěra) dle [sKD05] zní takto:

> *„Důvěra je vyjádření míry toho, jak hodně je subjekt ochotný být závislý na něčem nebo někom* (objekt) *v rámci dané situace* (kontext) *s pocitem relativního bezpečí i v případě možných negativních důsledků.“*

Takovými definicemi důvěry bychom mohli pokračovat ještě mnoho stran, neboť každý z autorů, zabývající se fenoménem důvěry ji vysvětluje a ohýbá svým směrem tak, aby zohlednil další a další aspekty, které lze v komplexní problematice chápání a významu důvěry nalézt. Není cílem této práce zde všechny tyto definice popsat a snažit se je pochopit. Ba co více, důvěra je extrémně komplexní oblast a není v silách jedince tuto problematiku celou řádně prostudovat a pochopit, na což upozorňuje již ve své disertační práci s názvem *Formalising Trust as a Computational Concept* [Mar94] z roku 1994 *Stephen P. Marsh*.

Výše popsané definice jsou tedy jakýmsi vybraným vzorkem mnoha různých prací, které jsou v oblasti *distribuované umělé inteligence* nejvíce citované, respektované a používané. V následující podkapitole se podrobněji podíváme na některé vlastnosti důvěry tak, jak jsou popsány ve významných pracích vztahujících se rovněž k oblasti DAI.

2.1.2 Vlastnosti

Vzhledem k tomu, že existuje celá řada vědeckých prací od různých autorů zabývajících se studiem a popisem důvěry, lze nalézt v těchto pracích i různé definice různých vlastností důvěry. Kolektiv autorů v práci [AMCR04] uvádějí tyto vlastnosti: *reflexivní*, *asymetrická*, *netranzitivní* nebo pouze *podmíněně transitivní* a *dynamická*. Podobný výčet vlastností uvádí i *Jennifer Golbeck* ve své disertační práci [Gol05], která je věnovaná aplikaci důvěry v sociálních sítích, následující vlastnosti: *ne plně tranzitivní*, *kombinovatelnost* (composability), *individuálnost* (personalization) a *asymetrická*. Autorský kolektiv *Wang & Vassileva* [WV03] uvádějí společné vlastnosti pro důvěru a reputaci takto: *kontextová závislost* (vícekontextovost), *mnohostrannost* (ve smyslu individuálnosti) a *dynamičnost*.

Autor *Abdul-Rahman* ve své disertační práci s názvem *A Framework for Decentralised Trust Reasoning* [AR05], která se věnuje oblasti návrhu systému pro decentralizované rozhodování na základě důvěry dokonce popisuje jedenáct vlastností důvěry, přičemž některé z nich jsou svázány s konkrétním multi-agentní přístupem. Tyto popisované vlastnosti jsou následující:

1. Důvěra je subjektivní.
2. Důvěra je situačně specifická.
3. Důvěra je závislá na charakteru agenta.
4. Důvěra není absolutní, existuje jako stupeň důvěry.
5. Důvěra zahrnuje očekávání budoucího výsledku.
6. V závislosti na situaci může důvěra způsobit pozitivní nebo negativní výsledek, tudíž je třeba zohlednit riziko, nejistotu a nevědomost.
7. Důvěra dává objektu schopnost řízení a příležitost k oklamání subjektu.
8. Neschopnost ověřit výsledek nějaké akce dokud není dokončena vyžaduje počáteční důvěru v objekt, že tuto akci zvládne.
9. Objekty jsou aktivní agenti, kteří mají schopnost něco vykonat s určitou mírou nezávislosti nad vlivem subjektů.
10. Důvěra není předpověď.
11. Důvěra není tranzitivní.

Všech jedenáct výše zmíněných vlastností jen ukazuje na fakt, že problematika chápání důvěry a jejich vlastností je komplexní a není pevně definována, tudíž velmi záleží na úhlu pohledu a konkrétní aplikaci jak bude popsána. V následujícím textu se podrobněji zaměříme na nejvýznamnější a nejčastěji zmiňované vlastnosti důvěry.

Reflexivita. Tato vlastnost důvěry udává, že entita vždy plně důvěřuje sama sobě.

Asymetrie. Asymetrie důvěry udává, že pokud nějaká entita A důvěřuje entitě B, pak není pravidlem, že entita B taktéž důvěřuje stejnou měrou entitě A. Asymetrie rovněž vyjadřuje, že vzájemná důvěra mezi entitami nemusí existovat (může být pouze jednosměrná).

Tranzitivita. Důvěra je obecně netranzitivní (tj. ne vždy tranzitivní). Tzn., že pokud A důvěřuje B a B důvěřuje C neplatí nutně, že A důvěřuje C. Podmínečná tranzitivita označuje vlastnost důvěry, kdy je splněn předpoklad *doporučení*. Tedy A důvěřuje B, B důvěřuje C a B doporučí A důvěřovat C, tzn. existence důvěry mezi A a C je *podmíněně tranzitivní* za předpokladu, že existuje doporučení mezi B a A vzhledem k C.

Individuálnost. Individualita (nebo také subjektivnost důvěry) vyjadřuje vlastnost důvěry, kdy důvěra různých entit A a B vzhledem k jednomu cíli C (pokud taková důvěra existuje) obecně nabývá různých hodnot.

Dynamičnost. Důvěra se v čase mění, tedy zapojíme-li do našeho vnímání důvěry časový prostor, pak lze říci, že důvěra mezi entitami A a B v čase nějakém t_0 je obecně odlišná od důvěry v čase t_1. V souvislosti s touto vlastností také popisuje [ARH00] vlastnost důvěry *ne monotónnost* – tedy důvěra se v průběhu jejího vývoje může měnit a to jak směrem vzhůru tak i směrem dolů v závislosti na vyhodnocení chování entity.

Další významnou vlastností je *kontextovost*. Tato vlastnost je jedním ze základních pilířů této disertační práce a proto jí věnujeme celou podkapitolu.

Kontextovost

Kontextovost důvěry znamená, že důvěra entity A vzhledem k entitě B je vždy závislá na nějakém kontextu c. Vzhledem k tomu, že tato práce si klade za cíl vytvořit model, který respektuje a využívá kontextovosti důvěry a reputace, podíváme se na tuto vlastnost podrobněji.

Kontextovost (nebo také kontextová závislost) důvěry je jednou z jejich základních přirozených vlastností a v reálném prostředí mezilidských vztahů je plně využívána. Jedná se o schopnost nebo možnost nahlížet na jednotlivce z různých úhlů pohledu. Tyto vlastnosti vychází z požadavků na zvýšení granularity důvěry v reálném světě. Na základě této vlastnosti jsme schopni důvěřovat jednotlivci (obecně jedné entitě systému, neboť pojem jednotlivec může evokovat představu, že lze věřit pouze lidem; nicméně obecná důvěra je zaměřena i vůči skupinám či organizacím – například důvěra v „peněžní ústav“ – konkrétní peněžní ústav tvoří mnoho jednotlivců a z hlediska sledovaného systému je to jedna entita) v různých aspektech jeho chování či jednání. Těmto aspektům pak říkáme kontexty. Pro ilustraci kontextovosti důvěry použijeme příklad z reálného prostředí.

> Náš lékař (řekněme zubní lékař) je pro nás důvěryhodný v kontextu „zubař“ a naše míra důvěry vzhledem k tomuto kontextu se v ideálním případě blíží k maximální důvěře. Nemáme tedy problém, v případě že nás bolí zub, vyhledat jeho odbornou pomoc. V momentě, kdy nás však začnou bolet záda, neexistuje důvod proč jít k tomuto lékaři, protože nám s největší pravděpodobností nemůže od bolesti zad pomoci. Proto vyhledáme jiného specialistu „ortopeda“, který je v kontextu „ortoped“ pro nás důvěryhodný nežli lékař „zubař“.

Shrneme-li tyto poznatky, tak máme dvě entity A a B, obě entity lze označit kontextem „lékař“, první z nich (entita A) je důvěryhodná jako „zubní lékař“, druhá (entita B) jako „lékař ortoped“. Nelze jednoduše říci, že lékař A je důvěryhodný nebo nedůvěryhodný, můžeme však říci, že lékař A je důvěryhodný v kontextu „zubní lékař“, důvěra v ostatních kontextech nemusí být zcela zřejmá.

Na výše popsaném příkladě si lze všimnout ještě jednoho zajímavého aspektu, který sebou vlastnost kontextovost přináší: schopnost *přenesení kontextu* na základě podobnosti. Pakliže víme, že obě entity lze označit společným kontextem „lékař“, lze bez přímé znalosti entity B částečně odvodit (ze znalosti entity A) schopnost entity B v kontextu „lékař“. Tedy, stane-li se např. dopravní nehoda, naše důvěra v „lékaře“, jako člověka, který je schopen poskytnout první pomoc zraněné osobě je podstatně větší nežli v jinou osobu (označme ji např. jako entitu C) se zcela odlišnou profesí a to i za předpokladu, že osobu C velmi dobře osobně známe a naše důvěra v něj v *různých kontextech* je vysoká. Schopností přenesení kontextu se budeme při návrhu modelu také zabývat.

Z výše popsaného je zřejmé, že vlastnost kontextovost rozšiřuje důvěru mezi dvěma subjekty o další prostor, neboť zavádí pojem *kontext* – dostáváme tak namísto binární relace ($subjekt \times objekt$) relaci ternární ($subjekt \times objekt \times kontext$) a nebo dokonce kvaternární s přihlédnutím k časovému rozměru a dynamičnosti důvěry ($subjekt \times objekt \times kontext \times čas$).

2.1.3 Reprezentace

Způsobů jakými lze vyjádřit míru nebo velikost důvěry či reputace je popsáno velmi mnoho. Podobně jak je tomu i u definice důvěry, různí autoři vytváří svoje vlastní reprezentace, další naopak přejímají a modifikují již používané. V této podkapitole popíšeme základní

prostředky pro popis důvěry, v dostupné literatuře pak lze nalézt pouze různé varianty těchto reprezentací.

Nejjednodušší způsob reprezentace důvěry, kterým lze agenta ohodnotit, je použití binární hodnoty, tedy např.: 0 – *nedůvěryhodný* (nedůvěra), 1 – *důvěryhodný* (důvěra). Tento způsob ohodnocování je využit například v pracích [MMH02b] a [SS02c]. Takováto reprezentace je použitelná pro některé konkrétní modely a případové studie, nicméně nemůže být obecněji použita pro složitější modely, neboť nereflektuje reálné prostředí a neumožňuje vyjádřit *částečnou důvěru* či *částečnou nedůvěru* [Sam07] v případě, kdy požadujeme vyšší granularitu reprezentace důvěry.

Dalším způsobem reprezentace je množina diskrétních hodnot (výše popsaná binární reprezentace je nejjednodušší podmnožinou této reprezentace) z oboru celých čísel. Tyto hodnoty jsou často svázány s nějakou slovní reprezentací vyjadřující míru důvěry. Globální aukční portál *eBay* [eBa10] stejně tak jako český aukční portál *Aukro* [Auk10] využívají v základní variantě tříhodnotovou reprezentaci ohodnocení důvěry jako zpětnou vazbu mezi prodejcem na nakupujícím odpovídající slovnímu hodnocení: *negativní*, *neutrální* a *pozitivní* (viz tabulka 2.1).

Hodnota	Význam
-1	negativní
0	neutrální
+1	pozitivní

Tabulka 2.1: Reprezentace důvěry jako zpětné vazby na aukčním portále eBay a Aukro.

Podobně i další webové portály jakými jsou *IMDb*[2], *ČSFD*[3] a další umožňující hodnocení uživatelům využívají stupnici v rozmezí celých čísel $[1, 10]$ v případě *IMDb* [IMD10] či $[0, 5]$ v případě *ČSFD* [Gro10]. Výsledek hodnocení jednotlivých uživatelů se pak agreguje do globální míry důvěry či v tomto smyslu spíše reputace. V práci [ARH97] můžeme nalézt diskrétní rozdělení důvěry v intervalu $[-1, 4]$, význam a popis jednotlivých diskrétních hodnot důvěry v tomto intervalu je zobrazen v tabulce 2.2.

Hodnota	Význam	Popis
-1	nedůvěra	kompletní nedůvěra v entitu
0	ignorance	nerozhodnost, nelze rozhodnut na základě důvěry
1	minimální	nejnižší možná důvěra v entitu
2	průměrná	střední důvěra, pro většinu známých entit
3	velká	vyšší důvěryhodnosti než většina entit
4	naprostá	naprostá důvěra v entitu

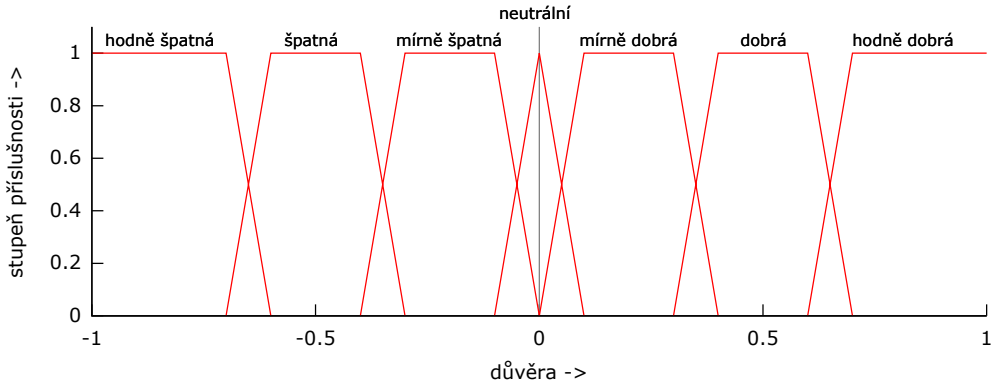
Tabulka 2.2: Reprezentace hodnoty důvěry dle [ARH97].

Další poměrně častou reprezentací důvěru je použití *fuzzy množin*, kde se jednotlivé významné úrovně důvěry mohou vzájemně překrývat. Způsob reprezentace důvěry pomocí

[2] *Internet Movie Database* je mezinárodní on-line databáze informací o filmech, televizních pořadech a vším ostatním, co s filmovou tvorbou souvisí. Jednotlivé firmy a pořady mohou být hodnoceny registrovanými uživateli.

[3] *Česko-Slovenská filmová databáze* je česko-slovenská obdoba databáze IMDb.

fuzzy množin lze nalézt například v práci [AMCR04, AMDS06] (kde je použito intervalu důvěry $[0, 1]$) nebo v práci [SS02a], kde je použito interval $[-1, 1]$ se sedmi významovými úrovněmi: *hodně špatná* (very bad), *špatná* (bad), *lehce špatná* (slightly bad), *neutrální* (neutral), *lehce dobrá* (slightly good), *dobrá* (good), *hodně dobrá* (very good). Příklad této fuzzy množiny pro reprezentaci důvěry dle [SS02a] je zobrazen na obrázku 2.1.



Obrázek 2.1: Příklad fuzzy množiny pro reprezentaci důvěry dle [SS02a].

Poslední obecněji používanou reprezentací je použití spojité veličiny v daném intervalu na množině reálných hodnot. Jedná se tak například o procentuální vyjádření míry důvěry na intervalu $[0, 100]$ (nebo $[0, 1]$ z oboru hodnot reálných čísel), kde míra důvěry je vyjádřena konkrétní hodnotou v rozsahu daného intervalu. Příklad takovéto reprezentace uvádí práce [JF01], kde důvěra mezi dvěma agenty je definována jako funkce $T : A \times A \rightarrow [0, 1]$. Konkrétním příkladem této reprezentace je zde hodnota důvěry 0.8 mezi agenty *Jules* a *Jim*: $T(Jules, Jim) = 0.8$, což odpovídá slovní reprezentaci "*Jules důvěřuje Jimovi z 80%*". Tento způsob reprezentace důvěry je použit také v práci [ZM00].

Poznámka k matematickému zápisu intervalů

V této kapitole jsme použily pro zápis intervalu závorky hranaté namísto špičatých. Pro všechny následující kapitoly a podkapitoly v celé práci bude při psaní matematických výrazů použita anglická syntaxe pro psaní uzavřeného intervalu. Tedy namísto špičatých závorek budeme používat hranaté závorky. Zápis intervalu $< 0, 1 >$ je pak tedy ekvivalentní se zápisem $[0, 1]$, který budeme nadále výhradně používat.

2.2 Reputace

Aby bylo možné provést efektivní a informované rozhodnutí na základě důvěry, je třeba se spolehnout nejen na vlastní znalosti, ale hledat i ostatní zdroje informací. Reputace je jedním z primárních mechanismů, který umožňuje získat další informace pro rozhodování jedince na základě důvěry. Na reputaci lze nejen nahlížet jako na důležitý zdroj informací pro rozhodování na základě důvěry, ale také jako na řídící mechanismus v nějakém sociálním společenství [Ras96, AR05].

V této kapitole se podíváme podrobněji na to, jak lze reputaci definovat a jakým způsobem ji lze použít pro podporu rozhodování na základě důvěry. Zaměříme se též na způsob,

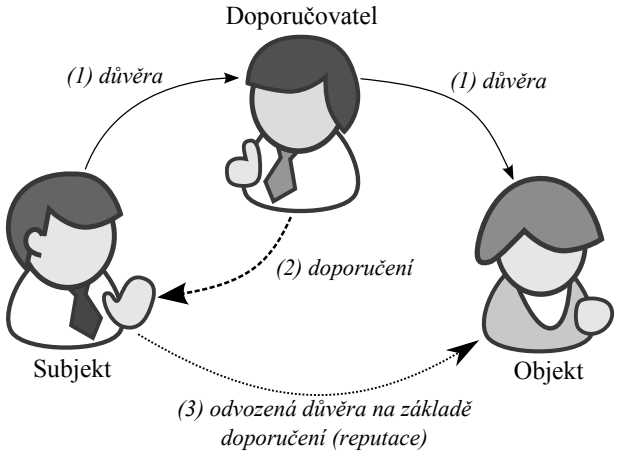
jak je reputace v rámci nějakého sociálního společenství distribuována s čímž také úzce souvisí pojem *doporučení*.

2.2.1 Doporučení

Ještě předtím, než si blíže popíšeme co přesně reputace je a jak ji lze chápat, je třeba se podrobněji podívat na pojem *doporučení*. Vzhledem k tomu, že tento pojem známe z našeho každodenního života, pokusíme se ho tedy vysvětlit tak, aby zapadal do kontextu této práce. Samozřejmě existuje mnoho variant definice tohoto pojmu, ty lze najít například v [KR03, AR05, Gra06], my však vzhledem k výše popsanému stanovíme vlastní:

> *Doporučení je předání subjektivní informace (názoru) nebo znalosti mezi dvěma subjekty vzhledem k nějakému aspektu (nebo aspektům, vlastnostem) konkrétního objektu.*

Subjekt, který provádí doporučení se anglicky nazývá slovem *recommender*, což by se dalo přeložit jako „doporučovatel“. Aby doporučovatel mohl o objektu doporučení provést, musí mít o objektu nějakou znalost. Doporučení je většinou provedeno na základě důvěry mezi doporučovatelem a objektem, přičemž musí existovat i vztah s nějakou mírou důvěry mezi subjektem a doporučovatelem. Tyto vztahy jsou demonstrovány na obrázku 2.2. Jak je vidět, nejdříve tedy musí existovat důvěra mezi jednotlivými prvky transakce, to je označeno (1). Následně doporučovatel provede doporučení objektu (2) a na toto doporučení vytváří *reputaci* a následně i důvěru subjektu vzhledem k objektu na základě reputace, respektive doporučení (3).



Obrázek 2.2: Princip doporučení.

Doporučení může mít různou formu i obsah, různé klasifikace doporučení lze nalézt např. v disertační práci *A Trust-Based Reputation Management System* [Gra06] autorky *Elizabeth L. Gray*. My zde popíšeme několik vybraných tříd klasifikace doporučení:

- *Obsah* – obsahem doporučení může být *kvalitativní* informace: „tuto restauraci si určitě oblíbíte, protože obsluha je naprosto úžasná“ nebo *kvantitativní*: „turistický průvodce může hodnotit tuto restauraci čtyřmi z pěti hvězd pro množinu parametrů

jako je cena, obsluha, nabídka, umístění, ...“. Kvalitativní doporučení je pro automatizované zpracovaní poměrně náročné, proto se pro *informatické* zpracování používá spíše kvantitativní exaktní vyjádření.

- *Explicitnost* – forma doporučení může být *explicitní* anebo *implicitní*. Explicitní by mohlo vypadat např. takto: „Doporučuji ti využít služeb tohoto kadeřníka“, implicitním se pak rozumí samotné oznámení faktu „Julia Roberts používá *konkrétního* kadeřníka v New Yorku“, čímž implicitně doporučuje všem jejím fanynkám *konkrétního* kadeřníka aniž by to explicitně vyjádřila.

- *Přímost* – doporučení může být založeno na základě *přímé zkušenosti* (nebo pozorování), na základě *nepřímé znalosti* nebo kombinace obojího. V případě přímé zkušenosti má doporučovatel přímou zkušenost s doporučovaným objektem a toto doporučení je zpravidla relevantnější než nepřímé doporučení. U nepřímého doporučení nemá doporučovatel přímou zkušenost s objektem a jeho doporučení tedy vychází např. z jiného předchozího doporučení. Vytváří se tím tak *řetěz doporučení*, kdy se jednotlivá doporučení od různých doporučovatelů řetězí.

- *Zdroj* – zdroj doporučení může být *anonymní*, *pseudonymní* nebo *identifikovaný*. Zdroj také může být člen lokální komunity anebo obecný globální. Zdrojem, ale i zároveň cílem doporučení může být také doporučovatel sám, pak se jedná o *autodoporučení*. Stanovení zdroje doporučení je důležité v momentě, kdy se provádí vyhodnocení doporučení a jeho transformace na reputace, protože na základě znalosti zdroje, případně jeho reputace lze určit váhu takového doporučení.

- *Kontext* – kontext doporučení může být obecný, vyjádřen například ve větě: „Jan je dobrý člověk“ anebo může být definován specifický kontext v doporučení tím, že přidáme do doporučení další aspekty doporučení, např. datum, čas, typ služby apod. Kontext je často specifikován jako součást samotného doporučení, např. „Alice doporučuje Roberta jako dobrého řidiče“. Nicméně pro většinu aplikací vychází kontext ze samotné podstaty aplikace (např. doporučení uživatelů na aukčním portálu je pevně svázáno s kontextem „prodávající“ či „kupující“) a proto není kontext často zahrnut ani v návrhu, přestože je známo, že reputace (stejně tak jako důvěra) je kontextově závislá [Mui03, SS05].

2.2.2 Co je reputace?

Podobně jako důvěra tak i reputace nemá jedinou univerzální obecně platnou definici. V rámci různých společenských a vědních disciplín se tyto definice liší. Navíc i v rámci jedné disciplíny jako je informatika a konkrétně pak oblast výzkumu důvěry a reputace se definice jednotlivých autorů reputace různí.

Tak například, již zmiňovaní autoři *Jøsang et al.* v odborném článku *Can We Manage Trust?* [sKD05] definují reputaci takto:

> *„Reputace je to, co je obecně známo o charakteru nebo postoji lidí nebo věcí.“.*

A následně ilustrují na dvou velmi jednoduchých a vypovídajících větách jejich chápání rozdílu mezi důvěrou a reputací:

1. *„Věřím ti na základě tvé dobré reputace.“*

2. *„Věřím ti i přesto, že máš špatnou reputaci.“*

Přičemž první věta demonstruje stav, kdy vzájemná důvěra dvou subjektů je stanovena na základě reputace jednoho z nich. Druhá věta pak demonstruje stav, kdy existuje nějaká *důvěrná* znalost mezi subjekty, stanovená například na základě přímé zkušenosti, která potlačí váhu pozitivní nebo negativní reputace, která je o druhém subjektu známa.

Další definici reputace lze nález v práci *Architecture and Algorithms for a Distributed Reputation System* [KR03] kolektivu autorů *Kinateder et al.*:

> *„Reputace nějaké entity A je z našeho pohledu průměrná důvěra všech ostatních entit vzhledem k A. Tedy na reputaci je pohlíženo z globálního hlediska, zatímco na důvěru je nahlíženo z lokálního a subjektivního hlediska.“.*

Pojem *průměrná důvěra* je závislý na kontextu použití této definice a může znamenat různé způsoby agregace znalostí v rámci sociální skupiny.

V disertační práci *A Framework for Decentralised Trust Reasoning* [AR05] definuje autor *Alfarez Abdul-Rahman* reputaci takto:

> *„Reputace je předpoklad o chování agentů stanovený na základě informací o nich nebo na základě sledování jejich chování v minulosti.“*

Reputace může být tedy chápána skrze *skupinový* nebo *společenský obecný* názor na objekt. Tento názor však nelze považovat za obecný globální názor celé komunity nebo skupiny v případě velkých sociálních skupin. Tedy i přesto, že se jedná o jakýsi skupinový názor, je vnímání reputace vždy subjektivní [AR05].

Stejně tak, jako důvěra může být založena na reputaci, tak i reputace může být stanovena na základě důvěry. Uvažujme příklad, kdy Alice by ráda využila služeb Carol, ale protože s ní nemá žádnou přímou zkušenost, musí se spolehnout na její reputaci, tu získá dotazem na Boba. Na základě doporučení od Boba a následně stanovené reputace Carol si Alice stanoví směrem ke Carol nějakou míru důvěry, díky které se rozhodne že využije nebo nevyužije jejích služeb. Takto lze velmi jednoduše popsat způsob, jak může být důvěra stanovena na základě reputace.

Druhý případ, kdy je reputace stanovena na základě důvěry si můžeme popsat takto: Bob má pouze přímé zkušenosti s Carol na základě kterých si stanovil míru důvěry. Nemá žádné jiné znalosti například z doslechu nebo doporučeních vzhledem ke Carol. Pakliže se Alice zeptá na názor Boba, co si myslí o Carol, tak jediné co může Bob odpovědět (v případě že nepředpokládáme záměrné lhaní) je doporučení stanovené na základě jeho míry důvěry vzhledem ke Carol. Toto doporučení si Alice agreguje společně s dalším případnými doporučeními od jiný členů skupiny a vzniká tak reputace Carol z pohledu Alice. Ta je založena mimo jiné na přímé důvěře Boba vzhledem ke Carol.

Je třeba si uvědomit, že význam reputace, stejně tak její definice, je závislá na tom, pro jakou konkrétní aplikaci je používána – a proto také vznikají její různé definice. Za zmínku tedy ještě stojí definice reputace autorů *Wang et al.* v práci [WV03], která se zabývá návrhem modelů důvěry a reputace pro *peer-to-peer* (P2P) sítě, kde jsou si všechny prvky rovny, tedy nerozlišujeme zde explicitně objekty a subjekty, všechny entity jsou označeny jako *peer* (v definici použijeme označení *klient sítě*).

„Reputace je víra klienta sítě ve schopnosti, poctivost a spolehlivost jiného klienta sítě založená na doporučeních přijatých od ostatních klientů sítě.“

Další různé definice reputace lze nalézt i v pracích [SS01, MMH02a, Gra06]. Zaměříme-li se na společné vlastnosti všech prostudovaných definic reputace, získáváme z nich jeden zásadní společný rys, který je spojuje:

Reputace nějakého objektu je znalost nebo informace o tomto objektu získána od ostatních subjektů z okolního prostředí.

Reputace je také jedním z hlavních zdrojů znalostí pro stanovení důvěry. Z toho plyne i hlavní rozdílnost mezi důvěrou a reputací. Zatímco důvěra je znalost získána hlavně z přímých zkušeností, tak reputace je znalost získaná nepřímo, nejčastěji na základě *doporučení* od ostatních subjektů. Důvěra je stanovena na základě mnoha faktorů, některé z nich mají vyšší váhu, některé nižší. Přímá interakce s objektem mívá často nejvyšší váhu, reputace o něco nižší. Pakliže však nejsou přímé zkušenosti k dispozici, hraje reputace významu roli. Reputace mezi subjekty a objekty ovlivňuje jejich interakci především dvěma významnými způsoby [sKD05]:

1. *Pozitivně ovlivňuje víru subjektů vzhledem k důvěře objektu.*
2. *Působí na sebekázeň objektu, protože objekt ví, že jeho špatné chování bude odhaleno a on bude komunitou potrestán.*

Vlastnosti reputace jsou velmi podobné těm, které jsou stanoveny u důvěry. Reputace, stejně jako důvěra je *kontextově závislá* [Mui03, WV03], je *subjektivní* [AR05] a z výše uvedených definicí rovněž vyplývá, že je *asymetrická* a *dynamická*.

Kapitola 3

Modely založené na důvěře

Původním název této kapitoly zněl *Modely důvěry a reputace*. Nicméně, tento název by nebyl zcela přesný, protože tato kapitola zahrnuje popis jak samostatných modelů důvěry a samostatných modelů reputace, tak i modely kombinující oba přístupy – tedy modely využívající jak důvěry tak i reputaci. Správný název kapitoly by tedy měl znít *Modely důvěry a/nebo reputace*, který postihuje všechny možnosti. Vzhledem k tomu, že pojmy důvěra a reputace spolu velice úzce souvisí, přičemž důvěra je vždy tím primárním aspektem, vžil se v anglické literatuře název *Trust-based models*, který lze přeložit jako *Modely založené na důvěře*. Tato kapitola tedy bude zahrnovat jak čistě modely důvěry a modely reputace, tak i modely, které používají kombinaci obou konceptů.

3.1 Úvod

V současnosti existuje velké množství modelů důvěry a/nebo reputace vycházejících z oblasti informačních technologií. Většina z nich se objevuje právě v posledním desetiletí, kdy už nejsou pojmy jako důvěra a reputace pouze doménou výzkumu v oblasti sociologie, ale pronikají právě do mnoha dalších vědních oborů, přičemž jedním z nich je také informatika.

Právě v oblasti informačních technologií lze tyto modely velmi obecně a „neformálně“ klasifikovat do dvou kategorií. První kategorii můžeme označit jako kategorii *teoretických modelů* a druhou jako kategorii *praktických modelů*. Toto označení nevychází ani tak z podstaty jejich používání jako z podstaty jejich původu.

Za teoretické modely lze označit takové modely, které vznikly především v akademické sféře na základě požadavku nalezení řešení konkrétního (třeba i praktického) problému. Tyto modely bývají poměrně sofistikované, komplexní a někdy i velmi složité. Jsou obvykle formálně matematicky definovány a používají různé informatické a matematické přístupy pro řešení dané problematiky. Jejich reálná aplikace bývá podložena pouze tvrzením autora (nebo autorů) a jejich konkrétní aplikace je v reálném světě jen těžko dohledatelná. Jedná se tedy spíše o teoretický popis možností využití principu důvěry a reputace obecně nebo pro konkrétní problémy.

Naproti tomu do kategorie praktických modelů lze zařadit takové modely, které jsou poměrně jednoduché a prakticky hojně využívané pro různé hodnocení obsahu např. v prostředí internetu. Tyto modely tak vycházejí spíše z praktického prostředí, přičemž často vznikaly evolucí jednoduchých hodnotících systémů a jejími autory jsou spíše programátoři než vědci. Příkladem takového reputačního modelu může být např. hodnocení uživatelů celosvětového aukčního portálu eBay [eBa10], případně českého aukčního portálu Aukro

[Auk10] a dalších. Základní princip těchto modelů je většinou na tolik jednoduchý, že jejich definice nevyžaduje matematické nebo informatické formalismy a vystačí si tak s neformálním popisem, případně jednoduchou matematickou rovnicí (typicky aritmetický nebo vážený průměr pro agregaci hodnocení více uživatelů).

Obě výše popsané kategorie však mají své opodstatněné uplatnění – *praktická* by měla být inspirována *teoretickou* a naopak – a vývoj ideálního modelu by měl kombinovat oba přístupy: jednoduchost a použitelnost praktického přístupu dále obecnost a formální pozadí teoretického přístupu. Především kategorie teoretických modelů nám poskytuje nezbytný obecnější nadhled na celou problematiku využití konceptu důvěry a reputace v oblasti informačních technologií. Tím do praktické roviny přináší nové myšlenky a přístupy, přičemž je pouze otázkou času, zda se v praxi uplatní přímo, nebo poslouží pouze v teoretické rovině k další evoluci přístupů nových.

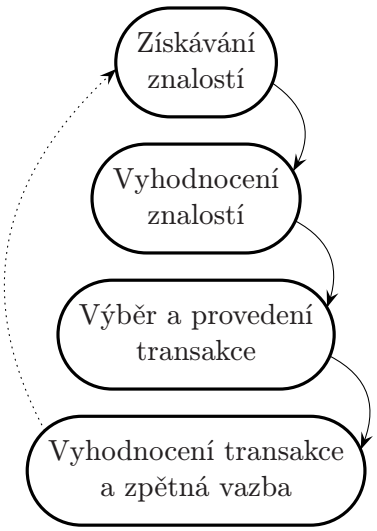
V této kapitole si v první části popíšeme základní obecný princip modelů založených na důvěře, konceptuální schéma jejich návrhu, principy a provedeme jejich klasifikaci dle různých kritérií. V další části se zaměříme na popis některých konkrétních modelů důvěry a reputace se zdůrazněním jejich hlavních charakteristik. Popisovat budeme především vybrané reprezentativní modely v dané oblasti (obecně uznávané modely komunitou, zabývající se výzkumem konceptu důvěry a reputace v oblasti informačních technologií a vědecké práce s velkým počtem citací) a modely významných autorů v dané oblasti.

3.2 Konceptuální schéma komponent obecného modelu

Následující popis obecného schématu komponent pro modely založené na důvěře byl stanoven na základě studia mnoha modelů důvěry či reputace a na základě poznatků vycházejících z prací [SS05, DLB08, SE08, MP09b, MP10]. Jednotlivé modely důvěry a reputace mají sice každý svá specifika, ale na základě jejich prostudování lze stanovit obecné schéma nebo vzor, který lze vždy alespoň částečně nalézt v každém modelu založeném na důvěře. Tyto komponenty by neměly chybět v žádném modelu důvěry a nebo reputace. Na obrázku 3.1 jsou tyto čtyři komponenty uvedeny včetně naznačení způsobu jejich návaznosti. Jednotlivé komponenty lze chápat jako na sebe navazující fáze stanovení a použití důvěry v obecném modelu. Tyto fáze si detailněji popíšeme v následujících podkapitolách.

V následujícím textu budeme používat termín *posuzovaný subjekt* (nebo zkráceně už jen *subjekt*) pro popis entity nebo obecně prvku systému, ke kterému je hodnota důvěry či reputace vztažena, tedy k prvku, který nějaký způsobem hodnotíme a na základě tohoto hodnocení se rozhodujeme zda s ním provedeme interakci či nikoliv. Dále použijeme termín *pozorovatel*, ve smyslu aktivního prvku systému, který pozoruje subjekty systému za účelem získávání znalostí o jejich chování. Termín *hodnotitel* pak rovněž označuje aktivní prvek systému, který na základě výsledku sledování pozorovatele provádí ohodnocení subjektů. Pozorovatel a hodnotitel bývá nejčastěji tentýž aktivní prvek, nicméně na základě principu komponentního přístupu lze jednotlivé kroky dekomponovat a delegovat tak různé úkoly různým prvkům systému.

Vzhledem k tomu, že popis jednotlivých komponent obecného modelu je platný pro modely založené na důvěře, jedná se tedy o popis společný jak pro modely důvěry, tak i pro modely reputace. Proto budeme v následujících podkapitolách používat pouze termíny *model důvěry* nebo *důvěra* ve smyslu *model důvěry a/nebo reputace* respektive *důvěra a/nebo reputace*.



Obrázek 3.1: Konceptuální schéma komponent obecného modelu.

3.2.1 Získání znalostí

První komponenta má za úkol o subjektu zájmu (případně o jeho okolí) získat informace tak, abychom měli znalosti na základě kterých se budeme rozhodovat. Sledování prostředí a získávání znalostí o potencionálních subjektech se kterými je možné provést interakci je tedy velmi důležitou komponentou, na jejíž výsledky jsou navázány bezpodmínečně všechny ostatní následující komponenty. Obecným výstupem této komponenty může být například seznam potencionálních subjektů, které nabízejí nějakou službu, jež je požadována.

Informace a znalosti o chování subjektů mohou mít různý původ – zdroj. Může se jednat o přímou zkušenost se subjektem (ať již přímou interakci nebo pouze pozorování v pozici „svědka“), může se jednat rovněž o předanou zkušenost z našeho okolí (ve smyslu doporučení od ostatních prvků systému), dále pak o předem stanovené znalosti o subjektu (fakta, axiomy) nebo o předpoklady a odvození na základě jiné znalosti (domýšlení). Tyto informace lze také často získat na základě kolektivní znalosti ve smyslu existence skupinové znalosti (znalost stanovená konsenzem několika zdrojů) nebo globální hodnotící entity v systému.

V případě, kdy získaná informace nepochází z přímé zkušenosti se subjektem, je třeba vzít v úvahu i riziko spojené s přijetím takové informace. Může se jednat o částečnou informaci nebo záměrně zkreslenou informaci a je tedy třeba vzít v úvahu důvěryhodnost zdroje, který takovou informaci poskytl. V takovém případě je vhodné ještě k získané informaci přiřadit určitou váhu, která udává její vliv v následném hodnotícím procesu do kterého získaná informace vstupuje. V práci [MP09a] jsou popsány některé bezpečnostní hrozby, které se mohou vyskytnout v případě, že model využívá různé zdroje informací.

3.2.2 Vyhodnocení znalostí

Tato komponenta navazuje bezprostředně na získávání znalostí a jejím cílem je interpretovat nebo také ohodnotit získané znalosti z komponenty *Získávání znalostí* a transformovat je na hodnotu důvěry vzhledem k subjektům. Tento proces transformace informací může být proveden různými přístupy. Cílem této komponenty je tedy ohodnocení subjektů (důvěrou či reputací) tak, aby bylo možné následně vybrat podmnožinu (nebo pouze jeden subjekt) jako nejvýhodnější subjekt(y) pro transakci/interakci. Výstupem této komponenty je seznam

nebo seřazená posloupnost vhodných subjektů s přiřazenou hodnotou důvěryhodnosti. V některých případech se může jednat o jeden konkrétní subjekt ohodnocený nejvyšší mírou důvěryhodnosti.

Metody vyhodnocení znalostí

Nejčastěji používanými přístupy jsou matematicko-informatické přístupy a metody jako např.: analytické řešení, Bayesova síť (využití podmíněných pravděpodobností založených na grafovém modelu), sociální síť (metody analýzy sociální sítě jakožto součástí problematiky teorie grafů), fuzzy množiny nebo biologií inspirované algoritmy (*bio-inspired*) jako např. ACO –*Ant Colony Optimalization*[1]. Tento proces, kde by měl být jasně definován vztah mezi vstupními znalostmi, parametry modelu a výstupními hodnotami důvěry, lze označit za jádro model důvěry. Některé výše zmíněné přístupy budou popsány v následujících kapitolách zaměřených na popis principu konkrétních modelů důvěry.

Důležité aspekty vyhodnocování

V této fázi je třeba vzít v potaz *váhu* (tak jak je popsána v předchozí komponentě *Získávání znalostí*) získané znalosti, která udává její vliv na výsledné hodnocení. Dalším aspektem, který by měl ve vhodně navrženém modelu ovlivňovat způsob vyhodnocení znalosti je její čas. To znamená, že v případě kdy hodnotíme subjekt na základě několika předchozích událostí, měla by mít poslední událost (poslední ve smyslu nejbližší k aktuálnímu času) největší váhu a naopak, nejstarší událost by měla mít váhu nejnižší. Tento přístup umožňuje zvýšit přesnost odhadu budoucího chování subjektu, protože přikládá vyšší váhu nedávným událostem na úkor událostem „dávno minulým“.

Je třeba zdůraznit, že tento proces vyhodnocení znalosti může být ovlivněn subjektivním přístupem hodnotitele. Na rozdíl od první fáze, kdy se pozorovatel snaží získat objektivní informace o sledovaném subjektu, tak v této fázi již může hrát mentální stav – především pak cíle a záměry – hodnotitele významnou roli a transformace jedné události, kterou sledovalo více pozorovatelů může každý z nich zcela odlišně interpretovat.

S touto komponentou také bezpodmínečně souvisí způsob, jak bude důvěra reprezentována. Tato komponenta by měla jasně definovat obor hodnot důvěry a význam těchto hodnot – zda bude použita například diskrétní množina hodnot, fuzzy množina nebo spojitá veličina v daném intervalu (viz popis způsobů reprezentace důvěry v kapitole 2.1.3). Ve všech výše zmíněných případech reprezentace důvěry je také nezbytné stanovit význam význačných nebo krajních hodnot reprezentace tak, aby bylo možné určit sémantický význam ohodnocení znalosti (informace). Tak tedy např. v práci [SS01] je použito pro reprezentaci reputace intervalu $[-1, 1]$ (z oboru reálných čísel), přičemž intuitivně hodnota -1 vyjadřuje nejhorší možné hodnocení, hodnota 0 vyjadřuje neutrální hodnocení a hodnota 1 nejlepší možné hodnocení. Z této reprezentace lze následně určit, že sémantický význam hodnot v intervalu $[-1, 0)$ je „negativní hodnocení“ a význam hodnot v intervalu $(0, 1]$ je považován za „pozitivní hodnocení“.

[1] Pojem *Ant Colony Optimization* překládán jako *Optimalizace pomocí mravenčích kolonií* je optimalizační metoda, inspirovaná komunikačními mechanismy reálných mravenčích kolonií. Tato metoda využívá biologického pozadí mravenčí kolonie, kdy se využívá (velmi zjednodušeně řečeno) principu pozitivní zpětné vazby při hledání nejkratších cest k potravním zdrojům drobnými a téměř slepými mravenci [DS04].

3.2.3 Výběr subjektu a provedení transakce

Tato komponenta by měla popsat fázi, ve které již máme ke konkrétním potencionálním subjektům pro interakci přiřazenu hodnotu důvěry. Chceme tedy vybrat jeden konkrétní subjekt (nebo množinu subjektů v případě implementace skupinové důvěry) se kterým provedeme zamýšlenou transakci/interakci.

Výběr subjektu

Nejčastěji je vybrán subjekt s nejvyšší mírou důvěry, v případě existence subjektů se stejným hodnocení je pak třeba definovat mechanismus, který vybere jeden z nich.

Před vlastním výběrem jednoho nejvhodnějšího subjektu bývá často nejdříve provedena „základní selekce“ vstupní množiny subjektů (z předchozí komponenty) tak, že na základě stanoveného minimálního *prahu* (anglicky *treshold*) vyřadíme z potencionální množiny všechny prvky s nižší (nebo rovnou) hodnotou důvěry. To má za cíl především to, že „vyfiltrujeme nekvalitní“ subjekty z rozhodovacího procesu, což může v některých případech snížit jeho (časovou) náročnost. Zvolený *práh* nemusí být nastaven jako globální nebo konstantní hodnota modelu, ale může se měnit lokálně nebo i v čase v závislosti na charakteristice množiny vstupních subjektů. Jako příklad stanovení lokálního prahu lze uvést situaci, kdy pro vstupní množinu subjektů stanovíme jejich průměrnou důvěryhodnost a pro další vyhodnocení vybereme jen takové subjekty, které jsou důvěryhodností „nadprůměrní“.

Výběr subjektu pouze na základě jeho důvěryhodnosti je nejčastějším přístupem, není však jediný. Například model TACS (*Trust Ant Colony System*) [MPS09] navržený pro P2P sítě inspirovaný ACO (*Ant Colony Optimalization*) používá pro výběr uzlu v síti, jakožto poskytovatele nějaké služby, princip „optimální cesty“. To znamená, že uzel je vybrán na základě nejlépe ohodnocené cesty v grafu, přičemž každá hrana grafu má v tomto modelu dvě *váhy* na základě kterých se „optimální“ cesta určuje.

Provedení transakce

V momentě, kdy hodnotitel stanoví „nejdůvěryhodnější“ subjekt pro provedení transakce, požádá ho o danou službu. Je třeba si uvědomit, že služba kterou subjekt nabízí a služba, která je ve finále subjektem poskytnuta může být zcela odlišná. V případě, že je subjekt „čestný“, provede službu o kterou byl požádán tak, jak nejlépe umí v závislosti na jeho množnostech a schopnostech. V případě, že je subjekt „zlomyslný“, poskytne s největší pravděpodobností službu odlišně (nejčastěji špatně), než jak ji nabízí. Způsobem vyhodnocení výsledku služby se zabývá následující fáze a budu popsána spolu s poslední komponentou.

V této fázi je ale třeba si také uvědomit, že od okamžiku kdy subjekt danou službu nabízel, do okamžiku kdy byl námi vybrán jako její poskytovatel mohlo uplynout relativně mnoho času a tudíž subjekt již nemusí být schopen tuto službu odpovědně provést. Jako příklad takového stavu lze uvést poskytování služby s omezenou kapacitou, kdy v momentě dotazu pozorovatele na dostupnost služby není její kapacita zcela naplněna a subjekt je tedy v tento okamžik vzhledem k této službě použitelný. Následně však v době, kdy probíhá fáze *Vyhodnocení znalostí* hodnotitelem se kapacita služby vyčerpá a subjekt musí další požadavky na poskytnutí služby odmítnout. To je třeba vzít v úvahu, neboť neposkytnutí služby subjektem z objektivních důvodů nemusí mít nezbytně nutně za následek negativní reakci protistrany jako v případě bezdůvodného odmítnutí služby nebo poskytnutí služby záměrně špatně.

3.2.4 Vyhodnocení transakce a zpětná vazba

Vyhodnocení provedené transakce se subjektem je velmi důležitou součástí celého modelu, neboť na základě výsledku transakce lze zavést do modelu zpětnou vazbu tak, aby mohla být důvěryhodnost subjektu aktualizována vzhledem k výsledku transakce. Obecným výstupem fáze vyhodnocení transakce z pohledu hodnotitele je informace o tom, jestli je s výsledkem „spokojený“ nebo „nespokojený“ (tedy pouze dvouhodnotová reprezentace výsledku). Výsledek transakce (anglicky nejčastěji označeno pojmem *outcome*) je ve většině modelů reprezentován vícehodnotovou reprezentací, typicky buď pomocí diskrétní množiny hodnot nebo pomocí spojité veličiny v daném intervalu (podobně jako hodnota důvěry).

Výsledek tohoto hodnocení, stejně jako zpětná vazba ohledně chování subjektu, je v různých modelech zpracován odlišně. Obecně lze tyto způsoby klasifikovat na dvě kategorie modelů: modely které použijí zpětnou vazbu pouze pro *interní zpracování v dalším hodnocení* a modely které provádí *rozšíření zpětné vazby v rámci systému*. Tyto kategorie si teď popíšeme detailněji.

Interní zpracování v dalším hodnocení

Tento způsob je nejčastěji použitý přístup a jeho princip spočívá v tom, že si hodnotitel výsledku transakce vytvoří událost nebo znalost o výsledku přímé interakce a tato znalost je pak jednou z hlavních vstupních proměnných výpočtu ve fázi *Vyhodnocení znalostí*. V případě, že se výsledek takto použije, má podstatně větší váhu nežli informace získaná nepřímo (např. z doporučení), neboť se jedná o přímo získanou znalost na základě přímé interakce se subjektem.

Rozšíření zpětné vazby v rámci systému

Princip tohoto přístupu lze velmi jednoduše popsat tak, že v rámci zpětné vazby provedeme „poděkování“ všem zúčastněným na základě toho, jakou měrou se podíleli na výběru subjektu, jež byl vybrán. Toto „poděkování“ může být jak v negativním, tak i v pozitivním smyslu.

Techničtěji řečeno, tento přístup spočívá v tom, že hodnotitel, který provádí vyhodnocení výsledku transakce provede distribuci informace o výsledku zpět do systému tak, že informace o výsledku transakce nezůstane pouze přiřazena subjektu se kterým byla transakce provedena, ale bude rozšířena i na všechny další prvky systému (nebo jejich podmnožinu), které měly vliv na výběr subjektu. To v praxi znamená, že v případě kdy jsme vybrali subjekt A na základě doporučení od jiné entity B a také C, pak nejen že provedeme aktualizaci důvěry vzhledem k subjektu A, ale „odměníme“ i entity B a C. Tento způsob „odměny“ je chápán jak ve smyslu kladném (kladné zpětná vazba) v případě pozitivně hodnoceného výsledku (zvýšení důvěry v A), tak i ve smyslu záporném (záporná zpětná vazba) v případě opačném (snížení důvěry v A).

Tento přístup se často používá v modelech, kdy je systém reprezentován jako sociální síť entit nebo v modelech, které jsou inspirované ACO principem. To proto, že samotný princip ACO je založen právě na principu zpětné vazby, která generuje hodnocení pro různé cesty (množiny hran a uzlů) grafu pomocí tzv. *feromonů*. Příkladem takových modelů jsou [ZZK06, MPS09]. V případě modelů používající principy sociální sítě je tento přístup velmi podobný a hodnotit se mohou opět hrany a uzly grafu (sítě), které byly v procesu výběru subjektu zainteresovány.

Zajímavým způsobem se řeší tato distribuce hodnocení v práci [MM02b], kdy se distribuce hodnocení ostatním entitám systému provádí pouze tehdy, kdy je hodnocení pozitivní. To je odůvodněno tím, že šíření negativního hodnocení by mohlo být zneužito k útokům typu DoS (*Denial Of Service*) na systému v momentě, kdy by záměrně „zlomyslné“ prvky systému distribuovaly negativní hodnocení.

Vybrané problémy vyhodnocování znalostí a transakce

Při interpretování získaných znalostí, jejich následném vyhodnocení a při vyhodnocení výsledku transakce mnohdy vznikají problémy, které je třeba si uvědomit již ve fázi návrhu modelu tak, aby jejich řešení, pokud existuje a je implementovatelné v rámci dané situace, mohlo být do modelu zapracováno. V následujících několika odstavcích popíšeme některé vybrané problémy, které byly identifikovány v práci [SE08].

Rozpoznání záměru. Jedna z důležitých otázek je, jakým způsobem lze rozhodnout a rozpoznat rozdíl mezi *záměrně zlomyslným*, *sobeckým* (ve smyslu upřednostňování vlastních cílů) a *nekvalifikovaným* (ve smyslu neschopnosti provést transakci tak, jak byla sjednána) chováním. Toto rozhodnutí je poměrně zásadní vzhledem ke stanovení odhadu budoucího chování subjektu. Aby se mohlo toto rozhodnutí provést, je třeba rozpoznat původní záměr.

Vedlejší efekt. Uvažujme situaci, kdy subjekt A požádá o provedení služby X jiný subjekt B. Subjekt B s úspěchem provede službu X tak, jak byla požadována. Nicméně, jako *vedlejší efekt* při provádění této služby provedl (např. v rámci mezikroku) něco, co negativně ovlivní subjekt A. Otázka zní, jak subjekt A vyhodnotí výsledek transakce s B v rámci X? Pozitivně či negativně? Subjekt B sice úspěšně splnil co měl, ale *vedlejší efekt* způsobil negativní vnímání celého výsledku z pohledu A. Navíc, *vedlejší efekt* může být za vlastní transakcí X zpožděn.

Proaktivní chování.[2] Analogicky jako problém s *vedlejším efektem* by měla být brána v potaz i situace, kdy se subjekt rozhodne chovat *proaktivně* a provede v rámci jedné transakce další kroky o které nebyl požádán. Uvažujme situaci, kdy subjekt A požádá jiný subjekt B o službu X_1. Subjekt B se však *proaktivně* rozhodne, že provede službu X_2 protože je přesvědčen, že bude více vhodná a splní tak požadavky subjektu A lépe, než kdyby provedl pouze služby X_1. Subjekt B tedy splní službu X_2 úspěšně a věří v to, že to provedl tak nejlépe jak mohl. Nicméně subjekt A může být s takto provedenou službou nespokojen, protože nebyla provedena přesně tak, jak požadoval a může vyhodnotit výsledek jako *negativní zkušenost*. Zároveň však může být potěšen tím, že subjekt B prokázal proaktivním chováním jistou míru „inteligence“, což může naopak vyhodnotit jako *pozitivní zkušenost*.

Neúplná informace. Získané znalosti nemusí být úplné nebo mohou být zkreslené. Rozhodnutí na základě takové znalosti je tedy vždy spojeno s nějakou mírou neurčitosti a potažmo i rizika. Základní myšlenkou řešení tohoto problému je minimalizace míry neurčitosti získané znalosti nebo zahrnutím míry neurčitosti do procesu rozhodování.

Změna chování. Náhlá změna v chování subjektu od „poctivého“ k „nepoctivému“ může mít obecně dvě příčiny. První z nich je ta, kdy subjekt byl od samého začátku „ne-

[2] Proaktivní chování je způsob chování, kdy se jedinec rozhodne jednat a reagovat promyšleně s předstihem vzhledem k plánování budoucí činnosti [Wik10c].

poctivý“ a choval se z počátku „poctivě“ jen proto, aby si vybudoval vysokou hodnotu důvěryhodnosti, kterou pak zneužije. Druhá příčina spočívá v tom, že subjekt byl „poctivý“, ale v průběhu vývoje se změnila jeho osobnost a stal se tak „nepoctivým“. Odhalením pravé příčiny změny chování by mělo napomoci při nalezení způsobu reakce na změnu chování.

Nerozhodnutelnost. Odvozování důvěry na základě zkušeností je z principu věci omezené. Uvažujme situaci, kdy jeden subjekt A požádá jiný subjekt B o provedení nějaké služby X, aniž by stanovil časový limit v kterém musí být služba provedena. V žádném časovém okamžiku není možné určit, zda subjekt B byl úspěšný nebo ne při provádění služby X (vyjma okamžiku, kdy B službu provede). Problém ohodnocení výsledku služby bez časového omezení je tedy *nerozhodnutelný*.

3.3 Obecná doporučení pro tvorbu modelů založených na důvěře

Autoři *Mármol & Pérez* se v práci *Towards pre-standardization of trust and reputation models for distributed and heterogeneous systems* [MP10] zabývají popisem a analýzou známých modelů založených na důvěře pro distribuované a heterogenní systémy. Na základě výsledků jejich studie stanovili deset obecných doporučení, které je vhodné vzít v potaz v případě definování nového modelu založeného na důvěře. Jak sami autoři v práci uvádějí, různé modely důvěry a reputace mají mnohdy své specifické využití a specifické atributy a proto není nezbytně nutné všechna doporučení do modelu zahrnout. Vždy je tedy třeba zvážit, zda konkrétní doporučení je vzhledem ke specifikám modelu relevantní a podle toho se zařídit.

Následuje výčet jednotlivých doporučení, která jsou převážně převzata z práce [MP10]. Posloupnost jednotlivých bodů doporučení je ponechána tak, jak je uvedeno v původní práci.

3.3.1 Identita a anonymita

Správně navržený model by měl být schopen splnit požadavek, kdy je vyžadována částečná anonymita entit v systému. Komponenta shromažďující znalosti o entitách systému by neměla spoléhat v existenci jednoznačných nepopiratelných identit entit. Model musí být schopen vyrovnat se s tím, že entity mohou systém opouštět a zpět do něj vstupovat pod jinou identitou (pod pseudonymem) z různých důvodů.

Model by měl rovněž garantovat, pakliže je to vyžadováno (např. zadáním), částečnou anonymitu entit z bezpečnostních důvodů. Nicméně čistě anonymní systém (zajišťující plnou anonymitu) není reálné implementovat, neboť různé zdroje transakcí by nebylo možné identifikovat a tyto transakce vzájemně provázat tak, aby bylo možné „budovat důvěru“ v entity. Generovaní unikátních identit, případně další bezpečnostní funkce mohou být zajištěny pomocí kryptografie nebo pomocí bezpečných hardwarových modulů.

3.3.2 Stárnutí událostí

Ve fázi vyhodnocení znalostí je třeba vzít v potaz *čas* (ve smyslu okamžiku získání znalosti nebo provedení transakce) jednotlivých znalostí/transakcí. Nedávno získané/provedené znalosti/transakce by měly mít větší váhu nežli ty starší. Zavedením tohoto přístupu lze v modelu umožnit zpřesnění odhadu budoucího chování entit a zmenšit důsledek *oscilací*

chování entit v čase. Autoři zároveň upozorňují, že způsob implementace této funkce může mít poměrně výrazný vliv na výkon a přesnost celého modelu.

3.3.3 Subjektivní hodnocení

Model by měl podporovat subjektivní hodnocení znalostí z pohledu hodnotitele tak, aby v hodnocení bylo možné zahrnout cíle a různá kritéria definována hodnotitelem. Stejně tak by měl model umožnit provádět hodnocení nejen dle kritérií hodnotitele, ale i dle názoru a dle kriterií definované konsenzem více entit systému. Zavedení subjektivního hodnocení umožňuje, že hodnota důvěry a/nebo reputace mnohem více koresponduje se specifickými cíli a požadavky jednotlivých entit v systému.

3.3.4 Možnost nápravy

Entity, které byly v minulosti klasifikovány jako „nepoctivé“ by měly dostat možnost nápravy v případě, že se „polepší“. V případě, že bychom na základě předchozí negativní klasifikace entitu trvale diskvalifikovali, pak ztrácíme možnosti ji někdy v budoucnu vybrat pro transakci a to i v případě, kdy následná transakce s ní může přinést větší přínos nežli transakce s nejvíce důvěryhodnou entitou. Z tohoto důvodu je vhodné zavést funkci *užitku*, neboť to může přispět k rozlišení nejlépe poskytované služby (v rámci služeb poskytovaných několika entitami) a tudíž jediným kritériem nemusí být hodnota důvěryhodnosti entity, která služby poskytuje.

3.3.5 Nově příchozí entity

Vhodně navržený a funkční model by měl poskytovat nově příchozím „poctivým“ entitám možnost spolupracovat v rámci systému a podílet se na transakcích mezi stávajícími entitami a to i za předpokladu, že jejich důvěryhodnost není tak vysoká jako důvěryhodnost jiných stávajících entit. V tomto případě by se rovněž mohla využít funkce *užitku*, která umožní výběr méně důvěryhodné entity v případě, kdy nabízí jistou spolehlivou (nebo více hodící se) službu. Tento princip by měl zabránit „poctivě“ chovající se entitě (nebo skupině entit) získat díky velké míře důvěryhodnosti „monopol“ na poskytování služeb tak, aby veškeré transakce byly provedeny přes ně. Takový „monopol“ skýtá jisté bezpečnostní riziko a vytváří tak zranitelné místo systému, které by mohlo být jednoduše zneužito k ovládnutí celého systému.

S tímto tématem tedy úzce souvisí problematika zvolení počáteční hodnoty důvěry a/nebo reputace, která bude přiřazena nově příchozí entitě do systému. Tento proces je velmi důležitý a ovlivňuje do značné míry použitelnost celého modelu, proto je třeba se na tuto problematiku při návrhu modelu zaměřit. Některé modely [ZM00] definují hodnotu důvěry nově příchozí entity jako minimální možnou hodnotu důvěry tak, že není možné se pod tuto hodnotu důvěry dostat v žádném případě, ať již se entita chová v systému sebehůře.

Navíc nové vytváření identit by měl systém nějakým způsobem „postihovat“ tak, aby v maximální možné míře zamezil „nepoctivým“ entitám s nízkou důvěryhodností systém opustit a zase se vrátit pod novou identitou s „čistým štítem“. Zároveň však nesmí být v rozporu s výše popsaným, tedy aby nově příchozí „poctivé“ entity měly možnost si důvěru vybudovat a v systému úspěšně fungovat a kooperovat. Tento problém souvisí i s doporučením 3.3.1, které se zabývá identitou a anonymitou entit v systému.

3.3.6 Změna v chování

Je třeba zamezit možnosti zneužití vysoké reputace, jinak může dojít k tomu, že entita dosáhne vysoké reputace a následně se může chovat po delší dobu „nepoctivě“. Model by tedy měl zajišťovat to, že rychle, přesně a efektivně odhalí opakované „nepoctivé“ chování entity, která dosáhla vysoké reputace. Jedním ze způsobu jak takové chování odhalit je implementace *historie chování*.

3.3.7 Nově příchozí *versus* stávající

Nově příchozí entita by neměla mít větší možnosti než stávající, v systému již působící entity. V opačném případě by mohla nově příchozí entita dosáhnout takové reputace, na jejímž základě by mohla provádět interakce s ostatními a následně se chovat „nepoctivě“ dokud by jí to její úroveň reputace dovolila. Poté by mohla systém opustit a opět se vrátit s jinou identitou a takto dále pokračovat. Model by tedy měl zajistit to, že nově příchozí entita nemůže být vybrána jako poskytovatel nějaké služby bezprostředně poté co průkazně důvěryhodná entita již do systému patří. Toto však musí být implementováno s ohledem na doporučení 3.3.5.

3.3.8 Více-kontextovost

Pro každou entitu by měla být stanovena různá hodnota důvěry v závislosti na typu poskytované služby. Jinak řečeno, entita může mít v rámci jedné její služby vysokou míru důvěryhodnosti a v rámci jiné služby zase nízkou. V případě, že by model toto doporučení nerespektoval, může vzniknout problém, že entita získá na důvěryhodnosti díky poskytování jedné služby, ale v případě poskytovaní jiné služby by informace o její důvěryhodnosti (vzhledem k této jiné službě) byla nerelevantní.

Tento bod doporučení v podstatě popisuje princip více-kontextového přístupu k důvěře nebo reputaci, tak jak je popisován v této disertační práci, kdy pro různé kontexty (aspekty, služby, atd.) jedné entity lze přiřadit zvlášť vlastní hodnotu důvěry nebo reputace.

3.3.9 Specifické prostředí

Přenosová rychlost, podobně jako energetická náročnost, je velice důležitý aspekt v některých typech systémů a prostředí – například v bezdrátových senzorových sítích (*WSN – Wireless Sensor Networks*). V případě, že model bude použit v takovémto prostředí, je třeba brát v úvahu to, aby vlivem časté komunikace mezi entitami nedošlo k přetížení sítě (přenosového pásma), což by mohlo mít za následek i energetické „vyčerpání“ entit v systému. Pakliže je model koncipován pro takovéto prostředí, musí být vhodně navržen tak, aby komunikace spojená se správou důvěry byla optimální vzhledem k možnostem sítě.

3.3.10 Důležitost transakce

Důležitost transakce nebo s ní spojená velikost rizika by měla mít bezprostřední vliv na způsob a míru „odměny“. Jinak řečeno, velmi důležitá transakce by měla být úměrně více „potrestána“ v případě že nebude splněna a naopak. Rovněž míra změny důvěryhodnosti by měla být úměrná míře důležitosti transakce, která změnu způsobila.

3.4 Klasifikace modelů založených na důvěře

V této kapitole budou popsány metody klasifikace modelů založených na důvěře, přičemž je v ní použito poznatků z klasifikace modelů popsaných v práci [SS05] a [MP10]. Výčet jednotlivých metod rozhodně není vyčerpávající a oproti původní klasifikaci ve zmíněních pracích je upraven nebo doplněn o vlastní poznatky stanovené na základě studia mnoha modelů důvěry a reputace. Tato klasifikace se tak zaměřuje na významné modely a snaží se nalézt jejich společné vlastnosti, podle kterých jsou navrženy jednotlivé metody klasifikace a jejich kategorie.

3.4.1 Typu modelu

Základním klasifikačním kritériem je tzv. typ modelu. Typem modelu rozumíme, zda model pracuje s důvěrou nebo s reputací. Kategorie dle typu modelu jsou následující:

- *Modely důvěry* – modely pracující pouze s důvěrou.
- *Modely reputace* – modely pracující pouze s reputací.
- *Hybridní modely* – modely pracující s důvěrou i reputací.

3.4.2 Původ informace

Dále lze modely klasifikovat na základě stanovení původu informací (znalostí), které slouží v modelu pro vyhodnocování důvěry. Přímé zkušenosti nebo znalosti „z doslechu“ (tzn. nepřímé znalosti) jsou hlavní zdroje původu informací pro většinu modelů založených na důvěře. V práci [SS05] jsou popsány ještě i další zdroje informací.

- *Přímá zkušenost* – tento zdroj je bezpochyby nejvíce významným a spolehlivým zdrojem informací. Přímé zkušenosti lze rozdělit na dva typy. Prvním typem jsou takové zkušenosti, které jsou stanoveny na základě přímé interakce s entitou. Druhým typem jsou zkušenosti stanovené na základě interakcí ostatních entit ve skupině (v systému).
- *Informace z doslechu* – také často označovány jako *nepřímé znalosti* nebo *doporučení* jsou takové informace, které získáme od ostatních entit v systému. Tyto informace mohou být založeny na základě jejich přímých zkušeností nebo mohou být získány od ostatních entit. Typickým rysem tohoto zdroje informací je to, že oproti přímým zkušenostem je jich podstatně více. Také komplexnost jejich zpracování v rámci modelu je vyšší. Hlavním důvodem je to, že tento zdroj informací má relativně vysokou míru neurčitosti (neúplná, zkreslená nebo skrytá informace) oproti přímé zkušenosti.
- *Sociologická znalost* – tyto znalosti jsou založené na základě sociálního vztahu mezi entitami a na základě jejich sociální role v systému. V reálném světě, jednotlivci, kteří patří do nějaké společnosti (sociální skupiny) vytváří mezi sebou různé typy vztahů. Navíc každý jedinec může patřit do více sociálních skupin, kde zaujímá různé role. Jak vztahy, tak i role ovlivňují jedince ve způsobu chování vzhledem k ostatním (jedincům, skupinám). Analýza sociálních vztahů mezi entitami systému pro získání znalostí je velmi komplexní problematika, která pracuje především s metodami analýzy sociální sítě. Existuje jen velmi málo modelů, které tento zdroj informací využívají.

- *Předsudek* – využití předsudku (anglicky *prejudice*) je další z možností jak stanovit důvěru či reputaci, nicméně tento přístup není v modelech založených na důvěře obvykle využívaný (na rozdíl od reálného světa). Využití tohoto zdroje informací předpokládá existenci znalostí o skupině entit, do kterých posuzovaný jedinec patří. Pojem předsudek je v reálném světě chápán převážně jako negativní postoj vůči nějaké sociální skupině. Přenesení tohoto negativního pojetí do virtuálního prostředí však není nezbytné a může tak být efektivním zdrojem informací pro budování důvěry či reputace.

3.4.3 Aplikace modelu

Další klasifikační kritérium je založeno na tom, pro jaký typ úlohy, aplikaci nebo prostředí je model navrhován. Sem kategorie modelů navržených pro *Peer-to-Peer* (P2P) sítě, kde se hodnota důvěry nejčastěji používá pro stanovení „kvality" jednotlivých uzlů sítě.

Druhou kategorií jsou modely navržené pro mobilní *Ad-hoc* sítě, kde se nejčastěji používá hodnota důvěry pro optimalizaci směrování (routování) paketů v rámci sítě, která je často charakterizována dynamicky se měnící topologií. Další častou aplikační sférou použití důvěry v rámci této kategorie je zajištění různých bezpečnostních mechanismů.

Poměrně novou a velmi aktuální oblastí aplikace jsou *bezdrátové senzorové sítě* (WSN). Tyto sítě mají na rozdíl od klasických LAN či WAN sítí poměrně omezené zdroje (přenosová rychlost, výpočetní a energetická kapacita apod.) a proto zde vzniká nový prostor pro nové sofistikované protokoly. Modely založené na důvěře mohou pro tuto oblast poskytnout řešení způsobu efektivního směrování dat v rámci sítě, případně i způsoby detekce kompromitovaných nebo porouchaných uzlů (tedy určitou formu bezpečnostních opatření).

Nejobecnější kategorií je oblast *agentních a multi-agentních systémů*. Tato kategorie zahrnuje nejvíce známých navržených modelů a jistým způsobem zahrnuje i kategorie popsané výše, neboť vhodně navržený model založený na důvěře pro multi-agentní systém lze aplikovat do různých aplikačních scénářů, jakými jsou např. zmíněně P2P sítě, WSN sítě či Ad-hoc sítě. Návrh modelů v rámci této kategorie tak může být obecnější nežli v kategoriích předchozích, přičemž v následné implementaci pak lze zvolit různé přístupy a agentní architektury k vlastní realizaci. Do této kategorie je rovněž zaměřena tato disertační práce.

3.4.4 Použité metody

Jako jednu z klasifikačních metod lze označit i způsob, jakým je hodnota důvěry či reputace vyhodnocována a zpracována. Jak již bylo zmíněno výše (v podkapitole 3.2.2), používají se především tyto čtyři přístupy:

- *Matematicko-analytické řešení* – metody založené na principu definování matematických vztahů (rovnic) a jejich postupných úprav pro nalezení řešení.

- *Formální logické systémy* – různé přístupy pro vyhodnocení důvěry založené na formálních matematických a logických systémech jako je např. predikátová logika, subjektivní logika (z anglického *Subjective logic*) [Jøs97], apod. Subjektivní logika je vytvořena přímo pro oblast AI (*Artificial Intelligence*), přičemž je aplikována v naší oblasti zájmu, konkrétní např. v práci *Trust Network Analysis with Subjective Logic* [JHP06].

- *Fuzzy množiny* – metody založené na principu aplikace fuzzy logiky (fuzzy pravidel `IF-THEN`), případně na definování hladin důvěry (nebo pro reprezentaci míry neurčitosti) pomocí fuzzy množin.

- *Bayesova síť (Bayesův vzorec)* – metody založené na použití Bayesova vzorce [BS94] (princip podmíněných pravděpodobností a princip tzv. *inverzní pravděpodobnosti*, kde místo pravděpodobnosti hypotézy za předpokladu experimentálních dat počítáme s pravděpodobností dat za předpokladu platící hypotézy) a na základě principu Bayesovi sítě [BS94] (použití pravděpodobnostního grafového modelu, který představuje soubor náhodných proměnných a jejich podmíněných zavilostí, kde jsou jednotlivé vztahy definovány pomocí grafové reprezentace).

- *Biologií inspirované algoritmy* – metody založené na principu reálných biologických procesů pozorovaných v přírodě. Jedná se především o algoritmy ACO [DS04] (*Ant Colony Optimalization* – optimalizace pomocí mravenčích kolonií) nebo také označované jako ACO metaheuristiky, případně o ACS [DG97] (*Ant Colony System*) algoritmy. Tyto algoritmy jsou založeny na biologickém pozadí mravenčí kolonie, kdy se využívá principu pozitivní zpětné vazby při hledání nejkratších cest k potravním zdrojům drobnými a téměř slepými mravenci pomocí zanechávaní pachových stop (feromonů) na cestě za potravou. Tento princip je v modelech založených na důvěře využíván např. pro stanovení nejdůvěryhodnější cesty, kterou tvoří posloupnost entit a jejich vzájemných vztahů v síti [MP08].

Nutno dodat, že většina modelů založených na důvěře nepoužívá výhradně jen jednu z výše popsaných metod, ale kombinuje více přístupů pro optimální dosažení požadovaných výsledků.

3.4.5 Granularita důvěry

Pod pojmem *granularita důvěry modelu* je v práci [SS05] myšleno to, zda je možné nějakým způsobem různé hodnoty důvěry vzhledem k jedné entitě (nebo objektu který chceme hodnotit). Jinak řečeno, tato klasifikace vytváří kategorie modelů, které můžeme rozdělit na *jedno-kontextové* a *více-kontextové*. Jedná se o klasifikaci dle zohlednění více-kontextového principu důvěry, kdy model zohledňuje nebo nezohledňuje závislost důvěry na kontextu. Kategorie rozeznávám primárně tedy dvě:

- *Jedno-kontextové (kontextově nezávislé)* – jsou takové modely, kde důvěra nebo reputace je stanovena binární relací mezi množinou objektů a oborem hodno důvěry, tedy pro jeden objekt je definována jedna konkrétní hodnota důvěry. Přičemž kontext pro který je důvěra definována není nejčastěji ani uveden (případně je definován implicitně na základě aplikace modelu).

- *Více-kontextové (kontextově závislé)* – jsou takové modely, kdy pro jeden objekt je možné definovat různé hodnoty důvěry vzhledem k nějakému specifickému kontextu. Jedná se tedy o zvýšení rozlišovací schopnosti důvěry (granularita důvěry) vzhledem k jednomu objektu.

3.4.6 Ostatní

Existuje i mnoho dalších potencionálních klasifikačních metod, ty však nejsou z našeho pohledu natolik důležité, aby si zasloužily samostatnou podkapitolu. Některé z nich zde vyjmenujeme bez detailnějšího popisu.

- *Reprezentace* – modely založené na důvěře je také možné klasifikovat podle toho, zda je hodnota důvěry a/nebo reputace reprezentována jako binární hodnota, množina diskrétních hodnot, spojitá veličina nebo fuzzy množina.

- *Viditelnost* – důvěra i reputace mohou být sdílené (viditelné) globálně v rámci skupiny subjektů nebo to mohou subjektivní lokální hodnoty přístupné pouze vyhodnocujícímu jedinci. Základní kategorie jsou tedy *subjektivní* a *globální*.

- *Viditelnost* – důvěra i reputace mohou být sdílené (viditelné) globálně v rámci skupiny subjektů nebo to mohou subjektivní lokální hodnoty přístupné pouze vyhodnocujícímu jedinci. Základní kategorie jsou tedy *subjektivní* a *globální*.

3.5 Formální modely důvěry a reputace

V současné době existuje v oblasti výzkumu důvěry a reputace nepřeberné množství modelů využívající důvěry a/nebo reputace. Vzhledem k tomu, že tato disertační práce se zaměřuje na modelování více-kontextové důvěry, popíšeme pouze nejdůležitější známé modely, které se rovněž nějakým způsobem více-kontextové podstaty důvěry dotýkají.

Kontextovost (nebo také kontextová závislost) důvěry je jednou z jejich základních přirozených vlastností a v reálném prostředí mezilidských vztahů je plně využívána. Jedná se o schopnost nebo možnost nahlížet na jednotlivce z různých úhlů pohledu. Tato vlastnost vychází z požadavků na zvýšení granularity důvěry v reálném světě. Na základě této vlastnosti jsme schopni důvěřovat jednotlivci (obecně jedné entitě systému, neboť pojem jednotlivec může evokovat představu, že v reálném světě lze věřit pouze lidem; nicméně obecná důvěra může být zaměřena na „cokoliv“ ve smyslu existence v reálném světě, stejně tak jako vůči skupinám či organizacím) v různých aspektech jeho chování či jednání. Těmto aspektům pak říkáme kontexty.

3.6 Více-kontextové modely důvěry a reputace

Definice kontextu a jeho použití, podobně jako vlastní definice důvěry a reputace, se v různých vědeckých článcích liší. Nicméně většina autorů se shoduje, že hodnota důvěry by měla být prezentována vždy s přihlédnutím ke kontextu [MMH02b, Mui03, SE08]. V několika pracích je rovněž kontext popisován jako *dimenze důvěry* [MC96, FSJ98], přičemž je tato dimenze důvěry v tom smyslu, jak zde popisujeme kontextovost důvěry. Rovněž v pracích [Mar94, MC96] je použito termínu *situačně-specifická* respektive *situační* důvěra, což je popsáno tak, že pro konkrétní *situaci* je důvěra v jednu entitu různá. Autoři *Kinateder & Rothermel* popisují v práci *Architecture and Algorithms for a Distributed Reputation System* [KR03] model důvěry ve kterém používají pojem *kategorie důvěry*, přičemž s každou kategorií je vždy spjata nějaká hodnota důvěry. Většina těchto prací tak popisuje více-kontextový přístup ve stejném nebo podobném významu tak, jak jej popisujeme v této práci, ale definují pro něj různé jiné názvy.

V následujících kapitolách se pokusíme popsat několik významnějších prací, které se zabývají popisem a formalizací více-kontextových modelů důvěry či reputace. Nicméně, existuje i mnoho dalších prací, které se o více-kontextovosti důvěry a/nebo reputace zmiňují, ale nevěnují ji dostatečnou pozornost, nebo problematiku více kontextů ve své práci nezohledňují [AM02, WV03, SS05].

3.6.1 Marsh (1994)

Jedna z prvních prací, která se zabývá formalizací modelu založeného na důvěře a principem využití důvěry oblasti informatiky je disertační práce s názvem *Formalising Trust as a Computational Concept* [Mar94] od autora *Stephena Marshe*. Tuto práci lze označit jako průkopnickou práci zaměřenou na fenomén využití důvěry v oblasti informatiky, konkrétně pak do podoblasti *distribuované umělé inteligence* a *multi-agentních systémů*. Jedná se tedy o totožnou cílovou oblast, do které směřuje i tato disertační práce.

Tato práce je zajímavá hned z několika aspektů. První z nich je ten, že vzhledem k tomu že se jedná o prakticky nejranější dílo svého druhu, autor čerpal poznatky ohledně principu důvěry převážně z filozofických, sociologických a psychologických oblastí výzkumu, díky čemuž popsal ve své práci velmi komplexně princip a fenomén důvěry tak jak jej známe z reálného světa. Druhým, velmi důležitým aspektem je fakt, že je to první práce svého druhu, která formálně popisuje sofistikovaný model důvěry, přičemž respektuje princip více-kontextového přístupu k důvěře zavedením pojmu *situační důvěra*.

Následující popis modelu a formalismů k němu použitých je převzat výhradně z práce [Mar94], přičemž jsou zde popsány pouze základní principy pro jeho pochopení a nejedná se o vyčerpávající přepis celého formalismu.

Základní princip modelu

Hlavním prvkem systému jsou agenti (jedná se tedy o multi-agentní systém), přičemž agenti mohou vytvářet společenství a mohou být členy různých komunit. Množina všech agentů je označena symbolem $\mathcal{A}$, agenti jsou označený malými písmeny $a, b, c, \ldots$. Skupiny, podskupiny a různá společenství agentů jsou označeny symboly $\mathcal{S}_1, \mathcal{S}_2, \ldots$, přičemž $\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{A}$.

Více-kontextový přístup je v modelu zaveden pomocí definice pojmu *situace*, který je zde chápán ve významu specifické situace pro konkrétního agenta v nějakém časovém okamžiku, přičemž situace je subjektivní vzhledem k agentu. Situace jsou v modelu označeny symboly řecké abecedy: $\alpha, \beta, \ldots$; zápis α_x označuje situaci α z pohledu agenta x.

Model důvěry je založen pouze na principu přímých interakcí (není zde použit princip doporučení a reputace), vzhledem k tomu je zde zaveden pojem *znalost*, která je dále označena písmenem K. Znalost zde reprezentuje informaci o tom, zda se dva různí agenti znají, jinými slovy, zdali spolu někdy měli přímou zkušenost a *pamatují si to*[3]. Znalost je reprezentována booleovskou hodnotou, přičemž zápis toho, že agent x zná agenta y je realizován pomocí notace $K_x(y)$. V případě že chceme popsat to, že agent x nezná agenta y použijeme notaci $\neg K_x(y)$.

V modelu jsou popsány tři typy důvěry:

1. *Základní důvěra.* Základní myšlenka pro pochopení tohoto typu důvěry je že „Agenti jsou považováni za důvěřující entity“. Základní důvěra, označována též jako *schopnost* (nebo dispozice – z původního *disposition*) *důvěřovat*, je obecná hodnota toho, jak je agent schopný důvěřovat bez ohledu na to komu důvěřovat. Jedná se tedy o agregovanou hodnotu veškeré důvěry, kterou agent má ve všechny ostatní agenty a zapisuje se T_x (agent x je schopen důvěřovat). Tato důvěra je reprezentována spojitou veličinou z intervalu $[-1, +1)$ (tedy $-1 \leq T_x < 1$).

2. *Obecná důvěra – důvěra v agenta.* Důvěra mezi dvěma agenty x a y ($x, y \in \mathcal{A}$), posuzovaná z pohledu agenta x bez určení situace je označena jako $T_x(y)$. Tato důvěra je

[3] Návrh modelu zahrnuje i princip „zapomínání“ znalostí a popis vlivu paměti na budování důvěry.

rovněž reprezentovaná spojitou veličinou v intervalu $[-1, +1)$. Varianta tohoto zápisu při zohlednění času je vyjádřena zápisem: $T_x(y)^t$ – důvěra agenta x vzhledem k agentu y v čase t.

3. *Situační důvěra.* Důvěra mezi dvěma agenty x a y v nějaké konkrétní situaci α je označena v modelu jako $T_x(y, \alpha)$, respektive $T_x(y, \alpha)^t$. Tato důvěra je opět definována na intervalu $[-1, +1)$.

Reprezentace důvěry

Jak je popsáno výše, hodnota důvěry je v navrženém modelu reprezentována jako spojitá veličina v intervalu $[-1, +1)$. Způsob této reprezentace není typický a to především díky tomu, že horní hranice intervalu je otevřená. Na tomto intervalu je třeba si uvědomit sémantický rozdíl mezi „neexistující důvěrou" a „nedůvěrou". V případě, že agenta vůbec neznáme nebo o něm nemáme znalosti, nelze říci, že je „nedůvěryhodný". Z tohoto pohledu má hodnota 0 význam spíše neznalosti důvěry nebo neschopnosti stanovení důvěry, než nedůvěry. Stejně tak je možné si pod touto hodnotou představit neutrální postoj: „ani důvěra ani nedůvěra". Naproti tomu hodnota -1 zcela jasně vyjadřuje význam „nedůvěry" nebo „absolutní nedůvěry" v negativní významu.

Jak je vidět z definice intervalu, tak hodnota -1 je přípustná, ale hodnota $+1$ již nikoliv. Autor v práci popisuje hodnotu $+1$ ve významu „slepé důvěry" (*blind trust*), přičemž zde argumentuje, proč není vhodné tuto hodnotu povolit. Hlavním argumentem (podloženým na základě několika citací jako např. [Luh88]) je tvrzení, že „slepá důvěra" ve skutečnosti není důvěra, protože odporuje základnímu principu důvěry ve smyslu „úrovně míry víry v něco (nebo někoho)", neboť zde objekt není posuzován z žádného racionálního hlediska (nehledáme důkazy k podložení naší víry, ale pouze „slepě" věříme). Autor zde dále poznamenává, že v případě použití hodnoty $+1$ pro nějakého agenta jako absolutní slepé důvěry, nemůže již být nikdo více důvěryhodnější, přičemž ponechání otevřeného intervalu důvěry lze na základě definice reálných čísel nalézt vždy hodnotu, která bude odpovídat důvěře vyšší.

Dále je v práci diskutováno, zda hodnota -1, jakožto „absolutní nedůvěra" není rovněž zatížena tímto problémem, neboť pakliže „nikdo není dokonalý, pak ani nikdo není nedokonalý". Autor však konstatuje to, že problém není shodný jako v případě „slepé důvěry", tedy že není v rozporu s principem stanovení důvěry, neboť před přiřazením „absolutní nedůvěry" je třeba konkrétního posuzování.

Důležitost a užitek

Pro definici situační důvěry jsou v navrženém modelu zavedeny další dva faktory, které mohou ovlivnit důvěru mezi dvěma agenty v rámci nějaké situace. První z těchto faktorů je definování *hodnoty užitku* (z anglického *utility*) pro konkrétní situaci z pohledu agenta, který provádí vyhodnocení. Hodnota užitku pro konkrétní situaci α z pohledu agenta x se označuje $U_x(\alpha)$ a může nabývat hodnot z intervalu $[-1, +1]$. Význam hodnoty užitku vychází z principu *teorie očekávaného užitku*[4] a zjednodušeně ho lze chápat jako zhodnocení „nákladů a výnosů" s přihlédnutím k riziku dané situace.

[4]Teorie očekávaného užitku (*expected utility theory*), nebo také někdy označována jako *hypotéza očekávaného užitku*, nám zjednodušeně říká, že užitek jedince čelícího nejistotě se počítá jako užitek v každé možné variantě, z kterého se pak udělá vážený průměr, kde váhami jsou jednotlivé pravděpodobnosti. Tento princip vychází z velké míry z ekonomicko-psychologických a sociálních věd a z principu *von Neumann-Morgensternovo* užitkové funkce.[Mon97]

Druhým faktorem ovlivňujícím situační důvěry je hodnota *důležitosti* (anglicky *importance*) situace. Důležitost v navrhovaném modelu se označuje $I_x(\alpha)$ a vyjadřuje subjektivní míru toho, jak je daná situace α pro agent x významná. Důležitost může nabývat hodnot z intervalu $[0, +1]$, tedy záporná hodnota se neuvažuje.

Vyhodnocení situační důvěry

Pro popis způsobu vyhodnocení situační důvěry na základě výše popsané notace je třeba navíc tuto notaci rozšířit tak, aby bylo možné reprezentovat důvěru v různých časových okamžicích, například pro potřeby simulace modelu. Jedná se tedy o rozšíření, které zavádí „temporální“ index, který umožňuje pro všechny výše popsané zápisy stanovit časový okamžik platnosti příslušné hodnoty. Např. tedy, situační důvěra agenta x v agenta y v situaci α je zapsána takto: $T_x(y, \alpha)$, přičemž tento zápis reprezentuje aktuální hodnotu. Zavedením temporálního indexu lze definovat, že situační důvěra agenta x v agenta y v situaci α pro čas t nabývá hodnoty $T_x(y, \alpha)^t$. Tento zápis lze analogicky aplikovat i na užitek $U_x(\alpha)^t$, důležitost $I_x(\alpha)^t$, znalost atd.

Vyhodnocení situační důvěry mezi agentem x a agentem y (z pohledu agenta x) v situaci α lze popsat následující rovnicí:

$$T_x(y, \alpha) = U_x(\alpha) \times I_x(\alpha) \times \widehat{T_x(y)} \,, \tag{3.1}$$

kde zápis $\widehat{T_x(y)}$ označuje odhad obecné důvěry, který bere v potaz všechny relevantní data vzhledem k situační důvěře $T_x(y, \alpha)$ v minulosti. Tedy např. pro aktuální čas t jsou to všechna agregovaná data odpovídající předchozí situační důvěře $T_x(y, \gamma)^T$ takové, že $T < t$ a γ je podobná nebo identická situace jako α.

Způsob exaktního výpočtu $\widehat{T_x(y)}$ je v práci definován několika způsoby s přihlédnutím na schopnost agenta „udržet informace v paměti“ (tedy, pro kolik časových okamžiků zpětně vezme hodnotu situační důvěry v potaz) a na jeho charakteru, konkrétně je-li *pesimista*, *realista* či *optimista*. Na základě těchto faktorů je pak statisticky stanoven způsob výpočtu agregované obecné důvěry jako minimální hodnota (pro pesimistu), maximální hodnota (pro optimistu) a průměrná hodnota (realistu). Konkrétně tedy pro stanovení průměrné hodnoty v daném časovém rozmezí $\theta < T < t$ je odhad obecné hodnoty důvěry pro výpočet situační důvěry stanoven rovnicí:

$$\widehat{T_x(y)} = \frac{1}{|A|} \sum_{\alpha \in A} T_x(y, \alpha)^T \,. \tag{3.2}$$

Situační důvěra, která je v navrženém modelu určena výhradně přímými interakcemi mezi agenty, je jedním z podkladů pro rozhodnutí, zda s nějakým agentem spolupracovat či nikoliv. Do tohoto rozhodnutí však vstupují i další aspekty jako např. riziko plynoucí z dané situace či vnímání kompetence protistrany k provedení dané akce v rámci dané situace. Pro tyto aspekty jsou zde popsány různé způsoby jejich výpočtu, které jsou založeny na využití základní, obecné a situační důvěry.

3.6.2 Abdul-Rahman, Hailes (2000)

V práci s názvem *Supporting Trust in Virtual Communities* [ARH00] popsali autoři *Rahman & Hailes* model důvěry pro agentní systémy, který respektuje kontextovou povahu důvěry. Formální definice nebo struktura kontextu není v této práci explicitně definována a je ponechána na konkrétní implementaci, používající tento model, jak bude stanovena. Tento model definuje způsob vyhodnocování přímé důvěry a důvěry v doporučovatele vždy z pohledu agenta x, přičemž používá množinu čtyř diskrétních hodnot pro stanování stupně důvěry. V následujících podkapitolách bude formálně popsán princip modelu včetně způsobu vyhodnocení důvěry, veškeré formalismy jsou převzaty z článku [ARH00].

Základní notace modelu

Model definuje zápis *přímé důvěry* pro agenta x vzhledem k jinému agentu a v určitém kontextu c a s výsledkem jako stupeň důvěry td takto:

$$t(a, c, td) \tag{3.3}$$

kde $td \in \{vd, d, n, vn\}$ označuje sémantiku stupňů důvěry dle tabulky 3.1.

Stupeň důvěry	Význam
vd	velmi důvěryhodný
d	důvěryhodný
n	nedůvěryhodný
vn	velmi nedůvěryhodný

Tabulka 3.1: Stupeň přímé důvěry a jeho význam.

Agent x v tomto modelu může také důvěřovat jinému agentu b vzhledem k jeho schopnosti poskytování *doporučení* vzhledem k nějakému kontextu c s výsledkem rtd:

$$rt(b, c, rtd) \tag{3.4}$$

kde hodnota rtd reprezentuje „sémantický rozdíl“ mezi hodnotou doporučení agenta b a vlastním vnímáním důvěryhodnosti doporučujícího agenta. Konkrétní popis vyhodnocení doporučení a určení „sémantického rozdílů“ je uveden v původním článku [ARH00] v kapitole 4.9.

Pro potřeby vyhodnocení *přímé důvěry* a *důvěry v doporučovatele* definují autoři další matematické struktury: množina agentů A, množina kontextů C, množina možných výsledků zkušeností (*experiences*) E, množina přímých zkušeností Q a množina pro doporučující agenty R. Jednotlivé množiny si nyní popíšeme podrobněji:

Množina A. Množina všech známých agentů z pohledu agenta x.

$$A = \{a_1, \ldots, a_n\} \tag{3.5}$$

Množina C. Množina všech známých kontextů z pohledu agenta x.

$$C = \{c_1, \ldots, c_n\} \tag{3.6}$$

Množina E**.** Uspořádaná množina možných výsledků zkušeností. Tato množina definuje, jakých hodnot mohou nabývat výsledky přímých zkušeností nebo zkušeností z doporučení a odpovídá hodnotám *velmi dobré* (vg – z původního anglického *very good*), *dobré* (g), *špatné* (b), *velmi špatné* (vb).

$$E = \{vg, g, b, vb\} \tag{3.7}$$

Hodnoty z množiny E odpovídají příslušným hodnotám stupňů důvěry (s mapováním: $vg \rightarrow vd$, $g \rightarrow d$ a analogicky) tak, jak jsou uvedeny v tabulce 3.1.

Množina Q**.** Množina přímých zkušeností, která obsahuje pro každého agenta a z množiny A se kterým má x přímou zkušenost a kontext c čtveřici definovanou jako $s = (s_{vg}, s_g, s_b, s_{vb})$, přičemž každá komponenta této množiny s_j, kde $j = e$ pro $e \in E$, odpovídá agregované hodnotě počtu hodnocení pro výsledek zkušenosti e. Tedy pro $S = \{(s_{vg}, s_g, s_b, s_{vb})\}$ je Q definováno takto:

$$Q \subseteq C \times A \times S \tag{3.8}$$

Množina R**.** Množina pro doporučující agenty, která implementuje schopnost modelu zpracovat *sémantický rozdíl* vnímání v poskytnuté doporučení agenta a jeho důvěryhodnosti. Pro lepší pochopení tohoto přístupu si ukážeme na jednoduchém příkladu tento princip. Mějme agenta a, který agentu x doporučí agenta b v nějakém kontextu c, přičemž stupeň doporučení agenta b je „velmi důvěryhodný“. Agent x na základě důvěryhodnosti a předchozích zkušeností s agentem b sníží tuto hodnotu o jeden stupeň (viz stupně důvěry v tabulce 3.1) a výsledek je „důvěryhodný“. Tedy *sémantický rozdíl* je roven hodnotě -1 (snížení o jeden stupeň).

Máme-li definovány celkem čtyři stupně hodnocení důvěry, pak množina možných hodnot, které může nabývat výsledek vyhodnocení sémantického rozdílů je v intervalu $[-3, 3]$ celých čísel (-3 pro snížení hodnocení ze stupně vd na stupeň vn, 3 pro zvýšení hodnocení ze stupně vn na stupeň vd a analogicky), tedy množina hodnot (označme ji symbolem G) sémantického rozdílů je $G = \{-3, -2, -1, 0, 1, 2, 3\}$.

Pro každého agenta a a kontext c získáváme čtyři množiny upravených hodnocení výsledků zkušeností T_{vg}, T_g, T_b a T_{vu}, kde každá množina T_e pro $e \in E$ reprezentuje upravené hodnocení pro každého agenta a a hodnocení doporučení e. Obor hodnot pro T_e je množina G. Nechť $T = \{T_{vg}, T_g, T_b, T_{vu}\}$, pak na základě výše popsaného se R definuje takto:

$$R \subseteq C \times A \times T \tag{3.9}$$

Vyhodnocení důvěry

Pro stanovení stupně *přímé důvěry* (td) agenta a vzhledem ke kontextu c využijeme relaci (c, a, s) z množiny Q, kde $s = \{(s_{vg}, s_g, s_b, s_{vb})\}$. Na základě této hodnoty se stupeň důvěry td určí dle indexu e komponenty s_e s maximální hodnotou z množiny s takto:

$$\exists td \in E \quad \forall s_e \in s, (s_e = max(s)) \Rightarrow (td = e) \tag{3.10}$$

V případě, že výsledkem je více shodně ohodnocených komponent, tedy funkce $max(s)$ vrátí více jak jednu hodnotu, je jako stupeň důvěry td přiřazena hodnota s *neurčitostí* podle tabulky 3.2 (kde symbol ? reprezentuje „nula nebo jiná hodnota“).

Pro vyhodnocení stupně *důvěry v doporučovatele* (rtd) agenta a vzhledem ke kontextu c využijeme relaci (c, a, t) z množiny R, kde $t = (T_{vg}, T_g, T_b, T_{vb})$. Hodnota rtd je určena na

e	td	Význam hodnocení
$vg \wedge g \wedge ?$	u^+	většinou dobré, občas špatné
$vb \wedge b \wedge ?$	u^-	většinou špatné, občas dobré
ostatní	u^0	stejně špatné jako dobré

Tabulka 3.2: Stupně důvěry s neurčitostí.

základě aplikace funkce **mod**[5] na absolutní hodnoty prvků z množiny T^a, která je stanovena jako sjednocením komponent z t, tedy $T^a = T_{vg} \cup T_g \cup T_b \cup T_{vb}$.

$$rtd = mod(\{\forall x \in T^a \big| |x|\}) \tag{3.11}$$

Tato hodnota pak udává *nejčastější vzdálenost* mezi hodnotami doporučení agenta a hodnotami stanovených na základě skutečných zkušeností vycházejících z jeho doporučení.

3.6.3 Sabater, Sierra (2001) – REGRET

Model navržený v práci *REGRET: A reputation model for gregarious societies* [SS01] nazvaný REGRET je jeden z nejvíce reprezentativních modelů důvěry a reputace pro multiagentní systémy vůbec (článek má stovky citací a je jedním z nejvíce citovaných prací v této oblasti). Reputace v tomto modelu je stanovena na základě třech různých dimenzí: *individuální dimenze*, *sociální dimenze* a *ontologická dimenze*. Tyto jednotlivé dimenze mají v modelu REGRET následující význam:

1. **Individuální dimenze.** Tato dimenze reprezentuje přímou individuální reputaci založenou na základě přímé interakce mezi dvěma agenty.

2. **Sociální dimenze.** Tato dimenze reprezentuje schopnost získání reputace na základě sociální interakce agentů v rámci skupiny a rovněž mezi skupinami, které jsou tvořeny na základě známostí.

3. **Ontologická dimenze.** Obě předchozí dimenze zohledňují reputaci pouze v jednom aspektu. Ontologická dimenze přidává možnost kombinovat reputaci na základě různých aspektů a získávat tak komplexnější schopnost stanovení reputace na základě subjektivního hodnocení aspektů jednoho subjektu. Kontexty jsou zde tedy označeny jako aspekty.

Model REGRET zahrnuje více-kontextovost reputace v pojmu *aspekt dohody* označený symbolem φ v následujících definicích. Tento aspekt se používá jako pojmenovaná proměnná *subjekt*, která je vždy součástí hodnocení v každé dimenzi reputace. Vzhledem k tomu, že REGRET je poměrně důležitým a velmi často citovaným modelem reputace v multiagentních systémech, přičemž zohledňuje více-kontextový přístup, popíšeme si tento model detailněji včetně všech důležitých komponent modelu a výpočetního procesu. Následující popis modelu a veškeré formalismy vychází výhradně z článku [SS01].

[5] Statistická funkce *mod* (z anglického *mode*, nebo také *modal value*, do češtiny přeloženo jako *modus*) je funkce, která vrací nejčastěji se vyskytující hodnotu z dané datové množiny nebo náhodné veličiny [Wik10b]. Tato funkce se také často označuje $\hat{X}$, kde X je právě množina prvků nebo náhodná veličina.

Dohoda a dojem

Autoři definují zpočátku dva pojmy: *dohoda* (*outcome*) a *dojem* (*impression*). Pojem *dohoda* je zde použit ve smyslu *výsledku dohody* mezi dvěma agenty, přičemž obsahem dohody jsou definované termíny (ve smyslu výrazu) a podmínky transakce včetně přiřazení konkrétních hodnot jednotlivým termínům. Forma zápisu dohody je reprezentována kombinací přiřazení mezi proměnnými a konstantami. Proměnné reprezentují vlastnosti dohody (*aspekty* dohody – *kontexty*, pozn. autora) a konstanty reprezentují hodnoty těchto vlastností. V dohodě se rozlišují dva typy proměnných, první z nich se nazývá *obecná* (*common*) a druhá *očekávaná* (*expected*). Proměnné označené jako *obecné* jsou součástí dohody mezi oběma partnery, proměnné *očekávané* jsou individuální pro každého partnera zvlášť. Z toho vyplývá, že na základě jednoho dialogu mezi dvěma agenty vznikají dvě různé subjektivní dohody, pro každého agenta jedna. Obecné proměnné dohody jsou označeny pomocí symbolu $_c$ u přiřazení $=$, tedy: $=_c$.

Příklad dohody z pohledu agenta b:

$o_b = (Datum\_doruceni =_c 10/02 \wedge Cena =_c 2000 \wedge Kvalita =_c A \wedge Datum\_doruceni = 15/02 \wedge Cena = 2000 \wedge Kvalita = C)$. Tento zápis nám říká, že agent b se dohodl s agentem a na dodání produktu do termínu 10/02 za cenu 2000 a s nějakou kvalitou „A“, nicméně agent b očekává, že produkt bude dodán pozdě v termínu 15/02 za dohodnutou cenu, ale v kvalitě „C“.

Příklad dohody pro stejný dialog z pohledu agenta a:

$o_b = (Datum\_doruceni =_c 10/02 \wedge Cena =_c 2000 \wedge Datum\_doruceni = 15/02 \wedge Cena = 2000)$, přičemž je vidět, že agent a nezahrnul do své dohody parametr $Kvalita$. Množinu všech dohod označujeme symbolem **O**.

Pro popis pojmu *dojem* si definujeme notaci pro označení skupiny agentů jako velká písmena $(\mathcal{A}, \mathcal{B}, \dots)$, pro označení agentů používáme malá písmena a indexy $(a_1, a_2, \dots)$. Agent pojmenovaný b_i je agent patřící do skupiny agentů $\mathcal{B}$, symbol **G** označuje množinu identifikátorů skupin agentů, symbol **A** označuje množinu identifikátorů agentů.

Nyní definujeme *dojem* jako subjektivní vyhodnocení nějakého aspektu z *dohody* agentem. *Dojem* označujeme symbolem ι a formálně ho definujeme jako n-tici:

$$\iota = (a, b, o, \varphi, t, W) \tag{3.12}$$

kde $a, b \in \mathbf{A}$ jsou agenti (a je vyhodnocující agent), $o \in \mathbf{O}$ je *dohoda*, φ je proměnná z dohody která je posuzována, t je čas kdy dojem vznikl a W je subjektivní hodnocení vztaženo ke specifickému aspektu φ, který je vyhodnocován z pohledu agenta a. Hodnocení W je z intervalu $[-1, 1]$množiny $\mathbb{R}$ přičemž intuitivně, hodnota -1 vyjadřuje absolutně negativní hodnocení, hodnota 0 vyjadřuje neutrální hodnocení a hodnota 1 pozitivní hodnocení.

Dále je definována množina všech možných dojmů jako I a databáze dojmů agenta $a \in \mathbf{A}$ jako $IDB^a \subseteq I$. Definujeme také databázi dojmů odpovídající vzoru p jako $IDB^a_p \subseteq IDB^a$, kde obecná formulace vzoru je definována takto:

$$(a, b, o, \varphi, t, W) : expr \tag{3.13}$$

kde $expr$ je logická formule z predikátové logiky prvního řádu nad komponentami *dojmu*. Symbol „_“ je použit pro reprezentaci „libovolné“ hodnoty. Například tedy, IDB^b_p se vzorem $p = (\_, a, \_, Datum\_doruceni, \_, \_) : true$ je množina všech dojmů agenta b vzhledem k agentu

a, které odpovídají aspektu *Datum_doruceni*. Pokud bychom chtěli stanovit množinu všech pozitivně hodnocených dojmů agenta a vzhledem k agentu b použili bychom vzor $p = (\_, b, \_, \_, \_, W) : W \geq 0$ nad databází IDB_p^a.

Subjektivní reputace a individuální dimenze

Subjektivní reputace je hodnota přímo vypočtená na základě databáze dojmů agenta. Subjektivní reputaci pro čas t a vzor p zapisujeme jako $R^t(IDB_p^a)$ a stanovujeme ji pomocí váženého průměru hodnocení dojmů s přihlédnutím k relevanci dojmu vzhledem k jeho výskytu v čase:

$$R^t(IDB_p^a) = \sum_{\iota_i \in IDB_p^a} \rho(t, t_i) \cdot W_i \tag{3.14}$$

kde

$$\rho(t, t_i) = \frac{f(t_i, t)}{\sum_{\iota_i \in IDB_p^a} f(t_j, t)} \tag{3.15}$$

kde funkce $f(t_i, t)$ je časově závislá funkce, která pro časy bližší hodnotě t vrací vyšší hodnoty. Jednoduchým typem takové funkce může být např. funkce: $f(t_i, t) = \frac{t_i}{t}$. Obecně funkce $f(t_i, t)$ zajišťuje to, že váha hodnocení pro různé interakce bude stanovena dle času výskytu, tedy pro výpočet reputace jsou důležité především aktuální (poslední) hodnocení na úkor starých (dále v minulosti). Výsledek reputace je vždy normalizován do intervalu $[-1, 1]$, tedy $R^t(IDB_p^a) \in [-1, 1]$.

Na základě definice subjektivní reputace zavedeme notaci $R_{a \to b}(subjekt)$, která reprezentuje subjektivní reputaci mezi agentem a a agentem b vzhledem k nějakému *subjektu* a odpovídá zápisu $R^t(IDB_p^a)$ pro $p = (a, b, \_, subjekt, \_, \_) : true$.

Individuální dimenze reputace v modelu REGRET je tedy definována právě takto zavedenou notací:

$$R_{a \to b}(subjekt) \tag{3.16}$$

Sociální dimenze

Vzhledem k tomu, že individuální dimenze je definována jako subjektivní reputace mezi dvěma agenty, musíme rozšířit pro definování sociální dimenze, která respektuje společenství agentů v rámci skupiny, pro definování skupinové reputace, výše popsanou notaci.

Definujeme si tedy *reputaci skupiny* z pohledu agenta a do které náleží agent b takto:

$$R_{a \to \mathcal{B}}(subjekt) = \sum_{b_i \in \mathcal{B}} \omega^{ab_i} \cdot R_{a \to b_i}(subjekt) \tag{3.17}$$

kde platí že $\sum_{b_i \in \mathcal{B}} \omega^{ab_i} = 1$. Tedy ω^{ab_i} udává váhu jedince b_i ze skupiny $\mathcal{B}$ z pohledu agenta a v rámci skupiny. Tato hodnota je stanovena na základě vnímání jednotlivých členů skupiny agentem a, přičemž v případě maximální nejistoty vzhledem ke členům této skupiny je tato váha stanovena pro všechny členy skupiny na hodnotu $\omega^{ab_i} = 1/|\mathcal{B}|$.

Dále si definujeme *skupinovou reputaci* z pohledu všech členů skupiny $\mathcal{A}$ vzhledem k jednomu agentu b takto:

$$R_{\mathcal{A} \to b}(subjekt) = \sum_{a_i \in \mathcal{A}} \omega^{a_i b} \cdot R_{a_i \to b}(subjekt) \tag{3.18}$$

kde opět platí že $\sum_{a_i \in \mathcal{A}} \omega^{a_i b} = 1$ a ω má podobný význam jako v předchozím případě. Poslední definice *reputace mezi skupinami* vyjadřuje vztah z pohledu všech členů nějaké skupiny $\mathcal{A}$ na skupinu jinou $\mathcal{B}$:

$$R_{\mathcal{A}\to\mathcal{B}}(subjekt) = \sum_{a_i \in \mathcal{A}} \omega^{a_i \mathcal{B}} \cdot R_{a_i \to \mathcal{B}}(subjekt) \tag{3.19}$$

kde opět platí že $\sum_{a_i \in \mathcal{A}} \omega^{a_i \mathcal{B}} = 1$.

Na základě výše popsané *individuální reputace* a sociálních reputací (*reputace skupiny*, *skupinová reputace* a *reputace mezi skupinami*) stanovíme sociální dimenzi takto:

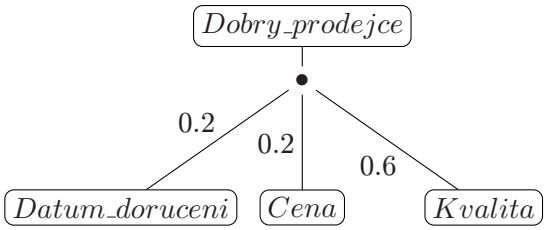
$$\begin{aligned} SR_{a\to b}(subjekt) = \quad & \xi_{ab} \cdot R_{a\to b}(subjekt) + \\ & \xi_{a\mathcal{B}} \cdot R_{a\to\mathcal{B}}(subjekt) + \\ & \xi_{\mathcal{A}b} \cdot R_{\mathcal{A}\to b}(subjekt) + \\ & \xi_{\mathcal{A}\mathcal{B}} \cdot R_{\mathcal{A}\to\mathcal{B}}(subjekt) \end{aligned} \tag{3.20}$$

kde $\xi_{ab} + \xi_{a\mathcal{B}} + \xi_{\mathcal{A}b} + \xi_{\mathcal{A}\mathcal{B}} = 1$ a tyto parametry určují význam (vliv) jednotlivých prvků (reputací) na celkovou reputaci sociální dimenze. Tyto parametry mohou být pro konkrétní implementace různé a zároveň to mohou být dynamické hodnoty, které se mění v čase.

Ontologická dimenze

REGRET pro popis ontologické dimenze používá grafovou reprezentaci pro propojení atomických aspektů *dohody* pomocí ohodnocených hran, přičemž jednotlivé atomické aspekty dohody vytváří s různou vahou (v závislosti na ohodnocení hrany) komplexní subjekty, na které lze rovněž aplikovat individuální nebo sociální dimenzi reputace tak, jak jsou popsány výše.

Příklad ontologické struktury je na obrázku 3.2. Tato struktura popisuje příklad propojení atomických aspektů (*Datum_doruceni*, *Cena*, *Kvalita*) dohody jež byla použita jako příklad na začátku popisu tohoto modelu. Propojením těchto atomických aspektů tak dostáváme subjekt *Dobry_prodejce*, kde kompozice tohoto subjektu je dána váhou jednotlivých atomických aspektů, kterou udává ohodnocení hrany mezi atomickým aspektem a komplexním subjektem.



Obrázek 3.2: Příklad ontologické struktury dle REGRET [SS01].

V případě ontologické dimenze se pak určuje reputace uzlu i ontologického grafu na základě jeho poduzlů takto:

$$OR_{a\to b}(i) = \sum_{j \in children(i)} w_{ij} \cdot OR_{a\to b}(j) \tag{3.21}$$

kde $OR_{a\to b}(j) = SR_{a\to b}(j)$ v případě kdy je j atomický aspekt. V případě vhodně zvolených parametrů $(\xi_{ab}, \xi_{a\mathcal{B}}, \dots)$ lze definovat ontologickou reputaci pro atomické aspekty jako individuální reputaci $R_{a\to b}(j)$. Pro konkrétní příklad na obrázků 3.2 lze ontologickou reputaci subjektu *Dobry_prodejce* s využitím sociální reputace atomických aspektů a jejich důležitosti (ohodnocení hrany) spočítat takto:

$$\begin{aligned} OR_{a\to b}(Dobry\_prodejce) = \quad & 0.2 \cdot SR_{a\to b}(Datum\_doruceni) + \\ & 0.2 \cdot SR_{a\to b}(Cena) + \\ & 0.6 \cdot SR_{a\to b}(Kvalita)\ . \end{aligned} \tag{3.22}$$

Autoři dále v článku uvádějí několik různých experimentálních scénářů, na nichž ukazují použití různých dimenzí důvěry. V závěru pak konstatují, že navržený reputační model má smysl pouze tehdy, je-li součástí vyjednávacího procesu mezi agenty. Vzhledem k tomu pak na tuto práci navazují v článku [SS02b], který popisuje integraci modelu spolu s vyjednávacím protokolem navrženým v [FSJ98].

3.6.4 Mui (2003)

Autor *Lik Mui* se v letech 2001 až 2004 (společně s kolegy) zabýval intenzivně výzkumem v oblasti důvěry a reputace. Důležité jsou především jeho vědecké články [MMH02b, MMH02a], které spolu s dalšími publikacemi vyústily v jeho disertační práci s názvem *Computational Models of Trust and Reputation: Agents, Evolutionary Games, and Social Networks* [Mui03]. V této disertační práci mimo jiné popisuje výpočetní modely důvěry a reputace pro multi-agentní systémy s ohledem na kontextovost důvěry a/nebo reputace. Kontextovou závislost reputace popisuje již v článku [MMH02b]. Při popisu tohoto modelu popíšeme jen jeho část, která je z našeho pohledu důležitá a zabývá se formalizací kontextu a popisu prostředí pro jeden a více kontextů. To především z toho důvodu, že pro výpočet reputace autor diskutuje množství různých přístupů a není zde prostor pro obsáhlý popis principu a formalizaci těchto přístupů. Následující definice kontextu a prostředí s více kontexty je převzato výhradně z disertační práce [Mui03].

V kapitole 3 a podkapitole 3.3 formalizuje *Mui* hodnotící (reputační) proces na množině *agentů* A a množině *objektů* O systému:

$$A = \{a_1, a_2, ..., a_M\} \tag{3.23}$$

$$O = \{o_1, o_2, ..., o_N\} \tag{3.24}$$

Přičemž agent je entita, která může hodnocení provádět a objekt je cílem hodnocení. V modelu je definováno, že agent může být zároveň i v množině objektů, tudíž může být hodnocen jinými agenty (nebo sebou samým). Dále je pro tento model definováno pouze dvouhodnotové hodnocení: „doporučit“ (reprezentováno hodnotou 1) a „nedoporučit“ (reprezentováno hodnotou 0). Funkce hodnocení je pak reprezentována takto:

$$\rho : A \times O \to \{1, 0\} \tag{3.25}$$

a zápis ρ_{ik} reprezentuje hodnocení agenta a_i vzhledem k objektu o_k. Pro modelování procesu sdílení názoru mezi agenty je použito události označované jako *interakce*. Událost *interakce* e je tedy prvkem z množiny všech událostí interakcí E, při kterém se dotazující agent a_i ptá na názor agenta a_j vzhledem k hodnocení objektu:

$$e \in E = A^2 \times O \times \{1, 0\} \cup \{\perp\} \tag{3.26}$$

Dále jsou v tomto modelu popsány dva přístupy k problematice sdílení hodnocení vzhledem ke kontextu reputace. V prvním případě jsou všechny objekty z množiny O hodnoceny v jednom kontextu. Tento případ je nazván jako *prostředí s jedním kontextem* (*uniform context environment*). V druhém případě je použito více kontextů pro sdílení doporučení mezi agenty a tento případ je nazván jako *prostředí s více kontexty* (*multiple contexts environment*).

V případě prostředí s jedním kontextem je reputace agenta a_j z pohledu agenta a_i stanovena jako pravděpodobnost toho, že v následující interakci agent a_j ohodnotí nějaký nový objekt o_k hodnocením „doporučit“. Reputace je pak reprezentována jako funkce:

$$R : A \times A \rightarrow [0, 1] \tag{3.27}$$

přičemž reputace agenta a_i vzhledem k sobě samému je rovna jedné: $R(a_i, a_i) = 1$.

V prostředí s více kontexty je definován jak kontext, tak i množina kontextů. Kontext je zde definován jako uspořádaná množina (seznam) inicializovaných atributů. Atribut (označený symbolem b) definuje (ne)existenci vlastnosti objektu a je definován jako funkce zobrazení objektu z množiny O do množiny $\{0, 1\}$:

$$b \in B \quad \text{a} \quad b : O \rightarrow \{0, 1\} \tag{3.28}$$

kde intuitivně „0“ reprezentuje neexistenci vlastnosti atributu objektu a „1“ označuje existenci vlastnosti atributu. Množina všech atributů B je definována takto: $B = \{b_1, b_2, \dots\}$, přičemž připouštíme, že tato množina může být nekonečná spočetná. Kontext je pak uspořádaná množina inicializovaných (přiřazených pro dané objekty s konkrétní hodnotou (ne)existence vlastnosti) atributů:

$$c = \{b_i, b_j, \dots, b_k\} \text{ kde } c \in C \tag{3.29}$$

kde C je množina všech kontextů a není povoleno, aby jeden kontext obsahoval duplicitní atributy. Na základě takto definovaného kontextu je reputace ve více-kontextovém prostředí definována jako čistě kontextově závislá:

$$R : A \times A \times C \rightarrow [0, 1] \tag{3.30}$$

Tuto notaci pak autor dále používá při tvorbě výpočetního modelu důvěry a reputace, který tvoří jádro jeho disertace. Tento výpočetní model je založen na využití aplikace podmíněných pravděpodobností a Bayesova vzorce [BS94] pro odhad chování agentů na základě jejich hodnocení (reputace). Autor zde používá velké množství různých hodnotících strategií a metod z oblasti analýzy sociálních sítí a je mimo rámec této práce zde tyto přístupy popsat.

Nicméně v práci se objevuje (podkapitola 3.3.2, strana 36) z našeho pohledu poměrně zajímavá zmínka o závislosti kontextů mezi sebou. Autor popisuje na příkladu z reálného světa, kde na základě reputace nějaké osoby i v kontextu „politik“ lze odvodit, že pokud je člověk i dobrý politik, pak lze předpokládat, že bude zároveň dobrý „řečník“. Autor konstatuje, že jednotlivé kontexty nejsou vždy zcela nezávislé a lze na základě některých společných atributů (v souvislosti s tím, jak on definuje kontext) stanovit úroveň podobnosti kontextů, na jejíž základě lze odvozovat reputace z jednoho kontextu na druhý. Nicméně tuto problematiku uzavírá konstatováním, že v jeho modelu jsou jednotlivé kontexty nezávislé a že není cílem jeho práce provádět toto stanovení podobnosti kontextů. Odvozováním reputace, na základě této podobnosti kontextů, se proto dále ve své práci již nezabývá.

3.7 Shrnutí

V této kapitole jsme se zaměřili na popsání stavu současné problematiky v oblasti návrhu a základních přístupů při vytváření modelů důvěry a nebo reputace. V úvodní části bylo popsáno základní konceptuální schéma obecného modelu důvěry včetně různých problémů, které s jednotlivými fázemi návrhu jednotlivých částí modelu souvisí. Byla rovněž popsána základní klasifikace modelů založených na důvěře.

Ve druhé části jsme popsali reprezentativní formální modely důvěry nebo reputace, přičemž jsme se zaměřili výhradně na ty, jež určitým způsobem zohledňují více-kontextovou povahu důvěry. Některé z těchto modelů explicitně definují význam pojmu kontext, jiné naopak ponechávají tento pojem v abstraktní rovině a záleží pak na konkrétní reálné aplikaci, jakého významu nabude.

Na základu studia nejen těchto modelů jsme stanovili jisté obecně platné předpoklady, které by měl více-kontextový model důvěry splňovat. Tyto předpoklady, včetně návrhu modelu důvěry založeného na nich, budou popsány v následující kapitole.

Kapitola 4

Hierarchický model důvěry s kontexty

Tato kapitola se detailně zabývá formulací modelu důvěry, který je navržen tak, že plně respektuje více-kontextovou povahu důvěry. Jedná se o koncept modelu, který není navržen pro jednu konkrétní aplikaci nebo účel, ale umožňuje využití potenciálu obecnějšího návrhu založeného na možnosti modelování důvěry s více kontexty.

Model, který bude popsán v následujících podkapitolách, je ve skutečnosti diskrétní systém, který je navržen jako komponenta pro podporu rozhodování agentů v heterogenním distribuovaném multi-agentním systému. Tento model zahrnuje strukturu báze znalostí agenta a popis chování systému včetně výpočetního modelu důvěry.

4.1 Základní vlastnosti modelu

Jak již bylo citováno v úvodní kapitole této práce: „existuje pouze velmi málo výpočetních modelů důvěry a reputace které, zohledňují více-kontextovou povahu důvěry a/nebo reputace a ještě mnohem méně z nich nabízí nějaký způsob řešení“ [SS05]. Námi navržený model je postaven především na dvou základních předpokladech, které byly stanoveny na základě studia velkého množství odborných prací zabývajících se modely důvěry a reputace:

1. Modely důvěry s jedním kontextem jsou pouze speciální případy modelů využívající více-kontextové vlastnosti důvěry. Proto modely s podporou více-kontextovosti jsou obecnější a využitelné pro více aplikací.

2. Ve více-kontextovém prostředí lze nalézt různé kontexty, mezi kterými lze pozorovat vzájemnou vazbu. Tedy jednotlivé kontexty ve více-kontextovém prostředí nejsou vzájemně zcela nezávislé.

Především pak právě pro druhý předpoklad lze konstatovat, že většina navrhovaných výpočetních modelů důvěry a/nebo reputace, které respektují kontextovou vlastnost, tento fakt zanedbává [Mui03]. Toto je také jeden z hlavní důvodů, proč jsme se na problematiku více-kontextových modelů důvěry zaměřili.

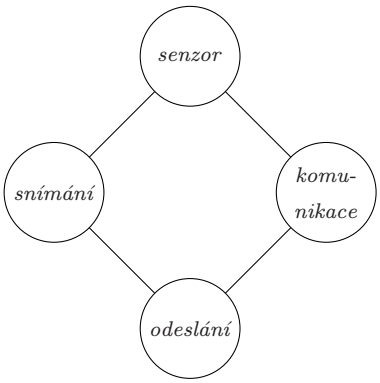
Pakliže jsme schopni mezi různými kontexty jedné entity nalézt vazbu, jsme schopni i v jisté omezené míře provést odvození z jednoho kontextu na druhý v případě, že je mezi nimi tato vazba. Tomuto odvození říkáme *přenesení kontextu* a je již popsáno (včetně příkladu) v kapitole 2.1.2. Jednou ze základních vlastností námi navrhovaného modelu je právě definice

způsobu jakým je důvěra na základě znalosti v jeden kontext přenesena pomocí vzájemné vazby na kontext druhý.

4.2 Hierarchický model důvěry s kontexty

Námi navržený model, který jsme pojmenovali *Hierarchický Model Důvěru s Kontexty* (anglicky *Hierarchical Model of Trust in Contexts* – zkratka HMTC) je založen na předpokladu, že vztahy mezi kontexty lze popsat jako *specializaci* nebo *kompozici*. To znamená, že jednotlivé kontexty jsou specializací jiných kontextů, případně jsou jejich komponenty.

Jako příklad specializace lze uvézt vztah mezi kontexty „lékař“ a „zubní lékař“ (analogie k příkladu uvedenému v kapitole 2.1.2), kde „zubní lékař“ je specializací kontextu „lékař“. Jako příklad kompozice lze pak uvést kontexty „automobil“, „brzdy“, „motor“ a „převodovka“ – kontext „automobil“ je kompozicí kontextů „brzdy“, „motor“ a „převodovka“, v každou tuto komponentu (kontext) lze vyjádřit určitou míru důvěry a následně pak lze určit celkovou míru důvěry v základní komponentu „automobil“.



Obrázek 4.1: Grafická reprezentace kontextů s cyklem.

Pro reprezentaci námi navrženého modelu a především pro vyjádření vztahu mezi kontexty jsme nejprve použili stromovou strukturu, ta se však záhy ukázala jako omezující, protože strom nesmí obsahovat cykly. V našem případě je však třeba cykly povolit, neboť jednotlivé kontexty mohou být specializací dvou nebo více různých nadkontextů (propojených kontextů na vyšší úrovni). Tento případ je ilustrován na obrázku 4.1, kde je uveden příklad senzorového uzlu, jež má dvě primární funkce: *snímaní veličin* a *komunikace* v rámci sítě s ostatními uzly. Specializací těchto dvou kontextů pak vzniká kontext *odeslání dat*, který reprezentuje schopnost uzlu odeslat nasnímané veličiny v rámci sítě.

Námi zvolená reprezentace je nejvíce podobná grafové struktuře pojmenované jako *vrstvený graf* (anglicky *layered graph*) [PY89, FFK+91], kde jsou jednotlivé uzly rozděleny do vrstev a jednotlivé hrany jsou povoleny pouze mezi uzly sousedních vrstev.

Definice 4.1. *Vrstvený graf* je takový souvislý graf, kde uzly jsou rozděleny do vrstev $L_0 = \{s\}, L_2, \ldots, L_n$ (každá vrstva přestavuje množinu uzlů) a hrany jsou pouze mezi uzly v sousedících vrstvách. Pro každou hranu je definována váha jako nezáporné celé číslo. V první vrstvě L_0 je pouze jeden uzel s nazývaný jako *výchozí* (anglicky *source*). [FFK+91]

Pro formulaci námi navrženého hierarchického modelu důvěry jsme vyšli z definice vrstveného grafu, přidali další komponenty a definovali jeho strukturální pravidla a funkce.

Definice 4.2. *Hierarchický model důvěry s kontexty* je pětice $HMTC = (N, E, T, \rho, w)$, kde:

1. N je konečná množina uzlů, které jsou rozděleny do vrstev $L = L_1 \cup L_2 \cup \cdots \cup L_{n-1} \cup L_n$ kde n definuje *úroveň grafu*. Ve vrstvě L_1 se nachází pouze jeden uzel, který nazýváme *kořenový*.

2. E je konečná množina hran. Hranu mezi uzly $u, v \in N$ zapisujeme jako $e_{u,v}$ (případně jako dvojici uzlů $(u, v) \equiv e_{u,v}$), přičemž hrany jsou povoleny pouze mezi uzly v sousedních vrstvách, musí tedy platit že:

$$\forall e_{u,v} : e_{u,v} \in E_i \rightarrow ((u \in N_i) \wedge (v \in N_{i+1})) \vee ((u \in N_{i+1}) \wedge (v \in N_i)) \tag{4.1}$$

   Množinu hran lze také pomyslně rozdělit do podmnožin dle uzlů, které je spojují a jim přiřazených vrstev. Pro graf úrovně n existuje právě $n - 1$ podmnožinu hran takto: $E = E_1 \cup E_2 \cup \cdots \cup E_{n-1}$, kde množina E_1 obsahuje všechny hrany mezi uzly z vrstvy L_1 do vrstvy L_2.

3. T je uspořádaná množina diskrétních časových okamžiků: $T = \{t_1, t_2, \ldots, t_{i-1}, t_i\}$, kde pro každé dva časové okamžiky t_x, t_y a celá čísla x, y taková $x < y$ platí, že časový okamžik t_x předchází časovému okamžiku. To zapisujeme také jako $t_x < t_y$. Význam časové množiny jako vlastní komponenty HMTC bude vysvětlen později.

4. ρ je *funkce důvěry*, která je zobrazením libovolného uzlu z N v čase z T na *rozšířený interval důvěry* označený τ:

$$\rho : N \times T \rightarrow \tau \tag{4.2}$$

   Rozšířený interval důvěry je definován jako klasický matematický interval (tedy podmnožina reálných čísel omezená dvěma mezními hodnotami, označenými jako meze intervalu) sjednocený s prázdnou množinou. Pro mezní hodnoty x a y rozšířeného intervalu platí že $x, y \in [0, 1]$ (přičemž interval $[0, 1]$ z množiny $\mathbb{R}$ označujeme jako *základní interval důvěry*) a $x \leq y$.

$$\tau \subset [0, 1] \cup \emptyset \tag{4.3}$$

   Význam základního a rozšířeného intervalu důvěry bude popsán v následujících podkapitolách.

5. w je *váhová funkce*, která přiřazuje každé hraně z E váhu z intervalu $[0, 1]$ z množiny $\mathbb{R}$.

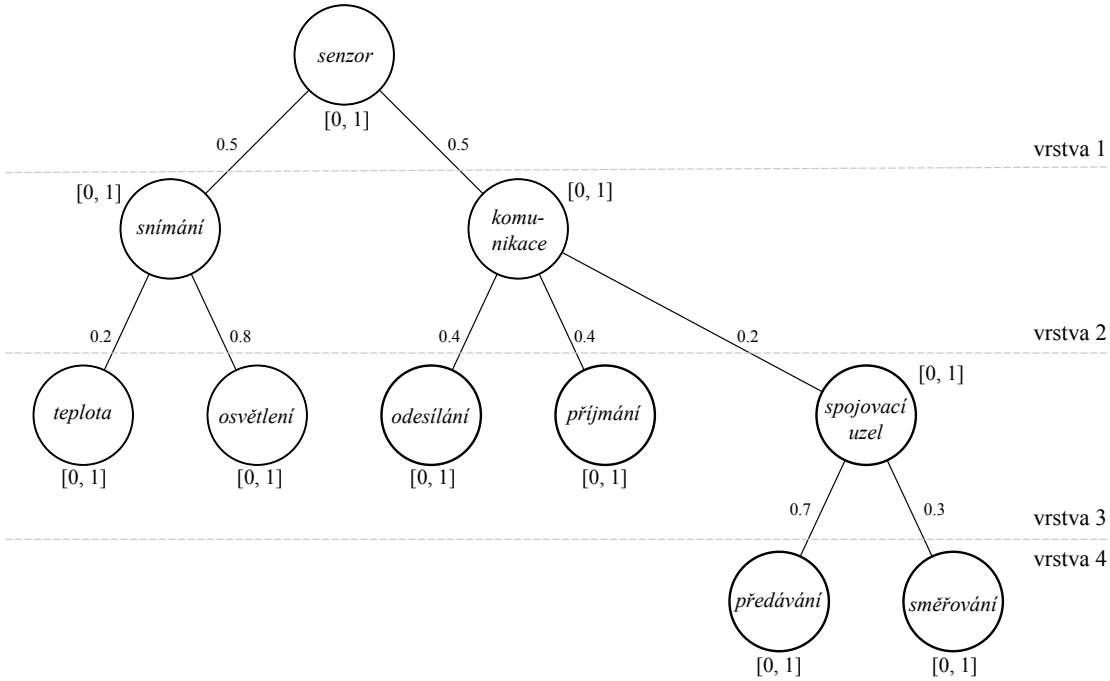
$$w : E \rightarrow [0, 1] \tag{4.4}$$

   Do definice této funkce je třeba také strukturální podmínku, že pro každý *neterminální* uzel je suma vah všech hran do uzlu přicházejících (směr z nižší vrstvy do vrstvy vyšší) rovna jedné:

$$\begin{gathered} \forall u^l \in N - N_T : \sum_{\forall v^{l+1}} w((u^l, v^{l+1})) = 1 \\ \text{pro } l = 1 \ldots (n-1), \text{ kde } n \text{ je úroveň grafu} \end{gathered} \tag{4.5}$$

Terminální uzel je takový uzel grafu, pro který neexistuje uzel v nižší vrstvě, jež by s ním byl spojen hranou. Množinu všech terminálních uzlů označujeme N_T. Označení hrany (u^l, v^{l+1}) je ekvivalentní označení $e_{u,v}$, přičemž je zde explicitně definováno, že uzel u se nachází ve vrstvě l a uzel v se nachází ve vrstvě o jednu nižší, tj. ve vrstvě $l + 1$.

Konkrétní příklad jednoduchého HMTC modelu je uveden na obrázku 4.2. Uvedený příklad reprezentuje kontexty stanovené pro případ pomyslného *senzorového uzlu* v bezdrátové senzorové síti, který můžeme hodnotit na základě dvou primárních kontextů: *snímání* veličin a *komunikace* v rámci senzorové sítě. Kontext *snímání veličin* se pak dále pomocí vztahu kompozice rozděluje na snímání *teploty* a intenzity *osvětlení*. Kontext *komunikace* je dekomponován na tři další podkontexty, jmenovitě: *odesílání*, *příjímání* a *spojovací uzel* – to ve smyslu schopnosti předávat data od jiných uzlů do sítě. Tento kontext *spojovací uzel* lze pak rozdělit na podkontexty *předávání* a *směřování*. Tyto dva kontexty jsou pro práci bezdrátové senzorové sítě poměrně důležité, neboť jednotlivé senzorové uzly nemají mnohdy představu o celé topologii sítě a naměřené veličiny jsou tedy předávaný do cílového uzlu mnohdy přes další uzly dostupné v jejich bezprostředním okolí.



Obrázek 4.2: Grafická reprezentace konkrétního příkladu jednoduchého HMTC modelu.

Na obrázku 4.2 je naznačen model v počátečním stavu a důvěra všech kontextů je inicializována výchozí hodnotou rozšířeného intervalu důvěry $[0, 1]$. Tato implicitní hodnota důvěry bude podrobněji vysvětlena v následující podkapitole. Kontexty jsou rozděleny do čtyř úrovní (vrstev), přičemž první úroveň obsahuje pouze kořenový kontext, na třetí a čtvrté úrovni máme terminální kontexty: *teplota*, *intenzita osvětlení*, *předávání* a *směřování*.

4.2.1 Reprezentace důvěry

Pro definici mezních hodnot důvěry v HMTC je použita spojitá veličina na intervalu $[0, 1]$ z množiny reálných čísel. Tento interval označujeme jako *základní interval důvěry*, jednotlivé hodnoty z tohoto intervalu vyjadřují míru důvěryhodnosti, kterou lze analogicky vyjádřit i procentuálně. Význačné hodnoty základního intervalu důvěry jsou tři:

1. Hodnota 0 (0%) je spodní krajní mez intervalu, vyjadřuje *absolutní nedůvěru* v kontext.

2. Hodnota 0.5 (50%) je střed intervalu, vyjadřuje *neutrální* postoj nebo také *nerozhodnost*: kontext není důvěryhodný ani nedůvěryhodný.

3. Hodnota 1 (100%) je horní krajní mez intervalu, vyjadřuje *absolutní* nebo také *maximální důvěru* v kontext.

Tento způsob reprezentace důvěry však vyžaduje exaktní znalost hodnoty důvěry v daný kontext a proto není zcela vhodný k reprezentaci *částečné důvěry*, která vzniká při odvozování na základě podobností (vztahů) mezi kontexty. Jinak řečeno, pakliže hodnotu důvěry v nějaký kontext odvodíme na základě vztahu mezi ostatními kontexty (kde pro některé z nich nemusí být hodnota důvěry známá) je třeba do tohoto odvození zahrnout jakousi *míru neurčitosti*, která je dána způsobem kompozice z více kontextů. Právě z tohoto důvodu definujeme pojem *rozšířený interval důvěry*, který umožňuje odvodit hodnotu důvěry v kontext na základě vztahů s ostatními kontexty, pro které nemusí být hodnota důvěry známá – tedy jsme schopni do odvozovacího procesu důvěry zahrnout neurčitost nebo lépe řečeno *neznalost*. Způsob, jakým je tato neznalost zahrnuta do výpočetního modelu odvozovaní důvěry, bude popsán v následujících podkapitolách. Zde jen předešleme, že základní princip vychází z postupu stanovení mezních hodnoty důvěry které mohou nastat, tedy jakýsi nejlepší a nejhorší možný případ ohodnocení.

Rozšířený interval důvěry tak reprezentuje minimální a maximální hodnotu důvěry, kterou může kontext nabývat. Dle definice uvedené výše (vzorec (4.3)) je rozšířený interval důvěry definován jako podmnožina množiny reálných čísel ohraničenými mezemi z množiny základního intervalu důvěry a je označen symbolem τ. Prvek ϑ z množiny τ zapisujeme jako interval s minimální a maximální hodnotou (jinak řečeno meze intervalu) takto: $\vartheta = [\vartheta_{min}, \vartheta_{max}]$, kde ϑ_{min} označuje minimální hodnotu důvěry a ϑ_{max} označuje maximální hodnotu důvěry, přičemž musí platit podmínka, že $\vartheta_{min} \leq \vartheta_{max}$ a $\vartheta_{min}, \vartheta_{max} \in [0, 1]$. Jednotlivé limity tedy reprezentují rozsah hodnot důvěry, která může být na kontext mapována. Jinak řečeno, daný kontext je důvěryhodný minimálně ϑ_{min} a maximálně však ϑ_{max}. Šíře tohoto intervalu konkrétního kontextu je stanovena na základě vztahů k ostatním kontextům a šíří jejich intervalu důvěry (mírou přesnosti znalostí důvěry). Základní myšlenka navrženého HMTC modelu při použití této intervalové reprezentace důvěry je taková, že při postupném získávání znalostí (přímé i nepřímé, jež s ohledem na předchozí znalosti transformujeme na hodnotu důvěry) se šíře tohoto intervalu redukuje, meze intervalu se přibližují a konvergují k jedné konkrétní hodnotě. Tento princip bude blíže popsán v podkapitole 4.3 *Události a aktualizace důvěry*.

Existují dvě význačné hodnoty této reprezentace, které mají v našem modelu speciální význam:

1. Hodnota $[0, 1]$, tedy $\vartheta_{min} = 0$ a $\vartheta_{max} = 1$: tato hodnota reprezentuje absolutní neznalost hodnoty důvěry pro daný kontext. Tato hodnota je použita při inicializaci modelu pro nastavení prvotní hodnoty důvěry kontextu.

2. $\emptyset$, tedy prázdná množina: tato hodnota je v našem modelu použita pro reprezentaci výjimečného stavu kontextu. Tento stav kontextu označujeme jako *konflikt*. Konfliktu a způsobu jeho využití bude věnován prostor v podkapitole 4.3.7.

Speciálním případem je rovněž situace, kdy $\vartheta_{min} = \vartheta_{max}$, tedy např. interval $[0, 0]$ který odpovídá pouze hodnotě 0 (absolutní nedůvěra bez neurčitosti). Takovýto případ je zcela legální a pracuje s tímto intervalem shodně, jako v případě kdy $\vartheta_{min} \neq \vartheta_{max}$.

Neurčitost

Jak již bylo řečeno výše, *rozšířený interval důvěry* (dále už jen *interval důvěry*) rovněž vyjadřuje míru neurčitosti (anglicky *uncertainty*). Tato míra neurčitosti je dána šíří intervalu důvěry a formálně ji definujeme jako rozdíl maximální a minimální složky intervalu důvěry takto:

$$uncert(\vartheta) = \vartheta_{max} - \vartheta_{min} \ , \tag{4.6}$$

kde $\vartheta \in \tau$. Přesný význam neurčitosti v intervalu důvěry bude popsán v kapitole 5 zabývající se konkrétním návrhem vyhodnocování důvěry pro HMTC model.

4.2.2 Operace s rozšířeným intervalem důvěry

Z důvodu potřeby zavedení výpočetního modelu odvozování důvěry v kontexty s použitím rozšířeného intervalu důvěry jsme zavedli vlastní význam pro některé operace nad množinou τ. Jedná se konkrétně o operace: *násobení*, *dělení*, *sčítání*, *odčítání* a *průnik*. Pro libovolné $\vartheta, \gamma \in \tau$ a pro libovolnou konstantu $k \in \mathbb{R}$ definujeme následující operace:

Násobení konstantou

$$\begin{aligned} \vartheta \cdot k = & \ [\vartheta_{min}, \vartheta_{max}] \cdot k \\ & = norm\ [\vartheta_{min} \cdot k, \vartheta_{max} \cdot k] \end{aligned} \tag{4.7}$$

Dělení konstantou

$$\frac{\vartheta}{k} = norm \left[\frac{\vartheta_{min}}{k}, \frac{\vartheta_{max}}{k} \right] \tag{4.8}$$

Sčítání

$$\vartheta + \gamma = norm\ [\vartheta_{min} + \gamma_{min}, \vartheta_{max} + \gamma_{max}] \tag{4.9}$$

Odčítání

$$\vartheta - \gamma = norm\ [\vartheta_{min} - \gamma_{max}, \vartheta_{max} - \gamma_{min}] \tag{4.10}$$

Průnik

$$\vartheta \cap \gamma = \{x : x \in \vartheta \wedge x \in \gamma\} \tag{4.11}$$

Jak vidno, sčítání a odčítání vychází z klasické intervalové aritmetiky [Wik10a] a operace průnik je definován v souladu s klasickou definicí známou z obecné teorie množin.

Funkce *norm* aplikovaná na interval zajišťuje, že výsledek je vždy normalizován do množiny τ, tedy že meze intervalu nikdy nepadnou mimo interval $[0, 1]$. Tuto funkci lze definovat například takto:

$$\begin{aligned} norm\ [x, y] = [min(max(0, x), 1), min(max(y, 0), 1)] \\ \text{pro } x \leq y \text{ a } x, y \in \mathbb{R} \ . \end{aligned} \tag{4.12}$$

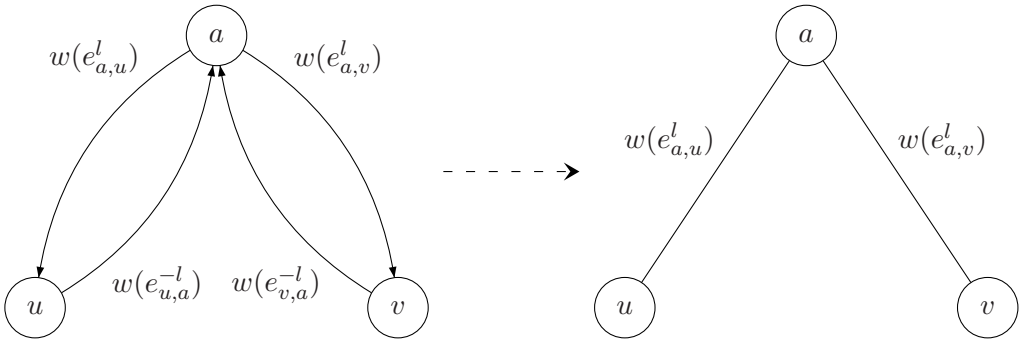
4.2.3 Funkce váhy hrany

Funkce váhy hrany – nebo také *váhová funkce* – je definována dle definice 4.2 a konkrétně dle vzorce (4.4) jako zobrazení hrany grafu na interval $[0, 1]$ z oboru $\mathbb{R}$. Toto zobrazení ve skutečnosti reprezentuje v našem modelu sílu vztahu mezi jednotlivými kontexty (uzly grafu). Váhy pro jednotlivé hrany mezi kontexty jsou stanoveny empiricky na základě znalostí a požadavků na modelovaný prvek. Naše definice modelu HMTC tedy nezahrnuje způsob jakým jsou jednotlivé hrany ohodnoceny, ale definuje pouze jisté omezující podmínky tak, aby bylo možné hodnotu váhy použít při odvozování důvěry mezi kontexty. Primární omezující podmínka pro stanovení hodnoty vah hran je popsána rovněž v definici 4.2 pomocí vzorce (4.5), který stanovuje maximální hodnotu součtu vah hran vstupující do neterminálního uzlu z uzlů v nižší vrstvě.

Vzhledem k této podmínce je třeba dodat, že námi definovaný graf je *orientovaný obousměrný graf*, což znamená že každá hrana má stanovenu orientaci a pro každou hranu $e^l_{u,v} = (u^l, v^{l+1})$ existuje i hrana $e^{-l}_{v,u} = (v^{l+1}, u^l)$ s opačnou orientací, kterou označujeme záporným znaménkem u hodnoty vrstvy. Pro úplnost definice váhové funkce pak musíme uvést, že váha hrany v jednom směru je rovna váze hrany ve směru opačném:

$$\forall e^l_{u,v} : w(e^l_{u,v}) = w(e^{-l}_{v,u}) \tag{4.13}$$

a z tohoto důvodu používáme při grafické reprezentaci těchto hran grafu zjednodušené zobrazení, kde obě hrany jsou reprezentovány hranou jednou (ve směru z vyšší vrstvy do vrstvy nižší) a není zde označena orientace. Toto zjednodušení je naznačeno na obrázku 4.3 (vlevo úplné značení, vpravo ekvivalentní zjednodušené značení).



Obrázek 4.3: Způsob zjednodušení značení hran.

4.2.4 Funkce důvěry a odvozování důvěry mezi kontexty

Funkce důvěry je zobrazením uzlu grafu v čase na interval důvěry (definice 4.2, vzorec (4.2)), tedy pro každý uzel grafu (kontext) jsme schopni pro daný časový okamžik z množiny T určit hodnotu důvěry (interval důvěry). Důvěru pro uzel u v čase t zapisujeme jako $\rho(u, t)$. Implicitní inicializace modelu je provedena tak, že v čase t_0 (počáteční čas – první prvek množiny T) je pro každý uzel nastavena hodnota důvěry na interval $[0, 1]$, což odpovídá absolutní neznalosti důvěry v daný kontext.

$$\forall n \in N : \rho(n, t_0) = [0, 1] \tag{4.14}$$

V následujících podkapitolách bude formálně popsán způsob, jak je důvěra pro jednotlivé uzly vzhledem k jejich poloze v grafu odvozována. Základní myšlenka výpočtu důvěry pro jednotlivé uzly spočívá v tom, že jednotlivé uzly jsou na sobě závislé úměrně vahou propojené hrany. Tedy, důvěra pro konkrétní uzel je určena na základě znalosti váhy hrany a důvěry uzlů s ním spojených. Vzhledem k tomu, že důvěru uzlu ovlivňují uzly jak z vyšší vrstvy tak i uzly z nižší vrstvy (vyjma uzlu *kořenového*, který nemá žádné naduzly a vyjma uzlů *terminálních*, které naopak nemají žádné poduzly), zavádíme v našem modelu dva typy odvození: odvození *směrem nahoru* a odvození *směrem dolů*.

Odvození směrem nahoru

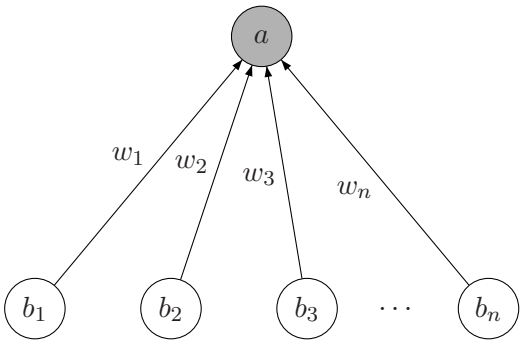
Odvození důvěry *směrem nahoru* lze použít pro všechny neterminální uzly, tedy uzly, které mají alespoň jeden poduzel v nižší vrstvě. Hodnota odvozené důvěry uzlu se určí jako suma násobku důvěry poduzlů a jim odpovídajících vah propojených hran. Pro tento výpočet zavedeme funkci *up*, kterou definujeme formálně takto:

$$up : N \times T \to \tau \tag{4.15}$$

a její vyčíslení se provádí podle následujícího předpisu takto:

$$up(a^l, t) = \sum_{\forall e^l_{a,b}, \forall b^{l+1}} w(e^l_{a,b}) \cdot \rho(b^{l+1}, t) \ , \tag{4.16}$$

kde (dle zavedené notace) a^l označuje uzel a ve vrstvě l a $e_{a,b}$ označuje hranu mezi uzly a a b. Operátor „$\cdot$“ odpovídá operaci *násobení konstantou* dle vzorce (4.7). Ukázka odvození důvěry směrem nahoru je zobrazena na obrázku 4.4. Uzel a je uzlem, pro který se odvození směrem nahoru provádí, uzly b_1 až b_n jsou propojené poduzly z nižší vrstvy.



Obrázek 4.4: Příklad odvození důvěry směrem nahoru.

Směr šipek na obrázku 4.4 pouze ilustruje směr odvození nikoliv orientaci hrany. Konkrétní výpočet hodnoty funkce *up* pro uzel a dle příkladu na obrázku 4.4 by pak vypadal takto:

$$up(a, t) = \sum_{i=1}^{n} w_i \cdot \rho(b_i, t) \ . \tag{4.17}$$

Odvození směrem dolů

Odvození důvěry *směrem dolů* lze použít pro všechny uzly, které mají alespoň jeden naduzel ve vyšší vrstvě – tedy vyjma uzlu kořenového. Hodnota odvozené důvěry směrem dolů je určena na základě hodnoty důvěry všech propojeným naduzlům v kombinaci s jejich přímými propojenými poduzly. Do odvození směrem dolů tedy vstupují nejen uzly z vyšší vrstvy pro uzel odvozovaný, ale i uzly ze stejné vrstvy – tedy sousedé, kteří jsou propojeni přes naduzel. Funkci pro odvození směrem nahoru označujeme *down* a má analogický předpis jako funkce *up*:

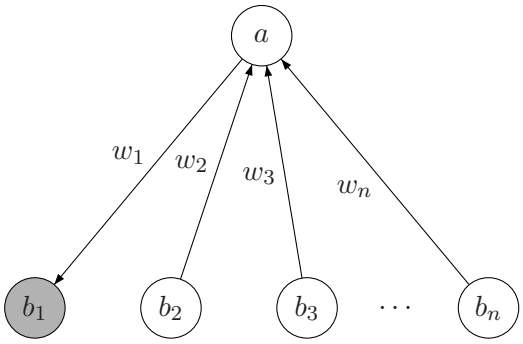
$$down : N \times T \to \tau \tag{4.18}$$

a její vyčíslení se provádí dle vzorce:

$$down(a^l, t) = \bigcap_{\forall b^{l-1}, \forall e^{l-1}_{b,a}} \frac{\rho(b^{l-1}, t) - \sum\limits_{\forall c^l, \forall e^{l-1}_{b,c}} w(e^{l-1}_{b,c}) \cdot \rho(c^l, t)}{w(e^{l-1}_{b,a})} . \tag{4.19}$$

Vyčíslení funkce pro odvození důvěry směrem dolů je složitější než odvození důvěry směrem nahoru. Tento závěr plyne z faktu, že do výpočtu vstupují nejen naduzly, ale i sousední uzly. Navíc je zde třeba zohlednit případ, kdy uzel pro který odvození provádíme má dva nebo více propojených naduzlů. Poté do výpočtu vstupují všechny poduzly těchto naduzlů.

První jednodušší příklad výpočtu směrem dolů je uveden na obrázku 4.5. V tomto příkladě se snažíme odvodit důvěru v uzel b_1, který má pouze jeden propojený naduzel a. Směr šipek opět naznačuje jakýsi pomyslný směr výpočtu, kde nejprve provádíme výpočet sumy součinů všech sousedních uzlů spojený přes naduzel a a následně ze znalosti této sumy postupujeme směrem dolů od uzlu a k uzlu b_1.



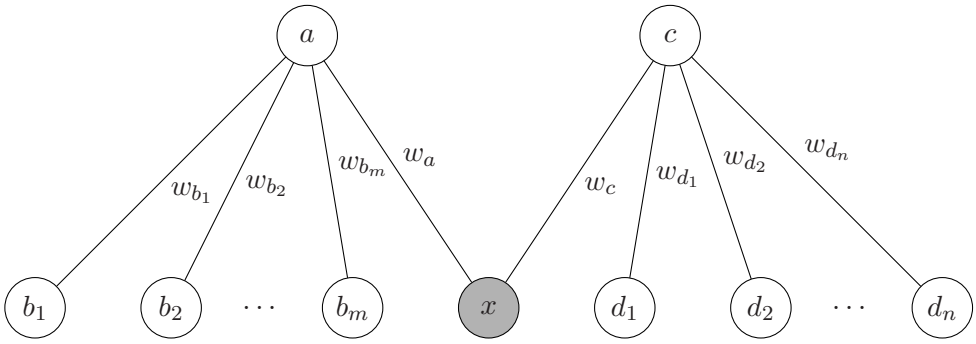
Obrázek 4.5: Příklad odvození důvěry směrem dolů (1).

Pro tento konkrétní příklad grafu na obrázku 4.5 pak uvádíme rovnici výpočtu (4.20), která odvozuje hodnotu důvěry uzlu b_1 směrem dolů.

$$down(b_1, t) = \frac{\rho(a, t) - \sum\limits_{i=2}^{n} w_i \cdot \rho(b_i, t)}{w_1} . \tag{4.20}$$

Druhý příklad výpočtu směrem dolů je komplikovanější v tom, že uzel x pro který provádíme odvození důvěry má dva propojené naduzly a a c. Tento příklad je naznačen

na obrázku 4.6. Odvození hodnoty důvěry uzlu x si lze pro postupné zpracování dekomponovat tak, že nejprve provedeme odvození důvěry směrem dolů přes naduzel a a jeho poduzly $b_1 \dots b_n$. Analogicky provedeme odvození důvěry směrem dolů přes naduzel c a oba výsledky zkombinujeme pomocí operátoru průniku (vzhledem k tomu, že operace průniku je komutativní, není důležité pořadí provádění jednotlivých odvození přes naduzly). Tímto postupem získáme výslednou hodnotu důvěry pro uzel x odvozený směrem dolů v případě, že má více propojených naduzlů.



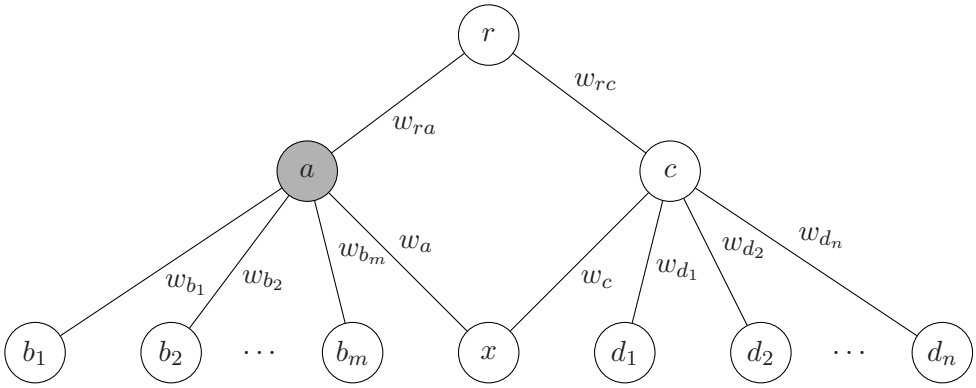
Obrázek 4.6: Příklad odvození důvěry směrem dolů (2).

Pro příklad popsaný výše a uvedený na obrázku 4.6 tedy můžeme použít rovnici (4.21), která demonstruje, jakým způsobem lze odvozování důvěry směrem dolů rozdělit na jednotlivé dílčí výsledky, které následně propojíme pomocí operátoru průniku.

$$down(x,t) = \frac{\rho(a,t) - \sum_{i=1}^{m} w_{b_i} \cdot \rho(b_i,t)}{w_a} \bigcap \frac{\rho(c,t) - \sum_{j=1}^{n} w_{d_j} \cdot \rho(d_j,t)}{w_c} . \tag{4.21}$$

Kombinace obou směrů odvození

Při odvozování důvěry v jednotlivé uzly rozeznáváme tři typy uzlů podle jejich pozice v grafu. Prvním typem je uzel kořenový, pro který je relevantní pouze odvození směrem nahoru, neboť nemá žádné naduzly. Druhým typem jsou pak uzly terminální, pro které se vždy provádí pouze odvození směrem dolů, neboť tyto uzly nemají žádné připojené poduzly. Posledním typem jsou pak uzly umístěné mezi dvěma vrstvami a propojené jak směrem nahoru na naduzly, tak i směrem dolů na poduzly. Pro tyto uzly je při odvozování důvěry třeba provést odvození důvěry jak směrem nahoru, tak i směrem dolů. Celé odvození důvěry pro tyto uzly se tak rozpadá na dva dílčí kroky, které generují dílčí výsledky. Ty se kombinují, podobně jak je tomu v případě odvození směrem dolů, pomocí operace průniku.



Obrázek 4.7: Příklad odvození důvěry v obou směrech.

$$up(a,t) = [w_a \cdot \rho(x,t)] + \sum_{i=1}^{m} w_i \cdot \rho(b_i,t) \tag{4.22}$$

$$down(a,t) = \frac{\rho(r,t) - [w_{rc} \cdot \rho(c,t)]}{w_{ra}} \tag{4.23}$$

$$\rho'(a,t) = up(a,t) \cap down(a,t) \tag{4.24}$$

Příklad kombinace odvození směrem nahoru i dolů je uveden na obrázku 4.7. Pro tento konkrétní případ provádíme odvození důvěry pro uzel a směrem nahoru pomocí poduzlů $b_1 \ldots b_m$ a x. Dále pak odvození směrem dolů přes naduzel r a sousední uzel (uzlu a) c. Postup konkrétního výpočtu je nastíněn v rovnicích (4.22), (4.23) a (4.24). Konečný výsledek odvození důvěry uzlu je označen jako $\rho'(a,t)$ – nejedná se tedy o hodnotu důvěry uzlu (ta musí být stanovena až na základě průniku s předchozí hodnotou, což bude popsáno v následující kapitole), ale o hodnotu důvěry odvozenou pro daný a uzel a čas t.

Obecná funkce odvození důvěry

Na základě výše uvedených funkcí pro odvození důvěry směrem nahoru a směrem dolů a na základě demonstrovaných příkladů můžeme na tomto místě stanovit obecnou funkci odvození pro libovolný uzel $a \in N$ ve vrstvě l. Tuto funkci nazveme *eval* a definujeme ji takto:

$$eval(a^l,t) = \begin{cases} up(a^l,t) & \text{pro kořenový uzel } (l=1) \\ down(a^l,t) & \text{pro terminální uzly} \\ up(a^l,t) \cap down(a^l,t) & \text{ostatní uzly} \end{cases} \tag{4.25}$$

4.3 Události a aktualizace důvěry

V předchozí kapitole 4.2 a jejích následných podkapitolách jsme položili základní formální definice nezbytné pro stanovení struktury námi navrhovaného hierarchického modelu důvěry s kontexty. Dále jsme podrobně vysvětlili a formálně definovali způsob reprezentace důvěry pro jednotlivé kontexty a popsali matematický aparát na jehož základě může být důvěra v nějaký kontext odvozena ze znalosti důvěry v kontexty s ním propojené.

Výsledná hodnota důvěry kontextu však není obecně rovna hodnotě vypočtené pomocí odvození, tedy pomocí funkce *eval* (vzorec (4.25)). To především z toho důvodu, že bychom ignorovali hodnotu důvěry kontextu, kterou měl kontext před odvozením. Naším cílem je však hodnotu důvěry v kontexty postupně zpřesňovat tím, že získáváme znalosti o důvěře v jednotlivé kontexty modelu. V takovém případě je pro nás hodnota důvěry kontextu před změnou důležitá neboť zahrnuje *historii* změn hodnoty důvěry kontextu.

Tato kapitola se tedy zabývá způsobem, jak lze provádět změny hodnoty důvěry v libovolný kontext modelu a jak tato změna ovlivní hodnoty kontextů okolních, s ním spojených. Pro bližší popsání způsobu aktualizace důvěry si nejdříve neformálně popíšeme princip fungování agenta, který využívá HMTC při rozhodování.

4.3.1 Konceptuální princip využití HMTC modelu pro podporu rozhodování agenta

Námi definovaný model HMTC reprezentuje část báze představ (*Belief Base*) agenta vzhledem k jedné entitě. V multi-agentním systému je však často třeba komunikovat mezi mnoha ostatními agenty a proto je třeba explicitně zdůraznit, že pro každou další entitu si agent vytváří vlastní další HMTC instanci modelu, přičemž jednotlivé instance se mohou lišit – grafická reprezentace jednotlivých kontextů pro různé entity může být rozdílná. Pro popis našeho návrhu se však zaměříme výhradně na případ, kdy jednotlivé instance HMTC modelů jednoho agenta mají shodnou strukturu. Každý agent může obsahovat libovolný počet instancí HMTC modelu, kde každá z nich je přiřazena právě jedné entitě (objektu) systému.

Základní koncept agenta využívající HMTC model při rozhodování je založen na rozšíření některé ze stávající agentní architektury o HMTC model a komponentu zajišťující řízení funkce HMTC modelu. Touto komponentou je prvek, který transformuje a směřuje *události*, jež mají vliv na vytváření a využití důvěry. Námi navrhovaný HMTC model je tak řízen událostmi, přičemž rozeznáváme dva typy těchto událostí: *externí události* a *interní události* (budou detailně popsány dále). Důležitým faktem je, že model HMTC ve skutečnosti neřeší vlastní rozhodování agenta (to je ponecháno na zvolené agentní architektuře a konkrétní implementaci), ale jeho výsledky nebo lepé řečeno znalosti agenta založené na jeho výsledcích jsou pouze jedním z kritérií, které vstupují do rozhodovacího a plánovacího algoritmu agenta. Detailní návrh prototypu agenta a agentního systému využívající HMTC model je popsán v kapitole 6.

4.3.2 Událost

Na základě výše popsaného principu modelu řízeného událostmi se teď pokusíme formálněji definovat *událost* ve vztahu k HMTC, nejprve si však blíže popíšeme jednotlivé typy událostí.

Externí událost je taková událost, která je vnímána agentem z prostředí. Může se jednat například o zprávu, která nese nějakou znalost anebo třeba nějaký podmět ze senzorů agenta (interakce, sledování, ...) – obecně tedy vjem z okolního prostředí agenta. Pro náš další popis budeme chápat externí událost konkrétně jako hodnotu doporučení v nějaký kontext jiného agenta nebo hodnotu důvěry vzhledem k ostatním agentům a jejich kontextům stanovená na základě přímého vjemu agenta (výsledek přímé interakce, sledování schopností, apod.).

Interní událost je pak taková událost, která je generována na základě vnitřního procesu agenta. To může být například aktualizace báze znalostí, vygenerování nebo změna

plánu apod. V této fázi popisu nás však pouze zajímá taková interní událost, která je generována na základě změny hodnoty znalosti důvěry v nějaký kontext, která je provedena na základě externí události. Takováto interní událost, podobně jako externí událost, která aktualizuje hodnotu důvěry kontextu vzhledem k nějaké jiné entitě v systému, musí být zpracována modelem HMTC.

Pakliže předpokládáme, že konkrétní instance HMTC modelu je vždy přiřazena právě jedné entitě systému, pak není nutné uvádět přiřazení události vzhledem entitě systému v případě, že události přiřadíme jednoznačný uzel (kontexty) konkrétní instance HMTC modelu (existuje zde tedy jednoznačná vazba $[entita] \rightarrow [instance\ HMTC] \rightarrow [uzel\ HMTC]$). Událost označíme písmenem u a definujeme ji jako dvojici uzel (kontext) a hodnotu důvěry takto:

$$u = (n, \delta) \tag{4.26}$$

kde $u \in U$, U označuje množinu všech událostí, uzel $n \in N$ reprezentuje kontext dle definice HMTC (viz definice 4.2) a δ je hodnota důvěry rozšířeného intervalu důvěry: $\delta \in \tau$. Množina U je množina všech událostí, jak externích tak interních. Můžeme si tedy označit množinu všech externích událostí jako U_e a množinu všech interních událostí jako U_i, pak $U = U_e \cup U_i$.

Událost v našem modelu reprezentuje změnu hodnoty důvěry konkrétního kontextu, přičemž hodnota δ reprezentuje nově získanou znalost o důvěře. Jak již bylo řečeno v úvodu této kapitoly, tak při změně hodnoty kontextu je třeba brát vždy v potaz jeho aktuální hodnotu důvěry, která reflektuje jakousi historii vývoje důvěry v tomto kontextu. Aplikace události na kontext tedy není provedena tak, že se původní hodnota důvěry přepíše nově získanou hodnotou důvěry události, ale provede se jejich kombinace pomocí operace průniku. Definujme si způsob aplikace konkrétní události $u_i = (n, \delta_i)$ a pro aktuální (čas t_i) hodnotu důvěry uzlu n takto:

$$\rho(n, t_i) = \rho(n, t_{i-1}) \cap \delta_i\ . \tag{4.27}$$

Tento způsob výpočtu hodnoty nové důvěry kontextu na základě původní hodnoty důvěry a hodnoty důvěry získané z události respektuje základní myšlenku našeho modelu, že hodnota důvěry je na základě každé další nové znalosti zpřesňována. To vyplývá z podstaty operace průniku aplikovaného na interval, kde výsledek operace je vždy interval menší nebo rovný než jeho operandy. Výjimkou je výsledek, který je roven prázdné množině. Takový stav po aplikaci události na libovolný kontext nazýváme *konflikt* a jeho popisem se budeme zabývat v podkapitole 4.3.7.

4.3.3 Externí událost a její šíření

Jak již bylo řečeno výše, externí událost je pro náš model chápána buď jako hodnota *doporučení* v nějaký kontext jiného agenta nebo jako hodnota důvěry vzhledem k ostatním agentům a jejich kontextům stanovená na základě přímého vjemu agenta – výsledek přímé interakce, sledování schopností, apod. Pakliže aplikujeme externí událost a aktualizujeme hodnotu důvěry nějakého kontextu (aplikace události, vzorec (4.27)) získáváme novou znalost do báze znalostí a na jejím základě provádíme odvození důvěry pro okolní kontexty, které jsou s aktualizovaným kontextem spojeny. To ve skutečnosti znamená, že na základě externí události generujeme další interní událost nebo události, které mají za úkol, na základě nové znalosti v bázi znalostí, provést aktualizaci hodnoty důvěry pro kontexty,

které mají na externě aktualizovaný kontext vazbu. Tento způsob šíření události si detailně ukážeme v následujících podkapitolách pomocí *algoritmu pro šíření události*.

Pro popis tohoto algoritmu však nejprve potřebujeme definovat jak hodnotu nově získané znalosti o důvěře δ externí události, tak i odpovídají hodnotu důvěry pro nově generované interní události. Vzhledem k tomu, že v této kapitole je naším cílem popsat způsob šíření externí události, nebudeme se zde zabývat popisem jejího vzniku a způsobem, jak byla agentem interpretována. To bude detailně popsáno v kapitole 5. Proto si definujeme získání hodnoty δ externí události jako abstraktní funkci *external*, která pro daný uzel a čas stanovuje hodnotu důvěry z rozšířeného intervalu důvěry:

$$external : N \times T \rightarrow \tau \ . \tag{4.28}$$

Máme-li takto stanovenu hodnotu externí události nad kontextem a:

$$\begin{gathered} u_e = (a, \delta_e) \\ \text{pak } \delta_e = external(a, t_e) \end{gathered} \tag{4.29}$$

je třeba stanovit i hodnotu interní události. Tu můžeme stanovit jako hodnotu vypočtenou pomocí funkce *eval*, která definuje obecné odvození důvěry pro libovolný kontext modelu a je definována pomocí vzorce (4.25). Na základě těchto skutečností tedy definujeme obecně hodnotu nově získané důvěry δ pro libovolnou událost (externí i interní) nad kontextem a takto:

$$\delta = \begin{cases} external(a, t) & \text{pro externí událost} \\ eval(a, t) & \text{pro interní událost} \end{cases} \tag{4.30}$$

4.3.4 Algoritmus šíření události

Algoritmus šíření události definuje způsob, jak jsou jednotlivé hodnoty důvěry v kontexty aktualizovány a odvozovány na základě příchodu externí události. Jedná se tedy o způsob generování a zpracování interních událostí. Základní myšlenka algoritmu je založena na principu aktualizace všech kontextů, které jsou přímo propojeny s kontextem, který je ovlivněn externí událostí. Pakliže je nějaký z nadkontextů nebo podkontextů (pojem *nadkontext* a *podkontext* je zde použit ve smyslu pojmů *naduzel* a *poduzel*, které byly používány při definici grafu pro HMTC – jedná se tedy o kontexty propojené na vyšší na nižší vrstvě) aktualizován, tedy jeho hodnota důvěry je změněna na základě *odvození důvěry*, je třeba rovněž algoritmus šíření důvěry aplikovat i pro jeho nadkontexty a podkontexty (vyjma kontextů, které aktualizaci vyvolaly).

Algoritmus šíření události (Algoritmus 1), pro který jsme zavedli zkratku ABE (z anglického *Algorithm of Bubbling Events*), je demonstrován pro příchozí externí události u_e pro kontext a v čase t_i. Pro funkci algoritmu je použita datová struktura fronta, která je pojmenována *Open*. Tato fronta *Open* obsahuje v průběhu algoritmu posloupnost kontextů určený pro případnou aktualizaci pomocí odvození důvěry funkcí *eval*. Vzhledem k tomu, že používáme datovou strukturu fronta, definujeme si pro demonstraci našeho algoritmu pomocné funkce pro práci s touto strukturou. Jednotlivé funkce a jejich význam včetně případných argumentů a návratové hodnoty jsou popsány v tabulce 4.1.

Vlastní algoritmus začíná aplikací externí události u_e, kdy je na základě rovnice (4.27) aplikována hodnota důvěry externí události δ_e na kontext a a jeho nová hodnota důvěry je tedy vypočtena pomocí průniku původní hodnoty důvěry s hodnotou důvěry události.

Zápis funkce	Význam a popis
$res \leftarrow$ `pop`(arg1)	vyjme prvek ze začátku fronty $arg1$ a přiřadí ho do res
`push`($arg1$, $arg2$)	vloží prvek $arg2$ na konec fronty $agr1$
`pushAll`($arg1$, $arg2$)	vloží uspořádanou množinu $arg2$ na konec fronty $arg1$
$res \leftarrow$ `getParents`($arg1$)	do množiny res vloží všechny propojené nadkontexty $arg1$
$res \leftarrow$ `getChilds`($arg1$)	do množiny res vloží všechny propojené podkontexty $arg1$

Tabulka 4.1: Význam a popis funkcí použitých v algoritmu šíření důvěry.

Spuštění hlavního bloku algoritmu má smysl pouze tehdy, změnila-li se hodnota důvěry kontextu a na základě aplikace externí události u_e. V případě, že se hodnota důvěry změnila, provedeme inicializaci fronty *Open* tak, že do ní vložíme všechny propojené nadkontexty a podkontexty kontextu a. Následně postupně vybíráme z fronty *Open* jednotlivé kontexty, přiřazujeme je do proměnné *node* a pro každý tento kontext provedeme odvození hodnoty důvěry aplikací funkce *eval*. Zároveň vytvoříme novou interní událost u_i pro tento kontext *node* s hodnotou získané důvěry δ_i odvozené právě pomocí *eval*. Následně provedeme aktualizaci důvěry tohoto kontextu *node* a v případě, že se hodnota důvěry změní, přiřadíme všechny jedno nadkontexty a podkontexty do fronty *Open* – vyjma kontextů, které již ve frontě *Open* jsou a vyjma kontextu a, který byl aktualizován jako prvotní na základě externí události. Algoritmus končí v momentě, kdy je fronta *Open* prázdná a tedy neexistuje žádný další kontext, pro který by výsledek odvození důvěry spolu s aktuální hodnotou důvěry změnil jeho hodnotu.

Výsledkem ABE je množina interních událostí U_i, která obsahuje všechny interní události, které byly generovány na základě prvotní externí události u_e a způsobily aktualizaci důvěry v daném kontextu. Pro každou událost z množiny U_i můžeme určit jinou událost z $U = U_i \cup u_e$, která ji způsobila. Říkáme tedy, že každá interní událost je vždy generována jinou interní nebo externí událostí. Na základě toho lze jednotlivé události uspořádat do posloupností, přičemž tyto posloupnosti budeme označovat jako *cesty*. Tímto konceptem se budeme zabývat v následující podkapitole.

Dalším důležitým aspektem ABE je to, že pro jeden konkrétní kontext, tedy pro jednu konkrétní externí událost v rámci jedné instance HMTC musí být jeho provedení atomické. To především z toho důvodu, že více souběžných externích událostí v rámci jedné instance HMTC by mělo za následek spuštění více souběžných procesů šíření události nad jednou bází znalostí, čímž by mohlo dojít ke konfliktům a opakovaným přepočítáváním hodnot kontextů, přičemž výsledné hodnoty důvěry dotčených kontextů by mohly nabývat chybných hodnot. Z toho důvodu je tedy vyžadována atomicita procesu šíření události nad jednou instancí HMTC.

Atomicita ABE též úzce souvisí s naší definicí časové množiny T (viz definice 4.2, bod 3), kdy každé externí události přiřazujeme právě jeden časový okamžik z množiny T. Všechny interní události vygenerované v závislosti na této externí události pomocí ABE jsou provedeny právě v tomto časovém okamžiku. Každé externí události tak lze jednoznačně přiřadit časový okamžik z množiny T, jehož index je o jedna větší nežli index posledního časového okamžiku poslední externí události dané instance HMTC.

Algoritmus 1: Algoritmus šíření události – ABE

Vstup: externí událost: $u_e = (\delta_e, t_i)$ nad uzlem a
Výstup: množina interních událostí generovaná na základě u_e: U_i

```
1  begin
2      ρ(a, t_i) ← ρ(a, t_{i−1}) ∩ δ_e
3      if ρ(a, t_i) ≠ ρ(a, t_{i−1}) then
4          Open ← ∅
5          pushAll (Open, getParents (a))
6          pushAll (Open, getChilds (a))
7          while Open ≠ ∅ do
8              node ← pop (Open)
9              δ_i ← eval(node, t_i)
10             u_i ← (node, δ_i)
11             ρ(node, t_i) ← ρ(node, t_{i−1}) ∩ δ_i
12             if ρ(node, t_i) ≠ ρ(node, t_{i−1}) then
13                 U_i ← U_i ∪ {u_i}
14                 Parents ← getParents (node)
15                 while Parents ≠ ∅ do
16                     p ← pop (Parents)
17                     if p ≠ a and p ∉ Open then
18                         push (Open, p)
19                 Children ← ∅
20                 Children ← getChilds (c)
21                 while Children ≠ ∅ do
22                     ch ← pop (Children)
23                     if ch ≠ a and ch ∉ Open then
24                         push (Open, ch)
```

4.3.5 Rozšířená definice HMTC o komponentu událostí

Na základě zavedení množiny událostí U je tak třeba rozšířit definici základního HMTC modelu na „rozšíření HTMC“ model, který budeme označovat $HMTC^2$. Původní pětici tak rozšíříme na šestici následovně:

$$HMTC^2 = (N, E, T, U, \rho, w) \tag{4.31}$$

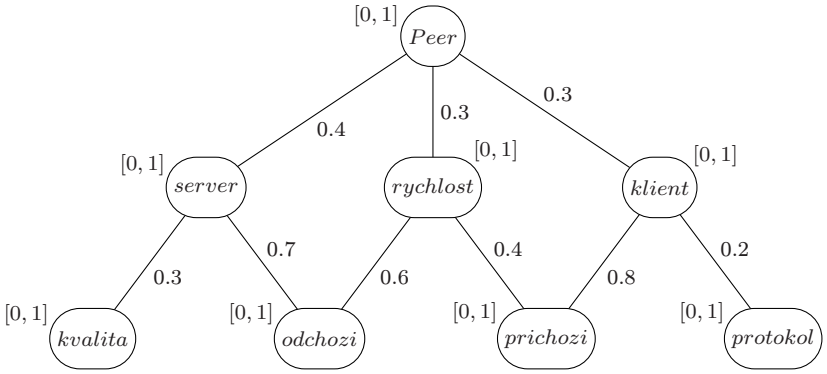
kde U označuje množinu všech aplikovaných událostí (interních i externích), přičemž ostatní komponenty zůstávají beze změny dle definice 4.2.

4.3.6 Příklad aplikace algoritmu šíření události

Pro vytvoření konkrétní představy toho, jak vlastní algoritmus šíření události funguje a jak lze na základě jedné znalosti důvěry v nějaký kontext odvodit znalost důvěry v kontext jiný, si aplikaci algoritmu ukážeme na konkrétním příkladě.

Příklad 1. Jako konkrétní příklad HMTC modelu jsme vytvořili grafovou reprezentaci kontextů jednoho uzlu P2P sítě, příklad takto navrženého modelu je na obrázku 4.8. Každého klienta P2P sítě lze posuzovat ze dvou základních aspektů: 1) jako poskytovatele dat, který sdílí data v síti – označení konkrétního kontextu v modelu jako *server*; 2) jako klienta, který data ze sítě stahuje – označení kontextu *klient*.

Pro důvěryhodnost poskytovatele dat (server) je důležité, zda poskytuje kvalitní obsah (např. poskytované soubory nejsou zavirované, atd.) a jak rychle je schopen tento obsah poskytovat do sítě – tedy významně záleží na rychlosti odesílání dat kterou disponuje. Na základě tohoto stanovíme dva terminální kontexty *kvalita* a *odchozí* (ve smyslu rychlosti odesílání dat), které definují nadkontext *server*. V případě kontextu klienta lze označit za důležité aspekty rychlost *příchozí* linky (rychlost stahování) a typ klienta, respektive typ *protokolu*, který pro připojení k ostatním uživatelům používá (např. zda umožňuje šifrovanou komunikaci apod.). Takto vytvoříme další dva terminální kontexty grafu, které definují kontext *klient*. Jak je vidět, tak kontexty *příchozí* a *odchozí* lze generalizovat a vytvořit obecnější kontext *rychlost* (přenosová rychlost) ve vyšší vrstvě.



Obrázek 4.8: Příklad HMTC modelu před aplikací externí události.

Celý model popsaný obrázkem 4.8 lze pomocí jednotlivých komponent *rozšířeného HMTC*, tedy $HMTC^2 = (N, E, T, U, \rho, w)$ popsat takto:

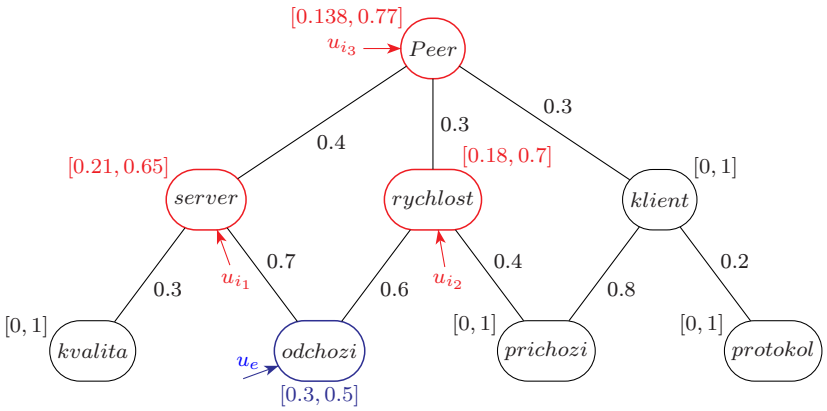
- $N = \{Peer, server, rychlost, klient, kvalita, odchozi, prichozi, protokol\}$,
- $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9\}$, kde:
  $e_1 = (Peer, server)$,
  $e_2 = (Peer, rychlost)$,
  $e_3 = (Peer, klient)$,
  $e_4 = (server, kvalita)$,
  $e_5 = (server, odchozi)$,
  $e_6 = (rychlost, odchozi)$,
  $e_7 = (rychlost, prichozi)$,
  $e_8 = (klient, prichozi)$,
  $e_9 = (klient, protokol)$.
- $T = \{t_0\}$, $U = \{\}$,
- $\forall n \in N : \rho(n, t_0) = [0, 1]$,
- $w(e_1) = 0.4, w(e_2) = 0.3, w(e_3) = 0.3,\ w(e_4) = 0.3,$
  $w(e_5) = 0.7, w(e_6) = 0.6, w(e_7) = 0.4, w(e_8) = 0.8, w(e_9) = 0.2.$

Inicializace modelu je zajištěna nastavením implicitní hodnoty $[0, 1]$ pro každý uzel v čase t_0. Pro první ukázku případu šíření událostí a odvozování důvěry vygenerujeme externí událost u_e pro uzel *prichozi* (řekněme např., že se jedná o externí událost získanou na základě sledování uzlu P2P sítě) s hodnotou $\delta_{u_e} = [0.3, 0.5]$, přičemž této události je přiřazen časový okamžik t_1. Nyní si popíšeme jednotlivé kroky algoritmu (čísla na začátku jednotlivých odrážek označují příslušný řádek/řádky Algoritmu 1):

- *Vstup:* událost $u_e = (\delta_e, t_1)$ aplikovaná na uzel *prichozi*.
- `2`: Aktualizace důvěry uzlu *prichozi* na základě aplikace události u_e:
  $u_e = ([0.3, 0.5], t_1)$,
  $\rho(prichozi, t_1) = \rho(prichozi, t_0) \cap \delta_e$, (dle 4.27)
  $\rho(prichozi, t_1) = [0, 1] \cap [0.3, 0.5] = [0.3, 0.5]$.
- `3`: Vykonání bloku algoritmu `4-25` v případě planosti podmínky:
  $\rho(prichozi, t_0) \neq \rho(prichozi, t_1)$,
  $\Rightarrow$ podmínka je platná, vstup do bloku kódu.
- `4`: Inicializace seznamu *Open*.
- `5`: Získání seznamu naduzlů a jejich vložení do *Open*:
  $Open \leftarrow Open \cup \{server, rychlost\}$.
- `6`: Získaní seznamu poduzlů a jejich vložení do *Open*:
  $Open \leftarrow Open \cup \{\}$.
- `7`: Vstup do bloku cyklu `while` v případě že množina *Open* je neprázdná:
  $Open = \{server, rychlost\}$,
  $\Rightarrow$ podmínka je platná, vstup do bloku kódu `while`.
- `8`: Odstraníme první prvek ze seznamu *Open* a přiřadíme ho do proměnné c:
  $c \leftarrow server$ a $Open = \{rychlost\}$.
- `9`: Stanovení hodnoty interní události δ_{i_1} dle aplikace odvození důvěry směrem nahoru i směrem dolů (funkce *eval*) pro uzel *node*:
  $\delta_{i_1} = eval(node, t_1)$, (dle 4.30)
  $\delta_{i_1} = up(node, t_1) \cap down(node, t_1)$, (dle 4.25)
  $up(node, t_1) = \big(w(e_4) \cdot \rho(kvalita, t_1)\big) + \big(w(e_5) \cdot \rho(odchozi, t_1)\big)$, (dle 4.16)
  $up(node, t_1) = [0.21, 0.65]$, (s využitím 4.7 a 4.9)
  $down(node, t_1) = \frac{\rho(Peer,t_1) - \Big(\big(w(e_2)\cdot\rho(rychlost,t_1)\big) + \big(w(e_2)\cdot\rho(klient,t_1)\big)\Big)}{w(e_1)}$, (dle 4.19)
  $down(node, t_1) = [0, 1]$, (s využitím 4.7–4.10)
  $\delta_{i_1} = [0.21, 0.65] \cap [0, 1] = [0.21, 0.65]$.
- `10`: Vytvoření interní události pro uzel *server* (proměnná *node*):
  $u_{i_1} = (\delta_{i_1}, t_1)$.
- `11`: Aktualizace důvěry uzlu *server* na základě nové události u_{i_1}:
  $\rho(server, t_1) = \rho(server, t_0) \cap \delta_{i_1}$, (dle 4.27)
  $\rho(server, t_1) = [0.21, 0.65]$.

- `12`: Vykonání bloku algoritmu v případě platnosti podmínky:
  $\rho(server, t_0) \neq \rho(server, t_1)$,
  $\Rightarrow$ podmínka je platná, vstup do bloku algoritmu od `14`.
- `13`: Přidání nové události do množiny U_i:
  $U_i \leftarrow U_i \cup \{u_{i_1}\}$.
- `14`: Inicializace seznamu $Parents$ na základě přidání všech naduzlů uzlu $server$:
  $Parents = \{Peer\}$.
- `15-18`: Přidání uzlů ze seznamu $Parents$ do seznamu $Open$ v případě, že již nejsou v seznamu $Open$ a nebo se nejedná o uzel nad kterým vznikla událost u_e:
  $Open = \{rychlost, Peer\}$.
- `19-24`: Inicializace seznamu $Children$ a jeho naplnění poduzly uzlu $server$ v případě, že tyto poduzly již nejsou v seznamu $Open$ a nejedná se o uzel nad kterým vznikla událost u_e:
  $Open = \{rychlost, Peer, kvalita\}$.
- Opakovaná aplikace bloku algoritmu `8--24` pro všechny uzly v seznamu $Open$, analogicky dle výše popsaného, dokud seznam $Open$ nebude prázdný.
- *Výstup:* Množina interních událostí $U_i = \{u_{i_1}, u_{i_2}, u_{i_3}\}$ vytvořených na základě externí události u_e a aktualizace důvěry uzlů $server$, $rychlost$ a $Peer$.

Na základě jedné externí události u_e bylo demonstrováno, jak vznikají události interní, které aktualizují pomocí odvozování důvěry směrem nahoru a směrem dolů důvěru v propojené kontexty. Pro tento konkrétní příklad byly aktualizovány, na základě jedné externí události, hodnoty důvěry pro čtyři různé kontexty. Na obrázku 4.9 je znázorněn konečný stav modelu pro aplikaci události u_e (modře) s naznačeným šířením interních událostí U_i (červeně).



Obrázek 4.9: Příklad HMTC modelu po aplikací externí události.

4.3.7 Konflikt

Navrhovaný model HMTC s rozšířením o události je koncipován tak, že se předpokládá konstantní chováni entit v systému, které generuje události, jejichž interval důvěry je stejný a nebo v ideální případě se zužuje, přičemž střední hodnota tohoto intervalu je konstantní

anebo se mění v jen velmi malém rozmezí. Interval důvěry v konkrétním kontextu se pak s příchodem takových událostí buď nemění a střední hodnota je konstantní anebo se interval zužuje a konverguje ke střední hodnotě intervalů příchozích událostí.

Zásadní problém ovšem nastává v případě, kdy různé externí události pro jeden kontext mají velmi odlišné hodnoty intervalu důvěry. Na základě výše popsaného způsobu aplikace události na kontext (rovnice (4.27)) a na základě výše popsaného algoritmu ABE vzniká *konfliktní stav*, který je způsoben aplikací operace průniku na dva intervaly jejíž výsledek je prázdná množina. Tento stav nazýváme jako *konfliktní stav* nebo jen zkráceně jako *konflikt*.

Velmi jednoduše lze vznik konfliktu demonstrovat na příkladu použitém v předchozí podkapitole (viz obrázek 4.9), kdy jsme aktualizovali hodnotu důvěry uzlu *odchozi* na hodnotu $[0.3, 0.5]$ na základě externí události u_e (kde $u_e^{t_1} = (odchozi, \delta_e)$ při $\delta_e = [0.3, 0.5]$). Uvažme případ, kdy provedeme na základě příchodu nové události $u_f^{t_2} = (odchozi, \delta_f)$ s hodnotou $\delta_f = [0.6, 0.9]$ pro uzel *odchozi* aktualizaci hodnoty důvěry v tomto uzlu (dle (4.27)):

$$\begin{aligned} \rho(odchozi, t_2) &= \rho(odchozi, t_1) \cap \delta_f \\ &= [0.3, 0.5] \cap [0.6, 0.9] \\ &= \emptyset \end{aligned} \tag{4.32}$$

Tento konfliktní stav kontextu *odchozi* je neakceptovatelný z hlediska dalšího použití této HMTC instance a musí být vyřešen. Uvědomme si však, co tento konflikt vlastně reprezentuje a jakým způsobem vznikl. Podíváme-li se na tuto situaci z obecnější roviny zjistíme, že se jedná o konflikt dvou zcela odlišných znalostí. Ty jsou v našem návrhu reprezentovány událostmi.

Obecný přístup k řešení konfliktu

Konflikt jako takový tedy není nic výjimečného a v reálném světě k němu také dochází. Uvažme příklad, kdy chceme koupit osobní automobil nějaké konkrétní značky a konkrétního typu. Dále mějme dva kolegy, kterých se zeptáme na jejich názor na vybraný automobil. První kolega nám odpoví „to je výborné auto, mám ho též, to ti doporučuji“, druhý nám odpoví „tento typ vozu je velmi poruchový, ten ti nedoporučuji“. V tuto chvíli vzniká v naší mysli konflikt, neboť jsme přijali dvě velmi odlišné informace. Navíc není zcela zřejmé, která z těchto informací je „lepší“ či více „pravdivá“ neboť obě mohou být založeny na reálných základech. V ideálním případě bude jedna z nich nepravdivá a druhá pravdivá, v takovém případě je řešením nalézt nepravdivou z nich a tu „odstranit“ a nechat „působit“ pouze tu pravdivou. V horším případě, jak již bylo zmíněno výše, mohou být obě informace pravdivé v tom smyslu, že obě jsou založeny na doložitelných faktech a tudíž nelze jednoznačně říci, že obě informace jsou pravdivé či nepravdivé. Řešením tedy může být zohlednění obou informací anebo žádné, případně výběr jedné z nich jako „více pravdivější“ např. na základě posouzení schopnosti jednotlivých kolegů v kontextu doporučení vozu. Poslední variantou je to, že obě informace jsou nepravdivé a v takovém případě je ideální (ne však vždy dosažitelné, neboť stanovení nepravdivosti informace může být velmi složité) řešení „odstranění“ obou informací.

Jakým způsobem se rozhodnout na základě znalosti o tom, která informace je pravdivá či nepravdivá je tedy zřejmé a nevyžaduje žádné složité uvažovaní. Popsaný příklad a způsob rozhodování však naráží na dva zásadní problémy. První a zásadnější problém je ten, jak rozhodnou pravdivost či nepravdivost získané informace. Tento problém je řešitelný pouze na základě získání všech možných faktů na jejichž základě by se toto rozhodnutí provedlo.

Toto řešení je však v našem prostředí nereálné. Druhým problém je, že na situaci pohlížíme pouze dvoustavovou logikou – tedy *je pravda* nebo *není pravda*. Tento pohled je však velmi zkreslený, protože ne vždy je možné informaci označit za *zcela pravdivou* nebo za *zcela nepravdivou*. Pokud bychom zavedli vyšší úroveň granularity při hodnocení pravdivosti získané informace, lze každou získanou informaci ohodnotit *mírou pravdivosti* informace a na základě této hodnoty rozhodnout, jakým způsobem bude daná informace přijata. První, výše zmíněný problém, jak stanovit míru pravdivosti je však stále otevřen.

Konkrétní řešení konfliktu v HMTC

Pokusíme-li se přenést výše popsané řešení problému konfliktu do našeho návrhu, musíme provést několik nezbytných kroků. První z nich bude stanovení metriky pro ohodnocení přijatých externích událostí a vygenerovaných interních událostí. Vzhledem k tomu, že toto hodnocení bude subjektivní vzhledem k agentu který ho provádí není vhodné ho označit jako *míru pravdivosti*, neboť pojem *pravda* je spíše chápán jako objektivní pojem. Z tohoto důvodu zavedeme pojem *váha události*, jež bude udávat subjektivní míru důvěryhodnosti události. Tuto váhu události označíme symbolem ω a zavedeme ji jako třetí prvek události. Rozšíříme tak původní definici (viz původní definice události (4.26)) následovně:

$$u = (n, \delta, \omega) \tag{4.33}$$

přičemž váhu události ω definujeme jako spojitou veličinu na intervalu $[0, 1]$. V následujícím textu budeme rovněž používat zápis $u^t = (n, \delta, \omega)$, který explicitně definuje odpovídající časový okamžik t ve kterém událost u nastala.

Způsob stanovení konkrétní váhy události ω bude detailně popsáno v kapitole 6. V tuto chvíli jen předestřeme, že váha externí události bude v případě přímé interakce (kdy se jedná o přímou zkušenost agenta) nabývat maximální hodnoty a v případě nepřímé zkušenosti, tedy např. doporučení, bude hodnota váhy události stanovena na základě důvěryhodnosti agenta, který doporučení poskytl. Hodnota váhy interní události, vygenerované na základě hodnoty váhy události externí, bude rovna hodnotě váhy, která ji způsobila.

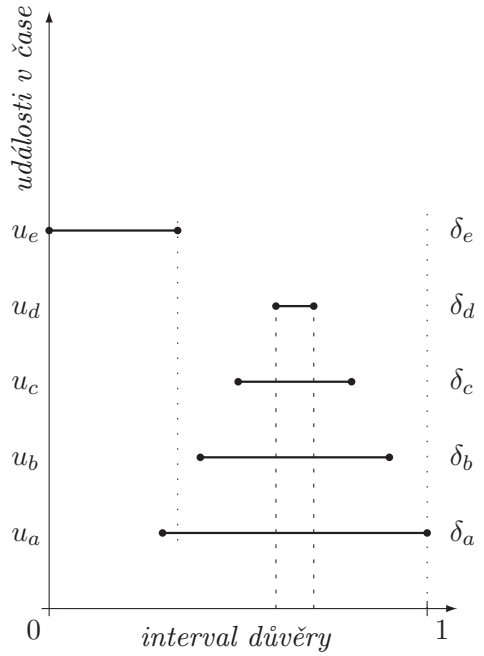
Konflikt mezi dvěma událostmi bude vyřešen tak, že se porovná jejich váha a událost s nižší vahou bude odstraněna. V případě, že váhy jednotlivých událostí, které způsobily konflikt budou shodné, bude odstraněna starší událost (to jsme schopni zajistit pomocí indexů časových okamžiků z časové množiny T). Tím se zajistí, že v případě nerozhodné situace při porovnání vah bude ponechána nejnovější událost, která by měla mít větší relevanci než událost starší. Toto řešení (kdy novější znalost má vyšší váhu nežli starší znalost) koresponduje s často používanými přístupy pro řešení stanovení váhy znalosti je detailněji popsáno v následujících podkapitolách.

Vícenásobný konflikt

Výše popsané řešení konfliktu zohledňuje pouze situaci, kdy nově příchozí externí událost u_e pro kontext n je v konfliktu pouze s jednou předchozí externí událostí u_d. V takovém případu se tedy jednoduše porovnají váhy těchto událostí a na základě toho se jedna z nich odstraní. Uvažujme však případ, kdy by byla příchozí událost u_e v konfliktu se dvěma nebo více předchozími událostmi. V takovém případě nestačí provést porovnání vah pouze dvou událostí, ale je třeba zohlednit i další potencionálně konfliktní události a jejich váhy.

Příklad takového vícenásobného konfliktu je naznačen na obrázku 4.10, kde postupně v čase přicházejí události $u_a, \ldots, u_d$, které nejsou konfliktní. Výsledný interval důvěry je

tak stanoven jako průnik všech těchto událostí a odpovídá hodnotě důvěry δ_d události u_d. Následná příchozí událost u_e je však v konfliktu hned se třemi předchozími událostmi u_b, u_c, u_d.



Obrázek 4.10: Příklad vícenásobného konfliktu způsobeného událostí u_e.

Řešení vícenásobného konfliktu je opět založeno na porovnání vah jednotlivých konfliktních událostí. V tomto případě však nelze porovnat váhy pouze dvou událostí, ale všech událostí jež jsou s příchozí událostí v konfliktu. Vyhodnocení se tedy provede tak, že stanovíme součet vah již aplikovaných událostí jež jsou s nově příchozí v konfliktu a porovnáme ho s váhou události příchozí. Pakliže váha nově příchozí události, která způsobila vícenásobný konflikt je větší nebo rovna sumě vah všech předchozích událostí jež jsou s ní v konfliktu, odstraníme všechny tyto předchozí události.

Příklad 2. Uvažujme případ popsaný na obrázku 4.10, kde událost u_e aplikovaná na kontext n vyvolala vícenásobný konflikt. Označme U_n množinu všech aplikovaných událostí na kontext n před příchodem události u_e v časovém okamžiku t_5, kde:

$$\begin{aligned} U_n &= \{u_a, u_b, u_c, u_d\}, \text{ kde:} \\ & u_a^{t_1} = (n, \delta_a, \omega_a), \\ & u_b^{t_2} = (n, \delta_b, \omega_b), \\ & u_c^{t_3} = (n, \delta_c, \omega_c), \\ & u_d^{t_4} = (n, \delta_c, \omega_d). \end{aligned}$$

Při rozhodování, které události odstranit, postupujeme takto:

1. Stanovíme množinu $U_{n_e^k}$ všech konfliktních událostí s $u_e^{t_5} = (n, \delta_e, \omega_e)$ pro kontext n:

$$U_{n_e^k} = \{u_k \in U_n : u_k = (n, \delta_k, \omega_k) \wedge \delta_k \cap \delta_e = \emptyset\}$$

konkrétně tedy, pro náš případ je $U_{n_e^k} = \{u_b, u_c, u_d\}$.

2. Stanovíme sumu vah událostí množiny $U_{n_e^k}$:

$$\omega_{n_e^k} = \sum_{u_k \in U_{n_e^k}} \omega_k$$

3. Stanovíme množinu událostí $U_{n_e^r}$, které odstraníme na základě porovnání vah:

$$U_{n_e^r} = \begin{cases} \{u_e\} & \text{pro } \omega_e < \omega_{n_e^k} \\ U_{n_e^k} & \text{pro } \omega_e \geq \omega_{n_e^k} \end{cases}$$

Vliv času události na řešení konfliktu

Jak si dokážeme v následující podkapitole 4.3.10, na pořadí aplikací jednotlivých nekonfliktních událostí nezáleží a výsledný interval důvěry bude vždy shodný. Nicméně čas události může ovlivnit dopad váhy události při řešení konfliktu. Obecně lze říci, že události by měly „stárnout“ tak, aby bylo možné více zohlednit poslední příchozí (aktuální) události na úkor událostí starých a potencionálně neaktuálních. Tento princip se nazývá obecněji „stárnutí událostí“ (nebo také „stárnutí znalostí“) a je částečně popsán v kapitole 3 a konkrétně v podkapitole 3.3.2.

Tento přístup stárnutí znalostí je rovněž použit mimo jiné v pracích [SS01] a [Mar94]. Konkrétně v práci [Mar94] je tento přístup diskutován v souvislosti se schopností agenta pamatovat si pouze určitý časový úsek v minulosti. To lze z obecného hlediska přirovnat ke schopnosti člověka, pamatovat si jen některé znalosti, neboť jeho paměť má omezené schopnosti a neaktuální údaje zapomínáme a postupně se vytrácí (respektive vytrácí se schopnost jejich zpětného vybavení). Stejně tak kapacita paměti umělého agenta je omezená, tudíž je vhodné velmi staré znalosti (ideálně pouze ty, které již nejsou relevantní) „zapomenout“ (odstranit).

Přístupů, jak „stárnutí událostí“ zohlednit lze najít hned několik, my si zde popíšeme jeden z nich, který je pro náš návrh nejvhodnější. Jeho princip je založen na definici funkce *stárnutí*, která pro každý časový okamžik události a aktuální časový okamžik definuje hodnotu v intervalu $[0, 1]$. Tuto hodnotu lze pak zahrnout do výpočtu vah událostí při řešení konfliktu. Jinými slovy, hodnoty vah jednotlivých událostí nebudou v průběhu času měněny, ale v případě výskytu konfliktu, kdy se váhy událostí porovnávají (jak bylo popsáno v předchozích odstavcích), vezmeme v potaz časový okamžik jejich výskytu a tyto váhy před porovnáním upravíme pomocí funkce *stárnutí*.

Konkrétní způsob definice funkce *stárnutí* záleží na příslušné implementaci. Je možné tuto funkci definovat tak, že pro velmi staré události bude její výsledek roven nule, čímž se tyto události zcela eliminují. To, co znamená „velmi stará“ událost rovněž záleží na konkrétní implementaci množiny časových okamžiků. Mohou to být konkrétní časová razítka s rozlišením na sekundy či milisekundy. Pak staré události lze definovat např. jako události, jejichž časové razítko je starší jak „jeden měsíc“ od poslední události (nebo aktuálního času). V případě, že časová množina jsou časové okamžiky indexované pomocí přirozených čísel, pak lze stanovit, že události s časovým okamžikem jejichž index je menší než nějaké číslo n, kde n je stanoveno rozdílem posledního indexu časového okamžiku (nebo aktuálního) a nějaké konstanty, jsou „velmi staré“.

Příklad 3. Mějme množinu časových okamžiků T, jednotlivé časové okamžiky zapisujeme jako t_i, kde index i je z množiny přirozených čísel $\mathbb{N}$ a hodnota $t_i = i$. S rostoucím časem index časového okamžiku roste, tedy časový okamžik t_1 předchází časovému okamžiku t_2 apod. Pak definujeme funkci *stárnutí* takto:

$$f(t_u, t_a) = \frac{t_u}{t_a} \,, \tag{4.34}$$

kde t_u je časový okamžik výskytu události u a t_a je aktuální (nebo poslední) časový okamžik – tedy musí platit podmínka že $t_u < t_a$.

Dále uvažujeme množinu událostí U_n na kontext n, kde $U_n = \{u_a, u_b, u_c, u_d\}$ a jednotlivé události jsou definovány takto: $u_a^{t_1} = (n, \delta_a, \omega_a), u_b^{t_2} = (n, \delta_b, \omega_b), u_c^{t_3} = (n, \delta_c, \omega_c), u_d^{t_4} = (n, \delta_c, \omega_d)$. Následně v časovém okamžiku t_5 vznikne nová externí událost u_e nad kontextem n s hodnotou získaného intervalu důvěry δ_e a vahou ω_e. Na základě této události vznikl konflikt mezi $u_b^{t_2}$ a $u_e^{t_5}$. Při řešení konfliktu bez zohlednění stárnutí událostí by se konflikt řešil tak, že provedeme porovnání vah konfliktních událostí a na jejich základě se rozhodneme, kterou z nich odstraníme:

$$\begin{array}{lll} \text{pro} & \omega_b \leq \omega_e & \text{odstraníme událost } u_b \,, \\ \text{pro} & \omega_b > \omega_e & \text{odstraníme událost } u_e \,. \end{array} \tag{4.35}$$

Při použití funkce *stárnutí* provedeme její vyhodnocení po obě události:

$$\begin{array}{l} u_b : f(t_2, t_5) = 2/5 = 0.4 \\ u_e : f(t_5, t_5) = 5/5 = 1.0 \end{array} \tag{4.36}$$

a následně zohledníme tyto výsledky při porovnání vah jednotlivých událostí takto:

$$\begin{array}{lll} \text{pro} & (0.4 \cdot \omega_b) \leq (1.0 \cdot \omega_e) & \text{odstraníme událost } u_b \,, \\ \text{pro} & (0.4 \cdot \omega_b) > (1.0 \cdot \omega_e) & \text{odstraníme událost } u_e \,. \end{array} \tag{4.37}$$

Obdobně jako v příkladu 3, kdy byly v konfliktu pouze dvě události, bychom aplikovali funkci *stárnutí* i na vícenásobný konflikt. Při stanovení časové množiny pomocí přirozených čísel a s využitím funkce (4.34) definujeme pro množinu $U_{n_e^k}$ konfliktních událostí s událostí u_e sumu vah takto:

$$\omega_{n_e^k} = \sum_{u_k \in U_{n_e^k}} \omega_k \cdot f(t_{u_k}, t_e)$$

kde t_{u_k} je časový okamžik konkrétní události $u_k \in U_{n_e^k}$ a t_e je časový okamžik události u_e. Tuto sumu vah bychom následně porovnali s vahou ω_e události u_e a na základě výsledku bychom provedli rozhodnutí, které události budou odstraněny.

4.3.8 Posloupnost událostí – Cesta

Na základě algoritmu šíření události (Algoritmus 1) a následném příkladu jsme ukázali, že každá externí událost může způsobit více interních událostí. Tyto interní události jsou způsobeny buď přímo (přímo následkem externí události) anebo nepřímo (následkem jiné interní události) externí události. V případě, že vznikne konflikt mezi dvěma externími událostmi v rámci jednoho kontextu a bude určeno, která z nich se odstraní (na základě řešení popsaného v předchozí podkapitole), je třeba rovněž odstranit i veškeré interní události, které odstraněná externí událost způsobila. To je logickým důsledkem odstranění externí události, neboť v případě že odstraníme z agentovy báze znalostí nějakou znalost, je třeba

také odstranit i všechny další znalosti, které tato znalost „vytvořila“. V případě, že bychom tyto znalosti (události) neodstranili, došlo by ke vzniku konfliktů mezi interními událostmi (pozůstatků odstraněné externí události s událostmi novými).

Abychom byli schopni odstranit všechny interní události vytvořené na základě odstraněné externí události, musíme být schopni příslušné interní události identifikovat. Z tohoto důvodu definujeme pojem *cesta*.

Definice 1. *Cestou* p označíme konečnou neprázdnou posloupnost, kde jednotlivé termy posloupnosti jsou události z U. Cestu p zapisujeme ve tvaru $p = u_0 u_2 \dots u_{k-1} u_k$, přičemž cestu z u_0 do u_k zapisujeme jako $p_{(0,k)}$. Událost u_0 označujeme jako počátek cesty, přičemž počátkem cesty je vždy externí událost. Událost u_k označujeme jako konec cesty.

Vzhledem k definici cesty zavedeme množinu všech cest a označíme jí symbolem P.

Definice 2. Libovolnou posloupnost událostí $u_i \dots u_j$ z cesty $p = u_0 u_1 u_2 \dots u_{k-1} u_k$ pro kterou platí že $i > 0$ a $j \leq k$ nebo $i \geq 0$ a $j < k$ nazýváme *podcestou* cesty p.

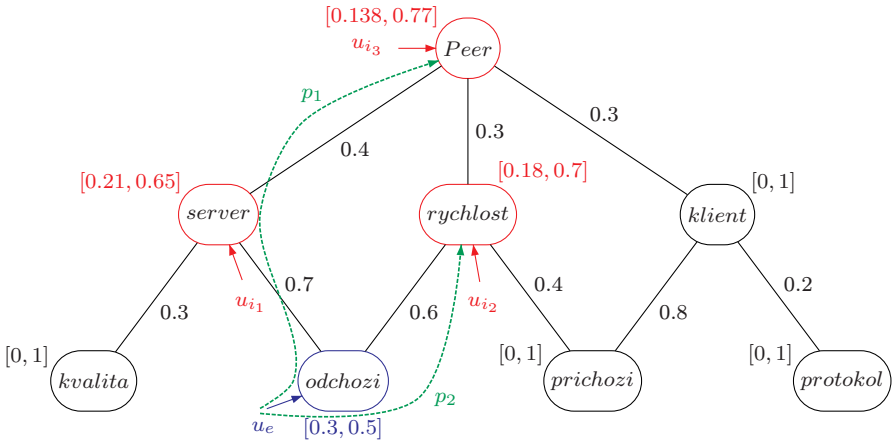
Příklad stanovení cesty a množiny cest

Pro demonstraci stanovení cesty navážeme na příklad demonstrace ABE a strukturu HMTC použitou v podkapitole 4.3.4, obrázky 4.8 a 4.9. V uvedeném případu byla vygenerována jedna externí událost u_e s hodnotou získané důvěry $\delta_e = [0.3, 0.5]$ pro kontext *prichozi*. Po aplikaci této události je množina událostí U a množina cest P stanovena takto:

- $U = U_e \cup U_i = \{u_e\} \cup \{u_{i_1}, u_{i_2}, u_{i_3}\}$,
- $P = \{p_1, p_2\}$, kde
  - $p_1 = u_e u_{i_1} u_{i_2}$
  - $p_2 = u_e u_{i_3}$.

Jedna externí událost tedy vyvolala tři interní události na jejichž základě byly stanoveny dvě cesty. Vzhledem k tomu, že kontext *prichozi* nemá žádné podkontexty, bylo jedinou možností provést odvození důvěry směrem nahoru. Kontext *prichozi* má dva nadkontexty a oba byly aktualizovány, směr šíření události se tedy rozdělil a vytvořil dvě cesty p_1 a p_2, ty jsou naznačeny na obrázku 4.11 (zeleně).

Pro tento příklad nyní uvažujme příchod další externí události u_f na kontext *klient* s hodnotou získané důvěry $\delta_f = [0.4, 0.6]$. Po aplikaci této události v čase t_2 budou následným procesem šíření události aktualizovány kontexty *klient*, *Peer*, *rychlost* a *prichozi* takto:



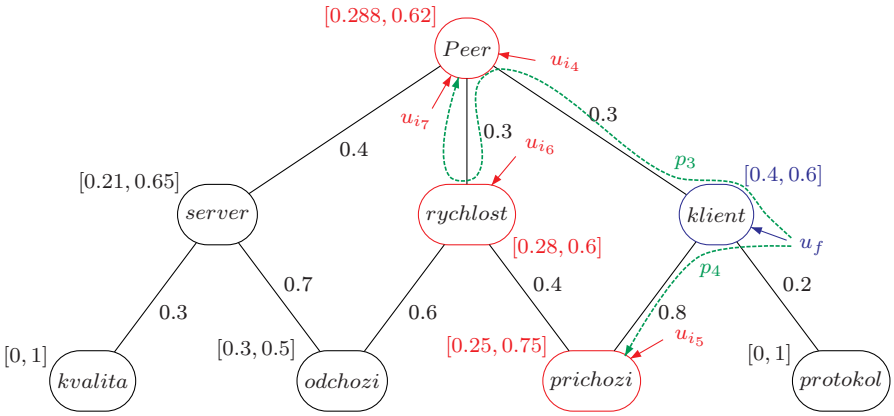
Obrázek 4.11: Příklad naznačení cesty v HMTC modelu.

$$
\begin{aligned}
u_f: &\quad \rho(klient, t_2) = \rho(klient, t_1) \cap \delta_f \\
&\quad \rho(klient, t_2) = [0,1] \cap [0.4, 0.6] = [0.4, 0.6] \\
\\
u_{i_4}: &\quad \rho(Peer, t_2) = \rho(Peer, t_1) \cap eval(Peer, t_1) \\
&\quad \rho(Peer, t_2) = [0.138, 0.77] \cap [0.258, 0.65] = [0.258, 0.65] \\
\\
u_{i_5}: &\quad \rho(prichozi, t_2) = \rho(prichozi, t_1) \cap eval(prichozi, t_1) \\
&\quad \rho(prichozi, t_2) = [0,1] \cap [0.25, 0.75] = [0.25, 0.75] \\
\\
u_{i_6}: &\quad \rho(rychlost, t_2) = \rho(rychlost, t_1) \cap eval(rychlost, t_1) \\
&\quad \rho(rychlost, t_2) = [0.18, 0.7] \cap [0.28, 0.6] = [0.28, 0.6] \\
\\
u_{i_7}: &\quad \rho(Peer, t_2) = \rho(Peer, t_2) \cap eval(Peer, t_2) \\
&\quad \rho(Peer, t_2) = [0.258, 0.65] \cap [0.288, 0.62] = [0.288, 0.62]
\end{aligned}
\tag{4.38}
$$

Na základě aplikace události u_f byl tedy změněn stav modelu, respektive námi sledované množiny U a P tak, že byla přidána jedna externí událost u_f a čtyři interní události $u_{i_4} \dots u_{i_7}$, které společně vytvořily další dvě cesty p_3 a p_4. V tomto okamžiku lze tedy popsat sledované množiny U, P takto:

- $U = U_e \cup U_i = \{u_e, u_f\} \cup \{u_{i_1}, u_{i_2}, u_{i_3}, u_{i_4}, u_{i_5}, u_{i_6}, u_{i_7}\}$,
- $P = \{p_1, p_2, p_3, p_4\}$, kde
  - $p_1 = u_e u_{i_1} u_{i_2}$
  - $p_2 = u_e u_{i_3}$
  - $p_3 = u_f u_{i_4} u_{i_6} u_{i_7}$
  - $p_4 = u_f u_{i_5}$

Na příkladu aplikace šíření události u_f si lze povšimnout ještě jednoho zajímavého aspektu. Hodnota důvěry kontextu $Peer$ byla procesem šíření události aktualizována celkem dvakrát (událostmi u_{i_4} a u_{i_7}) v jednom časovém okamžiku (během jednoho běhu algoritmu šíření události). Tato situace je z hlediska principu funkce validní, neboť explicitně



Obrázek 4.12: Příklad naznačení s příchodem nové externí události u_f.

nezakazuje, aby se jeden kontext do seznamu *Open* dostal vícekrát v případě, že byl aktualizován jeho nadkontext nebo podkontext. V našem případě byla aktualizace kontextu *Peer* způsobena nejprve v důsledku aktualizace důvěry uzlu *klient* a v druhém případě díky aktualizaci důvěry kontextu *rychlost*. To znamená, že dvojí aktualizaci důvěry kontextu *Peer* způsobila aktualizace důvěry jeho dvou různých podkontextů. Tato situace je demonstrována rovněž na obrázku 4.12 naznačením cesty p_3 a p_4 (zeleně).

V případě vícenásobné aktualizace hodnoty důvěry jednoho uzlu v rámci jednoho běhu ABE je třeba pouze zajistit, aby se aktualizace prováděla vždy s aktuální aktualizovanou hodnotou z předchozí aktualizace daného běhu ABE a nikoliv z aktualizace provedené na základě předchozího běhu ABE.

4.3.9 Rozšířená definice HMTC o komponentu množiny cest

Podobně jako v případě zavedení množiny událostí U, tak i v případě zavedení komponenty cest P je třeba rozšířit definici HMTC modelu. Toto rozšíření HMTC modelu budeme označovat $HMTC^3$ a pokud nebude v následujícím textu uvedeno jinak, tak se implicitně předpokládá použití právě tohoto rozšířeného modelu. Současnou šestici tak rozšíříme o komponentu P následovně:

$$HMTC^3 = (N, E, T, U, P, \rho, w) \tag{4.39}$$

kde P označuje množinu všech cest dle definice 1, přičemž ostatní komponenty zůstávají beze změny dle definice $HMTC^2$ (4.31).

4.3.10 Odstranění interních událostí s využitím cesty

Hlavní motivací pro definici *cesty* bylo stanovení posloupnosti interních událostí, které byly způsobeny událostí externí. Toto stanovení nově vzniknuvších posloupností interních událostí je důležité při řešení konfliktu. Konkrétně tedy v případě, kdy zvolíme příslušnou externí událost pro odstranění, musíme odstranit i všechny interní události, které aplikace externí události způsobila. Tyto události pak nazýváme *relevantní události* vzhledem k příslušné externí události. Princip vytváření interních událostí na základě události externí je popsán v kapitole 4.3.4 a rovněž v příkladu 1.

Pro následující demonstraci principu odstranění relevantních událostí si zavedeme funkci $path2set$, která převede cestu (posloupnost událostí) na množinu událostí:

$$path2set(p) = \{u_0, u_1, u_2, u_3\} \quad \text{pro} \quad p = u_0u_1u_2u_3. \tag{4.40}$$

Příklad 4. Uvažujme rozšířený $HMTC^3$ model a mějme příchozí externí událost u_e, která vyvolala konflikt v kontextu n. Na základě výše popsaného řešení konfliktu v kapitole 4.3.7 byla stanovena množina konfliktních událostí U_{k_n} jako podmnožina U, která obsahuje jednu nebo více externích událostí aplikovaných na kontext n, které se mají odstranit. Postup odstranění všech relevantních a externích událostí pro vyřešení konfliktu je následující:

1. Stanovíme množinu všech cest P_k, jež obsahují posloupnosti událostí, které mají být odstraněny:
$$P_k = \{p \in P : u_k \in U_{k_n} \text{ kde } u_k \text{ je počátek cesty } p\}. \tag{4.41}$$

2. Stanovíme množinu všech relevantních a externích událostí (události obsaženy v cestách množiny P_k):
$$U_r = \bigcup_{p_k \in P_k} path2set(p_k). \tag{4.42}$$

3. Odstraníme události U_r z množiny všech událostí U:
$$U = U \setminus U_r. \tag{4.43}$$

4. Provedeme aktualizaci intervalu důvěry všech kontextů, pro něž byly události odebrány. Výsledná hodnota důvěry každého kontextu je určena průnikem získané hodnoty důvěry všech jeho zůstavších událostí.

Věta 1. Libovolnou událost u_n, která způsobila aktualizaci hodnoty důvěry uzlu n lze z množiny U odebrat aniž bychom způsobili konflikt mezi propojenými kontexty, přičemž nezáleží na tom, kdy událost nastala.

Důkaz 1. Uvažujme množinu událostí $U_n = \{u_n^t, u_n^{t+1}, \ldots, u_n^{t+i}\}$, které způsobily aktualizaci důvěry uzlu n v časech $t, t+1, \ldots, t+i$, kde $t < (t+1) < \cdots < (t+i)$. Pak hodnota důvěry uzlu n, inicializovaného hodnotou intervalu $[0,1]$, je v čase:

$$\begin{array}{rrl} t\text{:} & \rho(n,t) & = [0,1] \cap \delta_n^t \\ t+1\text{:} & \rho(n,t+1) & = \rho(n,t) \cap \delta_n^{t+1} \\ t+2\text{:} & \rho(n,t+2) & = \rho(n,t+1) \cap \delta_n^{t+2} \\ & \vdots & \\ t+i\text{:} & \rho(n,t+i) & = \rho(n,t+i-1) \cap \delta_n^{t+i} \end{array} \tag{4.44}$$

tedy, hodnota uzlu n v čase $(t+i)$ je:

$$\rho(n,t+i) = [0,1] \cap \delta_n^t \cap \delta_n^{t+1} \cap \cdots \cap \delta_n^{t+i} \tag{4.45}$$

Vzhledem k tomu, že průnik je operace *komutativní*, musí platit následující rovnost:

$$\delta_n^t \cap \delta_n^{t+1} \cap \cdots \cap \delta_n^{t+i} = \delta_n^{t+i} \cap \cdots \cap \delta_n^{t+1} \cap \delta_n^t \tag{4.46}$$

a proto lze libovolnou událost u_n^x z množiny U_n odstranit a výsledná hodnota důvěry uzlu n bude stanovena průnikem hodnot událostí z množiny $U_n \setminus \{u_n^x\}$. □

4.4 Shrnutí

Tato kapitola se zabývala formálním popisem více-kontextového modelu důvěry, který tvoří jádro této disertační práce. Navržený model, nazvaný jako *hierarchický model důvěry s kontexty* (zkráceně HMTC), je založen na hierarchickém uspořádání jednotlivých kontextů jedné entity, jejíž důvěryhodnost je posuzována. Toto hierarchické uspořádání vytváří grafovou strukturu, na základě které je možné provádět mezi propojenými kontexty odvozování důvěry.

Entita reprezentovaná námi navrženým modelem je kompozicí různých, ne zcela nezávislých kontextů. Kontext v našem návrhu reprezentuje aspekt modelované entity, pro který může být stanovena hodnota důvěry na základě přímé nebo nepřímé zkušenosti s ní. Hierarchická struktura těchto aspektů je stanovena empiricky na základě znalosti modelované entity a na základě požadavků na konkrétní aplikaci, jež HMTC model používá. Každá posuzovaná entita systému je reprezentována právě jednou instancí HMTC modelu.

Navržený HMTC model je diskrétní systém, určený pro podporu rozhodování agentů v heterogenním distribuovaném multi-agentním systému. Model je řízen událostmi, které rozdělujeme na *externí* a *interní*, přičemž tyto události mohou způsobit aktualizaci důvěry kontextu, na který jsou aplikovány. Každá externí událost způsobí to, že je vyvolán algoritmus šíření události, který následně generuje události interní. S využitím tohoto principu je tak možné provést aktualizaci důvěry v různé kontexty jedné entity na základě jedné externí události, která představuje získání nové znalosti o důvěryhodnosti konkrétního kontextu. Navržený model zohledňuje i situaci, kdy dvě nebo více získaných znalostí o důvěryhodnosti konkrétního kontextu jsou vzájemně neslučitelné (protichůdné). Takový stav nazýváme *konflikt* a jeho řešením se zabývá poslední část této kapitoly.

Jednou ze specifických vlastností navrženého modelu je unikátní reprezentace důvěry pomocí matematického intervalu na množině reálných čísel. Tento *interval důvěry* nám umožňuje vyjádřit pomocí svých mezí minimální a maximální hodnotu důvěryhodnosti spolu s velikostí nejistoty, která je s hodnotou důvěry vždy určitým způsobem spojena. V následující kapitole velmi úzce na toto téma navážeme a popíšeme princip vyhodnocování důvěry, tedy stanovení intervalu důvěry, na základě výsledků přímých interakcí mezi entitami v systému.

Kapitola 5

Vyhodnocení důvěry pro model HMTC

Jak bylo popsáno v podkapitole 4.2.1, důvěra je v navrženém HMTC modelu reprezentována intervalem $[0, 1]$. Tento interval je chápán jako subjektivní vnímání schopnosti a spolehlivosti s jistou mírou neurčitosti nějakého agenta vzhledem jinému agentu v daném kontextu. Cílem každého agenta by tak mělo být to, aby vždy co nejpřesněji odhadl chování (vzhledem k danému kontextu) ostatních agentů v systému tak, aby si mohl vybrat na základě tohoto odhadu ideálního partnera k interakci v daném kontextu nebo kontextech.

V počáteční fázi návrhu systému předpokládejme, že agenti mají konstantní chování anebo, že jejich chování osciluje v určitém malém rozsahu v závislosti na jejich aktuálním mentálním stavu, případně na vlivech okolí. Schopnost agenta vzhledem k nějakému kontextu je stanovena hodnotou z intervalu $[0, 1]$, která udává jak kvalitně je schopen provést interakci v rámci tohoto kontextu. Tuto hodnotu označíme jako *střední hodnotu* a chápejme ji jako typické anebo nejpravděpodobnější chování agenta. Druhým parametrem, který definuje chování je hodnota *rozptylu* chování a tento parametr udává spolu se střední hodnotou maximální a minimální rozsah oscilací chování agenta. Tyto dvě hodnoty tak definují kvalitu agenta v daném kontextu jako interval, v němž jsou s *normálním rozložení pravděpodobnosti* generovány výsledky jednotlivých interakcí s agentem.

Intervalová reprezentace důvěry tedy stanovuje v našem návrhu to, s jakou pravděpodobností provede agenta nějakou interakci v daném kontextu, přičemž pro rozložení pravděpodobnosti na tomto intervalu je využito normálního rozložení. V následující podkapitole si tak přesněji popíšeme vztah intervalu důvěry vzhledem k normálnímu rozložení pravděpodobnosti náhodné veličiny. Z tohoto důvodu tak v úvodní části této kapitoly poněkud odbočíme a definujeme si některé základní pojmy z teorie pravděpodobnosti a matematické statistiky.

5.1 Normální rozložení pravděpodobnosti

Normální rozložení pravděpodobnosti je jedno z nejdůležitější a nejpoužívanějších rozložení pravděpodobností náhodné veličiny [MR03]. Normální rozložení je též někdy označováno jako *Gaussovo* (anglicky Gaussian) dle německého matematika *Carla Friedricha Gausse*. Pro termín *rozložení pravděpodobnosti* se rovněž často používá ekvivalent *rozdělení pravděpodobnosti*.

5.1.1 Spojitá náhodná veličina a její charakteristiky

Spojitá náhodná veličina je náhodná veličina, která může nabývat všech hodnot ze spojitého intervalu. Nyní si formálně definujeme pojem náhodná veličina její základní charakteristiky. Následující definice jsou převzaty z [MR03, Nov06]. Pro úplnost definic dodejme, že funkce P označuje *pravděpodobnost*.

Definice 3. Náhodná veličina je každé zobrazení $X : \Omega \to \mathbb{R}$ takové, že pro každé $x \in \mathbb{R}$ je

$$A = \{\omega | X(\omega) \leq x\} \in \mathcal{A} \,. \tag{5.1}$$

Jestliže $\mathcal{A}$ je systém všech podmnožin Ω, pak každá reálná funkce X definována na Ω je náhodná veličina. Náhodné veličiny budou v následujícím textu označeny X, Y, Z, jejich konkrétní realizace pak pomocí malých písmen x, y, z.

Definice 4. Distribuční funkce náhodné veličiny X je funkce označována jako $F : \mathbb{R} \to [0, 1]$ definována vztahem:

$$F(x) = P(X \leq x) \,. \tag{5.2}$$

Distribuční funkce je neklesající, v $-\infty$ má distribuční funkce nulovou hodnotu a v $+\infty$ má hodnotu jedna. Jestliže možné hodnoty náhodné veličiny X patří do intervalu (a, b) pak platí:

$$F(a) = 0, \quad F(b) = 1 \,. \tag{5.3}$$

Definice 5. Rozložení pravděpodobnosti pro spojitou náhodnou veličinu X určujeme funkcí $f(x)$, kterou nazýváme *hustota rozložení pravděpodobnosti* a pro niž platí:

1. $f(x) \leq 0$,
2. $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$,
3. $P(a \leq X \leq b) = \int_{a}^{b} f(x)\mathrm{d}x$.

Pro *spojitou distribuční funkci* F a pro funkci *hustoty rozložení pravděpodobnosti* $f(x)$ platí že:

$$F(x) = \int_{-\infty}^{x} f(u)du \quad \text{pro každé} \quad x \in \mathbb{R} \,. \tag{5.4}$$

5.1.2 Hustota rozložení pravděpodobnosti normálního rozložení

Rozdělení pravděpodobnosti spojité náhodné veličiny se určuje prostřednictvím funkce, kterou označujeme jako *hustota rozložení pravděpodobnosti*. Pro normální rozložení si nyní uvedeme definici této funkce.

Definice 6. Náhodná veličina X s hustotou rozložení pravděpodobnosti stanovenou funkcí:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \, e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{5.5}$$

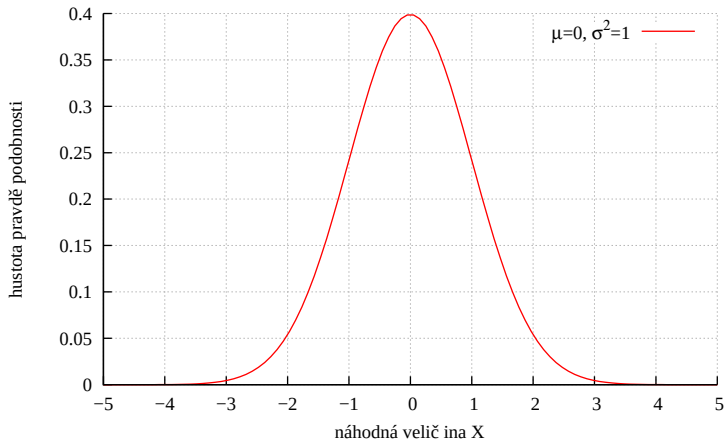
je **normální náhodná veličina** s parametrem μ a σ, kde $-\infty < \mu < \infty$ a $\sigma > 0$. Zároveň platí pro *střední hodnotu* a *rozptyl*:

$$\mathrm{E}(X) = \mu \quad \text{a} \quad \mathrm{V}(X) = \sigma^2 \tag{5.6}$$

a zápis $N(\mu, \sigma^2)$ označuje *normální rozložení pravděpodobnosti. Střední hodnota* a *rozptyl* jsou označeny jako μ a σ^2.

Pro úplnost této definice ještě uvedeme, že parametr σ je označován jako *směrodatná odchylka* (anglicky *standard deviation*), přičemž označení rozptylu σ^2 má anglický ekvivalent *variance*.

Rozložení $N(0, 1)$ označujeme jako normované (nebo standardní) normální rozložení a funkce hustoty tohoto rozložení se obvykle značí symbolem $\varphi(x)$. Grafické znázornění tohoto normovaného rozložení lze nalézt na obrázku 5.1.



Obrázek 5.1: Normované (standardní) normální rozložení.

Distribuční funkci normálního rozložení nelze vyjádřit elementárními funkcemi [Nov06]. Její hodnoty lze stanovit numericky (pomocí numerické integrace) nebo po transformaci náhodné veličiny X na normovanou normální veličinu Z stanovenou takto:

$$Z = \frac{X - \mu}{\sigma} \,. \tag{5.7}$$

Veličina Z má pak normované normální rozložení $N(0, 1)$ a distribuční funkci lze vyjádřit takto:

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \tag{5.8}$$

kde $\Phi(z)$ je distribuční funkce normovaného normálního rozložení $N(0, 1)$ definována takto:

$$\Phi(z) = \int_{-\infty}^{z} \varphi(u)\mathrm{d}u \quad \text{pro } z \in \mathbb{R} \,. \tag{5.9}$$

Vzorec pro numerické řešení obecné distribuční funkce normálního rozložení $N(\mu, \sigma^2)$ vypadá takto:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}}\mathrm{d}t. \tag{5.10}$$

Další možné stanovení distribuční funkce je možné pomocí takzvané *chybové funkce* (a anglické literatuře často označované jako *erf – error function*). V reálných aplikacích se pak používají různé aproximace této funkce nebo případně pouze odečtením hodnoty z předpočítané tabulky [MR03].

Dále je ještě vhodné zmínit často používané relevantní termíny v anglické literatuře: zkratka *PDF* (*Probability Density Function*) odpovídá funkci hustota pravděpodobnosti a zkratka *CDF* (*Cumulative Distribution Function*) odpovídá distribuční funkci. Tyto zkratky budou použity v následujících grafických reprezentacích společně s označením X, čímž se rozumí náhodná veličina s daným rozložením.

Empirické pravidlo pro normálně rozdělené náhodné veličiny

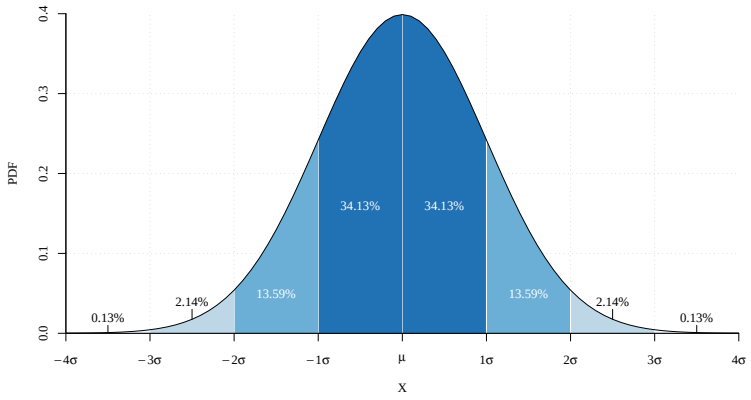
Pro každou normálně rozdělenou náhodnou veličinu X platí:

(a) $P(\mu - \sigma < X < \mu + \sigma) = 0.6826$,

(b) $P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544$,

(c) $P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9974$.

Toto pravidlo je známé jako *empirické pravidlo* (často také jako pravidlo *68-95-99.7* nebo *pravidlo tří-sigma* [DW92]) a je ve statistice a teorii pravděpodobnosti velmi často používané [Nov06]. Toto pravidlo nám říká, že většina hodnot (přibližně 68.26%) se neodlišuje od střední hodnoty μ o více než jednu směrodatnou odchylku σ a 99.74% hodnot jsou v pásmu do třech směrodatných odchylek od střední hodnoty (průměru). Grafické znázornění tohoto empirického pravidla je uvedeno na obrázku 5.2.



Obrázek 5.2: Vizualizace empirického pravidla a význam směrodatné odchylky normálního rozložení vzhledem k jeho funkci hustoty pravděpodobnosti.

5.2 Další důležitá rozložení

Pro následující podkapitoly a především pro formální popis intervalových odhadů je nezbytné popsat další dvě význačná rozložení: *Chí-kvadrát rozložení* a *t-rozložení*.

5.2.1 Chí-kvadrát rozložení – χ^2

Chí-kvadrát rozložení (označované jako χ^2) je rozložení odvozeno ze součtu nezávislých náhodných veličin s normovaným normálním rozložením. Tedy, pokud $X_1, X_2, \ldots, X_n$ jsou nezávislé náhodné veličiny, přičemž každá z nich má rozložení $N(0,1)$, pak náhodná veličina:

$$Y = X_1^2 + X_2^2 + \cdots + X_n^2 \tag{5.11}$$

má rozložení χ^2 (chí-kvadrát) o n stupních volnosti. Zkráceně ho pak označujeme jako χ_n^2. Alternativně lze rovněž toto rozložení definovat pomocí funkce hustoty rozložení pravděpodobnosti dle [MR03] následovně:

$$f(x) = \begin{cases} \dfrac{1}{2^{n/2}\,\Gamma(n/2)}\ \mathrm{e}^{-x/2}\ x^{(n/2)-1} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases} \tag{5.12}$$

kde $\Gamma(n/2)$ označuje *Gama funkci* (někdy také označován jako Eulerův integrál druhého druhu) s argumentem $n/2$.

5.2.2 t-rozložení

t-rozložení, nebo též často označováno jako *Studentovo* rozložení, je takové rozložení náhodné veličiny X pro kterou platí:

$$X = \frac{X_1}{\sqrt{\dfrac{X_2}{n}}} \tag{5.13}$$

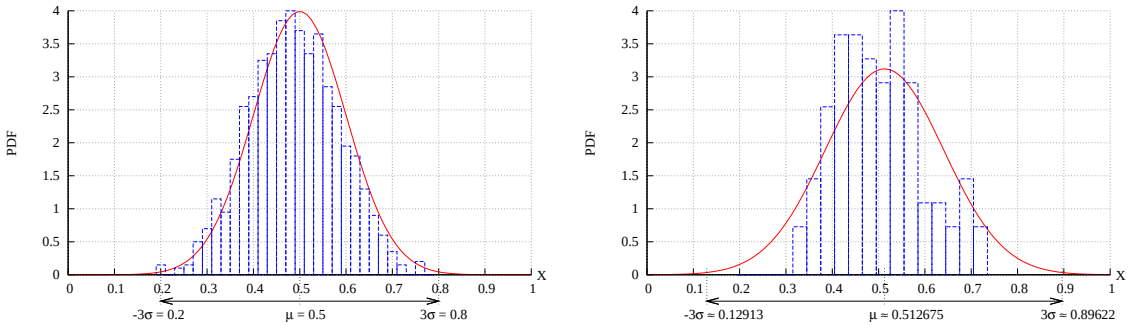
kde X_1 a X_2 jsou stochasticky nezávislé náhodné veličiny, přičemž X_1 je z normovaného normálního rozložení a X_2 je z χ_n^2 rozložení s n stupni volnosti. Toto rozložení pak označujeme jako $t(n)$. Funkce hustoty rozložení pravděpodobnosti t-rozložení o n stupních volnosti je dle [MR03] definována takto:

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\,\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \tag{5.14}$$

pro $x \in (-\infty, \infty)$.

5.3 Model a odhad chování agenta

Pakliže říkáme, že chování agenta v nějakém kontextu α je modelováno pomocí normálního rozložení $N(\mu, \sigma^2)$ a důvěra agenta v kontextu α je odhad tohoto chování, pak na základě dostatečného množství vzorků chování (výsledků interakcí / náhodného výběru) agenta v kontextu α jsme schopni stanovit interval důvěry na základě odhadu parametrů μ a σ^2. Odhad parametrů normálního rozložení z hodnot výběrového souboru se provádí dvěma způsoby: pomocí *bodového odhadu* a pomocí *intervalového odhadu* [Nov06]. V následující podkapitole se zaměříme právě na odhady parametrů μ a σ^2 normálního rozložení, nejdříve si však definujeme vztah mezi modelem chování agenta, chováním agenta, odhadem chování agenta a intervalem důvěry.



(a) Model chování agenta.

(b) Odhad chování agenta.

Obrázek 5.3: Model chování a odhad chování agenta.

5.3.1 Model chování agenta

Model chování agenta je definován pomocí normálního rozložení $N(\mu, \sigma^2)$, tedy parametry μ (*střední hodnota*) a σ^2 respektive jen σ (*rozptyl*, respektive *směrodatná odchylka*). Vzhledem k tomu že, používáme intervalu důvěry v rozsahu 0 až 1, stanovíme si pro parametry μ a σ současně následující podmínky:

$$\begin{aligned} 0 \leq \mu \leq 1 \ , \\ 0 \leq (\mu - 3\sigma) \ , \\ (\mu + 3\sigma) \leq 1 \ . \end{aligned} \tag{5.15}$$

Jinými slovy, střední hodnota musí být v intervalu $[0, 1]$ a trojnásobek směrodatné odchylky od střední hodnoty musí ležet rovněž v intervalu $[0, 1]$. Tyto podmínky zajistí, že z $99,74\%$ (vychází z empirického pravidla viz výše) bude chování agenta v intervalu $[0, 1]$.

Příklad modelu chování agenta je na obrázku 5.3(a) vyznačen červenou křivkou. Jedná se o normální rozložení s parametry $\mu = 0.5$ a $\sigma = 0.1$. Modrý histogram frekvenční analýzy náhodné spojité proměnné X, jež je vytvořena generátorem pseudonáhodných čísel normálního rozložení, odpovídá právě normálnímu rozložení pravděpodobnosti $N(\mu = 0.5, \sigma = 0.1)$ a reprezentuje reálné *chování agenta*. To znamená, že v okamžiku kdy je agent požádán o provedení nějaké interakce v daném kontextu, je na základě znalosti modelu chování agenta v tomto kontextu vygenerováno náhodné číslo dané modelem chování $N(\mu = 0.5, \sigma = 0.1)$. Pro zápis modelu chování agenta lze rovněž použít i intervalový zápis, který odpovídá v případě modelu $N(\mu = 0.5, \sigma = 0.1)$ intervalu chování $[0.2, 0.8]$. Tyto dva zápis tak budeme v následujícím textu považovat za ekvivalentní:

$$N(\mu, \sigma) \equiv \vartheta = [\mu - 3\sigma, \mu + 3\sigma] \ . \tag{5.16}$$

5.3.2 Odhad chování agenta

Pakliže víme, že model chování agenta vzhledem k nějakému kontextu je stanoven normálním rozložením pravděpodobnosti s parametry μ a σ, pak odhad chování takového agenta lze víceméně redukovat na odhad těchto dvou parametrů.

Vzhledem k tomu, že model HMTC používá pro reprezentaci důvěry interval, pak výsledkem odhadu chování agenta by měl být právě interval. Jak jsme předestřeli na začátku

této kapitoly, tak *interval důvěry je chápán jako subjektivní vnímání schopnosti a spolehlivosti s jistou mírou neurčitosti nějakého agenta vzhledem k jinému agentu v nějakém kontextu*. Vzhledem k této formulaci tak můžeme položit rovnost mezi odhadem chování agenta v nějakém kontextu a intervalem důvěry.

Pro model chování $N(\mu, \sigma)$ definujeme odhad chování jako interval důvěry $\hat{\vartheta}$ takto:

$$\begin{aligned} \hat{\vartheta} &= [\hat{\vartheta}_{min}, \hat{\vartheta}_{max}] \quad \text{kde} \\ \hat{\vartheta}_{min} &= \hat{\mu} - 3\hat{\sigma}, \\ \hat{\vartheta}_{max} &= \hat{\mu} + 3\hat{\sigma}. \end{aligned} \tag{5.17}$$

Symbolem $\hat{\vartheta}$ označujeme odhadnutý interval důvěry chování ϑ, $\hat{\mu}$ označuje odhad parametru μ a $\hat{\sigma}$ označuje odhad parametru σ modelu chování $N(\mu, \sigma)$. Jak je tedy vidět, odhad chování agenta je interval důvěry, stanovený na základě odhadů parametrů normálního rozložení, kde hranice intervalu $\hat{\vartheta}$ jsou ve vzdálenosti třech směrodatných odchylek $\hat{\sigma}$ od střední hodnoty $\hat{\mu}$.

Vlastní odhad parametrů a de facto chování agenta je stanoven na základě historie interakcí (což je ze statistického hlediska *konkrétní realizace náhodného výběru*) mezi agenty pomocí *intervalových odhadů* (bude popsáno podrobně v následující podkapitole). Na obrázku 5.3(b) je graficky znázorněn odhad chování agenta na základě výsledků po 100 interakcích (náhodný výběr o rozsahu 100) v daném kontextu pro model chování zobrazeném na obrázku 5.3(a). Na základě odhadnutých parametrů $\hat{\mu}$ a $\hat{\sigma}$ byl stanoven interval důvěry $[0.12913, 0.89622]$, přičemž tento interval by přibližně odpovídal chování definovaném normálním rozložením $N(\mu = 0.512675, \sigma = 0.127848)$ – je zde tedy patrná chyba odhadu (a to jak v případě μ tak i v případě σ).

5.3.3 Odhad parametrů μ a σ z množiny hodnot generovaných normálním rozložením

Jak již bylo řečeno výše, odhad neznámého parametru (obecného) rozložení lze provést dvěma způsoby. První způsob se nazývá *bodový odhad* a je založen na tom, že z hodnot výběrového souboru vypočítáme jedno číslo a to prohlásíme za odhad hledaného parametru. Druhý způsob se nazývá *intervalový odhad*, jehož výsledkem není jedna hodnota ale interval hodnot, přičemž zároveň stanovíme pravděpodobnost s jakou odhadovaný parametr leží v tomto intervalu.

Přesnost a chyba odhadu parametrů rozložení je výrazně ovlivněna počtem hodnot (náhodných výběrů) ze kterých se odhady provádí. Toto tvrzení vychází ze *Zákona velkých čísel*, který lze obecně formulovat takto: „Jestliže zvětšujeme počet nezávislých pokusů, přibližuje se empiricky zjištěná charakteristika, popisující výsledky těchto pokusů, charakteristice teoretické.“ [Nov06].

Toto tvrzení má samozřejmě i zásadní dopad pro námi navrhovaný model, neboť lze jednoznačně říci, že přesnost odhadu chování (tedy stanovení důvěry) agenta je tím vetší, čím více interakcí s ním provedeme. Teoreticky tak po nekonečném množství interakcí bude odhad chování agenta přesně korespondovat s jeho modelem chování. Naopak, v případě malého množství interakcí lze jen velmi těžko odhadnout skutečné chování agenta a s tím odpovídající interval důvěry. Nicméně, naše reprezentace důvěry – interval – nám umožňuje modelovat právě takovou situaci a vypořádat se s různou mírou neurčitosti. Neurčitost důvěry tak bude v ideálním případě s každou další interakcí klesat a efektivita agentů využívající takto stanovené důvěry bude s přibývajícími zkušenostmi vzrůstat.

V následujících podkapitolách si popíšeme oba způsoby odhadů parametrů μ a σ normálního rozložení a následně stanovíme způsob odvození intervalu důvěry z těchto odhadů. Následující definice a formulace jsou převzaty z [MR03, Nov06].

Bodový odhad parametrů

Bodový odhad parametru je hodnota *statistiky*, kterou použijeme pro odhad parametru. Pojem *statistika* (nebo také výběrová charakteristika), je charakteristika (jako např. *průměr*, *rozptyl* apod.) z daného náhodného výběru. Statistiky mají z pravděpodobnostního hlediska charakter náhodné veličiny, neboť jsou vypočteny z hodnot náhodného výběru, které jsou sami hodnotami náhodných veličin.

Střední hodnota

Předpokládejme, že $X = (X_1, X_2, X_3, \ldots, X_n)$ je náhodný výběr s normálním rozložením a $x = (x_1, x_2, x_3, \ldots, x_n)$ jsou její konkrétní realizace. Pak odhad střední hodnoty parametru μ statistického souboru x je definována jako průměr tohoto statistického souboru:

$$\hat{\mu} = \overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{5.18}$$

a $\hat{\mu}$ označujeme též jako *výběrový průměr*.

Rozptyl

Odhad rozptylu σ^2 je stanoven jako součet čtverců odchylek od průměru vydělený faktorem počtu prvků statistického souboru:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 \tag{5.19}$$

a *směrodatná odchylka* je pak jednoduše odmocnina tohoto rozptylu:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}\,. \tag{5.20}$$

Tento odhad $\hat{\sigma}^2$ je však *zkreslený (vychýlený) odhad* parametru σ^2 [Nov06]. To znamená, že se jedná o odhad jehož střední hodnota se liší od skutečné hodnoty odhadovaného parametru. Z tohoto důvodu se pro odhad parametru σ^2 používá *nestranný (nevychýlený, nezkreslený) odhad*, jež se označuje s^2 a jedná se jen o mírně upravený vzorec pro $\hat{\sigma}^2$ (jako dělitel je namísto faktoru n použito faktoru $n-1$):

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \tag{5.21}$$

a tuto hodnotu pak označujeme jako *výběrový rozptyl*. Z praktického hlediska a především pro algoritmický přepis rovnice (snížení počtu matematických operací a tím i snížení časové náročnosti výpočtu, především pro velké n) [MR03] se používá častěji tzv. *výpočetní vzorec výběrového rozptylu* který má tvar:

$$s^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i \right)^2}{n} \right). \tag{5.22}$$

Uvědomme si však, že při používání bodových odhadů parametrů se výsledná hodnota odhadnutých parametrů prakticky vždy liší od skutečných parametrů. Z tohoto hlediska je vhodné získat informaci o přesnosti odhadu, což můžeme provést pomocí tzv. *intervalového odhadu* parametrů.

Intervalový odhad parametrů

Intervalový odhad parametru je odhad pomocí intervalu, který získáme z bodového odhadu parametru a zadáním pravděpodobnosti s jakou parametr leží v tomto intervalu.

Definice 7. *Interval spolehlivosti* (t_{min}, t_{max}) je $100(1-\alpha)$ *procentním intervalem spolehlivosti* pro parametr t, kde $\alpha \in [0,1]$ a pro dvojici statistik t_{min} a t_{max} platí že:

$$P(t_{min} < t < t_{max}) = 1 - \alpha \tag{5.23}$$

pro libovolnou hodnotu parametru t. *Intervalový odhad* parametru t se *spolehlivostí* $1-\alpha$ je interval (t_{min}, t_{max}), kde t_{min}, t_{max} jsou hodnoty statistik na daném statistickém souboru $x = (x_1, ..., x_n)$.

Intervalový odhad nám tedy pomocí *intervalu spolehlivosti* (někdy též označován jako *konfidenční interval* z anglického *confidence interval* – dále už jen zkráceně CI) (t_{min}, t_{max}) vyjadřuje to, že parametr t se nachází v tomto intervalu s pravděpodobností $1 - \alpha$. Toto číslo $(1 - \alpha)$ se též často nazývá *koeficient spolehlivosti* a uvádí se v %.

Koeficient spolehlivosti vyjadřuje to, jak se můžeme spolehnout, že hodnota odhadovaného parametru skutečně leží uvnitř intervalu spolehlivosti. S rostoucím koeficientem spolehlivosti roste i rozpětí intervalového odhadu, platí tedy nepřímá úměra mezi *spolehlivostí* a *přesností odhadu* (intervalu). Při stanovování koeficientů spolehlivosti volíme hodnoty blízké jedné, podle konvence obvykle 0.90 (CI 90%), 0.95 (CI 95%) a 0.99 (CI 99%).

Obecně rozeznáváme dva typy intervalových odhadů: *jednostranné* a *dvoustranné*. Jednostranný odhad intervalu se rozděluje na *levostranný* a *pravostranný*. Pro tyto jednostranné intervaly platí, že horní respektive dolní hranice leží v nekonečnu a odhadujeme pouze dolní, respektive horní hranici. To znamená, že pro dané jednostranné intervaly platí různé podmínky (viz tabulka 5.1).

Typ odhadu	Interval	Podmínka
levostranný	(t_1, ∞)	$P(t > t_1) = 1 - \alpha$
pravostranný	$(-\infty, t_2)$	$P(t < t_2) = 1 - \alpha$

Tabulka 5.1: Jednostranný odhad intervalu a jeho rozložení.

Naproti tomu dvoustranné intervaly spolehlivosti (též označovány jako *symetrické intervaly spolehlivosti*) jsou takové intervaly spolehlivosti (t_1, t_2), které splňují následující podmínku:

$$P(t_1 \leq t) = P(t \leq t_2) = \frac{\alpha}{2}. \tag{5.24}$$

Vzhledem k tomu, že jednostranné intervalové odhady nejsou pro naši aplikaci použitelné (interval důvěry je vždy v rozsahu $[0,1]$ a nikdy ne v nekonečnu), dále už se jen zaměříme na oboustranné symetrické intervaly spolehlivosti.

Interval spolehlivosti pro střední hodnotu při neznámé směrodatné odchylce

Pro stanovení intervalu spolehlivosti normálního rozložení pro střední hodnotu při neznámé směrodatné odchylce se používá bodových odhadů výběrových odhadů ($\overline{x}$, s) a t-rozložení (viz 5.2.2) s $n-1$ stupni volnosti. Interval spolehlivosti pro μ je definován takto:

$$\overline{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} \leq \mu \leq \overline{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\,, \tag{5.25}$$

kde:

- n – je rozsah výběru,
- $\overline{x}$ – odpovídá průměrné hodnotě výběru a je ekvivalentní s bodovým odhadem $\hat{\mu}$,
- s – je *výběrová směrodatná odchylka* stanovená jako odmocnina výběrového rozptylu s^2 dle (5.22),
- $t_{\alpha/2}(n-1)$ – je hodnota určená z t-rozložení pro koeficient spolehlivosti $1-\alpha/2$ s $n-1$ stupni volnosti (často též označován jako $100(1-\alpha/2)$ nebo $(1-\alpha/2)\%$ *kvantil*[1] t-rozložení [MR03, MM02a]).

Kritické hodnoty kvantilů t-rozložení je poměrně komplikované spočítat, neboť se určují pomocí inverzní distribuční funkce tohoto rozložení, jejíž numerické vyčíslení není zcela triviální [Sha06], neboť se podobně jako v případě funkce hustoty tohoto rozložení (viz rovnice 5.14) používá funkce Gamma. Lze použít různé aproximace této funkce [Sha06], nicméně nejčastěji se kritické hodnoty kvantilů tohoto rozložení uvádí pro standardní α nebo $\alpha/2$ kvantily formou tabulky. Tato tabulka A.2 je součástí přílohy A.

Modelový příklad

Pro demonstraci výpočtu intervalových odhadů a pro názornost způsobu odhadu chování agentů jsme si stanovili následující případ, na kterém budou ukázány některé z příkladů využití intervalových odhadů. Uvažujme multi-agentní systém s libovolným počtem agentů, přičemž se zaměříme pouze na dva agenty. Agent A je agent, který poskytuje nějaké služby a agent B je agent, který jej o ně žádá. Model chování agenta A v kontextu β je stanoven normálním rozložením $N(\mu = 0.7, \sigma = 0.2/3)$, což odpovídá *intervalu chování* $\vartheta = [0.5, 0.9]$. Cílem tohoto příkladu je stanovit parametry rozložení pro odhad chování (tedy stanovení $\hat{\vartheta}$) agenta A z pohledu agenta B, který model chování agenta A nezná.

Příklad 5. V průběhu simulace systému agent B požádá agenta A o dvanáct interakcí v rámci kontextu β. Postupně je tedy vygenerována posloupnost výsledků interakcí: {0.754825, 0.655488, 0.699488, 0.715433, 0.724224, 0.754683, 0.613777, 0.692139, 0.739768, 0.676374, 0.693036, 0.625484}. Z této posloupnosti stanovil agent B bodový odhad střední hodnoty $\hat{\mu} = 0.69539325$. Sestrojte 95% interval spolehlivosti pro střední hodnotu μ.

Řešení

- Neznáme rozptyl σ^2, respektive směrodatnou odchylku σ, použijeme proto vzorec (5.25) pro sestrojení intervalu spolehlivosti.

[1] Kvantil, nebo též *výběrový* $\alpha/2$ *kvantil* je hodnota, která rozděluje výběr prvků na dvě části, jedna obsahuje $\alpha/2$ % prvků, které jsou menší (nebo stejné) než tento kvantil, druhá část $(1-\alpha/2)$ % prvků, které jsou větší (nebo stejné) než kvantil [MM02a].

- Výběrový rozptyl stanovíme z výsledku posloupnosti interakcí dle rovnice (5.22): $s^2 = 0.0021622654$, výběrová směrodatná odchylka je $s = \sqrt{(s^2)} = 0.0465001658$.
- Koeficient spolehlivosti je $0.95 = 1 - 0.05$, tedy $\alpha = 0.05$, $\alpha/2 = 0.025$.
- Počet interakcí bylo celkem dvanáct (rozsah výběru), tedy $n = 12$; počet stupňů volnosti $\nu = n - 1 = 11$.
- Stanovíme $t_{\alpha/2}(n-1)$ kvantil t-rozložení dle tabulky A.2 v příloze A: $t_{0.025}(11) = 2.201$.
- Dosazením do vzorce (5.25) dostáváme 95% interval spolehlivosti pro střední hodnotu:

$$0.665848 \leq \mu \leq 0.724938 \, .$$

Výsledek lze interpretovat tak, že pravděpodobnost toho, že interval $[0.665848, 0.724938]$ pokryje správnou hodnotu parametru μ (střední hodnota chování agenta), je rovna 95%. Jak vidíme, tak tento interval opravdu obsahuje skutečnou hodnotu μ, neboť ze zadaní modelového příkladu víme, že střední hodnota je 0.7.

Interval spolehlivosti pro rozptyl

Interval spolehlivosti pro rozptyl normálního rozložení se určuje pomocí takzvaného dolního $\chi^2_{\alpha/2}$ a horního $\chi^2_{1-\alpha/2}$ kvantilu Chí-kvadrátu rozložení (viz 5.2.1) o $n-1$ stupních volnosti a pomocí bodového odhadu výběrového parametru pro rozptyl s^2 takto:

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} \tag{5.26}$$

kde n je rozsah výběru.

Podobně jako v případě t-rozložení se kvantil Chí-kvadrát rozložení určuje pomocí s využitím výpočtu přes inverzní distribuční funkci anebo pomocí odečtu z tabulky s kritickými hodnotami tohoto rozložení. Tabulka tohoto rozložení je uvedena v příloze A, tabulka A.1.

Příklad 6. V návaznosti na předchozí příklad 5 sestrojte 95% interval spolehlivosti pro rozptyl, respektive směrodatnou odchylku chování agenta A.

Řešení

- Ze zadání příkladu víme, že koeficient spolehlivosti je 0.95, $\alpha = 0.05$, $n = 12$, počet stupňů volnosti $\nu = 11$.
- Z řešení předchozího příkladu 5 víme, že výběrový rozptyl je $s^2 = 0.0021622654$.
- Kvantil $\chi^2_{1-\alpha/2}(n-1)$ stanovíme odečtením hodnoty z tabulky A.1, tedy pro horní kvantil $\chi^2_{0.975}(11)$ dostaneme hodnotu $\chi^2_{0.975}(11) = 3.816$.
- Dolní kvantil $\chi^2_{\alpha/2}(n-1) = \chi^2_{0.025}(11) = 21.920$.
- A dosazením do vzorce (5.26) dostáváme 95% interval spolehlivosti pro rozptyl:

$$0.001085 \leq \sigma^2 \leq 0.006233$$

respektive, pod odmocnění i 95% interval spolehlivosti pro směrodatnou odchylku:

$$0.032940 \leq \sigma \leq 0.078952 \, .$$

Závěr je takový, že směrodatná odchylka chování agenta A leží s pravděpodobností 95% v intervalu $[0.032940, 0.078952]$. Vzhledem k tomu, že ze zadání modelového příkladu víme, že směrodatná odchylka je $0.2/3 \doteq 0.066667$, lze tento závěr považovat za platný, neboť skutečná směrodatná odchylka se opravdu nachází ve stanovém intervalu.

5.3.4 Stanovení intervalu důvěry na základě intervalového odhadu chování

Odhad chování agenta byl popsán výše v podkapitole 5.3.2, konkrétně pak pomocí rovnice (5.17). V tomto případě se však vycházelo z bodových odhadů parametrů, kde interval důvěry agenta byl stanoven na základě bodového $\hat{\mu}$ odhadu střední hodnoty a bodového odhadu směrodatné odchylky s, kde za použití *empirického pravidla* byl stanoven interval důvěry takto:

$$\hat{\vartheta} = [\hat{\mu} - 3s, \hat{\mu} + 3s] \ . \tag{5.27}$$

Spodní ϑ_{min} a horní ϑ_{max} mez intervalu důvěry leží tedy v případě bodových odhadů ve vzdálenosti třech směrodatných odchylek s od odhadu střední hodnoty $\hat{\mu}$. Tyto bodové odhady (a nakonec i stanovení intervalu důvěry na jejich základě) se téměř vždy liší od skutečných hodnot. V případě, že známe skutečné hodnoty, lze určit chybu bodového odhadu a následně i chybu odhadu intervalu důvěry. Při praktickém použití v námi navrhovaném modelu však agent, který se snaží odhadnout chování jiného agenta, jeho skutečného chování nezná a tudíž je prakticky nemožné určit chybu takového odhadu a provést jeho korekci.

Naproti tomu v případě intervalového odhadu jsme schopni získat informaci o přesnosti odhadu, respektive, přesnost odhadu si stanovíme jako počáteční podmínku odhadu parametrů pomocí *koeficientu spolehlivosti*. Tento přístup tedy dává agentům schopnost aby si sami určili, jak přesné budou jejich odhady chování ostatních agentů a mohou se tak dynamicky přizpůsobovat různým situacím a prostředím.

Výsledkem intervalového odhadu chování, tak jak bylo popsáno výše (rovnice (5.25) a (5.26)), jsou tedy dva intervaly: jeden pro střední hodnotu a jeden pro rozptyl (respektive směrodatnou odchylku). Označme si tedy spodní limit intervalu odhadnuté střední hodnoty jako $\hat{\mu}_{min}$ a horní limit jako $\hat{\mu}_{max}$. Analogicky poté spodní limit intervalového odhadu rozptylu $\hat{\sigma}^2_{min}$ a horní jako $\hat{\sigma}^2_{max}$ (respektive $\hat{\sigma}_{min}$ a $\hat{\sigma}_{max}$). Výsledný interval důvěry ϑ, respektive jeho dolní ϑ_{min} a horní limit ϑ_{max}, stanovíme jako *nejširší možný odhad* takto:

$$\begin{aligned} \hat{\vartheta}_{min} &= \hat{\mu}_{min} - 3\hat{\sigma}_{max} \\ \hat{\vartheta}_{max} &= \hat{\mu}_{max} + 3\hat{\sigma}_{max} \ . \end{aligned} \tag{5.28}$$

Jak je vidět, tak interval důvěry je rovněž stanoven pomocí empirického pravidla, přičemž pro daný koeficient spolehlivosti je spodní limit (nejnižší možná důvěra) stanoven ve vzdálenosti třech maximálních směrodatných odchylek od minimální střední hodnoty a analogicky, horní limit (nejvyšší možná hodnota důvěry) leží ve vzdálenosti třech maximálních směrodatných odchylek od maximální střední hodnoty.

Tím je zajištěno, že s danou přesností odhadu bude vždy odhadnutý interval důvěry širší, než je skutečný model chování. Tento přístup lze označit jako *pesimistický*, neboť pro danou přesnost máme v agenta vždy *menší důvěru* (ve smyslu větší šíře intervalu, tedy větší nejistoty) a chování agenta tak nebude pomocí tohoto odhadu nadhodnoceno (tzn., nevěříme agentu na základě odhadu důvěry více, než bychom měli). Rovněž je tento přístup stanoven s ohledem na základní myšlenku intervalové reprezentace důvěry v HMTC modelu, kdy říkáme, že s každou další interakcí se interval důvěry zužuje (klesá nejistota) a zpřesňuje (hodnota spodního a horního limitu konverguje ke střední hodnotě).

Příklad 7. Na základě zadání modelového příkladu a na základě výsledků intervalových odhadů střední hodnoty a směrodatné odchylky v příkladu 5 a 6 proveďte stanovení bodového a intervalového odhadu důvěry.

Řešení

Bodový odhad důvěry stanovíme s využitím vzorce (5.27), intervalový odhad s využitím vzorce (5.28).

- Ze znalosti zadání víme, že bodový odhad střední hodnoty $\hat{\mu} = 0.69539325$ a na základě výsledku z příkladu 5 víme, že výběrová směrodatná odchylka je $s = 0.0465001658$. Výsledný bodový odhad důvěry vypadá takto:

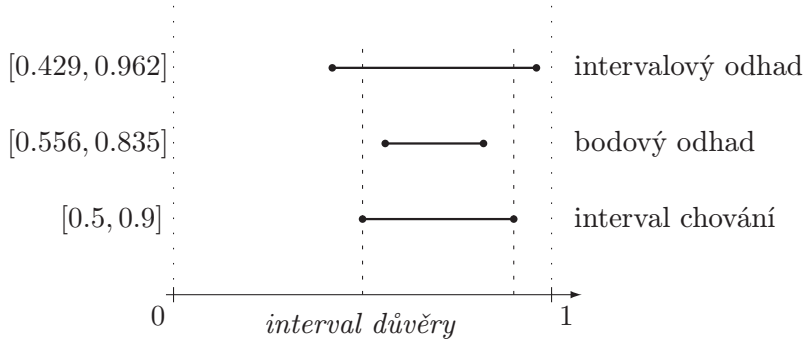
$$\begin{aligned} \hat{\vartheta} &= [\hat{\mu} - 3s, \hat{\mu} + 3s] \\ &\doteq [0.555893, 0.834894] \,. \end{aligned} \tag{5.29}$$

- Intervalový odhad vychází z intervalů stanovených intervalovými odhady pro koeficient spolehlivosti 95% v příkladu 5 a 6. Jednotlivé mezivýsledky jsou tedy následující:

$$\begin{aligned} \hat{\mu}_{min} &= 0.665848 \,, \quad \hat{\mu}_{max} = 0.724938 \\ \hat{\sigma}_{min} &= 0.032940 \,, \quad \hat{\sigma}_{max} = 0.078952 \end{aligned} \tag{5.30}$$

  a s využitím vzorce (5.28) stanovíme intervalový odhad důvěry takto:

$$\begin{aligned} \hat{\vartheta} &= [\hat{\mu}_{min} - 3\hat{\sigma}_{max}, \hat{\mu}_{max} + 3\hat{\sigma}_{max}] \\ &\doteq [0.42899, 0.96179] \,. \end{aligned} \tag{5.31}$$



Obrázek 5.4: Porovnání výsledků bodového a intervalového odhadu důvěry.

Na základě porovnání výsledků z tohoto příkladu vidíme, že výsledný bodový odhad důvěry vytváří podstatně užší interval důvěry než je skutečný interval chování. Tuto skutečnost vhodně demonstruje i obrázek 5.4. Naopak námi navržený pesimistický intervalový odhad je výrazně širší nežli skutečný interval chování. Výhodami, nevýhodami a podrobnou analýzou obou přístupů se zabývá následující podkapitola.

5.4 Analýza navržených metod odhadu intervalu důvěry

V této podkapitole se zaměříme na porovnání výše popsaných metod bodových a intervalových odhadů pro stanovení intervalů důvěry. Porovnání se zaměří především na chybu odhadu důvěry jednotlivých metod vzhledem k počtu provedených interakcí mezi agenty.

5.4.1 Simulační model a metoda analýzy odhadu

Pro analýzu a experimentální ověření navrhovaného způsobu odhadu důvěry na základě chování agenta byla vytvořena simulační aplikace, jejíž hlavní součásti tvoří:

- generátor pseudonáhodných čísel s normálním rozložením s parametricky nastavitelnou střední hodnotou a směrodatnou odchylkou,
- matematická knihovna funkcí pro práci s Chí-kvadrát a t-rozložením, která umožňuje stanovit pro libovolný stupeň volnosti n a koeficient spolehlivosti α příslušný kvantil daného rozložení,
- hlavní výpočetní část programu, která generuje množinu náhodných čísel v normálním rozložení s danými parametry a provádí bodové a intervalové odhady na základě náhodného výběru z množiny generovaných čísel,
- persistentní databáze, ve které jsou uloženy všechny vypočtené výsledky ze všech běhů simulační aplikace.

Experiment. Každý experiment obsahuje celkem 200 iterací[2] náhodných výběrů o daném rozsahu. Jedna iterace provádí celkem 99 náhodných výběrů, tedy v rozsahu $n = 2, \ldots, 100$ (výběr o velikosti $n = 1$ nemá z principu definice bodových a intervalových odhadů smysl). Při každém náhodném výběru jsou vypočteny bodové odhady základních parametrů výběru (střední hodnota, rozptyl, směrodatná odchylka), bodové odhady intervalu důvěry (dle vzorce (5.27)) a intervalové odhady důvěry (dle vzorce (5.28)). Tyto údaje jsou spolu s identifikátorem experimentu, identifikátorem iterace a velikostí rozsahu výběru uloženy do databáze. V rámci jedné iterace je realizováno 99 výběrů s různou velikostí, čemuž odpovídá 99 záznamů do databáze a v rámci jednoho experimentu je vytvořeno celkem 19800 záznamů.

Jednotlivé experimenty se jednoznačně rozlišují identifikátor experimentu, ke každému experimentu (respektive jeho identifikátoru) jsou uloženy informace o parametrech experimentu: model chování (střední hodnota a směrodatná odchylka) a koeficient spolehlivosti pro intervalový odhad. Na základě znalosti skutečného modelu chování a stanoveného odhadu chování jsme schopni určit chybu tohoto odhadu.

5.4.2 Stanovení chyby odhadu

Absolutní chyba odhadu

Základním přístupem pro stanovení chyby při známé hodnotě přibližného čísla a a přesného čísla A je použití tzv. *absolutní chyby*, která je vyjádřena jako absolutní hodnota rozdílu těchto čísel:

$$e = |A - a| \,. \tag{5.32}$$

V případě bodových a intervalových odhadů důvěry tak lze jednoduše stanovit absolutní chybu odhadu důvěry pro jednotlivé limity intervalů následovně:

$$\begin{aligned} e_{\hat{\vartheta}_{min}} &= |\vartheta_{min} - \hat{\vartheta}_{min}| \\ e_{\hat{\vartheta}_{max}} &= |\vartheta_{max} - \hat{\vartheta}_{max}| \,, \end{aligned} \tag{5.33}$$

[2]Hodnota 200 byla stanovena na základě sledování rozdílu střední hodnoty spodního a dolního intervalu důvěry mezi jednotlivými experimenty při stejných vstupních parametrech, přičemž odchylka střední hodnoty sledovaných parametrů byla vždy $< 1\%$ pro libovolnou velikost výběru $n = 2, \ldots, 100$ a koeficient spolehlivosti $(1 - \alpha) \geq 0.9$.

kde $\hat{\vartheta}_{min}$ je spodní (respektive $\hat{\vartheta}_{max}$ horní) limit odhadnutého intervalu důvěry a ϑ_{min} je spodní (respektive ϑ_{max} horní) limit skutečného intervalu důvěry, jež je stanoven na základě známého modelu chování $N(\mu, \sigma)$ takto:

$$\begin{aligned} \vartheta_{min} &= \mu - 3\sigma \\ \vartheta_{max} &= \mu + 3\sigma \ , \end{aligned} \tag{5.34}$$

kde pro parametry μ a σ musí platit podmínky (5.15).

Absolutní chybou jsme schopni určit chybu konkrétního náhodného výběru pro danou iteraci daného experimentu – rovnice (5.33). Vzhledem k tomu, že jednotlivé vstupní množiny výběrů jsou generovány pseudonáhodně, je vždy stanovení chyby jednoho výběru zatíženo nějakou nepřesností. Z tohoto důvodu jsme pro každou velikost náhodného výběru provedli více iterací a na jejich základě následně stanovili *průměrnou absolutní chybu* (*mean absolute error* – MAE) pro danou velikost výběru daného experimentu. Průměrnou absolutní chybu určíme následovně:

$$\overline{e} = \frac{1}{m} \sum_{i=1}^{m} |A_i - a_i| = \frac{1}{m} \sum_{i=1}^{m} e_i \ , \tag{5.35}$$

kde m je počet iterací provedených pro danou velikost náhodného výběru a e_i je absolutní chyba konkrétního náhodného výběru (pro A_i skutečnou hodnotu parametru a a_i odhadnutou hodnotu parametru).

Výsledný výpočet *průměrné absolutní chyby odhadu důvěry* pro velikost výběru n při provedení m iterací označíme jako E_n^a a vypočteme ho následovně:

$$E_n^a = \frac{1}{m} \sum_{i=1}^{m} e_i \ , \tag{5.36}$$

kde e_i je absolutní chyba odhadu důvěry pro daný výběr a iteraci i stanovená jako součet absolutní chyby spodního a horního odhadu v dané itaraci:

$$e_i = e_{i,\hat{\vartheta}_{min}} + e_{i,\hat{\vartheta}_{max}} \ . \tag{5.37}$$

Kritická chyba odhadu důvěry

Jak již bylo popsáno v předchozí podkapitole 5.3.4, tak stanovení intervalu důvěry na základě intervalových odhadů je dáno *nejširším možným odhadem* – vzorec (5.28) – tak, abychom odhadnutým intervalem důvěry „nenadhodnocovali“ chování agenta a zároveň aby byl respektován princip HTMC, kdy je postupně pomocí průniků jednotlivých intervalů chování agenta zpřesňováno.

Pro popis kritické chyby si nyní stanovíme význam pojmů *nadhodnocený* a *podhodnocený* odhad. Nadhodnoceným odhadem důvěry máme na mysli takový interval důvěry, pro který platí že:

$$\hat{\vartheta}_{min} > \vartheta_{min} \quad \text{nebo} \quad \hat{\vartheta}_{max} < \vartheta_{max} \ . \tag{5.38}$$

Jedná se tedy o odhad intervalu důvěry, kdy je spodní limit odhadu větší než skutečný spodní limit anebo kdy je horní limit odhadu menší nežli skutečný horní limit. Jinak řečeno, průnik odhadnutého intervalu se skutečným intervalem není ekvivalentní se skutečným intervalem: $\hat{\vartheta} \cap \vartheta \neq \vartheta$. Takový odhad důvěry může způsobit to, že důvěra v agenta je vyšší

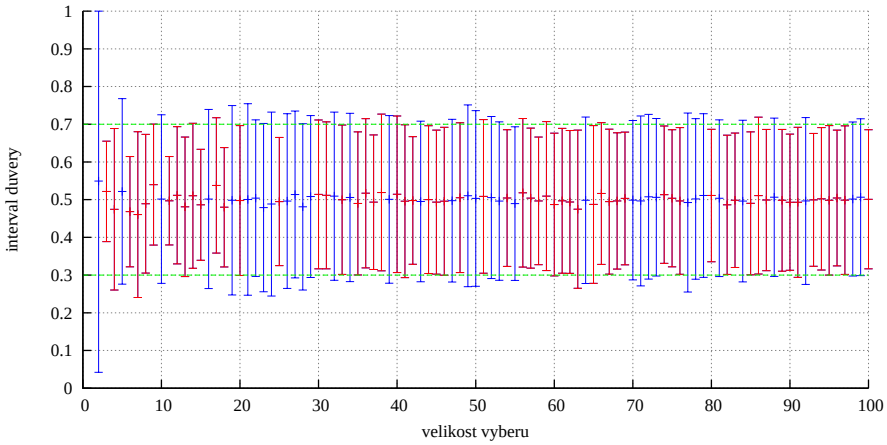
(buď na základě spodního nebo horního odhadu), než by měla být, přičemž bez vyvolání HMTC konfliktu ji není možné snížit.

Naopak *podhodnocený* odhad důvěry je takový odhad, pro který platí že:

$$\hat{\vartheta}_{min} \leq \vartheta_{min} \quad \text{a zároveň} \quad \hat{\vartheta}_{max} \geq \vartheta_{max} \,, \tag{5.39}$$

tedy průnikem odhadnutého intervalu se skutečným intervalem dostáváme interval skutečný: $\hat{\vartheta} \cap \vartheta = \vartheta$. Speciálním případem, který rovněž spadá i do množiny podhodnocených odhadů je tak i odhad intervalu, který je zcela přesný a ekvivalentní s intervalem skutečným.

Vzhledem k výše popsanému je použití absolutní chyby odhadu důvěry, tak jak je definována výše ne zcela vypovídající, neboť jak již sám název napovídá, jedná se o absolutní chybu bez ohledu na to, zda je chování agenta nadhodnoceno nebo podhodnoceno. Přičemž nadhodnocení je z našeho pohledu kriticky chybné a naopak podhodnocení je přípustné anebo dokonce žádoucí.



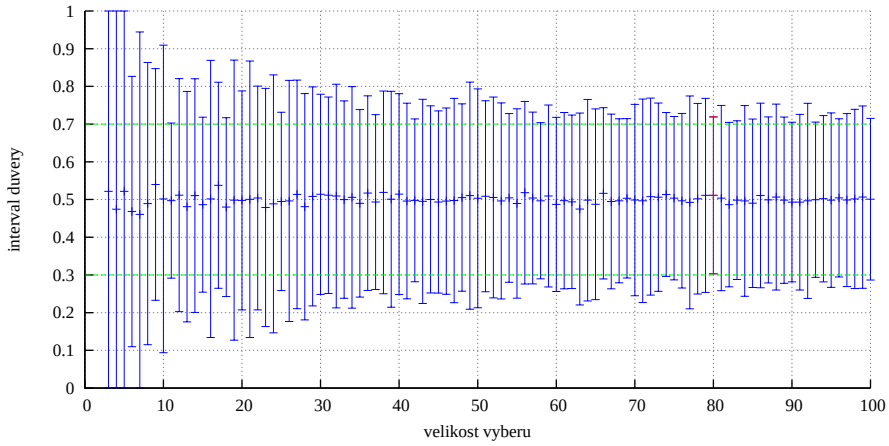
Obrázek 5.5: Bodový odhad důvěry.

Konkrétní případy *pod/nad*-hodnoceného odhadu důvěry jsou znázorněny na obrázcích 5.5 a 5.6. Jedná se o grafy zobrazující šíři odhadnutého intervalu důvěry jedné iterace s realizací 99 náhodných výběrů v rozsahu $n = 2 \ldots 100$ pro model chování $N(\mu = 0.5, \sigma = 0.2/3)$. Graf 5.5 reprezentuje zobrazení výsledných intervalů důvěry stanovených pouze bodovým odhadem a graf 5.6 reprezentuje výsledné intervaly důvěry stanovené na základě intervalového odhadu (pro $CI95\%$). Oba grafy odpovídají identické množině náhodných výběrů.

Modré intervaly důvěry odpovídají podhodnocenému, tedy přípustnému odhadu. Červené naopak odpovídají nadhodnocenému, tedy kriticky chybnému odhadu. Skutečná šíře intervalu důvěry odpovídá rozsahu $[0.3, 0.7]$ a jeho meze jsou vyznačeny v grafech zeleně. Jak je z grafů na první pohled patrné, tak bodový odhad se vyznačuje menší absolutní chybou ale podstatně větší kritickou chybou, neboť značná část výsledných intervalů důvěry je nadhodnocena.

Naopak v případě intervalového odhadu je vidět podstatně širší interval (tedy i větší absolutní chyba), ale pouze jeden nadhodnocený (kriticky chybný) odhad důvěry a to v případě výběru o velikosti $n = 80$.

Právě na základě těchto skutečností tak stanovíme *průměrnou kritickou chybu odhadu*



Obrázek 5.6: Intervalový odhad důvěry.

intervalu důvěry pro velikost výběru n při provedení m iterací:

$$E_n^k = \frac{1}{m} \sum_{i=1}^{m} (\varepsilon_{i,\hat{\vartheta}_{min}} + \varepsilon_{i,\hat{\vartheta}_{max}}) \,, \tag{5.40}$$

kde $\varepsilon_{i,\hat{\vartheta}_{min}}$ označuje spodní kritickou chybu odhadu v iteraci i pro velikost výběru n a $\varepsilon_{i,\hat{\vartheta}_{max}}$ horní kritickou chybu odhadu a jsou definovány takto:

$$\begin{aligned} \varepsilon_{i,\hat{\vartheta}_{min}} &= K(\hat{\vartheta}^i_{min} - \vartheta_{min}) \\ \varepsilon_{i,\hat{\vartheta}_{max}} &= K(\vartheta_{max} - \hat{\vartheta}^i_{max}) \,, \end{aligned} \tag{5.41}$$

kde funkce $K(x)$ označuje *funkci kritické chyby*, jež je definována takto:

$$K(x) = \begin{cases} x & \text{pro } x \geq 0 \\ 0 & \text{pro } x < 0 \,. \end{cases} \tag{5.42}$$

5.4.3 Experimentální výsledky

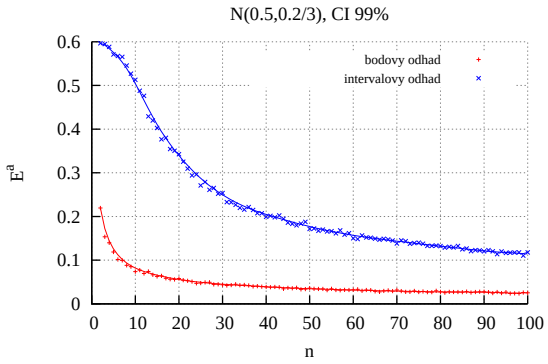
S využitím výše popsaných metod pro stanovení chyby odhadu je v následujících podkapitolách popsána sada experimentů, která pro různé koeficienty spolehlivosti a modely chování zobrazuje graficky průběh absolutní chyby a průměrné kritické chyby pro různé velikosti náhodných výběrů.

Jak již bylo popsáno výše, velikost náhodných výběrů byla zvolena v rozsahu $n = 2, \ldots, 100$, přičemž z praktického hlediska hodnota n odpovídá počtu jednosměrných interakcí mezi dvěma agenty v rámci jednoho kontextu. Pro naznačení spojitého průběhu (je-li použit) grafické reprezentace experimentálních dat byla zvolena Beziérova aproximace s využitím nástroje `gnuplot`.

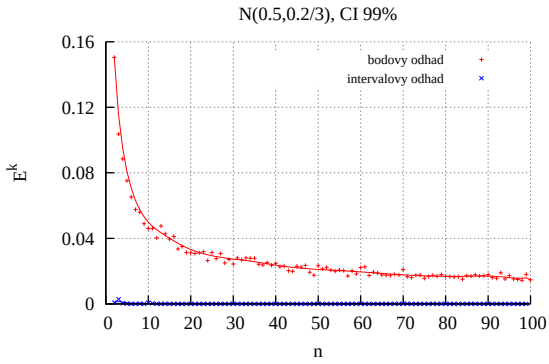
Experiment 1: Vliv koeficientu spolehlivosti na chybu odhadu

V tomto experimentu sledujeme průměrnou absolutní a kritickou chybu odhadů pro různé koeficienty spolehlivosti, parametry experimentu jsou následující:

- konstantní model důvěry $N(\mu = 0.5, \sigma = 0.2/3)$, což odpovídá skutečnému intervalu chování $[0.3, 0.7]$,
- různé koeficienty spolehlivosti $(1 - \alpha) \in \{0.99, 0.95, 0.90\}$.

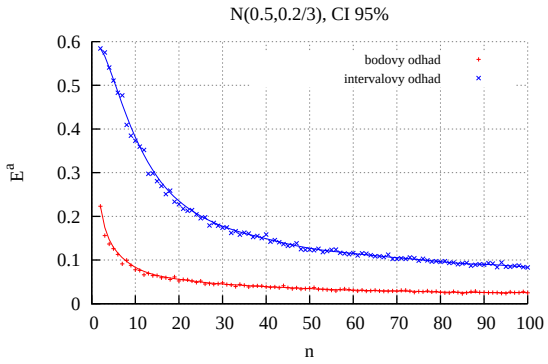


(a) Průměrná absolutní chyba – E^a.

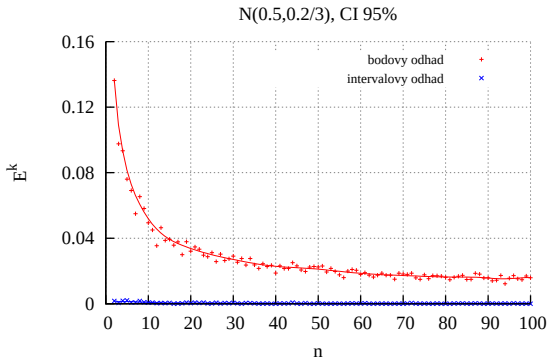


(b) Průměrná kritická chyba – E^k.

Obrázek 5.7: Chyby odhadu důvěry pro CI 99%.



(a) Průměrná absolutní chyba – E^a.

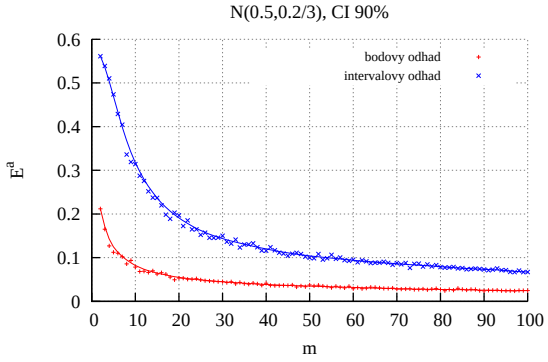


(b) Průměrná kritická chyba – E^k.

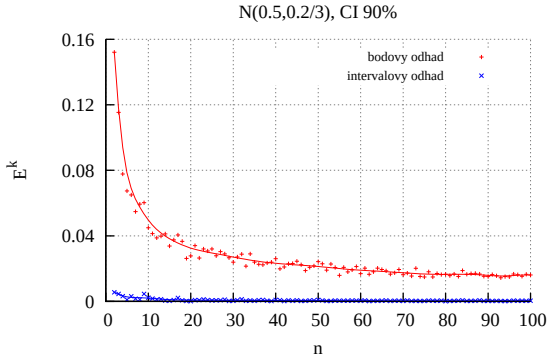
Obrázek 5.8: Chyby odhadu důvěry pro CI 95%.

Na obrázcích 5.7(a)–5.9(b) jsou zachyceny průběhy průměrné absolutní a průměrné kritické chyby pro bodové i intervalové odhady s různými koeficienty spolehlivosti. Jak je vidět v případě intervalového odhadu, tak koeficient spolehlivosti má vliv jak na absolutní tak i na kritickou chybu. Čím větší bude koeficient spolehlivosti (a tedy i šíře intervalového odhadu parametru rozložení), tím větší je absolutní chyba odhadu důvěry (tzn. odhadnutý interval důvěry pomaleji konverguje ke skutečné hodnotě), ale kritická chyba intervalového odhadu důvěry je minimální.

Zaměříme-li se tedy pouze na porovnání intervalových odhadů pro různé koeficienty spolehlivosti získáme graf na obrázku 5.10, který zobrazuje průběh průměrné absolutní chyby intervalového odhadu důvěry pro různé koeficienty spolehlivosti blízké hodnotě 1.



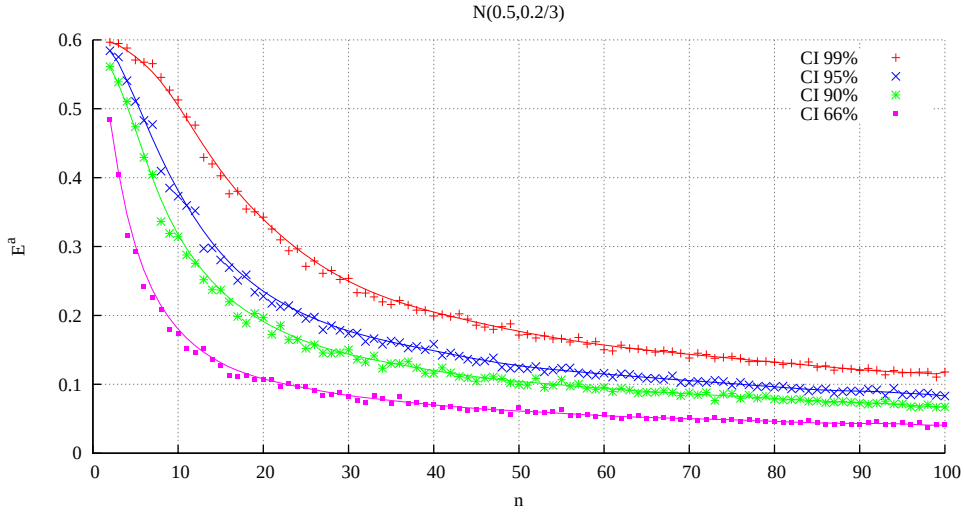
(a) Průměrná absolutní chyba – E^a.



(b) Průměrná kritická chyba – E^k.

Obrázek 5.9: Chyby odhadu důvěry pro CI 90%.

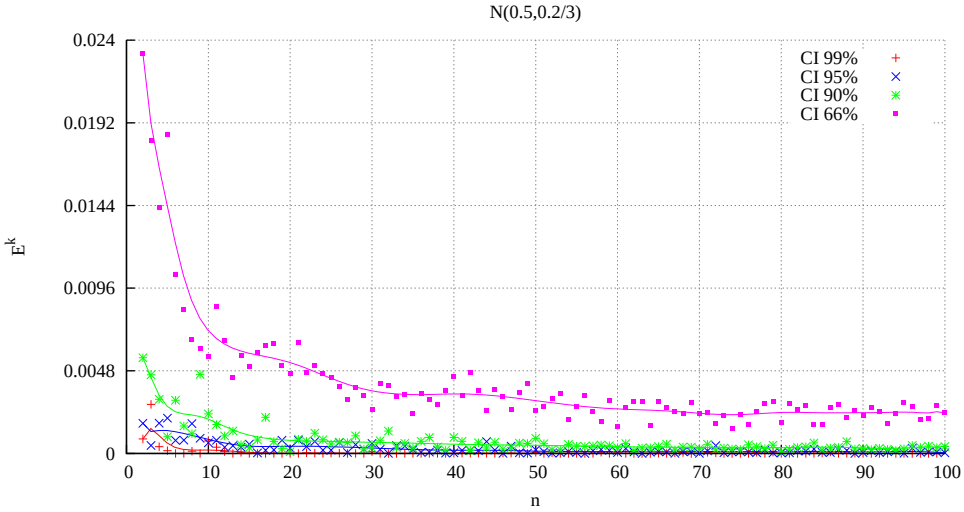
Pro představu o průběhu absolutní chyby s výrazně nižším koeficientem spolehlivosti byl doplněn koeficient $(1 - \alpha) = 0.66$.



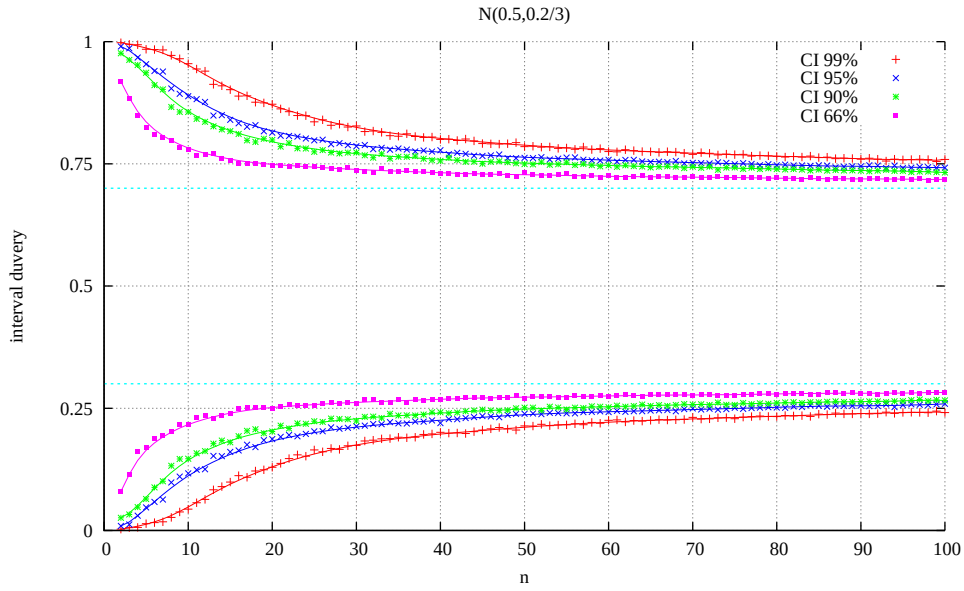
Obrázek 5.10: Průměrná absolutní chyba pro různé koeficienty spolehlivosti.

Další graf na obrázku 5.11 reprezentuje průběh průměrné kritické chyby intervalového odhadu důvěry pro různé koeficienty spolehlivosti, podobně jako předchozí graf, jsou zde zachyceny koeficienty spolehlivosti 0.99, 0.95, 0.90 a 0.66.

Na posledním grafu na obrázku 5.12 je znázorněn průběh průměrné hodnoty spodního a horního limitu ($\hat{\vartheta}_{min}$ a $\hat{\vartheta}_{min}$) odhadu intervalu důvěry pro různé velikosti výběru n stanovené pomocí intervalového odhadu pro různé koeficienty spolehlivosti. Osa Y reprezentuje interval důvěry v rozsahu $[0, 1]$, azurová čára v limitech 0.3 a 0.7 znázorňuje skutečný interval důvěry (ϑ_{min} a ϑ_{min}) daný modelem chování $N(0.5, 0.2/3)$.



Obrázek 5.11: Průměrná kritická chyba pro různé koeficienty spolehlivosti.

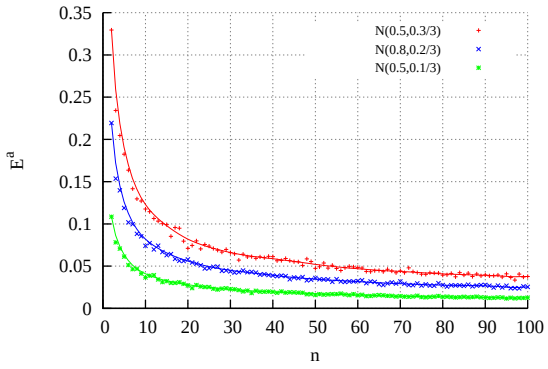


Obrázek 5.12: Intervalový odhad důvěry pro různé koeficienty spolehlivosti.

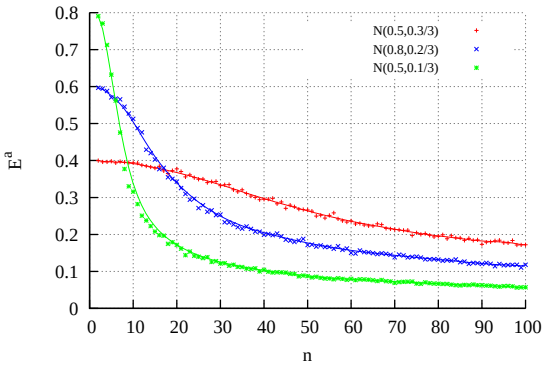
Experiment 2: Vliv velikosti rozptylu modelu chování na chybu odhadu

Velikost rozptylu (σ^2), respektive směrodatné odchylky (σ), u modelu chování je přímo úměrná šíři skutečného intervalu chování. Následující experimenty jsou zaměřeny na zjištění toho, jakým způsobem ovlivňuje velikost rozptylu, potažmo šíře intervalu chování, chyby bodových a intervalových odhadů pro různé velikosti výběrů.

Grafy na obrázcích 5.13(a) a 5.13(b) zachycují průběh průměrné absolutní chyby bodového odhadu a intervalového odhadu pro různé modely chování, přičemž střední hodnota je pro všechny modely shodná a mění se jen velikost rozptylu. Pro tento typ experimentu byly



(a) E^a pro bodový odhad.



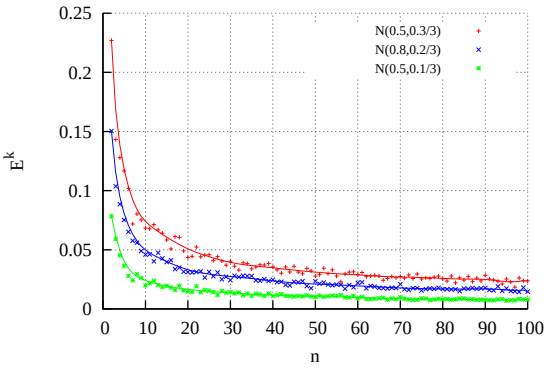
(b) E^a pro intervalový odhad.

Obrázek 5.13: Průměrná absolutní chyba odhadu důvěry pro různé šíře intervalu chování.

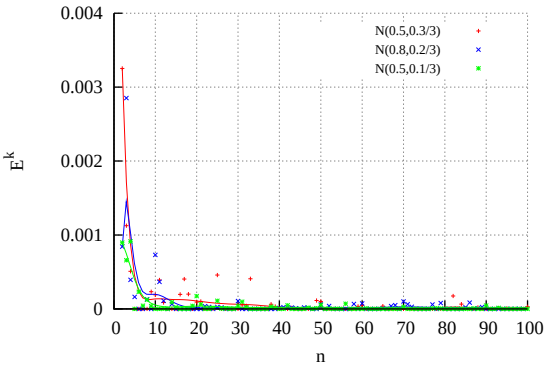
použity modely chování a jejich ekvivalentní intervalové reprezentace popsané v tabulce 5.2, koeficient spolehlivosti je pro všechny experimenty stanoven na hodnotu $(1 - \alpha) = 0.99$.

Model chování	Interval chování
$N(0.5, 0.3/3)$	$[0.2, 0.8]$
$N(0.5, 0.2/3)$	$[0.3, 0.7]$
$N(0.5, 0.1/3)$	$[0.4, 0.6]$

Tabulka 5.2: Modely chování a jejich intervalová reprezentace pro provedené experimenty.



(a) E^k pro bodový odhad.

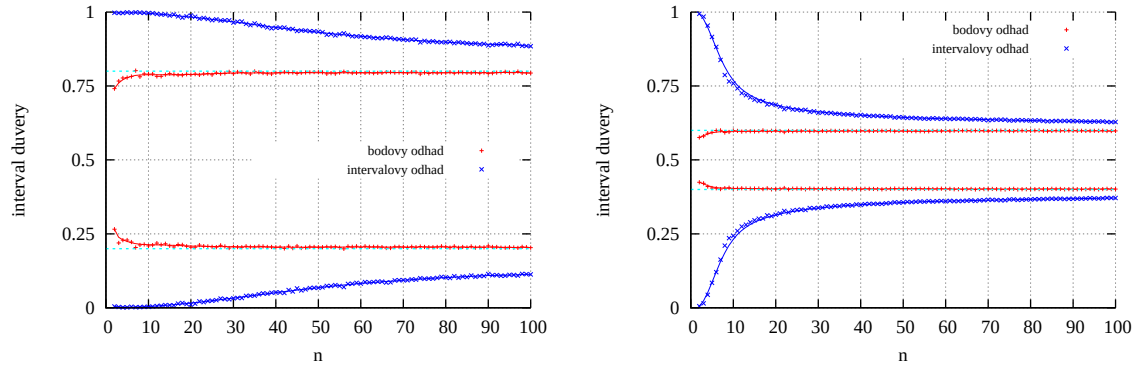


(b) E^k pro intervalový odhad.

Obrázek 5.14: Průměrná kritická chyba odhadu důvěry pro různé šíře intervalu chování.

Další dva grafy na obrázcích 5.14(a) a 5.14(b) zachycují průměrnou kritickou chybu pro různé modely chování. Zde je vhodné upozornit především na fakt, že kritická chyba je v případě bodových odhadů vyšší prakticky o dva řády s porovnáním s intervalovými odhady. Navíc při porovnání průměrné absolutní a kritické chyby bodových odhadů – grafy 5.13(a) a 5.14(a) – lze konstatovat, že jsou velmi podobné a ve stejných řádech (přičemž v případě

intervalových odhadů je situace zcela odlišná).



(a) Odhad intervalu důvěry pro model $N(0.5, 0.3/3)$. (b) Odhad intervalu důvěry pro model $N(0.5, 0.1/3)$.

Obrázek 5.15: Průměrná hodnota odhadu ϑ_{min} a ϑ_{max} pro různé šíře intervalu chování.

Poslední dva grafy tohoto experimentu na obrázcích 5.15(a) a 5.15(b) zachycují průměrnou hodnotu odhadu spodního a horního limitu intervalů důvěry pro dva různé modely chování: $N(0.5, 0.3/3)$ a $N(0.5, 0.1/3)$. Modrou barvou jsou zachyceny intervalové odhady chování a červenou naopak bodové odhady chování. Z výsledných grafů je zcela evidentní, že rozptyl, respektive šíře intervalu mají významný vliv především v případě intervalových odhadů na rychlost konvergence odhadované důvěry k intervalu chování.

5.4.4 Zhodnocení experimentálních výsledků

Z výše provedených experimentů, které postihují problematiku analýzy chyb odhadů intervalu důvěry vzhledem k různým koeficientům spolehlivosti intervalových odhadů, lze stanovit několik důležitých závěrů. Jednotlivé závěry jsou stanoveny s ohledem na princip navrženého HMTC modelu popsaného v kapitole 4.2.

V následujících bodech budou popsány zásadní závěry vycházející z výsledků experimentů provedených v předchozí kapitole na zkoumaném vzorku dat.

Závěr 1. *Stanovení intervalu důvěry na základě bodového odhadu parametrů se jeví spolu s použitím modelu HMTC jako naprosto nevhodné.*

Bodový odhad důvěry, stanovený dle vzorce (5.17), tedy pouze na základě vypočtených odhadů $\hat{\mu}$ a s, se na zkoumaném vzorku dat a oproti intervalovému odhadu vyznačuje vždy nižší průměrnou absolutní chybou, nicméně výrazně vyšší průměrnou kritickou chybou. Kritická chyba, tedy nadhodnocení důvěry, je v případě použití modelu HMTC poměrně zásadní, neboť výrazné zúžení intervalu důvěry pod hranice skutečného intervalu chování agenta má výrazný vliv na schopnost správného výběru partnera k další interakci. Jedinou možností jak interval důvěry v rámci daného kontextu opět rozšířit je odstranění událostí v rámci HMTC struktury vyvoláním konfliktu.

Navíc, podíváme-li se na průběh odhadu intervalu důvěry pomocí bodových odhadů s porovnáním intervalových odhadů (obrázky 5.15(a) a 5.15(b)) vidíme, že v případě malých rozsahů vytváří bodový odhad silně nadhodnocené (zúžené) intervaly důvěry, které se pak následně s větší velikostí výběru rozšiřují. Tento princip tedy jde přímo proti myšlence

HMTC která je založena na tom, že s každou další informací interval zužujeme směrem ke skutečné hodnotě intervalu chování.

Závěr 2. *Koeficient spolehlivosti v případě intervalového odhadu intervalu důvěry výrazným způsobem ovlivňuje průměrnou absolutní i průměrnou kritickou chybu.*

Jak je patrné z grafů na obrázcích 5.10 a 5.11, při volbě koeficientu spolehlivosti blízkému hodnotě 1.0 (tedy CI 100%) vzniká při intervalovém odhadu největší absolutní chyba. Naopak v případě kritické chyby je při takto zvoleném koeficientu spolehlivosti kritická chyba nejmenší – tzn., že odhadnutý interval důvěry je prakticky vždy podhodnocen. Tato skutečnost nicméně vychází ze samotného principu intervalového odhadu, neboť jak jsme se již zmínili dříve, mezi spolehlivostí a přesností odhadu existuje nepřímá úměrnost a s rostoucím koeficientem spolehlivosti roste i rozpětí intervalového odhadu.

Šíře odhadnutého intervalu důvěry v případě vyššího koeficientu spolehlivosti tak podstatně pomaleji konverguje ke skutečnému intervalu chování daným modelem chování. Zároveň však pravděpodobnost, že odhadnutý interval důvěry bude nadhodnocen je nižší než v případě menšího koeficientu spolehlivosti (což však neplatí v případě malých rozsahů výběru).

Závěr 3. *Velikost rozptylu modelu chování, respektive šíře intervalu chování, má velmi výrazný vliv především na absolutní chybu a v případě intervalového odhadu i na rychlost konvergence odhadnutého intervalu důvěry ke skutečnému intervalu chování.*

Rozptyl modelu chování má v případě intervalového odhadu zásadní vliv na to, jak rychle lze s malou chybou odhadnout velikost intervalu důvěry. Jinak řečeno, čím menší bude rozptyl chování, tím méně interakcí mezi agenty je potřeba k tomu, abychom přesně odhadli jeho chování. Jak je vidět na obrázku 5.13(b) a 5.15(b), velikost absolutní chyby intervalového odhadu je pro malý rozptyl a pro malé výběry (malé množství interakcí) velká, nicméně s každou další novou interakcí velmi rychle klesá a pro velké výběry ($n > 20$) je i pro velký koeficientu spolehlivosti $(1 - \alpha) = 0.99$ výrazně menší než pro velké rozptyly.

Závěr 4. *V případě malých výběrů je jak absolutní tak kritická chyba největší u obou způsobů odhadů intervalu důvěry.*

Tento závěr má pro celý návrh odhadu intervalu důvěry, jež je popsaný v této kapitole, veliký význam. Na základě výsledků experimentů bylo zjištěno, že pro rozsahy výběru $n = 2, \ldots, 4$ se jak bodové tak intervalové odhady téměř pro jakékoliv parametry modelu chování a koeficienty spolehlivosti vyznačují velmi velkou absolutní i kritickou chybou. To v důsledku znamená, že použití těchto odhadů je po dvou až čtyřech interakcích pro odhad intervalu důvěry pro model HMTC nevhodné.

Pro rozsah výběru $n = 5, \ldots, 10$ je stále absolutní i kritická chyba na významných hladinách, nicméně intervalový odhad, pro vysoké koeficienty spolehlivosti ($\alpha \geq 95\%$) a malé rozptyly ($\sigma^2 \leq (0.1/3)^2$), je možné efektivně použít při akceptování určité malé kritické chyby ($\varepsilon < 0.05$).

Výše popsané závěry jsou zahrnuty do výsledného návrhu multi-agentního systému využívajícího HMTC, intervalu důvěry a intervalových odhadů pro rozhodování agentů na základě důvěry. Konkrétně posledním, pro navržený systém zcela zásadním, závěrem č. 4 se navíc ještě zabývá následující podkapitola, která popisuje a navrhuje možnosti pro řešení situace v případě malých výběrů.

5.4.5 Analýza a řešení malých výběrů

V první řadě je třeba si uvědomit, že celý princip návrhu stanovení modelu chování a na jeho základě stanovení odhadů intervalu důvěry vůbec nepopisuje případ, který nastává bezprostředně po první interakci. Tedy situaci pro velikost výběru $n = 1$, pro který není ani bodový ani intervalových odhad definován, protože z jedné hodnoty není možné jakoukoliv relevantní statistiku (např. výběrový průměr či rozptyl) spočítat.

Tato podkapitola se zabývá návrhem řešení situace, kdy nelze bodový nebo intervalový odhad použít, neboť není pro daný rozsah definován (situace $n = 1$) a pro situace, kdy bodový nebo intervalový odhad nelze použít z důvodu vysoké kritické chyby ($\varepsilon \geq 0.05$), tedy pro výběry velikosti $n = 2, \ldots, 4$.

Analýza malých výběrů

U malých výběrů jsou závěry vždy zatíženy značnou mírou nejistoty. V odborné literatuře [MM02a] se setkáváme s různými robustními metodami odhadů parametrů pro zajištění maximální korektnosti výsledků. Nejčastěji používanou metodu je tzv. *Hornův postup* [HPC98], kdy pro malé výběry $4 \leq n \leq 20$ je intervalový odhad založen na *pořádkových statistikách* (seřadíme náhodnou veličinu vzestupně) a na základě *hloubky pivotu* a *pivotové polosumy* určíme interval spolehlivosti – více viz [HPC98, MM02a, Hol08].

Pro stanovení střední hodnoty a pro *zvláště malé výběry*, kdy $n = 2 \ldots 3$, se používají specifické funkce pro výpočet intervalového odhadu, které jsou pro normální rozložení založeny např. na goniometrických funkcích [MM02a].

Pro $n = 1$ je pak např. v práci [Hol08] pro výpočet intervalového odhadu střední hodnoty μ použito odvození na základě hustoty pravděpodobnosti normálního normovaného rozložení, přičemž výsledný vzorec má tvar:

$$x - \zeta|x| \leq \mu \leq x + \zeta|x| \,, \tag{5.43}$$

kde ζ je tabelována nebo numericky řešena dle [Hol08] pro různé hodnoty α (stupně spolehlivosti).

V literatuře lze najít i další robustní metody pro odhad rozptylů, jež jsou nezávislé na parametrech rozložení dat (neparametrické metody). Jako příklad lze uvést techniky *Bootstrap* a *Jackknife* [SvB84].

Výše popsané techniky však lze pro náš konkrétní případ považovat za příliš komplexní statistické techniky, neboť mimo jiné počítají i s tím, že známe pouze typ rozložení pravděpodobnosti a žádné další parametry. Nicméně z předchozích podkapitol je zřejmé, že námi používané normální rozložení pravděpodobnosti, jehož odhady parametrů hledáme má jistá omezení, které je dáno maximální šíří intervalu důvěry. Jak tedy víme, dle definice intervalu důvěry popsaného v kapitole 4.2.1 a na základě podmínky pro definici modelu chování (5.15), tak interval důvěry je stanoven v maximálním rozsahu $0 \ldots 1$, čemuž musí odpovídat i model chování, kde střední hodnota musí být v intervalu $[0, 1]$ a trojnásobek směrodatné odchylky od střední hodnoty musí ležet rovněž v tomto intervalu. Námi navrhované řešení tak bude zaměřeno na zohlednění této skutečnosti, přičemž namísto stanovení stupně spolehlivosti bude parametrem našeho odhadu intervalu *odhad neurčitosti* modelu chování, tj. odhad maximální šíře intervalu chování. Pakliže předem známe maximální šíři intervalu chování, kterou lze označit jako neurčitost u, známe na základě definice modelu chování

(5.17) i rozptyl σ^2, respektive směrodatnou odchylku σ:

$$\sigma = \frac{1}{6}u \ . \tag{5.44}$$

Na základě této znalosti lze pro řešení malých výběrů použít intervalový odhad pro střední hodnotu při známém rozptylu a velmi vysokém koeficientu spolehlivosti tak, abychom minimalizovali kritickou chybu.

Funkce pro řešení malých výběrů

Interval spolehlivosti pro střední hodnotu při známém rozptylu je stanoven na základě kvantilu normovaného normálního rozložení dle [Nov06] takto:

$$\underbrace{\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\hat{\mu}_{min}} \leq \mu \leq \underbrace{\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\hat{\mu}_{max}} \ , \tag{5.45}$$

kde $z_{\alpha/2}$ je $(\alpha/2)$-kvantil normálního normovaného rozložení stanovený na základě inverzní distribuční funkce normálního normovaného rozložení (vybrané kritické hodnoty těchto kvantilů lze nalézt v tabulce 5.3). Tento způsob řešení tedy předpokládá, že známe nebo jsme odhadli maximální hodnotu neurčitosti intervalu chování u. Odhad této neurčitosti budeme označovat $\hat{u}$ a z něho dle vzorce (5.44) stanovíme hodnotu σ pro intervalový odhad.

α	0.1	0.01	0.001	0.0001
z_α	1.2816	2.3263	3.0902	3.7190

Tabulka 5.3: Kritické hodnoty normovaného normálního rozložení N(0,1).

Postup pro nalezení intervalu důvěry $\hat{\vartheta} = [\hat{\vartheta}_{min}, \hat{\vartheta}_{max}]$ na základě intervalového odhadu střední hodnoty pro známý rozptyl a pro malé výběry $n = 1, \ldots, 4$ s předem odhadnutou maximální hodnotou neurčitosti chování $\hat{u}$ a 99.99% intervalem spolehlivosti:

1. Pro velikost výběru $n = 1$ (tedy jeden konkrétní výsledek x) stanovíme interval důvěry takto:

$$\begin{aligned} \hat{\vartheta}_{min} &= x - u \\ \hat{\vartheta}_{max} &= x + u \ . \end{aligned} \tag{5.46}$$

2. Pro velikost výběru $n = 2, \ldots, 4$ stanovíme interval důvěry (v souladu se vzorcem (5.45)) takto:

$$\begin{aligned} \hat{\vartheta}_{min} &= \hat{\mu}_{min} - 3\sigma \\ \hat{\vartheta}_{min} &= \hat{\mu}_{max} + 3\sigma \ , \end{aligned} \tag{5.47}$$

   kde $\sigma = 1/6u$ (dle (5.44)), $\hat{\mu}_{min}$ a $\hat{\mu}_{max}$ jsou stanoveny na základě intervalového odhadu dle (5.45).

Experimentálně jsme ověřili, že v případě kdy zvolíme α blízké nule, respektive interval spolehlivosti $(1-\alpha)$ pro tento kvantil blízký hodnotě jedna (tedy konkrétně 99.99% interval spolehlivosti, jemuž odpovídá kvantil $z_\alpha = 3.7190$), jsme schopni dosáhnout pro velikost výběrů $n = 1, \ldots, 4$ nulové kritické chyby ε (při provedení jednoho tisíce iterací pro každý výběr) pro všechny relevantní hodnoty neurčitosti $u \in [0, 1]$ za podmínky že $\hat{u} \geq u$.

5.5 Shrnutí

V této kapitole jsme přesně stanovili význam intervalu důvěry pro budoucí návrh prototypu distribuovaného multi-agentního systému využívající model důvěry HMTC. Interval důvěry tak udává rozložení pravděpodobnosti výsledku budoucího chování agenta v konkrétním kontextu. Chování agenta je stanoveno parametry normálního rozložení: *střední hodnotou* a *rozptylem*, přičemž střední hodnota udává typické chování agenta a rozptyl udává přípustnou odchylku od tohoto chování. Princip stanovení intervalu důvěry (v rámci konkrétního kontextu), respektive stanovení odhadu chování agenta, lze pak popsat statistickými metodami pro stanovení parametrů normálního rozložení na základě realizace náhodných výběrů (konkrétní výsledky interakcí mezi agenty) z tohoto rozložení.

Byly navrženy a popsány dvě metody stanovení intervalu důvěry, založené na *bodových* a *intervalových* odhadech parametrů normálního rozložení. Obě metody byly experimentálně ověřeny a s ohledem na princip HMTC modelu byly stanoveny metriky pro jejich vzájemné porovnání. Dosažené výsledky ukázaly, že vhodnější je použití metody založené na intervalovém odhadu parametrů. Tato metoda se v porovnání s bodovými odhady vyznačuje menší průměrnou kritickou chybou, která udává míru nadhodnocení intervalu důvěry.

Výsledky dosažené v této kapitole vytváří, společně s návrhem modelu HMTC popsaným v kapitole 4, ucelený soubor formalismů a metod, které umožňují vytvoření reálné aplikace multi-agentního distribuovaného systému založeného na principu více-kontextové důvěry. Následující kapitola se zabývá popisem návrhu a implementace prototypu takovéto aplikace.

Kapitola 6

Prototyp multi-agentního systému

Tato kapitola popisuje navržený a implementovaný prototyp multi-agentního systému, který je založen na navrženém HMTC modelu důvěry s využitím vyhodnocování důvěry agentů na základě jejich modelu chování a historie vzájemných interakcí tak, jak bylo popsáno v kapitole 5.

Cílem této kapitoly je popis způsobu využití navrženého modelu a experimentální ověření funkčnosti navrhovaného řešení a implementovaného prototypu na základních případech. Experimenty s modelem jsou zaměřeny především na porovnání efektivity jednokontextového a více-kontextového přístupu k důvěře, dále na možnost využití přenášení důvěry z jednoho kontextu na kontext druhý s využitím hierarchického propojení kontextů tak, jak bylo navrženo v kapitole 4.

6.1 Popis navrženého prototypu multi-agentního systému

Návrh a implementace prototypu multi-agentního systému založeného na modelu HMTC lze rozdělit do tří, vzájemně se překrývajících fází (budou stručně popsány níže). Výsledkem těchto fází bylo vytvoření robustního simulačního nástroje nazvaného `HMTCsim`, jehož jednotlivé komponenty budou detailně popsány v následujících podkapitolách.

Jednotlivé fáze vývoje simulačního nástroje `HMTCsim`:

1. **Model důvěry HMTC s vyhodnocováním intervalu důvěry na základě intervalových odhadů**. Tato část přímo vychází z jádra této disertační práce. Jedná se především o návrh modelu HMTC popsaného v kapitole 4 a jeho implementaci. Dále pak o návrh vyhodnocení intervalu důvěry v závislosti na historii interakcí mezi agenty v závislosti na jejich modelu chování, což je popsáno v kapitole 5.

2. **Rozhodovací a doporučovací protokol založení na intervalové reprezentaci důvěry.** Tato fáze, jejíž primárním autorem je *Ondřej Malačka*, vychází z práce jež byla publikována jako konferenční příspěvek v článku *Decision Making and Recommendation Protocol Based on Trust for Multi-Agent Systems* [MSZZ11]. Rozhodovací a doporučovací protokol (zkratka DMRP) není součástí této disertační práce a proto bude v následujících podkapitolách popsán jen stručně.

3. **Multi-agentní systém založený na spojení HMTC a DMRP včetně simulačního prostředí.** Poslední fáze integrovala obě předchozí fáze tak, aby mohl

být model HMTC použit společně s protokolem DMRP v multi-agentním systému, který byl implementován v jazyce *AgentSpeak(L)* [Rao96] pomocí frameworku *Jason* [HB04, BHW07]. Součástí této fáze bylo tedy nejen propojení HMTC a DMRP, ale i návrh a implementace multi-agentního a simulačního prostředí pro zobrazování a získávání výsledků simulací. Dále pak na vyšší úrovni i celkový návrh agentní architektury a komunikačních protokolů. Tato fáze vznikla na základě úzké spolupráce s Ondřejem Malačkou.

6.1.1 Základní architektura a použité technologie

Navržený a implementovaný multi-agentní systém je postaven s využitím nástroje *Jason* [HB04]. Jason je interpret rozšířené verze jazyka *AgentSpeak(L)* (dále už jen AgentSpeak), vytvořený v programovacím jazyce *Java*. Jazyk AgentSpeak je jeden z nejpoužívanějších agentně-orientovaných abstraktních programovacích jazyků založených na architektuře BDI (*Beliefs-Desires-Intentions*, což lze přeložit jako *představy-touhy-záměry*) [RG91], která představuje jeden z nejrozšířenějších přístupů k tvorbě agentů založených na praktickém usuzování [Bra87].

Nástroj Jason není jen interpret rozšířené (rozšíření se týká především možnosti komunikace agentů pomocí jazyka založeném na řečových aktech) verze jazyka AgentSpeak, ale poskytuje komplexní platformu pro tvorbu multi-agentních systémů, přičemž lze multi-agentní prostředí, agenty a jejich chování popsat jak v jazyce AgentSpeak, tak i v jazyce Java. Jason rovněž poskytuje přímo možnost propojení s různými dalšími agentními a multi-agentními systémy s využitím JADE (*Java Agent DEvelopment Framework*) [Lab10] a nebo SACI (*Simple Agent Communication Infrastructure*) [HS10] rozhraní tak, aby mohli být agenti jednoduše distribuováni v rámci sítě na různé platformy a architektury. Při použití jazyku Java lze zároveň jednoduše zajistit, že Jason pro každého agenta vytváří samostatné vlákno a agenti tak mohou běžet v rámci jednoho stroje paralelně (v případě více-jádrového nebo více-procesorového stroje).

Jason a jazyk AgentSpeak je tak v nástroji `HMTCsim` použit především pro definici agentního systému (architektura), definici základního chování agentů (cíle, akce a plány), komunikaci mezi agenty a část protokolu pro rozhodování a dotazovaní (DMRP). V jazyce Java je implementován hlavně model HMTC, část protokolu pro rozhodování a dotazování (DMRP), rozšířená verze báze představ (BB – *Belief Base*), multi-agentní prostředí (modely chování agentů) včetně implementace provádění interakcí mezi agenty. Ve spojení s databázovou vrstvou JDBC a databázovým systémem MySQL je pak Java rovněž použita pro ukládání dat výsledků simulačních experimentů.

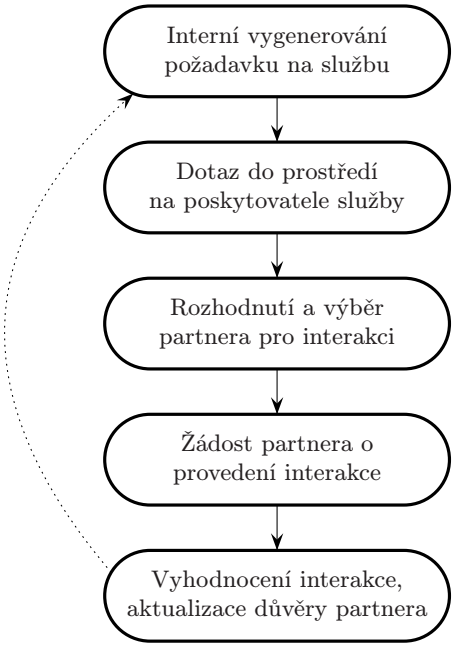
6.1.2 Interakce a interakční cyklus

Námi navržený model důvěry je založený na vyhodnocování výsledků přímých interakcí mezi agenty. V rozšířené verzi (která není součástí předložené disertační práce) umožňuje `HMTCsim` s využitím modelu HMTC (pomocí změny váhy události pro aktualizaci důvěry) a DMRP vyhodnocovat zároveň doporučování mezi agenty a pro vytváření důvěry použít i reputaci.

Jedním ze základních pojmů navrženého prototypu je *interakční cyklus*. Zde si popíšeme princip interakčního cyklu a identifikujeme jeho jednotlivé fáze. Detailní popis jednotlivých fázi, které realizují konkrétní komponenty `HMTCsim`, bude obsažen v následujících podkapitolách.

V rámci jednoho interakčního cyklu si každý agent (pojmenujme ho např. *Zdroj*) vygeneruje interně požadavek na provedení interakce v rámci nějakého kontextu (množina kontextů je předem daná, náhodně je jeden z nich vybrán). Pod pojmem kontext si lze zjednodušeně představit nějakou konkrétní službu, jež poskytuje jiný agent v systému a v rámci které lze provést vzájemnou interakci. Poté, co *Zdroj* ví, jakou službu požaduje, potřebuje zjistit, kdo takovou službu v systému poskytuje. Službu „žlutých stránek“, která udržuje seznam aktivních agentů a jimi nabízených služeb, poskytuje komponenta prostředí (bude popsána níže). Agent se tedy dotáže do *prostředí* a zjistí tak seznam aktivních agentů, kteří požadovanou službu poskytují. V rámci různých interakčních cyklů může být seznam aktivních agentů různý a navíc všichni agenti nemusí poskytovat všechny služby, proto je třeba v každém interakčním cyklu zjistit, kdo danou službu poskytuje.

Výsledkem dotazu do *prostředí* je odpověď obsahující neprázdnou množinu identifikátorů agentů (unikátních), kteří v daném cyklu požadovanou službu poskytují. V další fázi se tak *Zdroj* musí rozhodnout koho z této množiny agentů, kteří službu poskytují, si vybere. Tento výběr je proveden na základě důvěry potencionálních agentů z pohledu agenta *Zdroj* v rámci zvolené služby (kontextu), tuto část zajišťuje komponenta DMRP. Výsledkem této fáze je výběr jednoho konkrétního agenta, kterého si označíme jako *Partner* interakce.



Obrázek 6.1: Jednotlivé fáze jednoho interakčního cyklu agenta.

V následujícím kroku vyzve agent *Zdroj* agenta *Partner* k provedení interakce v rámci zvolené služby. Ten na základě tohoto požadavku „provede interakci“ tak, že pošle žádost do *prostředí* na provedení interakce pro agenta *Zdroj* v daném kontextu (o který byl požádán). V momentě kdy komponenta *prostředí* interakci provede (na základě modelu chování agenta *Partner*), je tento výsledek oznámen agentu *Zdroj*. Výsledek obsahuje hodnotu, která byla vygenerována *prostředím* na základě modelu chování agenta *Partner* s přihlédnutím na jeho individuální požadavky (agent *Partner* může ovlivnit výsledek interakce).

Na základě tohoto výsledku provede agent *Zdroj* zhodnocení interakce a aktualizaci důvěry v rámci své instance HMTC komponenty. V rámci poslední fáze vyhodnocení důvěry je také výsledek interakce a nová výsledná hodnota důvěry odeslána ostatním komponentám tak, aby se zajistilo uložení výsledku pro simulační a experimentální potřeby.

Tyto jednotlivé fáze jednoho interakčního cyklu jednoho agenta jsou naznačeny na obrázku 6.1. V momentě kdy je interakční cyklus agenta dokončen, začne se provádět další. Všichni aktivní agenti tak provedou v rámci simulace multi-agentního systému stanovený počet interakčních cyklů, v momentě kdy je ukončen poslední interakční cyklus, simulace systému končí. Typickým rysem pro distribuovaný systém je, že jednotlivé interakční cykly jednotlivých agentů nemusí být a v našem systému ani nejsou synchronizovány, tedy v jednom okamžiku tak různí agenti mohou provádět různé (ve smyslu identifikace cyklu v rámci celkové počtu) interakční cykly.

6.2 Konceptuální schéma nástroje `HMTCsim`

V této podkapitole popíšeme detailněji základní komponenty simulačního nástroje multi-agentního systému `HMTCsim`, jejich propojení a princip funkčnosti v případě, že se nejedná o komponenty založené na návrhu z předchozích kapitol (model HMTC a princip odhadu chování agentů na základě přímých zkušeností).

Celkové konceptuální schéma simulačního nástroje `HMTCsim` multi-agentního systému využívající model důvěry HMTC je uvedeno na obrázku 6.2. Jsou zde naznačeny základní komponenty nástroje, skupina agentů včetně naznačení jejich vzájemného propojení a způsobu komunikace jak na úrovni jednotlivých komponent, tak i různých agentů v systému.

6.2.1 Agenti

Hlavním prvek každého multi-agentního systému jsou agenti. Jak bylo popsáno výše, převážná část chování agentů je popsána v jazyce AgentSpeak, jež je interpretován interpretem Jason. Na obrázku 6.2 jsou agenti zvýrazněni šedou výplní. V rámci našeho systému rozlišujeme dvě základní role agentů: agenti zajišťující globální služby v rámci simulačního modulu (agent `BigBoss` a `DbManager`) a „základní agenti“ systému pro simulaci navrženého modelu využívající důvěru k rozhodování v rámci požadavků na provedení interakce.

Základní agenti

V rámci základního agenta systému je implementován v jazyce AgentSpeak interakční cyklus, jež byl popsaný v podkapitole 6.1.2. Definice agenta zároveň obsahuje statickou bázi představ a jednotlivé znalosti (představy) jsou z báze představ agenta dostupné buď pomocí interních akcí (implementační prvek nástroje Jason umožňující plné přizpůsobení chování agenta s využitím jazyka Java) anebo pomocí akcí s prostředím.

V rámci našeho simulačního prostředí jsou základní agenti pojmenování vždy `Agent`N, kde N je číslo počínaje hodnotou jedna do velikosti omezené maximální hodnotou datového typu `Integer` v jazyce Java na dané platformě. Reálný počet agentů, který je možné pro každou simulaci nastavit, je však omezený především výpočetními možnostmi stroje, na kterém je simulace spuštěna.

Agenti `BigBoss` a `DbManager`

Agent nazvaný `BigBoss` zajišťuje výpis stavu simulačního běhu v rámci grafického rozhraní, případně do logovacího souboru simulace. Tyto textové výpisy obsahují výsledky jednotlivých interakčních cyklů jednotlivých agentů, dále pak různé statistické informace o běhu simulace, stavu agentů, aktualizací důvěry a podobně. Tento agent rovněž zajišťuje korektní ukončení simulace v případě, že je ukončen poslední interakční cyklus všech základních agentů.

Agent se jménem `DbManager` zprostředkovává komunikační rozhraní pro zaznamenání stavu simulačního běhu do relační databáze tak, aby bylo možné simulační data automatizovaně a efektivně zpracovávat.

Základní důvod pro implementaci těchto agentů je abstrakce monitorování stavu simulačního modelu. Základní agenti pošlou pouze přes unifikované rozhraní informaci o svém stavu a ta je následně interpretována do textové podoby pro výstup na obrazovku anebo jako databázová položka. Tím se také rovněž snižuje složitost základních agentů a na praktické úrovni zvyšuje rychlost běhu simulace (jednotlivé nízkoúrovňové akce jsou delegovány a prováděny paralelně v samostatném vlákně).

6.2.2 Prostředí

Komponenta *prostředí* (anglicky *Environment*) zajišťuje v našem systému primárně tyto funkce: ve spolupráci s HMTC komponentou a bází znalostí zajišťuje poskytování informací o agentech a jejich kontextech, obsahuje definice modelů chování základních agentů dle navrženého principu v kapitole 5, zprostředkovává na základě žádostí od agentů provedení interakcí a generuje vjemy základním agentům o výsledku těchto interakcí.

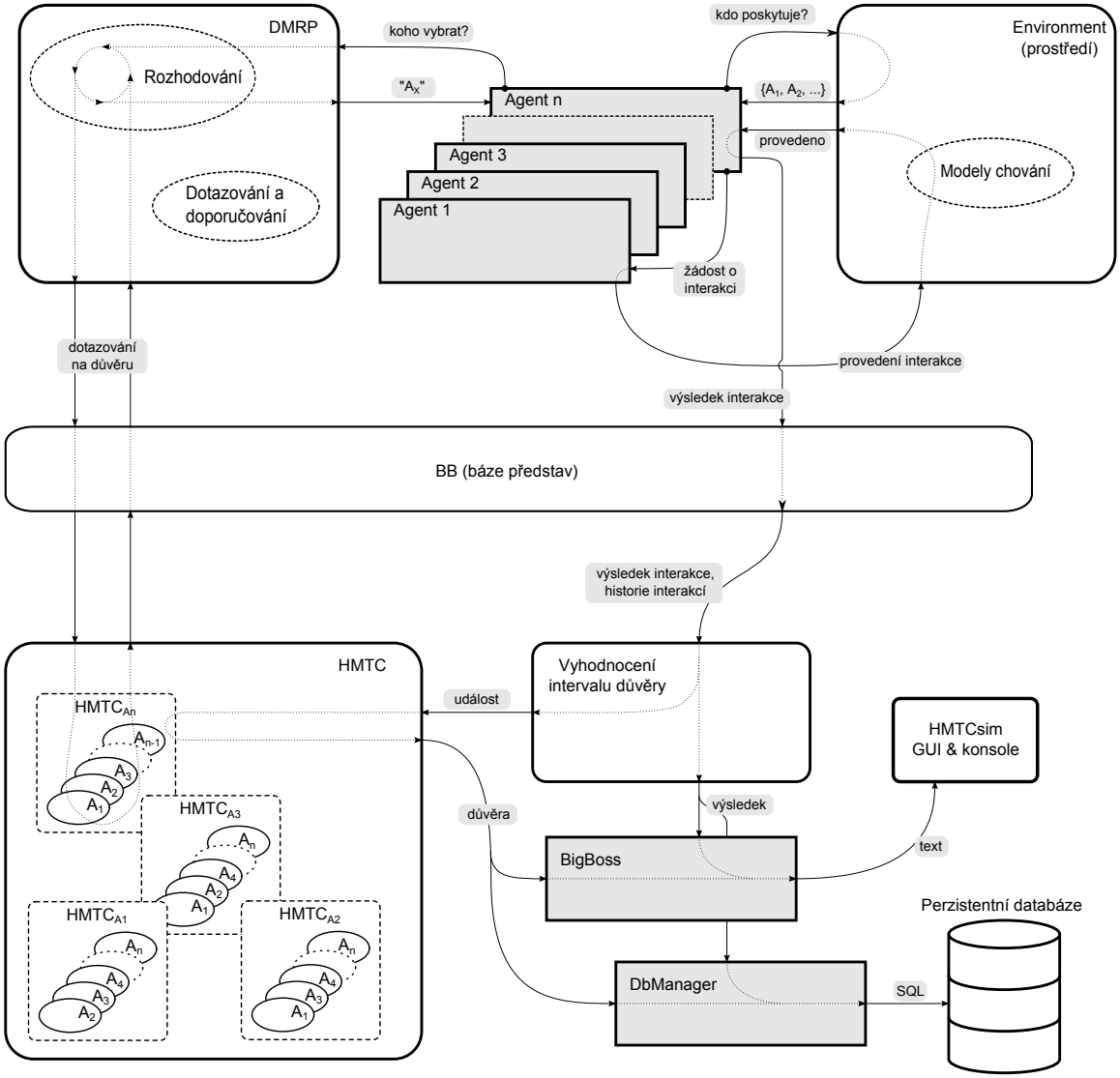
Princip použití této komponenty je částečně znázorněn diagramem sekvence na obrázku 6.3. Tento diagram zachycuje situaci, kdy agent *Zdroj* (pojmenování agentů je dle notace popsané v podkapitole 6.1.2) žádá agenta *Partner* o provedení interakce v rámci nějakého kontextu. Jak je vidět, tak důsledek tohoto požadavku je předání výsledku interakce přímo z *prostředí* jako vjem agentu *Zdroj*, přičemž *prostředí* tento výsledek generuje na základě modelu chování agenta *Partner* a na základě požadavku na kvalitu výsledku. Výsledek interakce je tedy dán funkcí $v(x, kvalita)$, kde x je hodnota vygenerovaná generátorem pseudonáhodných čísel normální rozložení z $N(\mu, \sigma)$ odpovídající modelu chování agenta *Partner* v daném kontextu a funkce v pro parametry x a $kvalita$ je definována takto:

$$v(x, kvalita) = \begin{cases} x & \text{pro} \quad x < kvalita \\ kvalita & \text{jinak} \end{cases} \tag{6.1}$$

6.2.3 Model HMTC, vyhodnocení intervalu důvěry a báze představ

Komponenta HMTC společně s komponentou *vyhodnocení intervalu důvěry* implementuje navržený model důvěry popsaný v kapitole 4 a způsob vyhodnocování důvěry na základě historie interakcí popsaný v kapitole 5. Jak je naznačeno na obrázku 6.2, tak komponenta HMTC obsahuje pro každého základního agenta v systému množinu ($HMTC_{A1}$, $HMTC_{A2}$, …, $HMTC_{An}$) instancí HMTC modelů ($HMTC_{A1} = \{A_2, A_3, \ldots, A_n\}$, kde A_2 až A_n jsou jednotlivé HMTC instance), přičemž přístup k jednotlivým instancím je zajištěn výhradně přes *bázi znalostí*, která zajišťuje rozhraní právě mezi HMTC modelem a agenty.

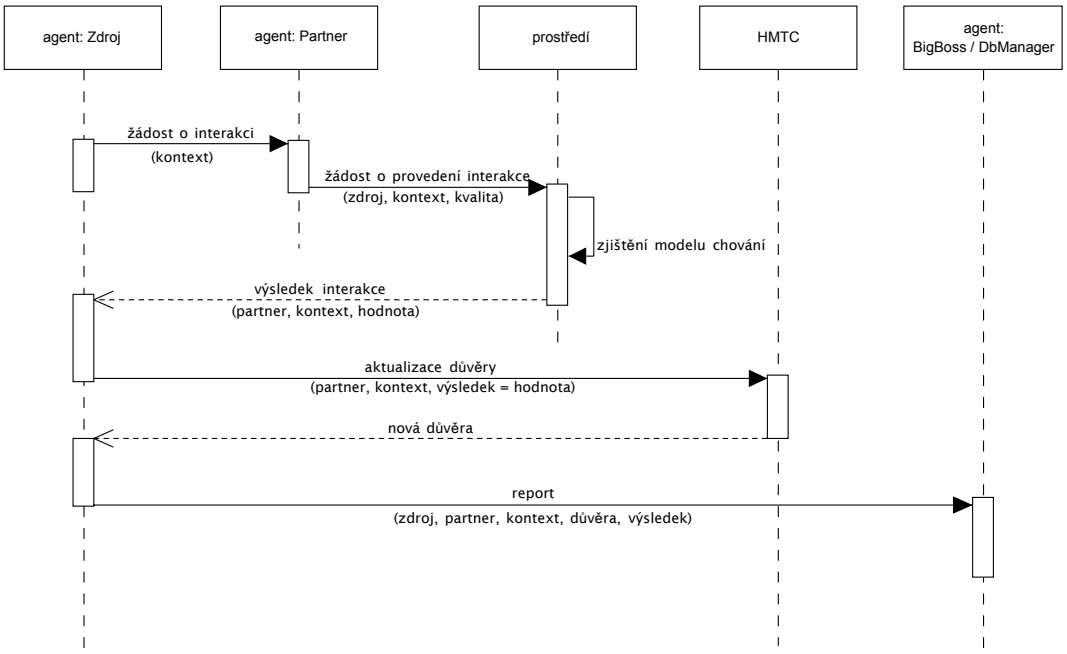
Komponenta *vyhodnocení intervalu důvěry* implementuje především robustní statistickou knihovnu na práci z různými rozloženími (normální rozložení, t-rozložení, a Chí-kvadrát



Obrázek 6.2: Konceptuální schéma prototypu „`HMTCsim`“ multi-agentního systému využívající model důvěry HMTC.

rozložení) a princip intervalových a bodových odhadů. Důležitou součástí komponenty je i generátor pseudonáhodných čísel s normálním rozložením, jež je využíván především komponentou *prostředí*. Vstupní množinu hodnot pro stanovení intervalového odhadu poskytuje *báze představ*, která pro každého agenta a interakci v daném kontextu uschovává informaci o výsledku. Výstupem je pak intervalový odhad důvěry, který spolu s typem interakce (v našem případě pouze přímé interakce) vytváří HMTC událost včetně stanovení její váhy, jež je HMTC komponentě předána a na základě které je hodnota důvěry aktualizována (pro daného agenta a kontext).

Jak bylo naznačeno výše, komponenta *báze představ* (zkráceně BB z anglického *Belief Base*) slouží primárně jako rozhraní mezi základním agentem, HMTC modelem a komponentou vyhodnocující interval důvěry. Z implementačního hlediska se jedná o propojení jazyků AgentSpeak (definice agentů) a Java (HMTC a vyhodnocování důvěry), přičemž takto vytvořené rozhraní umožňuje efektivní propojení různých komponent.



Obrázek 6.3: Diagram zachycující komunikaci mezi jednotlivými agenty a komponentami systému od okamžiku zaslání žádosti *Zdroje* na *Partnera*.

6.2.4 Komponenta DMRP

Komponenta DMRP implementuje protokol pro rozhodování agentů na základě důvěry a protokol pro vzájemné dotazovaní a doporučování agentů. Z tohoto pohledu ji tak lze chápat jako dva relativně nezávislé celky, přičemž pro účely této kapitoly je klíčová hlavně část, která se zabývá rozhodováním a výběrem agentů na základě jejich důvěryhodnosti. Ta bude zde stručně popsána, část zabývající se dotazováním a zpracováním doporučení, která nebude pro naše experimenty použita, je detailně popsána v původním článku [MSZZ11].

V následujícím textu popíšeme základní princip a algoritmy, které využívá komponenta DMPR v námi navrženém prototypu multi-agentního systému při rozhodování o výběru partnera pro transakci na základě jeho důvěryhodnosti v daném kontextu.

Základní notace a funkce nad intervalem důvěry

Množina agentů je označena písmenem A, jednotliví agenti jsou pak označeni symboly malé abecedy (s případnými indexy) $A = \{a, b, c_1, c_2, \dots\}$. Množina kontextů je označena písmenem C a jednotlivé kontexty zapisujeme pomocí symbolů řecké abecedy $C = \{\alpha, \beta, \gamma, \dots\}$. Důvěru z pohledu agenta a vzhledem k agentu b v kontextu α zapisujeme jako $\mathcal{T}_{a,b}(\alpha)$, přičemž tento zápis odpovídá „aktuální“ hodnotě důvěry a v případě, že chceme využít „temporálního“ přístupu, použijeme zápis s horním indexem, tedy pro stanovení důvěry v konkrétní časový okamžik t: $\mathcal{T}^t_{a,b}(\alpha)$. Hodnota důvěry je zde samozřejmě vyjádřena intervalem důvěry, pro jeho zápis používáme symbolu ϑ tak, jak bylo použito v předchozím textu: $\vartheta = [\vartheta_{min}, \vartheta_{max}]$.

Nad intervalem důvěry definujeme dále dvě funkce, funkci *nejistoty* a *transformační* funkci. Pro jejich zápis používáme jejich anglický zkrácený název *uncert* (z anglického *uncertainty*) a $trans$ (z anglického *transformation*).

Funkce *nejistoty* udává šíři intervalu chování agenta a je definována pro interval důvěry $\vartheta = [\vartheta_{min}, \vartheta_{max}]$ následovně:

$$uncert(\vartheta) = \vartheta_{max} - \vartheta_{min} \ . \tag{6.2}$$

Transformační funkce, založená na modifikovaném *Hurwiczově pravidle* [Hur52], se používá pro transformaci intervalu důvěry na hodnotu s využitím parametru *optimismu*. Tento parametr je lokálním (subjektivním) parametrem pro každého agenta, označujeme ho symbolem r a může nabývat hodnot v intervalu $[0, 1]$z množiny $\mathbb{R}$. Transformační funkce s využitím tohoto parametru pro interval důvěry ϑ je pak definována následovně:

$$trans(\vartheta, r) = \vartheta_{min} + (\vartheta_{max} - \vartheta_{min}) * r \ . \tag{6.3}$$

Další důležitou funkcí při definici rozhodování agenta je funkce *užitku*. Tato funkce vychází z principu ohodnocení možných stavů a stanovení subjektivního užitku agenta pro tyto stavy. Její použití v rámci rozhodování založené na důvěře je rovněž popsáno v pracích [Mar94, MP10] a vychází z obecného principu *teorie očekávaného užitku* (z anglického *expected utility theory*) [Mon97]. Funkci užitku označujeme symbolem μ a její definiční obor i její obor hodnot je z intervalu $[0, 1]$ z množiny $\mathbb{R}$. Funkce zároveň musí splňovat podmínku $\forall x, y \in [0, 1] : x < y \Rightarrow \mu(x) \leq \mu(y)$, tedy že je monotónní rostoucí.

V rámci této funkce rozeznáváme dva důležité body, pojmenované *minimální bod užitku* a *maximální bod užitku*, označené jako p_{min} a p_{max} pro které platí následující:

$$\begin{aligned} \mu(p_{min}) = \mu(0) \text{ kde } \forall x : x > p_{min} \Rightarrow \mu(x) > \mu(p_{min}) \text{ pro } p_{min}, x \in [0, 1] \\ \mu(p_{max}) = \mu(1) \text{ kde } \forall x : x < p_{max} \Rightarrow \mu(x) < \mu(p_{max}) \text{ pro } p_{max}, x \in [0, 1] \end{aligned} \tag{6.4}$$

Funkce užitku je lokálním (subjektivním) parametrem každého agenta v systému vzhledem k danému interakčnímu kontextu z množiny všech kontextů C. Vzhledem k tomu je pak zavedena notace μ_a^α pro zápis funkce užitku agenta a v kontextu α. Minimální a maximální bod užitku vzhledem k této notaci pak zapisujeme jako $p_{min}^{a,\alpha}$ a $p_{max}^{a,\alpha}$.

Rozhodování a výběr nejvhodnějšího agenta

Rozhodování a výběr nejvhodnějšího agenta pro provedení transakce v rámci nějakého kontextu je pak implementován algoritmem, jehož základem jsou funkce pro *výběr vhodných agentů* a *odstranění nevhodných agentů*. Princip těchto funkcí je založen na porovnání hranic intervalů důvěry všech potencionálních agentů s minimálním a maximálním bodem užitku.

Algoritmus 2: `filterUseful`

Vstup : kontext α, množina agentů A_α
maximální bod užitku: $p_{max}^{a,\alpha}$
Výstup: množina vhodných agentů

```
1 begin
2     A_u = ∅
3     for b ∈ A_α do
4         [ϑ_min, ϑ_max] = T_{a,b}(α)
5         if p^{a,α}_max ≤ ϑ_min then
6             A_u = A_u ∪ {b}
7     return A_u
```

Algoritmus 3: `filterUnuseful`

Vstup : kontext α, množina agentů A_α
minimální bod užitku: $p_{min}^{a,\alpha}$
Výstup: množina nevhodných agentů

```
1 begin
2     A_u = ∅
3     for b ∈ A_α do
4         [ϑ_min, ϑ_max] = T_{a,b}(α)
5         if p^{a,α}_min ≥ ϑ_max then
6             A_u = A_u ∪ {b}
7     return A_u
```

V případě, že funkce výběru vhodných agentů `filterUseful` popsaná algoritmem 2 vrátí neprázdnou množinu, je vybrán první agent s nejvyšší neurčitostí důvěry v daném kontextu (jež je určena funkcí *uncert*).

Množina agentů vyhodnocených jako nevhodných pomocí funkce `filterUnuseful`, jež je popsána algoritmem 3, je z množiny potencionálních agentů pro interakci odstraněna.

Posledním krokem (pakliže nebyl nejvhodnější agent vybrán již funkcí `filterUseful`) je výběr agenta na základě principu genetického algoritmu selekce nazývaného jako *vážená ruleta* nebo pouze jen *ruleta*, který pro každého jedince určí pomocí fitness funkce pravděpodobnost jeho výběru z množiny potencionálních agentů. Fitness funkce je stanovena na základě parametru optimismu r a s využitím transformační funkce *trans* pro daný interakční kontext. Tento výběr je tak založen na principu výběru jedince s pravděpodobností odpovídající jeho poměrné kvalitě (ta je stanovena na základě jeho důvěryhodnosti a parametru optimismu vyhodnocujícího agenta).

6.2.5 Simulační výstupy (GUI, konzole, databáze)

Nepostradatelnou komponentou celého nástroje je rovněž rozhraní pro získávání informací o stavu simulačního modelu. Tato kategorie komponent tak zahrnuje grafické uživatelské rozhraní pro spouštění a zastavení simulace, výpis stavu simulačního modelu na obrazovku, do textové konzole (jak pro systémy Windows tak Linux) a textového souboru. To je fyzicky realizováno s využitím standardního *Java Logging API* implementovaného v balíčku `java.util.logging`. Ukládaní stavu simulačního modelu po každé interakci mezi agenty do persistentní databáze postavené na systému MySQL (s využitím standardního *Java JDBC API*) nám rovněž umožnilo velmi efektivní vyhodnocování a zpracování simulačních výsledků. Jednotlivé stavy procesu simulace jsou z agentního systému na nižší úroveň (konzole, soubor, databáze) předávány přes rozhraní implementující dvěma agenty `BigBoss` a `DbManager`, kteří již byli popsáni výše.

6.3 Stanovení předpokladů a metodika jejich ověření

Cílem této disertační práce je nejen přínos nových přístupů a metod pro práci s důvěrou, ale i prokázání toho, že tyto nově navržené přístupy jsou v některých případech *více efektivní* a vhodné než přístupy známé a používané. Pojem „více efektivní“ či „více vhodný“ přístup je nepříliš konkrétní a lze si pod ním přestavit téměř cokoliv. Proto si v této kapitole přesně definujeme předpoklady, respektive cíle, kterých by námi navržený model měl dosáhnout tak, abychom mohli prohlásit, že je „více efektivní“ anebo „více vhodný“ pro konkrétní aplikaci. Nezbytnou součástí takového zhodnocení je i definice objektivní a relevantní metodiky měření efektivity modelu na základě experimentálních výsledků tak, aby bylo možné jednotlivé výsledky vzájemně porovnat.

6.3.1 Stanovení předpokladů

Na základě zadání cílů disertační práce jsme stanovili následující předpoklady, jež by měl navržený model splňovat:

1. Více-kontextový přístup stanovení důvěry je vhodný v případě, kdy agenti v systému poskytují různé služby s různou kvalitou anebo jsou hodnoceny v různých aspektech jejich chování. Důsledkem použití důvěry pro rozhodování agentů ve více-kontextovém

prostředí je to, že agenti využívající více-kontextový model jsou schopni přesněji určit vhodného partnera pro interakci než agenti využívající jedno-kontextový model.

2. Při použití více-kontextového modelu důvěry s možností odvozování důvěry (pomocí hierarchické struktury jak bylo v této práci navrženo) mezi kontexty je možné dosáhnout přesnějšího rozhodování agenta (ve smyslu výběru optimálního partnera pro interakci) v rámci konkrétního kontextu na základě odvozené důvěry v tento kontext i bez předchozí přímé zkušenosti. To však pouze za podmínky, že máme s agentem přímou zkušenost jež ovlivnila důvěru v některém z jeho propojených kontextů.

6.3.2 Metody vyhodnocení experimentálních výsledků

Pro vyhodnocení experimentálních výsledků a možnosti porovnání jednotlivých experimentů byly stanoveny dvě základní sledované veličiny, kterými jsou:

- *Počet použití agenta* – počet použití agenta (provedených interakcí s agentem) v různých kontextech s různými modely chování.
- *Průměrný užitek* – průměrný užitek agenta(ů) v závislosti na čase (interakčním cyklu) experimentu stanovený na základě výsledku interakce.

V případě sledování *počtu použití agenta* nebo skupiny agentů v nějakému interakčním kontextu (tj. kontext, v němž je možné s daným agentem provést interakci) lze na základě modelu chování pro tento kontext jasně rozhodnout, zda je tento počet odpovídají či nikoliv. Pro kontext a agenta s „dobrým“ chováním by měl být počet použití maximální a pro kontext se „špatným“ chováním minimální.

Vzhledem k tomuto si neformálně definujeme pojem „dobré“ a „špatné“ chování agenta v rámci nějakého kontextu. Pojem „dobré“ chování (respektive „dobrý“ model chování) je takové nastavení modelu chování v rámci kontextu, že většina výsledků interakce v tomto kontextu po aplikaci funkce užitku nabývá hodnot blízkých jedné. Naopak, analogicky „špatné“ chování generuje výsledky, které po aplikaci funkce užitku nabývají hodnot blízkých nule.

Druhou základní sledovanou veličinou je hodnota užitku agenta na základě výsledku transakce. Výsledný užitek lze sledovat pro daného agenta, kontext a interakční cyklus. To je však v případě většího množství agentů a kontextů v systému přehledně neinterpretovatelné. Z tohoto důvodu sledujeme především globální průměrný užitek všech agentů a všech jejich kontextů vzhledem k interakčnímu cyklu. Porovnáním jednotlivých průběhů takto stanoveného průměrného užitku pro různé konfigurace experimentů jsme ověřili, že tato metrika má vysokou vypovídající hodnotu při stanovení efektivity rozhodování agentů.

6.4 Experimentální výsledky

V této podkapitole provedeme experimentální ověření stanovených předpokladům výše a to včetně zhodnocení dosažených výsledků a potvrzení či vyvrácení stanoveného předpokladu.

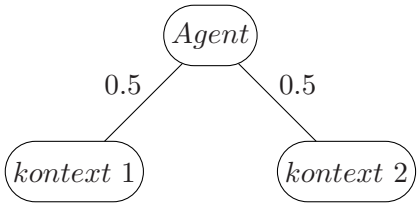
6.4.1 Experiment 1: Jedno- versus více-kontextový přístup

První experiment byl zaměřen na ověření základní funkčnosti a rozdílu mezi jedno- a více-kontextovým přístup. Pro tento experiment byla zvolena nejjednodušší HMTC struktura,

která obsahuje kořenový kontext *Agent* a dva podkontexty označené jako *kontext* 1 a *kontext* 2. Tato struktura, včetně vah jednotlivých hran, je naznačena na obrázku 6.4.

V následujícím textu budeme používat zkratku SC pro jedno-kontextový přístup (z anglického *Single-Context*) a zkratku MC pro více-kontextový přístup (*Multi-Context*) ke stanovení důvěry.

Experiment probíhal tak, že v rámci multi-agentního systému s deseti agenty (označenými Agent1, ..., Agent10) bylo provedeno 100 interakčních cyklů, přičemž v každém z nich si každý agent náhodně generoval požadavek na provedení transakce v jednom ze dvou kontextů *kontext* 1 a *kontext* 2. V případě MC přístupu se aktualizoval vždy příslušný kontext a při rozhodování o výběru agenta se rovněž brala v potaz důvěra agenta v rámci interakčního kontextu. V případě SC přístupu se aktualizoval pouze kořenový kontext *Agent* a rozhodování probíhalo pouze na základě důvěry v tento kořenový kontext. Zároveň bylo v rámci tohoto experimentu vypnuto odvozování důvěry mezi kontexty tak, aby se navzájem neovlivňovaly.



Obrázek 6.4: Struktura HMTC modelu pro základní porovnání jedno- a více-kontextového přístupu.

Množina deseti agentů byla rozdělena do třech pomyslných „kategorií“ na základě jejich modelu chování v jednotlivých kontextech. Tyto tři kategorie agentů lze popsat jako *dobří agenti*, *napůl dobří/špatní agenti* a *špatní agenti*. Kategorie *dobrých agentů* obsahuje takové agenty, jejichž model důvěry v oba kontexty je stanoven rozložením $N(\mu = 0.85, \sigma = 0.05)$. Kategorie *napůl dobří/špatní agenti* obsahuje takové agenty, jejichž model chování pro jeden z kontextů je stanoven $N(0.85, 0.05)$ a pro druhý kontext je stanoven $N(0.15, 0.05)$ – jedná se tedy o intervaly chování $[0.7, 1.0]$ a $[0.0, 0.3]$ (kde první lze označit jako „dobré“ chování druhé jako „špatné“ chování). Třetí kategorie *špatní agenti* obsahuje pouze takové agenty, jejichž model chování je dán $N(0.15, 0.05)$. Konkrétní rozdělení agentů a jejich modelů chování pro tento experiment je popsáno v tabulce 6.1.

	Model chování $N(\mu, \sigma)$	
Agent	kontext 1	kontext 2
Agent1	$N(0.85, 0.05)$	$N(0.85, 0.05)$
Agent2	$N(0.85, 0.05)$	$N(0.85, 0.05)$
Agent3	$N(0.15, 0.05)$	$N(0.85, 0.05)$
Agent4	$N(0.85, 0.05)$	$N(0.15, 0.05)$
Agent5	$N(0.15, 0.05)$	$N(0.15, 0.05)$
⋮	⋮	⋮
Agent10	$N(0.15, 0.05)$	$N(0.15, 0.05)$

Tabulka 6.1: Modely chování jednotlivých agentů (Exp.1).

Hlavní parametry a sledované veličiny experimentu

Pro provedení experimentu v nástroji `HMTCsim` byly zvoleny následující parametry:

- koeficient spolehlivosti intervalového odhadu: $(1 - \alpha) = 0.95$ ($CI = 95\%$),
- odhad max. neurčitosti pro stanovení intervalového odhadu malých výběrů: $\hat{u} = 0.3$,
- parametr *optimismus* byl nastaven globálně pro všechny agenty na hodnotu $r = 0.2$,
- funkce užitku byla stanovena globálně pro všechny agenty a všechny kontexty následovně:

$$\mu(x) = \begin{cases} 0 & \text{pro} \quad 0 \leq x \leq 0.2 \\ \frac{x-0.2}{0.6} & \text{pro} \quad 0.2 < x < 0.8 \\ 1 & \text{pro} \quad 0.8 \leq x \leq 1\ , \end{cases} \tag{6.5}$$

  tedy $p^{\alpha}_{min} = 0.2$ a $p^{\alpha}_{max} = 0.8$ pro libovolný kontext α.

S těmito parametry bylo provedeno celkem 10 experimentů pro každý přístup (jedno- a více-kontextový). Pro každý přístup bylo následně statisticky analyzováno celkem 10000 interakcí (10 experimentů × 10 agentů × 100 interakčních cyklů), z nichž byla stanovena průměrná hodnota použití jednotlivých agentů v daném cyklu a průměrná hodnota užitku všech agentů v daném interakčním cyklu. Výstupy těchto statistických dat byly zpracovány graficky a jsou prezentovány v následující podkapitole. Pro naznačení spojitého průběhu (je-li použit) grafické reprezentace průměrných výsledků byla zvolena Beziérova aproximace s využitím nástroje `gnuplot`.

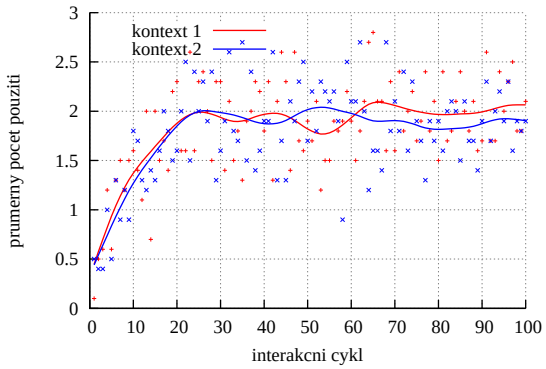
Výsledky

První čtveřice grafů, na obrázcích 6.5(a)–(d), ukazuje porovnání SC a MC přístupu pro agenta Agent1, který reprezentuje skupinu *dobrých agentů*. Výsledky pro agenta Agent2, který rovněž patří do skupiny dobrých agentů, jsou graficky velmi podobné a proto jsou grafy vždy zobrazeny pouze pro jednoho vybraného agenta z dané kategorie.

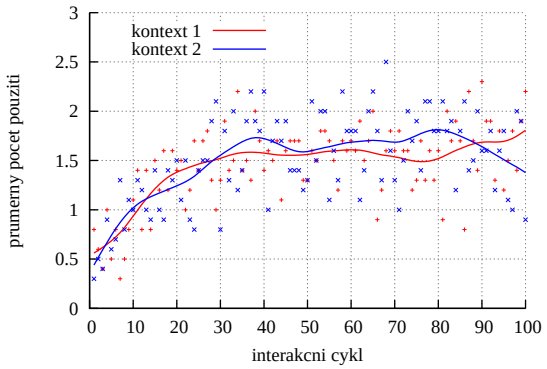
Jak je vidět na obrázcích 6.5(c) a 6.5(d), tak počet použití agenta Agent1 je vyšší v případě SC přístupu (celkově zhruba o 17% interakcí více než v případě MC přístupu). To je dáno tím, že agent v obou kontextech poskytuje „dobré“ služby a pokud jsou výsledky agregovány pouze do jednoho kontextu, je pro zjištění jeho „dobrého“ chování potřeba méně interakcí. V případě MC přístupu je potřeba prakticky 2× tolik interakcí, abychom dosáhli stejné hodnoty důvěry v obou dvou „dobrých“ kontextech.

V případě analýzy použití agenta Agent3, jež patří do skupiny *napůl dobří/špatní* je situace zcela odlišná a zde se již naplno ukazuje nevýhoda SC přístup v případě různého chování v různých kontextech. Jak je vidět na obrázku 6.6(a), tak v případě agregace dvou různých chování do jednoho kontextu nejsou ostatní agenti schopni rozpoznat skutečné chování tohoto agenta a ten je jen zřídka používán v obou dvou kontextech (zhruba 1× v každém cyklu celkem pro oba kontexty). To je způsobeno tím, že v případě vzniku konfliktu jsou staré události odstraněny a nahrazeny novými, aktuálními. Ty jsou střídavě generovány na základě „dobrého“ a „špatného“ chování a výsledkem je to, co lze vidět na obrázcích 6.6(c) a 6.6(d) tedy, že počet použití agenta Agent3 v případě MC přístupu je větší o téměř 60% než v případě SC přístupu.

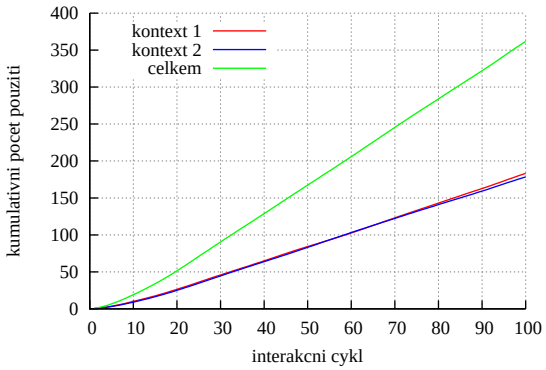
Poslední zkoumanou skupinou jsou agenti z kategorie *špatní*, kde byl jako reprezentativní vzorek agenta vybrán Agent5. Zde jsou výsledné průběhy průměrného i kumulativního



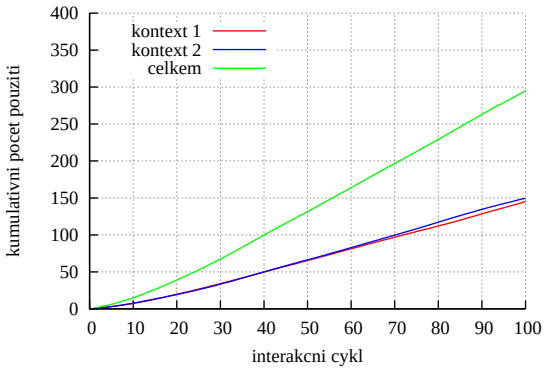
(a) Agent1 - jedno-kontextový přístup.



(b) Agent1 - více-kontextový přístup.



(c) Agent1 - jedno-kontextový přístup.



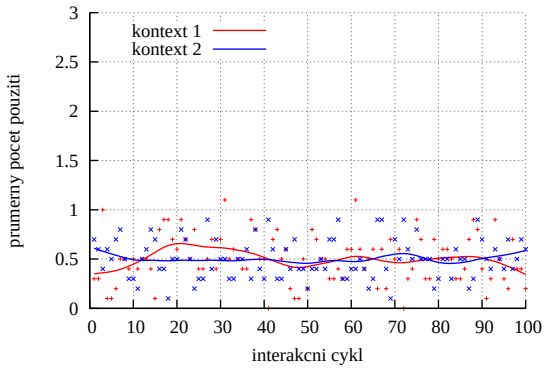
(d) Agent1 - více-kontextový přístup.

Obrázek 6.5: Průměrný a kumulativní počet použití agenta Agent1 vzhledem k interakčnímu cyklu.

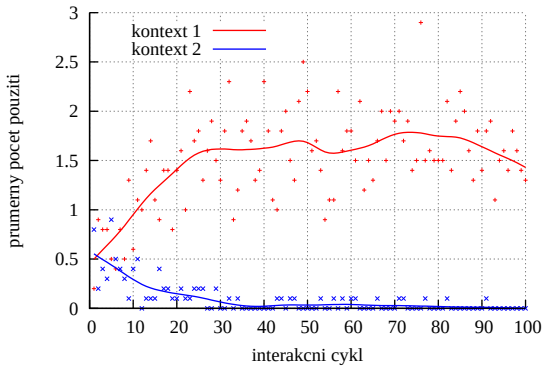
počtu použití agenta velmi podobné, jen v případě MC přístupu se opět projevuje fakt, že je třeba provedení více interakcí tak, aby „špatné“ chování agenta bylo odhaleno v obou kontextech.

Posledním grafickým výstupem experimentu je průběh průměrného užitku všech agentů v systému vzhledem k interakčnímu cyklu. V grafu na obrázku 6.8 je vidět, že průměrný užitek agentů je při menším počtu interakcí (až zhruba do 30 interakčního cyklu) větší v případě SC než v případě MC přístupu. To je dáno tím, že dvě skupiny agentů *dobří* a *špatní* jsou v případě SC přístupu velmi rychle „odhaleni“ (stanovení důvěry intervalovým odhadem velmi rychle konverguje ke skutečnému intervalu chování) právě díky tomu, že důvěra je agregována pouze do jednoho kontextu a oni mají v obou dvou kontextech stejný model chování.

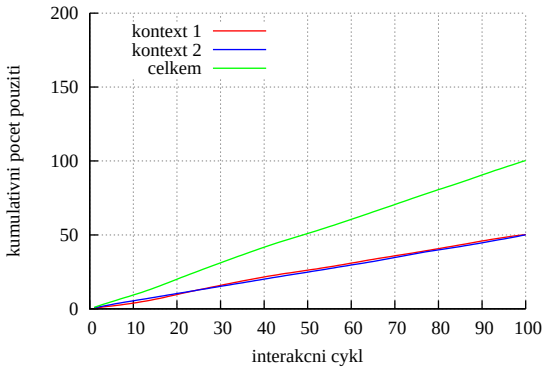
Naproti tomu v případě MC přístupu je třeba provést 2× více interakcí pro dosažení stejné hodnoty důvěry, neboť se vždy aktualizuje pouze jeden ze dvou kontextů a „odhalení“ skutečného intervalu chování je pomalejší. Nicméně zhruba po třicátém interakčním cyklu už je průměrný užitek agentů větší pro MC přístup a to především díky tomu, že agenti v kategorii *napůl dobří/špatní* jsou podstatně efektivněji využíváni než v přístupu SC.



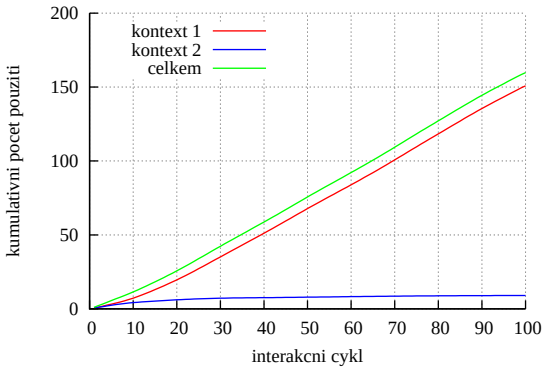
(a) Agent3 - SC přístup.



(b) Agent3 - MC přístup.



(c) Agent3 - SC přístup.



(d) Agent3 - MC přístup.

Obrázek 6.6: Průměrný a kumulativní počet použití agenta Agent3 vzhledem k interakčnímu cyklu.

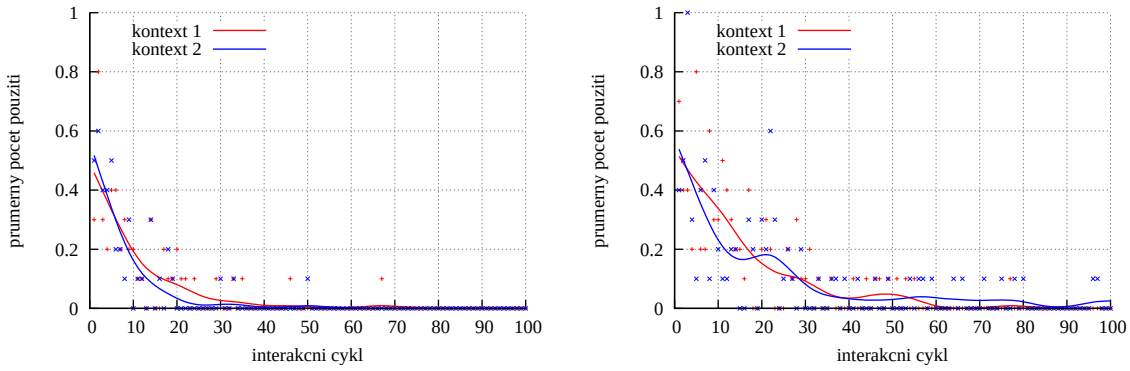
Zhodnocení „Experimentu 1“

Z výše prezentovaných průběhů sledovaných veličin simulace multi-agentního systému lze pro zvolenou více-kontextovou strukturu říci, že první předpoklad stanovený v podkapitole 6.3.1 byl potvrzen. Bylo ověřeno, že více-kontextový přístup je vhodný v případě, kdy je třeba agenty hodnotit v různých aspektech přičemž jejich chování (nebo kvalita) v rámci těchto aspektů je různé. Pakliže agenti nejsou hodnoceni ve více kontextech anebo je jejich chování ve všech kontextech shodné, je vhodnější použít jedno-kontextový přístup.

6.4.2 Experiment 2: Více-kontextový přístup s přenášením důvěry mezi kontexty

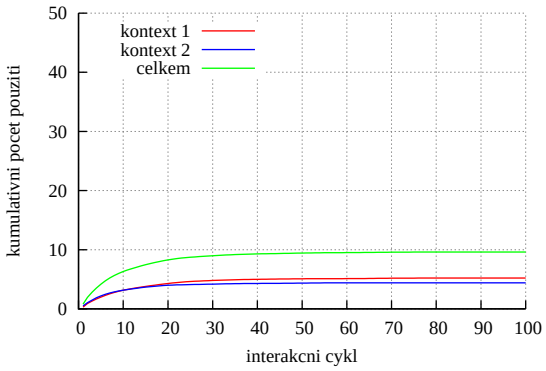
Tento experiment je zaměřen na ověření druhého předpokladu, který lze zjednodušeně interpretovat tak, že při použití přenášení důvěry mezi kontexty pomocí *algoritmu šíření události* agenti dosahují větší efektivity při rozhodování, neboť na základě jedné externí události může být aktualizováno více kontextů. Důvěra v kontext, jež byla stanovena na základě přenesení důvěry z kontextu jiného může být úspěšně použita v procesu rozhodování aniž by agent měl v tomto kontextu s protějškem přímou zkušenost.

Pro základní ověření principu přenášení důvěry mezi kontexty byla zvolena jednoduchá

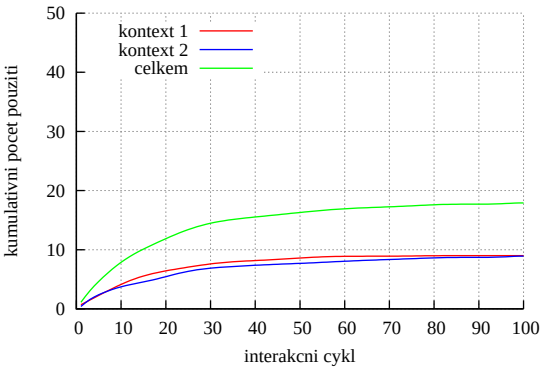


(a) Agent5 - jedno-kontextový přístup.

(b) Agent5 - více-kontextový přístup.

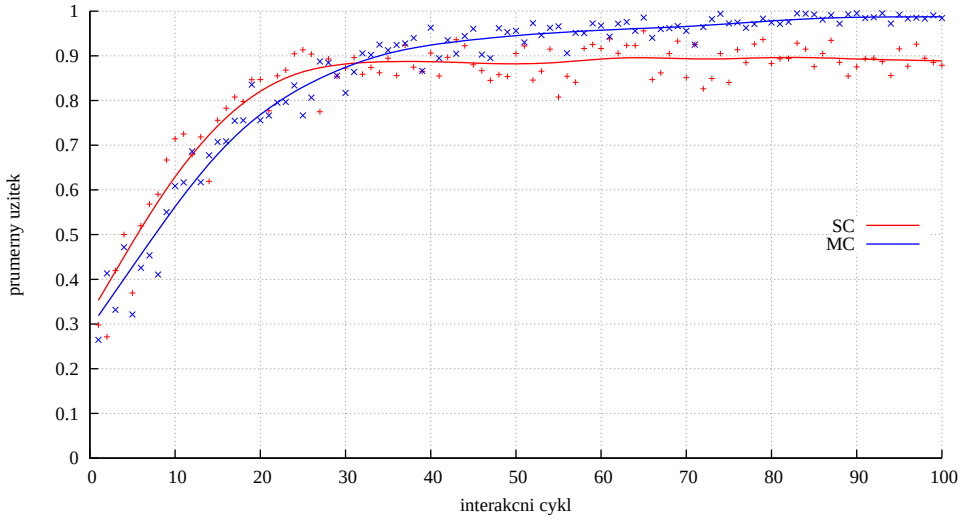


(c) Agent5 - jedno-kontextový přístup.

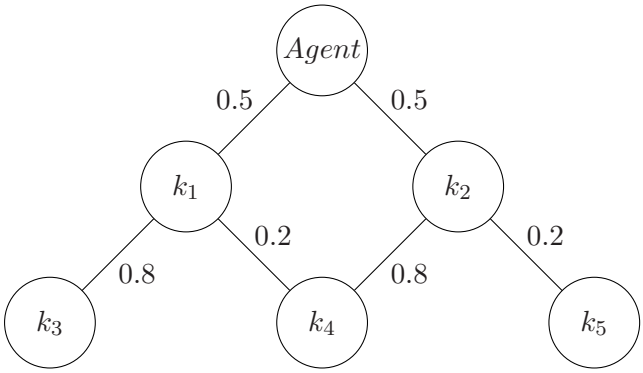


(d) Agent5 - více-kontextový přístup.

Obrázek 6.7: Průměrný a kumulativní počet použití agenta Agent5 vzhledem k interakčnímu cyklu.



Obrázek 6.8: Porovnání průběhu průměrného užitku pro SC a MC přístup všech agentů vzhledem k interakčnímu cyklu.



Obrázek 6.9: Jednoduchá struktura HMTC modelu pro ověření principu přenášení důvěry mezi kontexty.

HMTC struktura s pěti interakčními kontexty označenými $k_1, \ldots, k_5$ s propojením, jež je naznačeno na obrázku 6.9. Při použití této struktury s takto stanovenými váhami jednotlivých hran bude docházet k odvození důvěry směrem dolů hlavně mezi interakčními kontexty $k_1 \to k_3$, $k_1 \to k_4$, $k_2 \to k_4$ a $k_2 \to k_5$. K odvození důvěry směrem nahoru pak dochází hlavně mezi interakčními kontexty $k_3 \to k_1$, $k_4 \to k_1$, $k_4 \to k_2$ a $k_5 \to k_2$.

Hlavní parametry a sledované veličiny experimentu

Pro takto navrženou struktury byly zvoleny shodné parametry jako v případě „Experimentu 1“, tedy následovně: $CI = 95\%$, $\hat{u} = 0.3$, $r = 0.2$, funkce užitku $\mu(x)$ je definována ekvivalentně se (6.5): $p^{\alpha}_{min} = 0.2$ a $p^{\alpha}_{max} = 0.8$.

Pro simulaci takto navrženého experimentu bylo rovněž použito 10 agentů (Agent1, ..., Agent10), kteří provedli každý celkem 100 interakčních cyklů. Modely chování jednotlivých agentů byly nastaveny s ohledem na strukturu HMTC a váhu jednotlivých hran mezi kontexty tak, abychom agenty rozdělili do dvou skupin (kde A1–A5 je skupina {Agent1, ..., Agent5} a A6–A10 je skupina {Agent6, ..., Agent10}), přičemž model chování agenta z první skupiny v konkrétním kontextu je jiný než model chování ve stejném kontextu agenta druhé skupiny. Tyto modely chování jednotlivých agentů byly stanoveny na základě hodnot váhy hran propojených kontextů a jsou popsány v tabulce 6.2.

	Model chování $N(\mu, \sigma)$				
Skupina	k_1	k_2	k_3	k_4	k_5
A1–A5	$N(0.73, 0.05)$	$N(0.23, 0.05)$	$N(0.85, 0.05)$	$N(0.25, 0.05)$	$N(0.15, 0.05)$
A6–A10	$N(0.27, 0.05)$	$N(0.77, 0.05)$	$N(0.15, 0.05)$	$N(0.75, 0.05)$	$N(0.85, 0.05)$

Tabulka 6.2: Modely chování jednotlivých agentů (Exp.2).

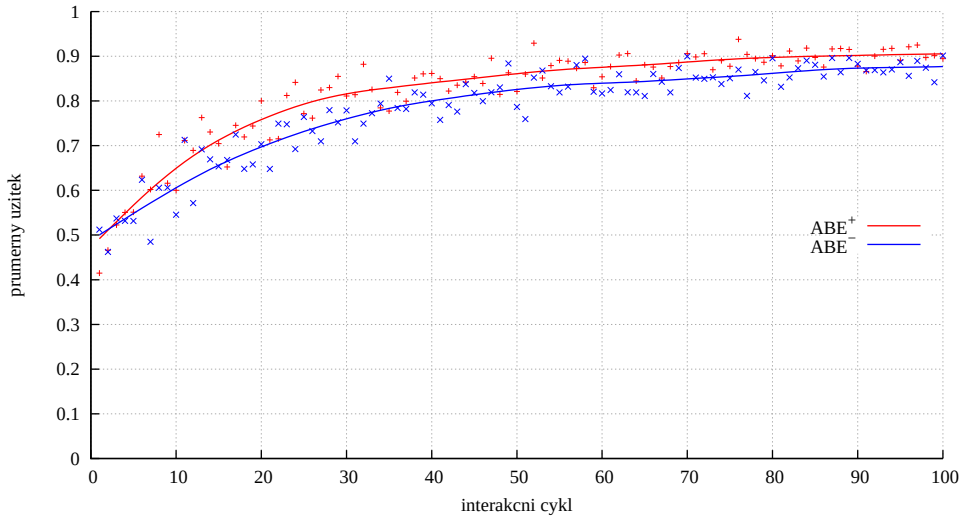
Abychom ověřili rozdíl mezi stavem, kdy je důvěra mezi kontexty odvozována a kdy ne, byly vytvořeny dvě konfigurace experimentů. V první konfiguraci bylo odvozování důvěry mezi kontexty povoleno tak, že při příchodu externí události byl proveden standardní *algoritmus šíření události* (ABE). V druhém případě byl ABE zakázán a po příchodu externí události byla provedena jen aktualizace konkrétního interakčního kontextu.

Podobně jako v předchozím „Experimentu 1“ bylo provedeno s danou konfigurací celkem 10 experimentů, přičemž hlavní sledovanou veličinou byla průměrná hodnota užitku všech agentů v daném interakčním cyklu. Tato hodnota do značné míry udává schopnost agentů provést optimální rozhodnutí na základě znalosti důvěry v jednotlivé kontexty potencionálních partnerů k interakci.

Výsledky

První graf na obrázku 6.10 znázorňuje průběh průměrné hodnoty užitku všech agentů v systému vzhledem k interakčnímu cyklu. Červená křivka reprezentuje situaci, kdy je ABE povolen (označeno jako ABE^{+}) a probíhá odvozování důvěry mezi kontexty na základě principu HMTC. Modrá křivka znázorňuje situaci, kdy ABE není povolen (označeno jako ABE^{-} a aktualizace důvěry probíhá pouze u jednotlivých kontextů tak, jako kdyby nebyly vůbec propojeny.

Jak je vidět, tak v případě kdy je ABE povolen, je průměrný užitek prakticky v celém průběhu mírně vyšší než v případě zakázaného ABE (jedinou výjimku tvoří několik prvních interakčních cyklů, neboť experimenty začínají s naprosto totožným nastavením počáteční důvěry).



Obrázek 6.10: Porovnání průběhu průměrného užitku s povoleným a zakázaným ABE.

Počty použití jednotlivých skupin agentů a jejich kontextů pro různé situace (ABE^{+} a ABE^{-}) jsou zachyceny v tabulce 6.3. Počty použití jsou vyjádřeny procentuálně, sloupec *skupina* označuje skupinu agentů, sloupec ABE^{+} zachycuje procentuální použití dané skupiny agentů v daném kontextu pro povolené ABE a sloupec ABE^{-} analogicky pro zakázané ABE. Poslední sloupec zachycuje rozdíl v procentních bodech mezi ABE^{+} a ABE^{-}.

Při porovnání modelů chování jednotlivých skupin agentů v různých kontextech, jež je uvedeno v tabulce 6.2, s výsledným rozdílem počtu použití jednotlivých agentů, dle tabulky 6.3 lze vyvodit následující:

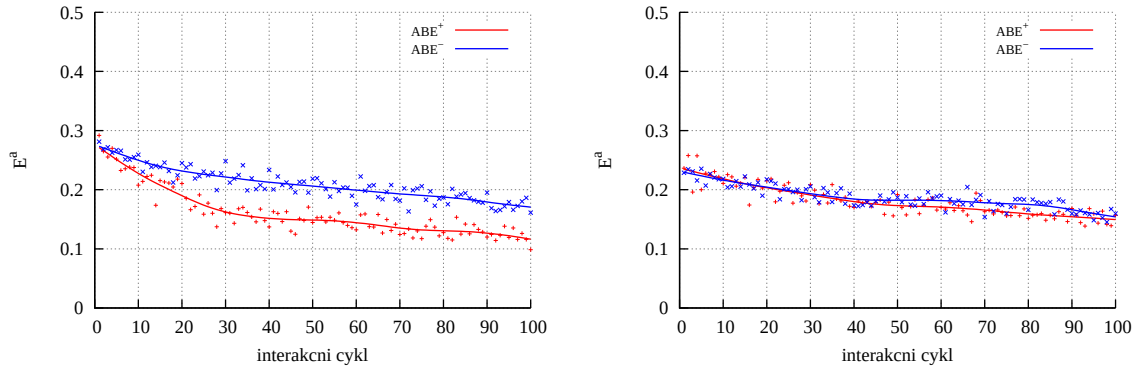
1. v případě „dobrého“ modelu chování skupiny agentů pro daný kontext je rozdíl ABE^{+} - ABE^{-} vždy pozitivní, tedy agenti jsou v tomto kontextu častěji používaní v případě povoleného ABE,

Kontext	Skupina	ABE$^+$	ABE$^-$	ABE$^+$ - ABE$^-$
k_1	A1–A5	17.18%	16.68%	0.50%
	A6–A10	2.48%	4.07%	-1.59%
k_2	A1–A5	2.11%	3.48%	-1.37%
	A6–A10	18.14%	16.17%	1.97%
k_3	A1–A5	17.89%	16.44%	1.45%
	A6–A10	2.14%	3.45%	-1.31%
k_4	A1–A5	2.86%	3.78%	-0.92%
	A6–A10	17.29%	16.03%	1.26%
k_5	A1–A5	3.38%	3.44%	-0.06%
	A6–A10	16.53%	16.46%	0.07%
		100%	100%	

Tabulka 6.3: Porovnání počtu použití jednotlivých skupin agentů vzhledem k interakčnímu kontextu s povoleným a zakázaným ABE.

2. naopak v případě „špatného“ modelu chování je rozdíl negativní a agenti jsou při povoleném ABE méně využívání než v případě zakázaného ABE.

Poslední sledovanou veličinou v tomto experimentu byla velikost průměrné absolutní E^a a kritické chyby E^k odhadu hodnoty důvěry pro daný interakční cyklus v jednotlivých kontextech. Jednotlivé průměrné chyby byly stanoveny na základě hodnot důvěry, které byly zaznamenány bezprostředně po provedení interakce v daném kontextu. Na obrázku 6.11(a) je zachycen průběh průměrné absolutní chyby odhadu důvěry pro kontext k_1 a skupinu agentů A1–A5 pro různé situace ABE$^+$ a ABE$^-$. Ostatní analyzované průběhy (ostatní kontexty a skupiny agentů) mají buď průběh E^a podobný (ABE$^+$ se vyznačuje nižší průměrnou absolutní chybou) anebo jsou jednotlivé křivky pro situace ABE$^+$ a ABE$^-$ téměř shodné (jako v případě E^a pro kontext k_2 a skupinu agentů A6–A10 na obrázku 6.11(b)).



(a) E^a pro kontext k_1 a skupinu agentů A1–A5. (b) E^a pro kontext k_2 a skupinu agentů A6–A10.

Obrázek 6.11: Průměrná absolutní chyba odhadu důvěry pro různé situace ABE$^+$ a ABE$^-$.

Při analýze průměrné kritické chyby E^k bylo zjištěno, že je pro danou množinu experimentů a pro téměř všechny interakční cykly nulová. Pokud v některém případě nabývá

nenulové hodnoty, jedná se o zanedbatelnou velikost $E^k \ll 0.01$, z tohoto důvodu zde nejsou jednotlivé průběhy prezentovány.

Zhodnocení „Experimentu 2“

Na základě výsledků prezentovaných výše lze konstatovat, že druhý předpoklad stanovený v kapitole 6.3.1 byl potvrzen. Při použití principu odvozování důvěry mezi kontexty se schopnost agentů ve smyslu přesnějšího rozhodování zvýšila a agenti si při rozhodování o výběru vhodného partnera počínali lépe než v případě, kdy odvozování důvěry neprobíhalo.

Tento závěr však platí pouze za předpokladu, že je použita taková HMTC struktura, která odvozování důvěry mezi kontexty umožňuje. Dále je třeba konstatovat, že výsledné zvýšení efektivity při rozhodování agentů s použitím odvozování důvěry, nebylo v případě provedených experimentů nijak zvlášť výrazné. Změny se projevovaly v případě průměrného užitku v průměru v jednotkách procent a v případě počtu použití v jednotkách anebo v desetinách procentního bodu.

6.5 Shrnutí

Navržený více-kontextový modelu důvěry HMTC, včetně způsobu stanovení intervalu důvěry na základě výsledků přímých zkušeností, byl v této kapitole převeden do reálné aplikace, prototypu distribuovaného multi-agentního systému s názvem `HMTCsim`. V první polovině této kapitoly je popsán základní koncept navrženého prototypu včetně jednotlivých komponent jak z funkčního hlediska, tak i z implementačního hlediska. Velká část této kapitoly rovněž popisuje princip vzájemných interakcí mezi agenty a způsob rozhodování agentů na základě důvěry a výběr optimálního partnera pro následující interakci pro zvolený interakční kontext. Součástí vytvořeného prototypu je rovněž komplexní simulační rozhraní jak pro spouštění experimentů s navrženým modelem, tak i pro ukládání simulačních výsledků pro následující zpracování.

Druhá část kapitoly se zaměřuje na experimentální ověření funkčnosti prototypu, přičemž pro vzájemné porovnání jednotlivých experimentů jsou zde použity dvě hlavní metriky: *počet použití* a *průměrný užitek*. Experimenty byly koncipovány tak, abychom ověřili dva základní předpoklady, které vycházejí z hlavních přínosů disertační práce, to jest: 1) jedno- versus více-kontextový přístup k důvěře a 2) možnost přenášení důvěry mezi kontexty na základě definice vztahů mezi nimi pomocí grafové struktury.

Experimentálně bylo ověřeno, že více-kontextový přístup je vhodný v případě, kdy je třeba agenty hodnotit v různých kontextech, přičemž jejich chování (ve smyslu kvality) v rámci těchto kontextů je různé. V případě, že agenti mají v různých kontextech shodné nebo velmi podobné chování, pak je z důvodu rychlosti odhadu jejich chování výhodnější použít jedno-kontextový přístup. Rovněž bylo experimentálně ověřeno, že v případě přenášení důvěry mezi kontexty na základě definice jejich vzájemné vazby je možné dosáhnout přesnějšího rozhodování agentů ve smyslu výběru vhodnějšího partnera pro interakci.

Kapitola 7

Závěr

Tato disertační práce se zabývá problematikou modelování více-kontextové důvěry v prostředí distribuovaných agentních a multi-agentních systémů. Tato kapitola shrnuje zvolené přístupy k řešení problematice a dosažené výsledky. Nakonec jsou zde popsány možné směry dalšího výzkumu.

7.1 Dosažené výsledky

Problematika modelování důvěry a reputace je multidisciplinární obor a výzkum těchto jevů probíhal a probíhá v celé řadě oblastí. Tato práce se zaměřila na modelování důvěry s více-kontextovým přístupem do oblasti umělé inteligence, konkrétně pak do oblasti distribuovaných agentních a multi-agentních systémů. Více-kontextovým přístupem se zde rozumí stanovením důvěry pro různé aspekty jedné důvěryhodné entity tak, aby bylo možné takovou entitu posuzovat subjektivně z mnoha různých pohledů. Zde je třeba také podotknout, že důvěra i reputace je vždy subjektivní a proto neexistuje žádná universální objektivní metoda, jak je stanovit.

V předložené disertační práci byl navržen model důvěry, pojmenovaný jako *hierarchický model důvěry s kontexty* (HMTC), jež vychází nejen s faktu, že důvěryhodná entita může být posuzována z různých hledisek a aspektů, ale navíc i z faktu, že jednotlivé aspekty entity jsou vždy součástí celku a nelze je posuzovat zcela odděleně. Vzhledem k tomuto byl v práci navržen způsob, který pro více-kontextový model důvěry umožňuje definovat vzájemné propojení jednotlivých kontextů v hierarchickou strukturu a vytvořit tím tak celek definující důvěryhodnou entitu z množiny různých aspektů. Pro takto navrženou strukturu kontextů je možné provádět aktualizaci důvěry pro různé kontexty, přičemž vzájemné propojení kontextů umožňuje provést odvození důvěry na základě znalosti důvěry v jeden kontext pro kontext druhý. Formální návrh HMTC modelu včetně strukturálních pravidel byl publikován v práci [SZ10c], způsob aktualizace jednotlivých kontextů a princip odvozování důvěry na základě událostmi řízeného modelu byl publikován v práci [SMZ10] a [SZ10a].

Pro navržený model důvěry byla stanovena unikátní reprezentace hodnoty důvěry pomocí intervalu, která reprezentuje nejen důvěryhodnost entity ale i nejistotu, která je s hodnotou důvěry vždy nějakým způsobem spojena. Význam této intervalové reprezentace byl pro použití s navrženým způsobem vyhodnocování explicitně definován v kapitole 5 jako míra pravděpodobnosti, se kterou entita provede interakci v daném kontextu s kvalitou ve stanoveném rozsahu. Jako pravděpodobnostní rozložení bylo zvoleno normální rozložení pravděpodobnosti, které je definováno pomocí střední hodnoty a rozptylu, přičemž střední

hodnota udává *typické* chování agenta a rozptyl udává přípustnou odchylku od tohoto chování. Na základě statistických metod odhadů parametrů normálního rozložení byl navržen robustní způsob odhadu důvěry na základě výsledku interakce mezi entitami v různých kontextech. Tento způsob definuje, spolu s principem vytváření a aplikace událostí v modelu HMTC, postup pro aktualizaci důvěry v jednotlivé kontexty. Princip statistického vyhodnocení důvěry na základě přímých interakcí mezi agenty byl publikován v práci [SMZH11].

Navržený více-kontextový model důvěry a s ním související principy odhadu důvěry, její aktualizace a šíření v rámci hierarchické struktury byl implementován v rámci prototypu multi-agentního systému, který společně s algoritmem pro rozhodování agentů na základě důvěry vytvořil simulační nástroj `HMTCsim`. Pomocí tohoto nástroje bylo provedeno základní ověření funkčnosti navrženého modelu a principu více-kontextového přístupu. Bylo experimentálně ověřeno, že v případě kdy je možné agenty posuzovat v různých aspektech různých kvalit je výrazně výhodnější použít více-kontextový přístup na rozdíl od jednokontextového. Princip odvozování důvěry mezi propojenými kontexty jednoho agenta umožnil zvýšení efektivity rozhodování agentů ve smyslu schopnosti optimálního výběru partnera pro transakci. Experimentální výsledky prezentované v kapitole 6 však také jednoznačně ukázaly na fakt, že ne vždy je více-kontextový přístup vhodný a vždy je třeba zvážit použití takového přístupu pro řešení konkrétního problému. Experimentální výsledky vytvořené s použitím implementovaného simulačního nástroje `HMTCsim`, spolu s principem rozhodování agentů na základě více-kontextové důvěry, byly publikovány v práci [MSZZ11]. Jedním z dílčích výstupů disertační práce je rovněž simulační nástroj `ContextGraph`, publikovaný v práci [SZ10b] a využitelný pro simulaci funkce algoritmu šíření události nad hierarchickou strukturou HMTC.

Koncept více-kontextového přístupu, který je popsán v této disertační práci, se v obecné rovině chápání a využití důvěry prokázal jako použitelný a přínosný z hlediska nových poznatků obohacující velmi širokou a komplexní problematiku fenoménu důvěry a reputace.

7.2 Možné směřování dalšího výzkumu

Pevně věřím, že výsledky popsané v této disertační práci přispěly k rozšíření poznatků v oblasti budování modelů důvěry. Současně s tímto jsme však odhalili řadu problémů, otázek a nových výzev, na které by bylo vhodné se zaměřit v budoucím výzkumu. Jmenujme tedy zde některé zásadní z nich, které rozhodně stojí za zmínku a jimiž by bylo vhodné se v dalším výzkumu zabývat:

- Zahrnutí doporučení a reputace v rámci aktualizace důvěry v HMTC struktuře a jejich využití při rozhodování. Tento problém je již z velké míry, v rámci výzkumného týmu na FIT VUT v Brně, zpracován jak v HMTC modelu (formou stanovení váhy HMTC událostí a způsobem jejich odstranění při vzniku konfliktů), tak i v rozhodovacím algoritmu (navržen a implementován protokol pro dotazování a doporučování) je téměř vyřešen. Bylo experimentálně ověřeno, že poskytování doporučení a z nich vzniknuvší reputace má pozitivní vliv na budování důvěry a schopnosti agentů při rozhodování. V tomto směru tak zbývá zpracovat pouze návrh řešení na šíření zpětné vazby agentům, kteří doporučení poskytli.

- Zohlednění problematiky dynamického chování agentů. Námi navržený způsob odhadu důvěry a následná aktualizace důvěry v jednotlivé kontexty HMTC struktury vychází z předpokladu, že chování agenta je víceméně konstantní a na základě každé

další interakce provádíme s využitím důvěry pouze zpřesnění odhadu chování. Předběžné experimenty ukázaly, že současný návrh si s dynamickým chováním částečně poradí (interval důvěry je možné ve velmi omezené míře postupně rozšířit díky vznikajícím konfliktům), ale pro efektivní řešení tohoto problému je třeba některé navržené přístupy zcela revidovat.

- Způsob stanovení hierarchické struktury kontextů včetně určení jednotlivých vah propojených hran. Jak bylo ukázáno na provedených experimentech v podkapitole 6.4.2, tak výsledná efektivita odvozování důvěry mezi kontexty je výraznou měrou ovlivněna především zvolenou strukturou propojených kontextů a váhou jednotlivých hran. Proto se tak nabízí otázka, jakým způsobem optimálně tuto strukturu definovat tak, aby bylo dosaženo větší efektivity šíření důvěry, než v případě empiricky zvolené konfigurace. V tomto směru by bylo rovněž vhodné ověřit možnosti dynamicky se měnících vah hran propojených kontextů v čase, případně možnost definovat tato propojení s udáním směru (pouze jednosměrné, obousměrné asymetrické apod.).

- Alternativní významy intervalu důvěry. V této práci je interval důvěry chápan jako odhad modelu chování agenta v nějakém kontextu, přičemž model chování je popsán parametry normálního rozložení. Intervalová reprezentace důvěry však obecně nabízí celou řadu dalších alternativních způsobů definice jejího významu. Jednou z množností je použití *Beta rozložení* $B(\alpha, \beta)$ namísto normálního rozložení pravděpodobnosti pro definici modelu chování. Další možností, která se nabízí, je použití ve významu spojeném s *fuzzy logikou* a důvěře v ní reprezentované, kdy je třeba vyjádřit jak fuzzy hodnotu tak i stupeň příslušnosti.

Literatura

[AM02] Farag Azzedin and Muthucumaru Maheswaran. Towards trust-aware resource management in grid computing systems. In *Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid*, pages 452–457. IEEE Computer Society, 2002.

[AMCR04] Florina Almenárez, Andrés Marín, Celeste Campo, and Carlos García R. PTM: A pervasive trust management model for dynamic open environments. In *In First Workshop on Pervasive Security, Privacy and Trust PSPT'04*, 2004.

[AMDS06] Florina Almenárez, Andrés Marín, Daniel Dáz, and Juan Sanchéz. Developing a model for trust management in pervasive devices. In *Proc. of the 3rd IEEE International Workshop on Pervasive Computing and Communication Security (PerSec)*, pages 267–272. IEEE Computer Society Press, 2006.

[AR05] Alfarez Abdul-Rahman. *A Framework for Decentralised Trust Reasoning.* PhD thesis, University of London, Department of Computer Science, 2005.

[ARH97] A. Abdul-Rahman and S. Hailes. Using recommendations for managing trust in distributed systems. In *IEEE Malaysia International Conference on Communication '97 (MICC'97)*, 1997.

[ARH00] Alfarez Abdul-Rahman and Stephen Hailes. Supporting trust in virtual communities. *Hawaii International Conference on System Sciences*, 6:6007, 2000.

[Auk10] Aukro. Aukro - Aukce OnLine, 2010. http://www.aukro.cz [Online; cit. 3.5.2010].

[BHW07] Rafael H. Bordini, Jomi F. Hübner, and Michael Wooldridge. *Programming Multi-Agent Systems in AgentSpeak using Jason.* Wiley, 2007.

[Bra87] Michael E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987.

[BS94] José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory.* Wiley, 1994.

[Cof07] Piotr Cofta. *Trust, Complexity and Control: Confidence in a Convergent World.* John Wiley & Sons, Ltd, 2007.

[DG97] Marco Dorigo and Luca Maria Gambardella. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evolutionary Computation*, 1(1):53–66, 1997.

[DLB08] Pierpaolo Dondio, Luca Longo, and Stephen Barrett. A translation mechanism for recommendations. In *Proceedings of the 2nd Joint iTrust and PST Conferences on Privacy, Trust Management and Security (IFIPTM'08)*, pages 87–102, 2008.

[DS04] Marco Dorigo and Thomas Stützle. *Ant Colony Optimization*. MIT Press, Cambridge, MA, 2004.

[DW92] Shu-Ho Dai and Ming Wang. *Reliability Analysis in Engineering Applications*. Van Nostrand Reinhold, 1992.

[eBa10] eBay. eBay - Online auction, 2010. http://www.eaby.com [Online; cit. 3.5.2010].

[FC01] Rino Falcone and Cristiano Castelfranchi. The socio-cognitive dynamics of trust: Does trust create trust? In *Proceedings of the workshop on Deception, Fraud, and Trust in Agent Societies held during the Autonomous Agents Conference: Trust in Cyber-societies, Integrating the Human and Artificial Perspectives*, pages 55–72, London, UK, 2001. Springer-Verlag.

[FFK+91] Amos Fiat, Dean P. Foster, Howard Karloff, Yuval Rabani, Yiftach Ravid, and Sundar Vishwanathan. Competitive algorithms for layered graph traversal. In *SFCS '91: Proceedings of the 32nd annual symposium on Foundations of computer science*, pages 288–297, Washington, DC, USA, 1991. IEEE Computer Society.

[FSJ98] Peyman Faratin, Carles Sierra, and Nick R. Jennings. Negotiation decision functions for autonomous agents. *International Journal of Robotics and Autonomous Systems*, 24:159–182, 1998.

[Gam00] Diego Gambetta. Can we trust trust? In Diego Gambetta, editor, *Trust: Making and Breaking Cooperative Relatioins*, chapter 13, pages 213–237. Published Online, 2000.

[Gam10] Diego Gambetta. Nuffield college, diego gambetta, 2010. [Online; cit 5.12.2010].

[GBS08] Saurabh Ganeriwal, Laura K. Balzano, and Mani B. Srivastava. Reputation-based framework for high integrity sensor networks. *ACM Transactions on Sensor Networks*, 4(3):1–37, 2008.

[Gol05] Jennifer Ann Golbeck. *Computing and applying trust in web-based social networks*. PhD thesis, University of Maryland at College Park, College Park, MD, USA, 2005.

[Gra06] Elizabeth L. Gray. *A Trust-Based Reputation Management System*. PhD thesis, Trinity College, University of Dublin, April 2006.

[Gro10] POMO Media Group. Česko-slovenská filmová databáze (CSFD), 2010. http://www.csfd.cz [Online; cit. 3.5.2010].

[GS03] Tyrone Grandison and Morris Sloman. Trust management tools for internet applications. In *1st International Conference on Trust Management (iTrust 2003)*, pages 91–107, April 2003.

[HB04] Jomi F. Hübner and Rafael H. Bordini. Jason: a Java-based interpreter for an extended version of AgentSpeak, 2004. http://jason.sourceforge.net [Online; cit. 3.5.2011].

[Hol08] Lukáš Holčák. Statistická analýza souboru s malým rozsahem. Master's thesis, Vysoké Učení Technické v Brně, Fakulta Strojního Inženýrství, 2008.

[HPC98] Paul S. Horn, Amadeo J. Pesce, and Bradley E. Copeland. A robust approach to reference interval estimation and evaluation. *Clinical Chemistry*, 44(3), 1998.

[HS10] Jomi F. Hübner and Jaime S. Sichman. SACI - simple agent communication infrastructure, 2010. http://www.lti.pcs.usp.br/saci/ [Online; cit. 3.5.2010].

[Hur52] Leonid Hurwicz. A criterion for decision-making under uncertainty. Technical Report 355, Cowles Commission, 1952.

[IMD10] IMDb.com. The Internet Movie Datase (IMDB), 2010. http://www.imdb.com [Online; cit. 3.5.2010].

[JF01] Andrew J. I. Jones and Babak Sadighi Firozabadi. *On the characterisation of a trusting agent – aspects of a formal approach*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[JHP06] Audun Jøsang, Ross Hayward, and Simon Pope. Trust network analysis with subjective logic. In *Proceedings of the 29th Australasian Computer Science Conference - Volume 48*, ACSC '06, pages 85–94. Australian Computer Society, Inc., 2006.

[JIB07] Audun Jøsang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision support systems*, 43(2), 2007.

[Jøs97] Audun Jøsang. Artificial reasoning with subjective logic. In *Proceedings of the Second Australian Workshop on Commonsense Reasoning*, 1997.

[KR03] Michael Kinateder and Kurt Rothermel. Architecture and algorithms for a distributed reputation system. In *Proceedings of the 1st international conference on Trust management (iTrust 2003)*, pages 1–16. Springer-Verlag, 2003.

[KSGM03] Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina. The EigenTrust algorithm for reputation management in P2P networks. In *Proceedings of the 12th international conference on World Wide Web*, pages 640–651. ACM, 2003.

[Lab10] Telecom Italia Lab. JADE (java agent development framework), 2010. http://jade.tilab.com [Online; cit. 3.5.2010].

[Luh88] Niklas Luhmann. Familiarity, confidence, trust: Problems and alternatives. *D. Gambetta, editor, Trust: Making and Breaking of Cooperative Relations, Oxford, 1988*, pages 94–107, 1988.

[Mar94] Stephen Paul Marsh. *Formalising Trust as a Computational Concept.* PhD thesis, University of Stirling, April 1994.

[MC96] D. Harrison Mcknight and Norman L. Chervany. The meanings of trust. Technical report, Carlson School of Management, University of Minnesota, 1996.

[MM02a] Milan Meloun and Jiří Militký. *Kompedium Statistického Zpracování Dat.* Academia, 2002.

[MM02b] Pietro Michiardi and Refik Molva. CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. In *Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security: Advanced Communications and Multimedia Security*, pages 107–121, 2002.

[MMH02a] L. Mui, M. Mohtashemi, and A. Halberstadt. A computation model of trust and reputation. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02)*, volume 7, page 188, 2002.

[MMH02b] L. Mui, M. Mohtashemi, and A. Halberstadt. *Evaluating Reputation in Multi-agents Systems*, volume 2631/2003. Springer Berlin / Heidelberg, 2002.

[Mon97] Philippe Mongin. Expected utility theory. In *Handbook of Economic Methodology*, pages 342–350. Edward Elgar Publishing, 1997.

[MP08] Félix Gómez Mármol and Gregorio Martínez Pérez. Providing trust in wireless sensor networks using a bio-inspired technic. In *Networking and Electronic Commerce Research Conference (NAEC 08)*, 2008.

[MP09a] Félix Gómez Mármol and Gregorio Martínez Pérez. Security threats scenarios in trust and reputation models for distributed systems. *Computers & Security*, 28(7):545–556, 2009.

[MP09b] Félix Gómez Mármol and Gregorio Martínez Pérez. TRMSim-WSN, trust and reputation models simulator for wireless sensor networks. In *IEEE International Conference on Communications (IEEE ICC 2009), Communication and Information Systems Security Symposium*, 2009.

[MP10] Félix Gómez Mármol and Gregorio Martínez Pérez. Towards pre-standardization of trust and reputation models for distributed and heterogeneous systems. *Computer Standards & Interfaces*, 32(4):185–196, 2010.

[MPS09] Félix Gómez Mármol, Gregorio Martínez Pérez, and Antonio F. Gómez Skarmeta. TACS, a trust model for P2P networks. *Wireless Personal Communications, Special Issue on "Information Security and data protection in Future Generation Communication and Networking"*, 51(1):153–164, 2009.

[MR03] Douglas C. Montgomery and George C. Runger. *Applied Statistics and Probability for Engineers*. John Wiley & Sons, 3 edition, 2003.

[MSZZ11] Ondřej Malačka, Jan Samek, František Zbořil, and František Vítězslav Zbořil. Decision making and recommendation protocol based on trust for multi-agent systems. In *Proceedings of Trust, Reputation and User Modeling Workshop (TRUM 2011)*, pages 33–40, 2011.

[Mui03] L. Mui. *Computational Models of Trust and Reputation: Agents, Evolutionary Games, and Social Networks*. PhD thesis, Massachusetts Institute of Technology, 2003.

[Net09] Arnoštka Netrvalová. *Modelovaní důvěry*. PhD thesis, University of West Bohemia, 2009.

[Nov06] Jana Novovičová. *Pravděpodobnost a matematická statistika*. Vydavatelstvý ČVUT, 2006.

[PY89] Christos H. Papadimitriou and Mihalis Yannakakis. Shortest paths without a map. In *ICALP '89: Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, pages 610–620, London, UK, 1989. Springer-Verlag.

[Rao96] Anand S. Rao. AgentSpeak(L): BDI agents speak out in a logical computable language. In *Proceedings of the 7th European workshop on Modelling autonomous agents in a multi-agent world*, pages 42–55, Secaucus, NJ, USA, 1996. Springer-Verlag New York, Inc.

[Ras96] Lars Rasmusson. Socially controlled global agent systems. Master's thesis, Royal Institute of Technology, Stockholm, Department of Computer and Systems Sciences, 1996.

[RG91] Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI-architecture. In James Allen, Richard Fikes, and Erik Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484. Morgan Kaufmann publishers Inc.: San Mateo, CA, USA, 1991.

[RHJ04] S. D. Ramchurn, D. Hunyh, and N. R. Jennings. Trust in multi-agent systems. *Knowledge Engineering Review*, 19:1:1–25, 2004.

[Sam07] Jan Samek. A trust-based model for multi-agent systems. In *Proceedings of the 6th EUROSIM Congress on Modelling and Simulation*, page 6, 2007.

[SE08] Eugen Staab and Thomas Engel. Combining cognitive and computational concepts for experience-based trust reasoning. In *Proc. of the 11th Int. Workshop on Trust in Agent Societies (TRUST '09)*, pages 41–45, May 2008.

[Sha06] William T. Shaw. Sampling student's T distribution – use of the inverse cumulative distribution function. *The Journal of Computational Finance*, 9(4), 2006.

[SK09] Xiaoyuan Su and Taghi M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009, 2009.

[sKD05] Audun Jøsang, Claudia Keser, and Theo Dimitrakos. Can We Manage Trust? In *Proceedings of the Third International Conference on Trust Management (iTrust), Versailes*, pages 93–107. Springer-Verlag, 2005.

[SMZ10] Jan Samek, Ondřej Malačka, and František Zbořil. Event driven multi-context trust model. In *The 10th International Conference on Intelligent Systems Design and Applications (ISDA 2010)*, pages 911–917. IEEE Computer Society, 2010.

[SMZH11] Jan Samek, Ondřej Malačka, František Zbořil, and Petr Hanáček. Multi-agent experimental framework with hierarchical model of trust in contexts for decision making. In *Proceeding of the 2nd International Conference on Computer Modelling and Simulation*, pages 128–136. Department of Intelligent Systems FIT BUT, 2011.

[SS01] Jordi Sabater and Carles Sierra. REGRET: reputation in gregarious societies. In *Proceedings of the fifth international conference on Autonomous agents*, pages 194–195, 2001.

[SS02a] Jordi Sabater and Carles Sierra. Reputation and social network analysis in multi-agent systems. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 475–482, New York, NY, USA, 2002.

[SS02b] Jordi Sabater and Carles Sierra. *Social aspects of ReGreT, a reputation model based on social relations*, pages 336–343. Lecture Note in Artificial Intelligence. Springer, 2002.

[SS02c] S. Sen and N. Sajja. Robustness of reputation-based trust: boolean case. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 288–293, 2002.

[SS05] Jordi Sabater and Carles Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, 2005.

[SvB84] Eric P. Smith and Gerald van Belle. Nonparametric estimation of species richness. *Biometrics*, 40(1), 1984.

[SZ10a] Jan Samek and František Zbořil. Algorithmic evaluation of trust in multilevel model. In *EUROSIM 2010 - 7th EUROSIM Congeres on Modelling and Simulation*. Czech Technical University in Prague, 2010.

[SZ10b] Jan Samek and František Zbořil. ContextGraph: Simulation tool for hierarchical model of trust in context. In *Proceedings of CSE 2010 International Scientific Conference on Computer Science and Engineering*, Vol. 1, pages 265–270. The University of Technology Košice, 2010.

[SZ10c] Jan Samek and František Zbořil. Hierarchical model of trust in contexts. In *Networked Digital Technologies*, volume 88 of *Communications in Computer and Information Science (CCIS)*, pages 356–365. Springer Verlag, 2010.

[WE05] 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.

[Wik10a] Wikipedia. Interval arithmetic – Wikipedia, The Free Encyclopedia, 2010. [Online; cit. 5.12.2010].

[Wik10b] Wikipedia. Mode (statistics) — Wikipedia, The Free Encyclopedia, 2010. [Online; cit 5.12.2010].

[Wik10c] Wikipedia. Proactivity — Wikipedia, The Free Encyclopedia, 2010. [Online; cit 19.12.2010].

[WP97] J. Winkler and M. Petrusek. *Velký sociologický slovník*. Karolinum Praha, 1997. ISBN 80-7184-164-1.

[WS06] Yonghong Wang and Munindar P. Singh. Trust representation and aggregation in a distributed agent system. In *Proceedings of the 21st national conference on Artificial intelligence - Volume 2*, pages 1425–1430. AAAI Press, 2006.

[WV03] Y. Wang and J. Vassileva. Trust and reputation model in peer-to-peer networks. In *P2P '03: Proceedings of the 3rd International Conference on Peer-to-Peer Computing*, page 150, 2003.

[ZM00] G. Zacharia and P. Maes. Trust management through reputation mechanisms. *Applied Artificial Intelligence*, pages 881–907, 2000.

[ZZK06] Tang Zhuo, Lu Zhengding, and Li Kai. Time-based dynamic trust model using ant colony algorithm. *Wuhan University Journal of Natural Sciences*, 11:1462–1466, 2006.

Vybrané publikace autora

[MSZ10] Ondřej Malačka, Jan Samek, and František Zbořil. Increasing profit in agent business model with trust. In *Proceedings of the 7th EUROSIM Congress on Modelling and Simulation*, Vol. 2, page 6. Czech Technical University Publishing House, 2010.

[MSZZ11] Ondřej Malačka, Jan Samek, František Zbořil, and František Vítězslav Zbořil. Decision making and recommendation protocol based on trust for multi-agent systems. In *Proceedings of Trust, Reputation and User Modeling Workshop (TRUM 2011)*, pages 33–40, 2011.

[Sam06] Jan Samek. Security model of information systems. In *Proceedings of XXVIIIth International Autumn Colloquium ASIS 2006*, pages 101–105. MARQ, 2006.

[Sam07] Jan Samek. A trust-based model for multi-agent systems. In *Proceedings of the 6th EUROSIM Congress on Modelling and Simulation*, page 6, 2007.

[SMZ10] Jan Samek, Ondřej Malačka, and František Zbořil. Event driven multi-context trust model. In *The 10th International Conference on Intelligent Systems Design and Applications (ISDA 2010)*, pages 911–917. IEEE Computer Society, 2010.

[SMZH11] Jan Samek, Ondřej Malačka, František Zbořil, and Petr Hanáček. Multi-agent experimental framework with hierarchical model of trust in contexts for decision making. In *Proceeding of the 2nd International Conference on Computer Modelling and Simulation*, pages 128–136. Department of Intelligent Systems FIT BUT, 2011.

[SZ08] Jan Samek and František Zbořil. Multiple context model for trust and reputation evaluating in multi-agent systems. In *Proceedings of CSE 2008*, pages 336–343. Faculty of Electrical Engineering and Informatics, University of Technology Kosice, 2008.

[SZ09] Jan Samek and František Zbořil. Agent reasoning based on trust and reputation. In *Proceedings MATHMOD 09 Vienna - Full Papers CD Volume*, pages 538–544. ARGE Simulation News, 2009.

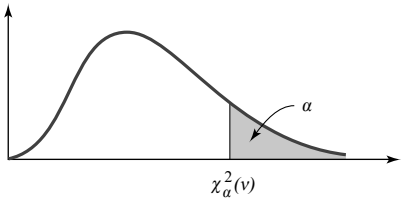
[SZ10a] Jan Samek and František Zbořil. Algorithmic evaluation of trust in multilevel model. In *EUROSIM 2010 - 7th EUROSIM Congeres on Modelling and Simulation*. Czech Technical University in Prague, 2010.

[SZ10b] Jan Samek and František Zbořil. ContextGraph: Simulation tool for hierarchical model of trust in context. In *Proceedings of CSE 2010 International Scientific Conference on Computer Science and Engineering*, Vol. 1, pages 265–270. The University of Technology Košice, 2010.

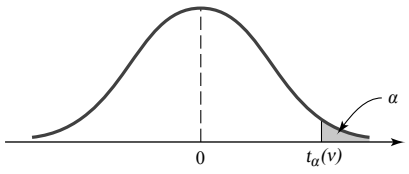
[SZ10c] Jan Samek and František Zbořil. Hierarchical model of trust in contexts. In *Networked Digital Technologies*, volume 88 of *Communications in Computer and Information Science (CCIS)*, pages 356–365. Springer Verlag, 2010.

Příloha A

Tabulky kritických hodnot důležitých rozložení



Obrázek A.1: Naznačení průběhu χ^2-rozložení a významu hodnoty α.



Obrázek A.2: Naznačení průběhu t-rozložení a významu hodnoty α.

ν \ α	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.191	55.003
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	43.745	47.400	50.725	54.776	57.648
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	46.059	49.802	53.203	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	48.363	52.192	55.668	59.893	62.883
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.996	21.426	23.654	25.695	28.196	50.660	54.572	58.120	62.428	65.476
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Tabulka A.1: Kritické hodnoty χ^2-rozložení, ν je stupeň volnosti.

ν \ α	0.2	0.1	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	1.376	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	1.061	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.978	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.941	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.920	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.906	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.896	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.889	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.883	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.879	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.876	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.873	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.870	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.868	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.866	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.865	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.863	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.862	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.861	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.860	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.859	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.858	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.858	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24	0.857	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.856	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.856	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.855	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.855	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.854	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.854	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
31	0.853	1.309	1.696	2.040	2.453	2.744	3.022	3.375	3.633
32	0.853	1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622
33	0.853	1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611
34	0.852	1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601
35	0.852	1.306	1.690	2.030	2.438	2.724	2.996	3.340	3.591
36	0.852	1.306	1.688	2.028	2.434	2.719	2.990	3.333	3.582
37	0.851	1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574
38	0.851	1.304	1.686	2.024	2.429	2.712	2.980	3.319	3.566
39	0.851	1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558
40	0.851	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.849	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.848	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
70	0.847	1.294	1.667	1.994	2.381	2.648	2.899	3.211	3.435
80	0.846	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
90	0.846	1.291	1.662	1.987	2.368	2.632	2.878	3.183	3.402
100	0.845	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
∞	0.842	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Tabulka A.2: Kritické hodnoty t-rozložení, ν je stupeň volnosti.