ZÁKLADY DEMPSTER-SHAFEROVY TEORIE A JEJÍ APLIKACE PRO MODELOVÁNÍ BEZPEČNOSTI A SPOLEHLIVOSTI (I.)

BERÁNEK L.
Jihočeská univerzita, Katedra aplikované matematiky a informatiky, České Budějovice, beranek@ef.jcu.cz

V tomto třídílném seriálu budeme demonstrovat aplikaci Dempster-Shaferovy teorie na problémy z oblasti bezpečnostního a spolehlivostního inženýrství. Tato teorie se používá zejména v expertních systémech, má však uplatnění i jinde, zejména u metod, jejichž součástí je expertní odhad. V tomto prvním díle příspěvku popíšeme základy Dempster-Shaferovy teorie, v druhém díle budeme tyto základy ilustrovat na příkladech. Ve třetím díle ukážeme aplikaci na analýze stromu událostí. Cílem celého miniseriálu je ukázat, že Dempster-Shaferova teorie oproti teorii pravděpodobnosti lépe pracuje s neurčitostí vyplývající např. z nedostatku dostupných informací nebo neznalostí a že ji lze s výhodou použít pro modelování bezpečnosti a spolehlivosti.

1. Úvod

Základní přístup k modelování nejistoty pomocí domněnkových funkcí (také pomocí Dempster-Shaferovy teorie) lze demonstrovat na jednoduchém příkladě. Písmenem h označíme jev, že určitý specialista označí jako příčinu havárie nefunkční hladinoměr v kotli. Pak podle pravidel pravděpodobnosti platí:

$$P(h) + P(\neg h) = 1 \quad (1)$$

kde $\neg h$ je negace h.

Předpokládejme, že máme druhého specialistu, který však nezná konstrukci daného zařízení. Potom nemůžeme stanovit pravděpodobnost toho, že tento specialista označí nefunkční hladinoměr za příčinu havárie, ale ani toho, že za příčinu havárie nefunkční hladinoměr neoznačí. Důvodem je to, že tento druhý specialista nemá představu (bez obeznámení se s příslušnými výkresy a technologickými postupy), co by vlastně mohlo být příčinou havárie. Tento problém lze lépe popsat pomocí Dempster-Shaferovy (DS) teorie [1], která je považována za obecnější přístup k reprezentaci nejistoty než Bayesovský přístup. V DS teorii pro vyjádření postoje druhého specialisty týkajícího se jevu označení nefunkčního hladinoměru za příčinu havárie kotle používáme termín domnění (belief, přesvědčení, důvěra). Jeho míru domnění označíme $m(h)$ a její negaci za $m(\neg h)$, v našem případě se obě hodnoty budou blížit nule. To je scénář, který klasický pravděpodobnostní přístup neumožňuje. Ten se soustředí spíše na objektivní ohodnocení výskytu jevu, kdežto přístup pomocí domněnkových funkcí je spíše založen na přesvědčení týkajícím se výskytu daného jevu.

Rozdíl mezi Bayesovským pravděpodobnostním modelem a DS teorií je tedy konceptuální. V pravděpodobnostním modelu se předpokládají Booleovské jevy, které buď existují nebo neexistují. Důsledkem těchto předpokladů je, že přesvědčení o správnosti určitého jevu (výroku, hypotézy) je spojeno s přesvědčením o neplatnosti negace tohoto jevu (výroku, hypotézy). V DS teorii není žádný příčinný vztah mezi výrokem (hypotézou, jevem) a jeho negací. Z toho důvodu nedostatek domnění o určitém výroku (hypotéze, jevu) neříká nic o přesvědčení o negaci tohoto výroku (hypotézy, jevu). Spíše nedostatek domnění o určité specifické hypotéze naznačuje domnění vztahující se k celé množině výroků, což nazýváme stav nejistoty (nevíme, který z výroků z dané množiny je ten správný, a proto přiřadíme domnění k celé množině výroků). Jestliže označíme nejistotu C, pak ve výše uvedeném příkladu $m(C) = 1$, což lze znázornit pomocí vztahu: $m(h) + m(h) + m(C) = 1$ (viz také níže rovnicí (1)).

2. Základní pojmy Dempster-Shaferovy teorie

Dempster-Shaferova teorie [2, 3, 4] definuje na počátku celkovou oblast úvah (frame of discernment) nazvanou také rámec zjištění nebo také rámec domnění, což je úplný systém vzájemně disjunktních (hypotézy se významově nepřekrývají) základních hypotéz $\Theta = \{h_1, \ldots, h_n\}$. Hypotézy mohou vyjadřovat všechny možné stavy (například příčiny poruchy) uvažovaného systému. 2^Θ označuje množinu všech podmnožin množiny Θ. Jedná se o potenční množinu množiny Θ. Libovolná podmnožina 2^Θ může být jednotlivá hypotéza h_i, nebo množina hypotéz, např. $\{h_1, h_2\}$ apod.

V tomto kontextu jsou důkazy symptomy nebo události, které se vyskytly nebo mohou vyskytnout. Důkaz se vztahuje k jednotlivé hypotéze nebo množině hypotéz (např. určitý symptom (hodnota měřicího přístroje) ukazuje na určitou poruchu). Předpokládáme přitom, že různé důkazy nepovedou ke stejné hypotéze nebo množině hypotéz. Kvalitativní vztah mezi důkazem a hypotézou odpovídá vztahu příčina-důsledek. Důkaz vede k formulaci hypotézy nebo množiny hypotéz. Stupeň této implikace je určen údajem datového zdroje.

Jako datové zdroje označujeme osoby, organizace nebo nějaké jiné entity, která poskytují informace pro vyhodnocení. V bezpečnostním a spolehlivostním inženýrství jako datové zdroje slouží obvykle výsledky empirických studií nebo závěry expertů. Avšak z pohledu datového zdroje (například experta nebo operátora, který posuzuje určitou závadu) nemusí být jisté, která hypotéza odpovídá objektivní realitě. Výhodou použití Dempster-Shaferovy teorie je to, že umožňuje uvažovat více

– jednotlivých důkazů v rámci vztahu jednotlivé hypotézy nebo
– jednotlivých důkazů v rámci vztahu více hypotéz

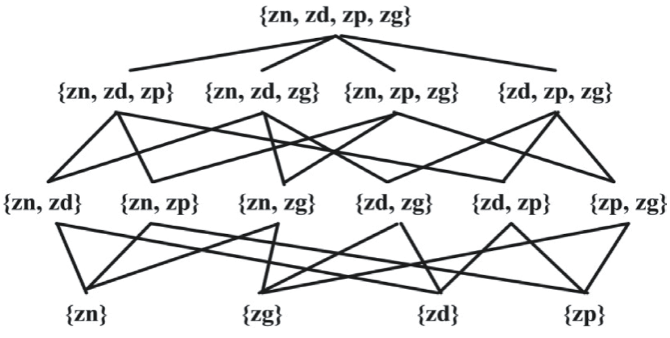
jako neurčité ohodnocení systému, v kterém je právě jen jedna hypotéza objektivně pravdivá.

Příklad 1

Počítač nepracuje správně. Možné příčiny jsou závada napájení (*zn*), závada na hlavní desce (*zd*), závada operační paměti (*zp*) nebo závada grafické karty (*zg*). Definujte rámec domnění.

Odpověď: Uvedené příčiny tvoří množinu vzájemně disjunktních základních (atomických) hypotéz. Rámec domnění tedy je množina $\Theta = \{zn, zd, zp, zg\}$.

Obr. 1 – Množina hypotéz (potenční množina 2^Θ)



V DS teorie je každá podmnožina hypotéz z 2^Θ považována za specifickou hypotézu, viz obr. 1. Nedostatek domnění k jednotlivým hypotézám se nahrazuje domněním k celé množině všech hypotéz 2^Θ, což nazýváme stav nejistoty.

2.1 Domnění

Uvedli jsme, že v DS teorii důkazu používáme pojem domnění [4] jako přesvědčení nebo stupeň důvěry k určité specifické hypotéze. Míra domnění ke specifické hypotéze se v DS teorii označuje reálným číslem z intervalu <0, 1>. Toto číslo indikuje také stupeň, se kterým jsme přesvědčeni, že důkazy (evidence) podporují hypotézu. V DS teorii se důkaz proti specifické hypotéze považuje za důkaz pro její negaci (například důkaz proti {*zn*} je považován za důkaz pro {*zd, zp, zg*} a podle toho bude přiřazen stupeň domnění).

2.2 Základní přiřazení

Základní přiřazení (*basic belief assignment, bba*) je funkce *m* definovaná na množině všech podmnožin množiny Θ nabývající hodnot z intervalu <0, 1>, tj. m: $2^\Theta \in$ <0, 1>, která má tuto vlastnost:

$$\sum_{A \subseteq \Theta} m(A) = 1 \qquad (2)$$

Tato rovnost znamená, že všechna tvrzení jednotlivých datových zdrojů musí být normalizována, aby se zajistilo, že důkazy uvedené datovými zdroji mají stejnou váhu (to znamená, že žádný datový zdroj není důležitější než jiný).

Platí tedy, že základní přiřazení (*bba*) reprezentované funkcí *m* přiřazuje každému prvku množiny 2^Θ míru domnění pomocí reálného čísla z intervalu <0, 1> tak, že součet těchto čísel je roven 1. Hodnota *m(A)* reprezentuje míru domnění, která je přiřazena každému prvku *A* z množiny 2^Θ. Na tomto místě je třeba rozlišit mezi pravděpodobností a základním přiřazením. Pravděpodobnostní distribuční funkce jsou definované na množině Θ. Základní přiřazení je ale na rozdíl od pravděpodobnosti definováno na potenční množině 2^Θ.

Podle těchto podmínek nemusí být domnění přiřazeno jen základním (atomickým) hypotézám h_i, ale i libovolné množině $A = \{h_1,\ldots,h_n\}$, kde A ´ 2^Θ. Hodnota *m(A)* představuje míru domnění, že platí hypotéza *A*, přičemž nevypovídá nic o míře domnění k složkám množiny *A*. Nemusí tedy platit $m(B) \leq m(A)$ pro $B \subseteq A$. Pro $A \neq h_i$ (kde h_i je libovolná atomická hypotéza *i*), je *m(A)* míra podpory, kterou jsme ochotní přidělit složené hypotéze *A* na úkor podpory $m(h_i)$ atomických hypotéz h_i.

Například jestliže pro určité specifické prostředí Θ bychom měli množinu $m(h_i) \neq 0$ pouze pro atomické hypotézy, pak $m(h_i)$ přejde v pravděpodobnosti *pi* , kde platí:

$$\sum_{i} m(h_i) = 1 \qquad (3)$$

Důsledkem částečné neznalosti (ignorance) spojené s *A* je následující nerovnost: $m(A)+m(\neg A) \leq 1$, kde $\neg A$ je komplement *A*. Jinými slovy, DS teorie nám umožní reprezentovat jen naši faktickou znalost (domnění o hypotéze), aniž bychom se museli zabývat souvislostí s komplementem (negací) hypotézy.

Každé nepřiřazené domnění zbývající po přiřazení domnění ke vhodným podmnožinám (viz následující příklad) je označeno jako *m*(Θ) a je přiřazeno k celé množině všech hypotéz 2^Θ. Nevztahuje se tedy k negaci podmnožin (jako by tomu bylo v případě teorie pravděpodobnosti).

Příklad 2

Domníváme se, že závada na počítači je způsobena napájením nebo chybou na hlavní desce ({*zn, zd*}). Naše přesvědčení vyjádříme stupněm 0,6 (existují určité indicie této závady, které ohodnotíme stupněm 0,6). Neexistuje žádný důkaz pro podporu výběru mezi závadou na hlavní desce a závadou napájení. Jaká jsou základní přiřazení *bba*?

Odpověď: *m*({*zn, zd*}) = 0,6, a *m*(Θ) = *m*({*zn, zd, zp, zg*}) = = 1 – 0,6 = 0,4.

Příklad 3

Test vyvrátil, že by porucha počítače byla způsobena závadou na napájení, a to stupněm 0,7. Jaká jsou bba?

Odpověď: Důkaz proti závadě na napájení se považuje za důkaz pro negaci (*zn*). Tedy m({*zd, zp, zg*}) = 0,7 a *m*(Θ) = 0,3.

2.3 Domněnková funkce

Domněnková funkce *Bel (A)* je funkce, která přiřazuje hodnotu z <0, 1> každé neprázdné podmnožině *A* z 2^Θ takto:

$$Bel(A) = \sum_{B \subseteq A} m(B) \qquad (4)$$

Domněnková funkce platnosti hypotézy *A* je tedy definována jako součet základních přiřazení všech podmnožin množiny *A*. Na rozdíl od základního přiřazení vyjadřuje *Bel(A)* míru domnění, kterou máme k *A* z důvodu podpory podmnožin $B \subseteq A$.

Když se podíváme na definici, vidíme, že existuje shoda mezi mírou domnění a základním přiřazením. Jestliže *A* je základní (atomickou) hypotézou, pak *Bel(A)* = *m(A)*. Pro funkci *Bel(A)* platí:

$$\begin{aligned} & Bel(A) + Bel(\neg A) \leq 1 \\ & Bel(A) < Bel(B)\ pro\ A \subseteq B \\ & \mathrm{Bel}(\Theta) = 1 \\ & Bel(\emptyset) = 0 \end{aligned} \qquad (5)$$

Příklad 4

Jaká je velikost *Bel*({*zn, zd, zg*})?

Odpověď: *Bel*({*zn, zd, zg*}) = *m*({*zn, zd, zg*}) + *m*({*zn, zd*}) + + *m*({*zn, zg*}) + *m*({*zd, zg*}) + *m*({*zn*}) + *m*({*zd*}) + *m*({*zg*}).

Příklad 5

m({*zn*}) = 0,6. Jaká je velikost *Bel*({*zn*}) ?

Odpověď: Pro množinu s jediným prvkem platí, že *Bel(A)* = *m(A)*, tedy *Bel*({*zn*}) = 0,6.

Bel(A) odpovídá celkové velikosti domnění přiřazené podmnožině *A*. Znalost *Bel*(¬*A*) (domnění ke komplementu *A*) je také užitečná informace. Vedle funkce *Bel(A)* existují jiné funkce, které vyjadřují stejnou informaci, ale mají jinou interpretaci, například míra věrohodnosti (measure of plausibility, upper probability function) [2, 3, 4] :

***2.4 Míra věrohodnosti** (measure of plausibility) :*

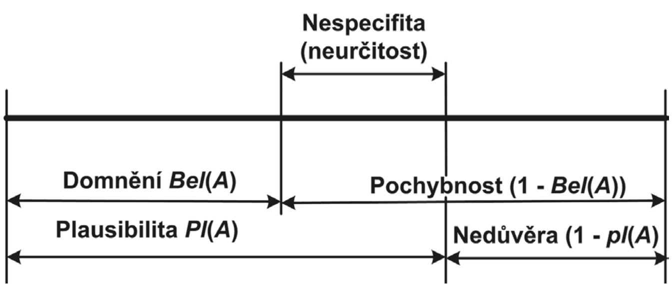
Míra věrohodnosti je definována vztahem:

$$Pl(A) = 1 - Bel(\neg A) \qquad (6)$$

Hodnota *Pl(A)* vyjadřuje, nakolik bychom měli věřit hypotéze *A*, jestliže i všechna dosud neznámá fakta by *A* podporovala (viz obr. 2).

Míra věrohodnosti vyjadřuje tedy horní interval naší důvěry k hypotéze *A*. Rozdíl *Pl(A)* – *Bel(A)* se nazývá nejistota o hypotéze *A* nebo nespecifičnost (neurčitost). Velikost intervalu <*Bel(A)*, 1- *Bel*(¬*A*)> reprezentuje sumu domnění přiřazenou elementům, jejichž průnik s *A* je nenulový, ale které nejsou podmnožinou *A*.

Obr. 2 – Intervaly domnění



2.4 Kombinace funkcí domnění

DS teorie umožňuje kombinovat základní přiřazení pomocí operátoru slučování (kombinace) [2, 3, 4]. Fungování tohoto operátoru je založeno na myšlence, že součin $m_1(A)$. $m_2(B)$ podporuje (ve smyslu domnění) průnik $A \cap B$.

Dokončení na další straně

Operátor slučování (také ortogonální suma) označovaný ⊕ funguje následujícím způsobem: předpokládejme, m_1 a m_2 jsou domněnkové funkce indukované příslušnými datovými zdroji. Pak lze tyto obě domněnkové funkce m_1 a m_2 na 2^Θ sloučit (kombinovat). Výsledná domněnková funkce $m_{12}(C) = m_1 \oplus m_2$ je definována takto:

$$m_1 \oplus m_2(C) = m_{12}(C) = \frac{\sum_{A,B:A\cap B=C} m_1(A)\cdot m_2(B)}{K}, \tag{7}$$

kde

$$K = 1 - \sum_{A,B:A\cap B=\phi} m_1(A)\cdot m_2(B) = \sum_{A,B:A\cap B\neq\phi} m_1(A)\cdot m_2(B) \tag{8}$$

Toto pravidlo pro kombinaci domněnkových funkcí se označuje jako Dempsterovo pravidlo.

Faktor K je indikátorem toho, do jaké míry jsou m_1 a m_2 protikladné (konfliktní, odporující si). Jestliže $K = 0$, pak říkáme, že zdroje jsou úplně neslučitelné (zcela si odporují).

Kombinace více základních přiřazení je jednoduchá, protože ortogonální suma ⊕ je komutativní a asociativní. Pro použití ortogonální sumy je nutno, aby domněnkové funkce byly nezávislé a spolehlivé. Pokud nejsou, je nutno použít některou z modifikací rovnice (4) uvedenou například v [2, 5].

Aplikaci domněnkových funkcí a jejich kombinace si ukážeme v druhém dílu příspěvku.

3. Literatura

[1] SHAFER, G. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press, 1976, 292 s.

[2] DANIEL, M. Several Notes on Belief Combination. In *Proceedings of the Workshop on the Theory of Belief Functions (Belief 2010)* [online]. Brest : ENSIETA, 1.−2. April 2010 [cit. 2010-10-27]. Dostupné z WWW: <http://www.ensieta.fr/belief2010/>.

[3] DEMPSTER, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics*. 1967, 38, 5, s. 325-339.

[4] MAŘÍK, V., et al. *Umělá inteligence 2.* Praha: Academia, 1997. 373 s. ISBN ISBN 80-200-0504-8, Kapitola 2.3 Neurčitost v expertních systémech.

[5] YAGER, R. R. Hierarchical aggregation functions generated from Belief structures. *IEEE Transactions on Fuzzy Systems*. 2000, 8, s. 481–490.

Abstract

Summary: *We will demonstrate in this three-part series the application of Dempster-Shafer theory to problems from the field of safety and reliability engineering. This theory is mainly used in expert systems, but it has application also in other areas, especially for methods involving expert estimation. In the first part, we describe the basics of the Dempster-Shafer theory and present some simple illustrative examples. In the second part, we present examples in which we use combination of information from various resources. It is the simple example from safety and reliability engineering. In the third part, we show an application of the Dempster-Shafer theory to the event tree analysis. The aim of the mini-series is to show that Dempster-Shafer theory copes with uncertainty arising from such lack of available information or ignorance compared to the probability theory and that it can be used with advantage for modeling in safety and reliability engineering.*

Keywords: safety and reliability engineering, belief function, Event Tress Analysis, uncertainty modeling