



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

ÚSTAV AUTOMATIZACE A INFORMATIKY

INSTITUTE OF AUTOMATION AND COMPUTER SCIENCE

**VYUŽITÍ OVLÁDÁNÍ DOMÁCÍCH SPOTŘEBIČŮ
HLASEM**

USING OF VOICE FOR HOME EQUIPMENT CONTROL

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Lukáš Bubla

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. Branislav Lacko, CSc.

BRNO 2019

ZADÁNÍ VŠKP 1

(tento list nahradíte oficiálním zadáním práce)

ABSTRAKT

Bakalářská práce je zaměřena na domácí automatizaci, konkrétně na ovládání hlasovými pokyny. Popisuje způsoby komunikace a ovládání automatizovaných soustav, a především problematiku počítačového rozpoznávání řeči pomocí různých metod. Zobrazuje konkrétní výrobek pro hlasové ovládání, vyvinutý českou společností Elko ep, s.r.o., dále pak v krátkosti nahlíží do problematiky ovládání gesty. V závěru práce porovnává výrobek od společnosti Elko ep, s.r.o. se zahraničními výrobky jiných společností a diskutuje výhody a nevýhody ovládání hlasem či gesty.

ABSTRACT

The bachelor thesis is focused on home automation, specifically on voice control. It describes methods of communication and control of automated systems and especially problematics of computer voice recognition using different methods. The thesis displays specific product for voice controlling developed by Czech company Elko ep, s.r.o. and then briefly looks at the issue of gesture control. At the end compares product from Elko ep with other products from foreign companies and discuss advantages and disadvantages of voice or gestures control.

KLÍČOVÁ SLOVA

Automatizované soustavy, rozpoznání řeči, komunikační protokoly, API, domácí automatizace, hlasové ovládání

KEYWORDS

Automated systems, speech recognition, communication protocols, API, application programming interface, home automation, voice control

BIBLIOGRAFICKÁ CITACE

BUBLA, L. *Využití ovládní domácích spotřebičů hlasem*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2019. 29 s. Vedoucí bakalářské práce doc. Ing. Branislav Lacko, CSc..

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu mé bakalářské práce doc. Ing. Branislavu Lackovi, CSc. za jeho ochotu, rady a věcné poznámky při konzultacích a vypracovávání práce.

ČESTNÉ PROHLÁŠENÍ

Prohlašuji, že tato práce je mým původním dílem, zpracoval jsem ji samostatně pod vedením doc. Ing. Branislava Lacka, CSc. a s použitím literatury uvedené v seznamu literatury.

V Brně dne 25. 4. 2019

.....

Lukáš Bubla

OBSAH

ÚVOD	9
1 CÍLE	10
2 OVLÁDÁNÍ AUTOMATIZOVANÝCH SOUSTAV	11
2.1 TECHNOLOGIE OVLÁDÁNÍ DOMÁCÍCH SPOTŘEBIČŮ A DOMÁCNOSTI OBECNĚ.....	12
2.1.1 X10	12
2.1.2 KNX.....	13
2.1.3 Zigbee	14
2.1.4 Z-Wave	14
2.1.5 Bluetooth Low Energy (BLE)	14
3 METODY POČÍTAČOVÉHO ROZPOZNÁNÍ ŘEČI	15
3.1 Simple pattern matching	15
3.2 Akustické modelování	16
3.2.1 Skrytý Markovův model (HMM)	16
3.3 Jazykové modelování	18
4 AUTOMATICKÝ PROSTŘEDEK S MOŽNOSTÍ HLASOVÉHO OVLÁDÁNÍ OD SPOLEČNOSTI ELKO EP	19
4.1 eLAN-RF-Wi-003 – smart krabička	19
5 VYUŽITÍ OVLÁDÁNÍ POMOCÍ GEST	22
6 ZHODNOCENÍ A DISKUZE	23
ZÁVĚR	24
SEZNAM POUŽITÉ LITERATURY	25
SEZNAM OBRÁZKŮ	27
SEZNAM TABULEK	28
SEZNAM ZKRATEK	29

ÚVOD

Obecně se pod termínem automatizace skrývá používání technologií, které umožňují vypustit z procesu aktivní lidské řízení. Hlavní význam spočívá ve využití při průmyslových operacích, kde se opakují stále stejné pracovní úkony, jako například obráběcí centra, výrobní a montážní linky, kotle, pece a mnoho dalších. Například automaticky řízené stroje vykonávají práci opakovaně, přesně dle zvoleného programu, což šetří čas i náklady, a právě to je účel automatizace.

Jelikož došlo během několika málo let k obrovskému rozvoji v oblasti automatizace, začali se o tyto prostředky čím dál více zajímat i samotné domácnosti, a to i díky tomu, že se již nejedná jen o drahou záležitost průmyslu. Tento zájem je pochopitelný, neboť automatizovaný smart spotřebič dokáže v domácnosti ušetřit množství energie čili i peněz. S automatizací se už teď setkává člověk prakticky denně. Typickým příkladem může být systém ABS, který je součástí každého automobilu nebo elektronický termostat řídící vytápění či klimatizaci místnosti za pomoci teplotního čidla.

Problém automatizovaných soustav je v ovládní a řízení. Obecně nejjednodušší je řízení pomocí jedniček a nul (symbolicky zapsané stavy *zapnuto / vypnuto*). Dále je možné ovládat spojitě systémy pomocí otočných regulátorů, jaké můžeme vidět např. u tzv. stmívačů osvětlení, anebo pomocí vícepolohových prepínačů. Ovšem některé současné automatizované soustavy jsou mnohonásobně složitější, a proto musí být ovládané větším množstvím regulátorů, prepínačů, otočných spínačů, kde si jako příklad lze představit kokpit letadla.

Snaha usnadnit ovládní těchto složitých soustav, ale i usnadnění každodenních činností přivádí člověka na myšlenku využití mluveného slova. Ovládní systémů deseti prsty se nahradí hlasem, který je přirozeným prostředkem komunikace člověka s člověkem a může tedy být i prostředkem komunikace člověka a stroje. Tato myšlenka však mohla vzniknout jen díky tomu, že dříve bylo vyřešeno, jak zpracovat lidský hlas, což je rozebráno v druhé části práce.

1 CÍLE

Cíle závěrečné práce jsem převzal ze zadání mé bakalářské práce:

1. Popsat problematiku řízení s využitím lidského hlasu.
2. Uvést technologie ovládání domácnosti a přehled možností počítačového rozpoznání řeči.
3. Vybrat automatický prostředek, který lze aplikovat pro tyto účely v domácnosti a charakterizovat jeho funkce a používání.

Jako automatizovaný prostředek, na kterém jsem se rozhodl demonstrovat aplikaci ovládání hlasem jsem si vybral produkt od společnosti ELKO EP s.r.o, konkrétně *eLAN-RF-003*. V závěru jsem pro komplexní pohled na ovládání strojů a zařízení začlenil kapitolu o řízení pomocí gest.

2 OVLÁDÁNÍ AUTOMATIZOVANÝCH SOUSTAV

V souvislosti s ovládním automatizovaných soustav, nelze opominout pojem ASCII (*American Standard Code for Information Interchange*). V mnoha případech se totiž hlasově vydávané pokyny převádějí počítačově do alfanumerického textu, který je zakódován ve standardu ASCII.

ASCII je kódová tabulka čítající 128 znaků využívaných v informatice. Z její znakové sady vychází většina současných standardů pro kódování textu. Pokud je alfanumerický dokument zaznamenaný v kódu ASCII, plyne z toho možnost dokument jednoduše tisknout nebo upravovat, jelikož neobsahuje žádné speciální znaky.

ASCII TABLE

Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char
0	0	[NULL]	32	20	[SPACE]	64	40	@	96	60	`
1	1	[START OF HEADING]	33	21	!	65	41	A	97	61	a
2	2	[START OF TEXT]	34	22	"	66	42	B	98	62	b
3	3	[END OF TEXT]	35	23	#	67	43	C	99	63	c
4	4	[END OF TRANSMISSION]	36	24	\$	68	44	D	100	64	d
5	5	[ENQUIRY]	37	25	%	69	45	E	101	65	e
6	6	[ACKNOWLEDGE]	38	26	&	70	46	F	102	66	f
7	7	[BELL]	39	27	'	71	47	G	103	67	g
8	8	[BACKSPACE]	40	28	(72	48	H	104	68	h
9	9	[HORIZONTAL TAB]	41	29)	73	49	I	105	69	i
10	A	[LINE FEED]	42	2A	*	74	4A	J	106	6A	j
11	B	[VERTICAL TAB]	43	2B	+	75	4B	K	107	6B	k
12	C	[FORM FEED]	44	2C	,	76	4C	L	108	6C	l
13	D	[CARRIAGE RETURN]	45	2D	-	77	4D	M	109	6D	m
14	E	[SHIFT OUT]	46	2E	.	78	4E	N	110	6E	n
15	F	[SHIFT IN]	47	2F	/	79	4F	O	111	6F	o
16	10	[DATA LINK ESCAPE]	48	30	0	80	50	P	112	70	p
17	11	[DEVICE CONTROL 1]	49	31	1	81	51	Q	113	71	q
18	12	[DEVICE CONTROL 2]	50	32	2	82	52	R	114	72	r
19	13	[DEVICE CONTROL 3]	51	33	3	83	53	S	115	73	s
20	14	[DEVICE CONTROL 4]	52	34	4	84	54	T	116	74	t
21	15	[NEGATIVE ACKNOWLEDGE]	53	35	5	85	55	U	117	75	u
22	16	[SYNCHRONOUS IDLE]	54	36	6	86	56	V	118	76	v
23	17	[ENG OF TRANS. BLOCK]	55	37	7	87	57	W	119	77	w
24	18	[CANCEL]	56	38	8	88	58	X	120	78	x
25	19	[END OF MEDIUM]	57	39	9	89	59	Y	121	79	y
26	1A	[SUBSTITUTE]	58	3A	:	90	5A	Z	122	7A	z
27	1B	[ESCAPE]	59	3B	;	91	5B	[123	7B	{
28	1C	[FILE SEPARATOR]	60	3C	<	92	5C	\	124	7C	
29	1D	[GROUP SEPARATOR]	61	3D	=	93	5D]	125	7D	}
30	1E	[RECORD SEPARATOR]	62	3E	>	94	5E	^	126	7E	~
31	1F	[UNIT SEPARATOR]	63	3F	?	95	5F	_	127	7F	[DEL]

Obr. 1: Ukázka ASCII tabulky se 128 znaky [24]

ISO/IEC 8859-1:1998, Information technology — 8-bit single-byte coded graphic character sets — Part 1: Latin alphabet No. 1, je součástí skupiny ASCII kódů podle ISO/IEC 8859 ke standardnímu kódování, poprvé publikována roku 1987. [10] ISO 8859-1 kóduje to, co je označováno jako „Latinské abeceda č.1,“ a skládá se ze 191 znaků latinského písma. Toto schéma kódování je používané v celé Americe, západní Evropě, Oceánii, a i ve větší části Afriky. Jedná se o základ pro nejpůlárnější 8-bitové znakové sady a první blok znaků v Unicode.

Kódová stránka Windows-1252, která byla používaná v dřívějších verzích operačního systému Microsoft Windows, se shoduje s ISO-8859-1 ve všech kódech s výjimkou kódů 128 až 159 (hex 80 až 9F), kde se málo užívané C1 prvky nahrazují doplňkovými znaky včetně všech znaků, které chybí v ISO-8859-15. Text Windows-1252

bývá běžně chybně označován za ISO-8859-1. Výsledkem chybného rozpoznání bylo, že všechny citáty a apostrofy, vytvořené „inteligentními uvozovkami“ v softwaru pro zpracování textu, byly na operačních systémech, jiných než Windows, nahrazeny otazníky nebo čtverci, které způsobovaly obtížné čtení textu. Většina současných webových prohlížečů a e-mailových klientů používá ISO-8859-1 jako Windows-1252 tak, aby vyhovovala tomuto chybnému označování. Toto je nyní standardní chování specifikace HTML5, která požaduje, aby dokumenty zadané jako ISO-8859-1 byly skutečně analyzovány pomocí kódování Windows-1252.

K září 2018 prohlašuje 3,8 % ze všech webových stránek, že používají ISO 8859-1. To však nezahrnuje stránky, které používají Windows-1252 a UTF-8, oba jsou prohlížeči běžně rozpoznávány. [23]

Za nástupce ASCII tabulky lze považovat rozhraní API (*Application Programming Interface*) – česky „aplikační programovací rozhraní“. API obsahuje množství protokolů, funkcí, částí kódů a poskytuje programátorovi digitální stavební kameny pro jednodušší a účinnější vývoj softwarových aplikací. Umožňuje komunikaci a výměnu informací mezi jednotlivými aplikacemi. [20]

Člověk se s aplikací rozhraní API setkává denně, vnímá to jako samozřejmost a většinou se ani nepozastaví nad způsobem jakým bylo dosaženo daného úkonu. Například aplikace pro mobilní telefony, která umožňuje porovnávání restaurací v blízkém okolí, využívá několik API najednou, jedno poskytuje informace o denním menu a otevírací době, další získává data o poloze a další umožňuje zajistit si rezervaci na danou hodinu. Další využití je při přihlašování do aplikací nebo webových stránek například pomocí Facebook účtu. Facebook poskytuje své API, které může využít například internetový obchod a potenciální zákazník se zaregistruje a přihlásí jednoduše jedním kliknutím.

2.1 TECHNOLOGIE OVLÁDÁNÍ DOMÁCÍCH SPOTŘEBIČŮ A DOMÁCNOSTI OBECNĚ

Technologií pro ovládání smart domácích spotřebičů existuje mnoho. Mezi nejpopulárnější patří KNX nebo X10 anebo relativně nové alternativy jako Zigbee, Z-Wave, WiFi, INSTEON nebo Bluetooth Low Energy.

2.1.1 X10

Jedná se o komunikační protokol, který ovládá domácí spotřebiče prostřednictvím v domě již existující elektrické sítě. Protokol byl vytvořen roku 1978 a představuje první prostředek domácí automatizace. I přes jednoduchost a omezenější funkce je stále hojně používán i v dnešní době. X10 využívá nízkou přenosovou rychlost v řádech bitů, která však postačuje pro přenos řídicích příkazů a je limitován na 256 zařízení. [4]

Pro ovládání využívá X10 šest základních funkcí –

- zapnout
- vypnout
- snížit
- zesílit
- vše zapnuto
- vše vypnuto

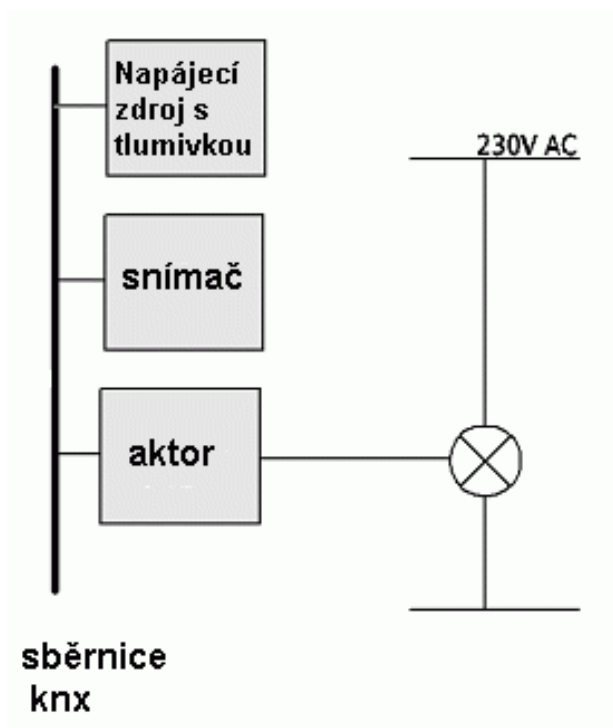
Příkladem ovládání může být světlo, které pomocí protokolu X10 lze rozsvítit/zhasnout, ale lze ovládat i intenzitu daného osvětlení. [11]

2.1.2 KNX

KNX je standardní síťový protokol (nebo také sběrnice), který vznikl spojením tří evropských asociací (European Home Systems Protocol, BatiBUS, European Installation Bus). KNX využívá pět komunikačních kanálů a to

- kroucenou dvojlinku
- elektrickou síť (podobně jako X10)
- radiové vlny
- infračervené vlny
- ethernet (známý také jako EIBnet/IP nebo KNXnet/IP). [4]

Využití KNX se nejvíce vyskytuje při ovládání topení, klimatizace nebo monitorování poplachu a bezpečnostních zařízení, případně měření energetické náročnosti některých spotřebičů.



Obr. 2: Nejmenší instalace systému KNX [22]

2.1.3 Zigbee

Zigbee představuje bezdrátový jazyk, sloužící ke komunikaci zařízení mezi sebou. Skládá se z univerzálních a otevřených protokolů. Motivací ke vzniku tohoto standardu a založení Zigbee aliance, byla existence průmyslových operací, u kterých není vhodné používat Bluetooth. Zigbee aliance čítá velké množství firem, především z oblasti automatizace domácnosti, které se podílí na dalším rozvoji a mezi nejznámější patří Huawei, Comcast, Philips, Amazon Lab126, Honeywell, Intel aj. [12] Zigbee zařízení jsou malá, s nízkou spotřebou energie a nízkými náklady na infrastrukturu a mohou být rozdělena do tří tříd:

1. koordinátor
 - komplexní zařízení ovládající síť, v každé síti pouze jeden
2. router
 - propojuje samostatná zařízení v síti
3. koncové zařízení
 - nemůže přenášet informace určené pro další zařízení. [4]

2.1.4 Z-Wave

Jedná se o bezdrátový protokol, který na rozdíl od Zigbee není otevřený a je k dispozici pouze zákazníkům společností Zensys a Sigma Designs. Protokol je podporován aliancí Z-Wave, do které spadá více než sto společností z oblasti domácí automatizace [4]. Z-Wave je založen na nepravidelné síťové topologii (*mesh network topology*), kdy každé zařízení instalované do sítě se stává opakovačem signálu, proto čím více zařízení je v síti instalováno, tím vyššího pokrytí lze dosáhnout. V každé síti Z-Wave lze mít instalováno až 232 zařízení. [13]

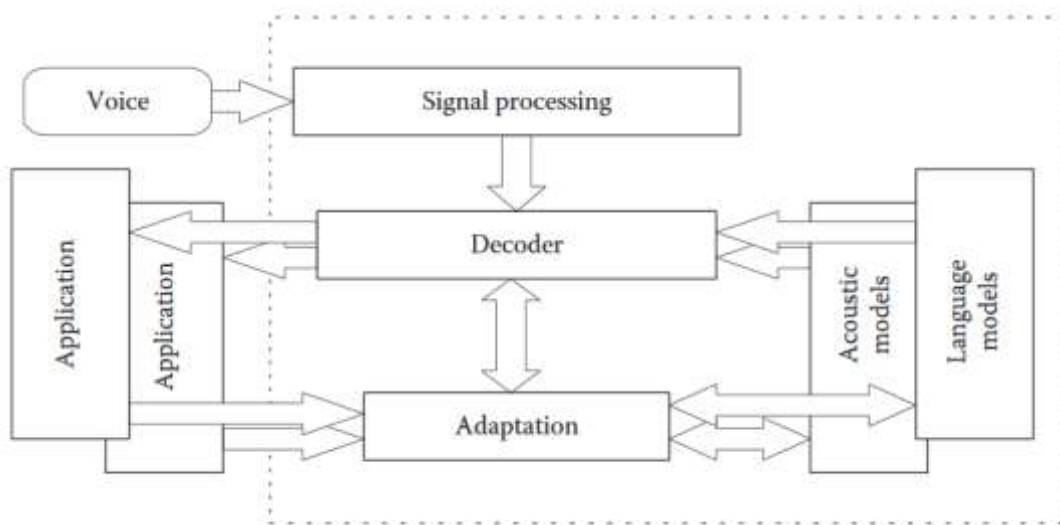
2.1.5 Bluetooth Low Energy (BLE)

Jedná se o variantu bluetooth, která je sice omezena na maximální rychlost 1Mb/s, avšak spotřebovává o polovinu méně energie. Není kompatibilní s klasickým Bluetooth a umožňuje bezdrátové připojení na principu krátkých pulzů. Technologii BLE můžeme najít dnes například v přenosných zařízeních pro fitness, sport nebo zdravotní péči, v zařízeních, kde není potřeba přenášet nepřetržitě velké množství dat, ale stačí data posílat v krátkých dávkách. [14]

3 METODY POČÍTAČOVÉHO ROZPOZNÁNÍ ŘEČI

Samotné rozpoznávání řeči, je složitá záležitost a přístup zpracování mluveného slova se odvíjí od požadavků na danou aplikaci. V oblasti rozpoznání řeči existuje mnoho vlivů a skutečností, které systémům na rozpoznání řeči komplikují jejich práci. Mezi takové patří například rychlost řeči, akcenty, kvalita reproduktorů či mikrofону, rušivé elementy, pro systém neznámá slova nebo také gramatika aj.

Typický systém na rozpoznávání řeči je proto složitější a skládá z několika komponent, jak je vyobrazeno na obr. 4.1. Aplikační rozhraní, jehož součástí je dekodér, přijímá po zpracování hlasového signálu data, která mohou být použita k přizpůsobení dalších částí systému [5]. Ke zpracování lidského hlasu jsou zde využity dva modely. Prvním je model akustický, který zpracovává znalosti akustiky, dialektů, ale i rozdíly mezi pohlavími. Druhým modelem je model jazykový, jenž k rozpoznání řeči využívá systém, který odhaduje, jaké slovo týkající se příkazu, který chtěl uživatel pravděpodobně provést, bude následovat. Zmíněné dva modely patří k základním metodám počítačového rozpoznávání lidské řeči.



Obr. 3: Základní schéma systému rozpoznávání řeči [5] – str. 340

3.1 Simple pattern matching

Simple pattern matching je metoda, která využívá k rozpoznání pokynů předem definované vzory. Nejčastěji se jedná o jednoslovné pokyny, kdy počítač zaznamená celé slovo a následně ho porovnává s jednotlivými zvukovými nahrávkami frází ze své databáze. Uvést lze příklad počítačového automatu při telefonování do call centra, kde pro přepojení na správnou telefonní linku stačí jednoslovné odpovědi jako „ano“, „ne“

nebo možnosti „jedna“, „dva“ atd. Automat přiřazuje pokyn k nahraným předlohám, porovná a vybere tu s největší shodou.

Nevýhodou tohoto systému je jeho omezené použití. Zaměřuje se především na slova, která nezní podobně, aby nedocházelo k záměně dvou pokynů. [8]

3.2 Akustické modelování

Existuje celá řada známých faktorů, které ovlivňují přesnost systémů rozpoznávání řeči. Nejčastějšími podmínkami, které mají vliv na přesnost rozpoznávání, jsou změny kontextu, změny mluvčího, fonetika nebo vlivy prostředí (akustika). S jistotou lze prohlásit, že akustické modelování je ústřední částí každého systému rozpoznávání řeči. [5]

Akustické modelování řeči je proces přijímání vlnových signálů řeči a pro jejich zpracování se využívá nejčastěji tzv. Skrytých Markovových modelů (HMM), které pracují na principu pravděpodobnosti a napomáhají tak například eliminovat nežádoucí vnější vlivy. Součástí akustického modelování je také modelování výslovnostní, které rozděluje slovo na jednotlivé sekvence zvuků a popisuje jejich použití při vytváření větších řečových jednotek (slova, fráze).

3.2.1 Skrytý Markovův model (HMM)

HMM je v dnešní době jeden z nejúspěšnějších a nejčastěji užívaných technik počítačového rozpoznávání řeči, avšak nejen řeči. Jedná se o nástroj rozpoznávání ručně psaného textu, pro modelování dat časových řad a využití nachází také v oblasti výpočetní molekulární biologie, dalších oblastech umělé inteligence a v poslední době také v oblasti tzv. počítačového vidění, které lze využít k ovládnutí zařízení pomocí gest. [9]

U HMM se zpracovávají 3 základní problémy: [18]

1. Vyhodnocení (Evaluation)

- Vyhodnocuje pravděpodobnost, že daný model generoval sledovanou sekvenci. Pomocí algoritmu Forward určuje pravděpodobnost každého stavu, který poskytuje nějaký výstupní stav.

V praxi si to lze představit například na modelování chování zákazníka, kdy se hledá nejpravděpodobnější profil zákazníka nebo například ve fyzice při vybírání nejpravděpodobnějšího termodynamického jevu. [19]

2. Dekódování (Decoding)

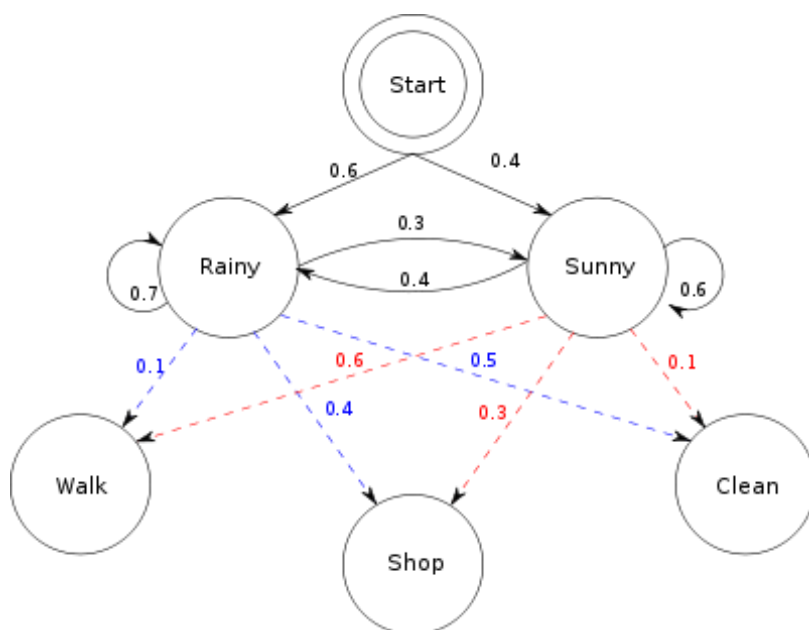
- HMM počítá nejpravděpodobnější posloupnost skrytých stavů pro daný model a sledovanou sekvenci a nalezne nejlepší cestu, za použití Viterbiho algoritmu.

Typické aplikace mohou být například rozpoznávání ručně psaných dokumentů, kde pozorovány jsou vzory písmen a skryté jsou písmena.

3. Trénink (Training)

- Využívá se Baum-Welchova algoritmu pro určení nejpravděpodobnějšího modelu, který generoval sledovanou sekvenci (tzv. bez dozoru) nebo pozorovanou a skrytou sekvenci (tzv. pod dohledem).

Názorný příklad toho, jak fungují HMM můžeme nalézt na anglické verzi webu Wikipedia [15]. Uvažujeme dva přátele (Oliver a Nina), kteří bydlí ve velké vzdálenosti od sebe a komunikují spolu pouze po telefonu. Oliver si chce povídat pouze o tom, co během dne dělal. Ovšem Ninu by zajímalo, jaké je u Olivera počasí, a proto se to pokouší odhadnout. Existují dvě možnosti – deštivo (rainy) a slunečno (sunny), ale Nina je nemůže určit sama z indicií, které vyvodí během rozhovoru. Oliver má pouze tři aktivity, které během týdne provozuje – chodí na procházku (walk), na nákup (shop) nebo uklízí svůj byt (clean) – a pravděpodobnost každé z nich je závislá na počasí. Jakmile Oliver řekne Nině, co přes den dělal, stává se ze systému HMM. Nina zná pravděpodobnosti všech okolností a činností – viz obr. 4.2, přičemž pravděpodobnost na počátku reprezentuje odhad Niny, jaké bylo počasí, když poprvé Oliver telefonoval.



Obr. 4: Pravděpodobnostní schéma daného příkladu. [15]

Pokud si tento příklad převedeme na potřeby akustického modelování, bude to vypadat následovně. Po vyslovení první slabiky, bude program uvažovat statisticky nejpravděpodobnější slabiky, které budou následovat. Pokud například vyslovíme slabiku *te-*, pak bude software pracovat s velkou pravděpodobností, že bude následovat slabika *-ta*, protože se jedná o často používané slovo *teta*, ale uvažuje i další varianty s různými pravděpodobnostmi. Když však bude následovat slabika *-plo-*, rozptýlí možnosti se zužuje na různé varianty s kořenem slova *teplo*.

Svým způsobem na tomto principu fungují i našeptávače vyhledávacích polí Google nebo Seznam, které už po napsání prvního písmena zobrazí seznam možných hledaných slov či frází.

3.3 Jazykové modelování

Akustické modelování, popsané v předchozí kapitole, je jednou ze dvou základních složek, které jsou potřeba k rozpoznávání řeči. Druhou složkou je modelování jazykové, které, stejně jako akustické, využívá statistickou pravděpodobnost. [5]

Doteď se uvažovalo víceméně spisovného jazyka, ale může nastat situace, kdy bude počítač muset rozpoznat jazyk, ovlivněný například nářečím. V takovém případě přichází na řadu jazykové modelování. V tomhle ohledu velmi pokročila společnost Google, která využívá k rozšiřování databáze slov přepis YouTube videí a dat z hlasového vyhledávání. Využití získaných dat k rozšíření slovníků vedlo, dle zprávy společnosti Google, ke snížení chyb v určování slov až o desítky procent¹.

Samotné jazykové modelování funguje podobně jako modelování akustické, pouze s tím rozdílem, že nevyužívá slabiky, ale rovnou celá slova. V komunikaci existují určitá ustálená spojení nebo jsou některá slova používána v kombinaci s jinými častěji než ostatní. Proto se počítač, při rozpoznávání řeči, pokouší přiřadit zvuk k sekvenci slov. Jazykový model poskytuje jakýsi kontext pro rozlišení slov, která znějící podobně. Pokud si počítač není jistý některým slovem, využije ta slova, která se ve větě nachází před a za hledaným výrazem. Pomocí sousedních slov se statisticky určí nejpravděpodobnější varianta. Jako příklad, kdy se využije jazykové modelování, lze uvést fráze „*recognize speech*“ a „*wreck a nice beach*“, které se v americké angličtině vyslovují téměř stejně, ale jejich význam si nemůže být vzdálenější. [16] Nejlepších výsledků se však dosahuje, pokud se využijí poznatky jazykového modelu v kombinaci s modely akustickými, resp. výslovnostními.

¹ Viz abstrakt <http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/40491.pdf>

4 AUTOMATICKÝ PROSTŘEDEK S MOŽNOSTÍ HLASOVÉHO OVLÁDÁNÍ OD SPOLEČNOSTI ELKO EP

V dnešní době už lze konstatovat, že hlasoví asistenti v mobilních zařízeních, počítačích a jiných zařízeních fungují spolehlivě. Stejně tak lze s jistotou říci, že lidé čím dál častěji využívají prostředků pro zautomatizování domácnosti. Z těchto důvodů se logicky zvedá poptávka po možnostech ovládání domácnosti pomocí hlasu. Na trhu se dá dnes nalézt už několik produktů, které lze k ovládání domácnosti pomocí hlasu využít, například od společnosti Amazon nebo Apple. Pokud však člověk hledá prostředek, pomocí kterého by mohl hlasem ovládat svou domácnost, nemusí chodit ani za hranice. Společnost ELKO ep, s.r.o. je český výrobce, který patří k předním světovým výrobcům elektronických zařízení pro bytovou a průmyslovou automatizaci již 26 let.

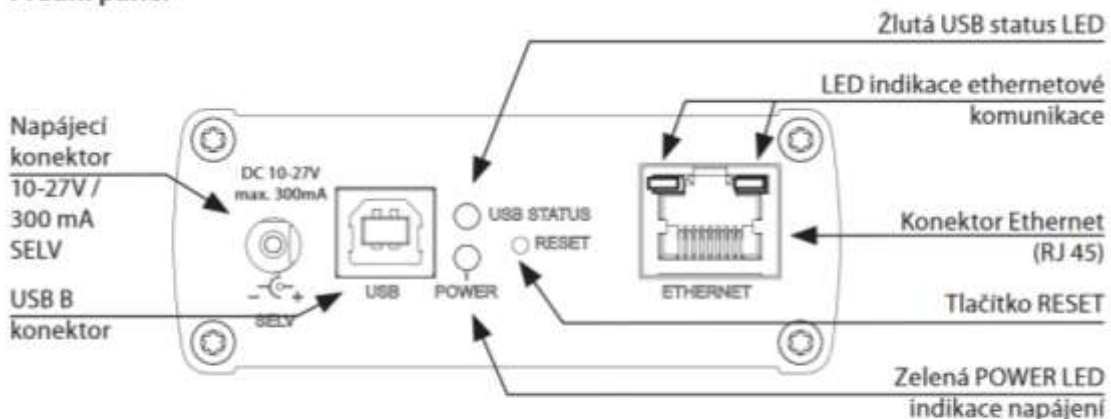
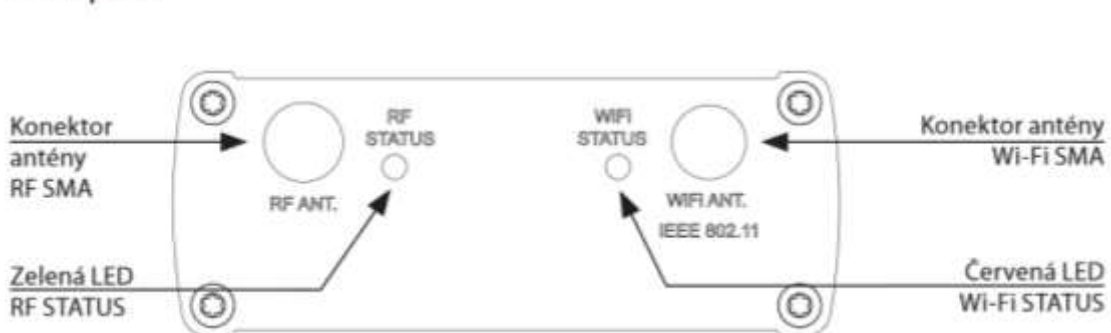
4.1 eLAN-RF-Wi-003 – smart krabička

Smart RF krabička umožňuje člověku ovládat elektroinstalaci pomocí mobilní aplikace a hlasových asistentů v mobilních telefonech, ale také pomocí smart TV nebo smart hodinek.

Technické parametry	eLAN-RF-Wi-003
Rozhraní RF Control	
Komunikační protokol:	RF Touch Compatible
Vysílací frekvence:	866 MHz, 868 MHz, 916 MHz
Způsob přenosu signálu:	obousměrně adresovaná zpráva
Výstup pro anténu RF:	SMA konektor
Anténa RF:	1 dB
Indikace RF komunikace:	1 x zelená RF status LED
Dosah ve volném prostranství:	do 100 m
Rozhraní Ethernet	
Indikace provozního stavu ETH:	zelená LED
Indikace komunikace ETH:	žlutá LED
Komunikační rozhraní:	100 Mbps
Přednastavená IP adresa:	192.168.1.1
Rozhraní Wi-Fi	
Standard:	IEE 802.11 b/g/n / 2.4 GHz
Zabezpečení Wi-Fi:	WEP, WPA-PSK, WPA2-PSK
Výstup pro anténu Wi-Fi:	R-SMA
Anténa Wi-Fi:	1 dB

Indikace Wi-Fi komunikace:	1x červená Wi-Fi status LED
Dosah:	do 200 m
Napájecí napětí/proud:	10-27 V DC / 300 mA SELV
Napájení:	adaptér s konektorem jack \varnothing 2.1 mm nebo konektor USB-B
Indikace napájecího napětí:	zelená LED POWER
Tlačítko RESET	nastavení do výchozích hodnot
Napájecí zdroj:	230 VAC / 12 V DC / 300 mA SELV
Další údaje	
Pracovní teplota:	-20 .. +50 °C
Skladovací teplota:	-25 .. +70 °C
Stupeň krytí:	IP20
Stupeň znečištění:	2
Pracovní poloha:	libovolná
Rozměry:	90 x 52 x 65 mm
Hmotnost:	145 g

Tab. 1: Tabulka technických parametrů. [17] – str. 20

Přední panel**Zadní panel**

Obr. 5: Popis přístroje eLAN-RF-Wi-003 [17] – str.20

RF krabička může přijímat povely až ze 40 prvků. Díky tomu, že vytváří vlastní Wi-Fi síť, je člověk schopen ovládat jednotlivé prvky v domácnosti bez připojení k internetu, stačí se pouze připojit ke krabičkou vytvořené síti.

Pomocí mobilní aplikace je možno regulovat podlahové vytápění, osvětlení, spínat spotřebiče, nastavit si jednotlivé scény nebo například ovládat spotřebiče odjinud než z domu aj. V případě Smart TV nebo smart hodinek jsou možnosti ovládání o něco omezenější.

5 VYUŽITÍ OVLÁDÁNÍ POMOCÍ GEST

Ovládání smart zařízení jak v domácnosti, tak kdekoli jinde, by mělo být dostupné všem, i lidem s handicapem. Němí lidé nemohou ovládat domácnost pomocí hlasových pokynů, ale mohli by využít ovládání gesty.

Ovládání zařízení pomocí pohybů lidského těla lze uskutečnit umístěním čidel na několik míst na těle a pomocí snímače určovat jejich polohu v prostoru. Tato metoda je však pro běžné použití drahá a nepraktická. Častější a praktičtější je způsob, kde je snímání prováděno pomocí kamery, která bude obrazově snímat lidské pokyny.

Použit lze například kameru 2D, což je jednoduchý přístup, kdy je kamera schopna počítat prsty, rozlišit otevřenou / zavřenou dlaň a další gesta, která jsou uložena v sadě šablony. Kamera s 2D technologií však rozliší pouze statický obraz, nedokáže pracovat s pohybem těla, navíc spolehlivost je silně ovlivněna vnějšími faktory osvětlení nebo barvou pleti.

V dnešní době jsou však stále dostupnější i 3D kamery, což demonstruje i fakt, že existují přídavné kinecty k herním konzolám. Tyto kinecty umožňují trojrozměrné snímání pohybu, což poskytuje možnost hrát hry bez klávesnice nebo joysticku pouze za pomoci pohybů těla. Oproti tomu máme software od Leap Motion, který umožňuje například malování nebo manipulaci s malými objekty ve virtuálním prostoru, což nalézají zajímavá uplatnění např. v CAD aplikacích. [22]

Další způsoby, jak lze rozpoznávat gesta mohou být například řešení založená na ultrazvuku, infračerveném záření anebo využití elektrického pole, které je výhodné především pro průmyslové aplikace.

6 ZHODNOCENÍ A DISKUZE

Český jazyk obecně je složitým jazykem, proto i rozpoznávání řeči je pro počítače mnohem náročnější než například angličtina. Tato náročnost je způsobena ohebností, čili skloňováním a časováním jazyka. Ohebnost jazyka způsobuje, že má čeština větší slovníky s více možnými variantami slov. Tyhle skutečnosti zapříčiňují to, že v současné době jsou u nás hlasově ovládané automatizované soustavy, ale i hlasoví asistenti například v mobilních telefonech, ovládané převážně angličtinou.

Jak již bylo zmíněno, smart RF krabička společnosti Elko ep využívá pro hlasové ovládání hlasové asistenty, kteří jsou zabudováni přímo v mobilních telefonech. Čili nelze brát RF krabičku jako hardware sloužící k rozpoznávání řeči, ale jedná se v podstatě jen o prostředníka spojujícího dva spotřebiče, které by spolu jinak daným způsobem nemuseli komunikovat.

Výše uvedené můžeme porovnat s jinými společnostmi, které se hlasovým ovládáním domácnosti zabývají. Společnosti Amazon (produkt Amazon Echo), Google (Google Home) nebo společnost Apple (Apple HomePod). Produkty těchto společností mají podobu reproduktoru a v sobě mají přímo zabudovaného hlasového asistenta, jsou připojené k wifi síti a pomocí mikrofonu přijímají pokyny a předávají je cílovému spotřebiči, který je připojen ke stejné síti nebo úkon sami realizují.

Hlasové ovládání však nelze využít vždy, a proto se využívá i ovládání gesty. Praktický příklad využití gest lze nalézt v dopravě. Výrobci automobilů se snaží ovládání vozů zjednodušit tak, aby se řidič co nejvíce věnoval pohledu na cestu. V tomto směru lze sice využít ovládání hlasem, ale to za některých okolností nemusí být ideální (problém s hlasivkami, rušení z rozhovoru dalších pasažérů aj.). Z tohoto důvodu např. společnost BMW vyvinula bezdotykový systém ovládání, který snižuje potřebu řidiče natahovat se k ovládacímu panelu auta. S výhledem plně autonomních vozidel se však gesta musí přesouvat i ven z automobilu, kdy je potřeba vzájemné komunikace vozidel a chodců, kde téměř nebude jiná možnost než využití určitých typů gest.

Z pohledu ovládání domácnosti tedy mají obě možnosti ovládání (hlas i gesta) své výhody i nedostatky. Ze strany hlasu se jedná například o problémy okolního hluku, problémy s velkými slovníky, ale naopak se dá použít, pokud má člověk plné ruce nebo i ve tmě, což ovládání gesty neumožňuje. Do budoucna lze předpokládat, že nejvýhodnější bude kombinace obou metod, při které se budou hlas a gesta vzájemně doplňovat.

ZÁVĚR

Cílem této práce bylo seznámit čtenáře s problematikou řízení pomocí hlasu a jejich principem a technologiemi. K uvedení do tématu však byly v první kapitole popsány obecně způsoby řízení automatizovaných soustav pomocí nejpoužívanějších protokolů a kódů, jako například ASCII.

Dále se práce zabývala přímo metodami, které se využívají k počítačovému rozpoznávání hlasu. Metody byly rozděleny na 3 základní způsoby modelování, a to simple patter matching, akustické a jazykové modelování, kde podrobněji bylo rozebráno akustické modelování pomocí Skrytých Markovových modelů.

Ve čtvrté kapitole byl rozebrán konkrétní výrobek eLAN-RF-Wi-003, který lze využít k ovládání automatizované domácnosti, a to nejen s využitím hlasu. Byly popsány jeho parametry a možnosti využití.

Na závěr bylo ve stručnosti zahrnuto ovládání pomocí gest, jeho využití a způsoby rozpoznávání a bylo porovnáno s ovládáním pomocí hlasových pokynů.

SEZNAM POUŽITÉ LITERATURY

- [1] BALÁTĚ, J.: *Automatické řízení*. BEN 2003 Praha.
- [2] ŠVARC, I., ŠEDA, M., VÍTEČKOVÁ, M.: *Automatické řízení*. CERM 2007 Brno.
- [3] MATERNA, P., PALA, K., ZLATUŠKA, J.: *Logická analýza přirozeného jazyka*. Academia 1989 Praha.
- [4] SÁNCHEZ, Antonio a Belén CARRO. *Digital services in the 21st century: a strategic and business perspective*. Hoboken, New Jersey: Wiley, 2017. The ComSoc guides to communications technologies. ISBN 978-1-119-31485-1.
- [5] EDITED BY NITIN INDURKHIA a FRED J. DAMERAU. *Handbook of natural language processing*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC, 2010. ISBN 9781420085938.
- [6] EDITED BY MARK JOHNSON, SANJEEV P. KHUDANPUR, Mari OSTENDORF a Roni ROSENFELD. *Mathematical foundations of speech and language processing*. S.l.: Springer, 2013. ISBN 9781461264842.
- [7] POTŮČEK, M. *Bezdrátový systém pro ovládání domácích spotřebičů hlasem*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 64 s. Vedoucí diplomové práce Ing. Ondřej Sajdl, Ph.D.
- [8] PORT, Jake. Hey Siri, how does voice recognition software understand me?. *Cosmos: The science of everything* [online]. 14. prosinec 2016 [cit. 2018-10-31]. Dostupné z: <https://cosmosmagazine.com/technology/hey-siri-how-does-voice-recognition-work>
- [9] GHAMRANI, Zoubin. An Introduction to Hidden Markov Models and Bayesian Networks. *International Journal of Pattern Recognition and Artificial Intelligence*. 2001, **15**(1), 25.
- [10] ISO/IEC 8859-1: *Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1*. 1987.
- [11] What is X10?. *Smarthome*, [online]. [cit. 2018-10-30]. Dostupné z: <https://www.smarthome.com/sc-what-is-x10-home-automation>
- [12] Global leaders creating innovative products and services for the Internet of Things, *Zigbee alliance*, [online]. [cit. 2018-10-30]. Dostupné z: <https://www.zigbee.org/zigbeealliance/our-members/>
- [13] How Z-Wave works. *Z-Wave: Safer. Smarter*. [online]. [cit. 2018-10-31]. Dostupné z: <https://www.z-wave.com/learn>
- [14] Co je to Bluetooth Low Energy?. *@365tipu* [online]. 6. září 2017 [cit. 2018-10-31]. Dostupné z: <https://365tipu.cz/2017/09/06/tip883-co-je-to-bluetooth-le-ble-bluetooth-low-energy-bluetooth-smart-bluetooth-5/>
- [15] Hidden Markov model: A concrete example. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-11-05]. Dostupné z: https://en.m.wikipedia.org/wiki/Hidden_Markov_model#A_concrete_example
- [16] Language model. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-11-06]. Dostupné z: https://en.wikipedia.org/wiki/Language_model
- [17] ELKO EP, s.r.o. *Bezdrátová elektroinstalace: technický katalog*. 1. Holešov, 2018. Dostupné z: https://www.elkoep.cz/media/files/download/item/files-188/11_iNELS_RF_technical_2018_CZ_print_version.pdf

- [18] Prerna Rai, Moirangthem Goldie Meitei, Ferdousi Khatun, Abhijit Choudhury and Udit Kr. Chakraborty. Article: Commentary on Application of Hidden Markov Model in Google Page Ranking. *IJCA Proceedings on International Conference on Computing and Communication ICC* 2016(1):35-41, September 2016. Dostupné z: <https://pdfs.semanticscholar.org/3eb3/40d9605a6e46d49302f0be945afc43da1b.pdf>
- [19] HEIMEL, Max. An Introduction to Hidden Markov Models. Berlin, 2010. Dostupné z: <http://isabel-drost.de/hadoop/slides/HMM.pdf>
- [20] What is an API?. MuleSoft [online]. [cit. 2019-03-13]. Dostupné z: <https://www.mulesoft.com/resources/api/what-is-an-api>
- [21] LEE, Frank. How Gesture Control Will Transform Our Devices. Medium [online]. 1.2.2018 [cit. 2019-04-25]. Dostupné z: <https://medium.com/iotforall/how-gesture-control-will-transform-our-devices-32d4527a6d25>
- [22] MICHALEC, Libor. Úvod do KNX. In: Automatizace.hw [online]. 10.7.2013 [cit. 2019-05-07]. Dostupné z: <https://automatizace.hw.cz/teorie-a-praxe/knx.html>
- [23] ISO/IEC 8859-1. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2019-05-14]. Dostupné z: https://cs.wikipedia.org/wiki/ISO/IEC_8859
- [24] ASCII. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2019-05-15]. Dostupné z: <https://simple.wikipedia.org/wiki/ASCII#/media/File:ASCII-Table-wide.svg>

SEZNAM OBRÁZKŮ

Obr. 1: Ukázka ASCII tabulky se 128 znaky [24]

Obr. 2: Nejmenší instalace systému KNX [22]

Obr. 3: Základní schéma systému rozpoznávání řeči [5] – str. 340

Obr. 4: Pravděpodobnostní schéma daného příkladu. [15]

Obr. 5: Popis přístroje eLAN-RF-Wi-003 [17] – str.20

SEZNAM TABULEK

Tab. 1: Tabulka technických parametrů. [17] – str. 20

SEZNAM ZKRATEK

2D	dvourozměrné, dvoudimenzionální
3D	trojrozměrné, trojdimenzionální
ABS	Anti-Lock Brake System
API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
BLE	Bluetooth Low Energy
CAD	Computer-Aided Drafting
ETH	Ethernet
HMM	Hidden Markov Models
IEC	International Electrotechnical Commission
IP	Internet Protocol
ISO	International Organization for Standardization
LED	Light-Emitting Diode
RF	Radio Frequency
SELV	Safety Extra-Low Voltage
s.r.o.	Společnost s ručením omezeným
WiFi	Wireless Fidelity