Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra informačních technologií

Studijní program: Aplikovaná informatika

Obor: Informační systémy a technologie

SOUVISLOST VÝVOJE HODNOTY KRYPTOMĚN S DYNAMIKOU UŽIVATELI TVOŘENÉHO OBSAHU NA INTERNETU

DIPLOMOVÁ PRÁCE

Student : Bc. et Bc. Mariya Timofeeva

Vedoucí : Ing. Michal Šebesta

Oponent : Ing. Tomáš Bruckner, Ph.D.

2015

Prohlášení

Prohlašuji, že diplomovou práci jsem vypracovala samostatně a veškerou použitou literaturu a další prameny jsem řádně označila a uvedla v přiloženém seznamu.

……...

V Praze 29. dubna 2015 Mariya Timofeeva

Poděkování

Rada bych tímto způsobem poděkovala především svému vedoucímu diplomové práce Ing. Michalu Šebestovi za vstřícný přístup, cenné připomínky a za poskytnutý čas při zpracování diplomové práce. Touto cestou bych rada poděkovala Ing. Ivanu Jelínkovi za poskytnutou možnost seznámení s využitými nástroji v této diplomové práci a za přínosné konzultace a cenné rady. Také bych chtěla poděkovat Bc. Petru Školovi za pomoc při úpravě crawleru.

Abstrakt

Diplomová práce se zabývá souvislostmi vývoje hodnoty kryptoměn s dynamikou uživateli tvořeného obsahu na Internetu. Hlavní cíle práce jsou zanalyzovat nestrukturovaná data se zaměřením na vývoj témat jednotlivých kryptoměn v čase a propojit klíčová témata a souvislosti s vývojem hodnot jednotlivých kryptoměn.

Pro úspěšné dosažení hlavních cílů jsou v práci popsány metody a proces analýzy nestrukturovaných dat, vysvětlená problematika kryptoměn, identifikován zdroj stažení dat a zvolené kryptoměny pro následnou analýzu.

Přínosem práce je identifikace klíčových témat v čase u zvoleného zdroje, nalezení odpovědi na otázku o čem uživatelé nejvíce diskutují v souvislosti s pěti vybranými kryptoměnami a zjištění existence a případný popis reakcí uživatelů na vývoj kurzů zvolených kryptoměn. Dílčím přínosem práce je praktická ukázka využití zvolených nástrojů při analýze nestrukturovaných dat. Jako další dílčí přínos vidím způsob analýzy dat, využitý v analytické části.

Práce je rozdělená na tři základní části. První část se zabývá teoretickým zakotvením analýzy nestrukturovaných dat a problematikou kryptoměn. Druhá část se věnuje návrhu řešení pro zpracování analytické části. Třetí část je zaměřená na zmapování nejdiskutovanějších kryptoměn ve zvoleném zdroji, identifikaci klíčových témat v čase u pěti zvolených kryptoměn (Bitcoin, Nxt, Dash, Monero, Counterparty) a propojení klíčových témat těchto kryptoměn s vývojem jejich kurzu vůči americkému dolaru.

Klíčová slova: data, analýza, kryptoměna, Bitcoin, Nxt, Dash, Monero, Counterparty.

Abstract

This thesis analyzes a continuity between the changes of the exchange rates of the cryptocurrencies and the dynamics of the created content on the Internet by their users. The main goals of this thesis are to analyze unstructured data focusing on the development of the topics for different cryptocurrencies and to interconnect the key topics with the changes of the exchange rates of this cryptocurrencies.

In order to successfully obtain goals I describe process and methods of the analysis unstructured data, explain the issues of the cryptocurrencies, identify a source of the downloading data and choose some cryptocurrencies for subsequent analysis.

The contribution of this thesis is the identification of the development key topics in the chosen source, finding topics, which are the most under discussion about the chosen five cryptocurrencies. Also this thesis detects the existence of the reaction of the users to the changes of exchange rates of the chosen cryptocurrencies and describes it. Particular contribution of this thesis is a practical demonstration of the application the elected technology for analyzing unstructured data. Also additional contribution is the methods using in the analytical part.

Structure of this thesis is divided to the three main parts. The first part describes the analysis of unstructured data theoretically and explain the issues of cryptocurrencies. The second part devotes a draft of the solution for creating an analytical part. The third part maps the most discussion cryptocurrencies in the chosen source, identifies the development of the key topics for five chosen cryptocurrencies Bitcoin, Nxt, Dash, Monero and Counterparty and interconnects the key topics this cryptocurrencies with changes their exchange rates in relation to American dollar.

Keywords: data, analysis, cryptocurrency, Bitcoin, Nxt, Dash, Monero, Counterparty.

Obsah

Úvod 9
Rešerše zdrojů 12
1 Data 13
1.1 Strukturovaná data 14
1.2 Nestrukturovaná data 14
1.3 Big Data 15
1.3.1 Volume (Objem) 16
1.3.2 Velocity (Rychlost) 17
1.3.3 Variety (Rozmanitost) 18
1.3.4 Veracity (Věrohodnost) 18
2 Proces analýzy nestrukturovaných dat 19
3 Metody analýzy dat 24
3.1 Statistické metody 25
3.1.1 Shluková analýza 25
3.1.1.1 Rozdělovací metody 25
3.1.1.2 Hierarchické metody 26
3.1.1.3 STC 28
3.1.1.4 Lingo 29
3.1.2 Regresní analýza 31
3.1.3 Korelační analýza 32
4 Kryptoměna 33
4.1 Definice kryptoměny 33
4.1.1 Bitcoin 33
4.2 Peněženka a klíče 36
4.3 Transakce 37
4.4 Blok a síť bloků 37
4.5 Těžba 39
4.6 Shrnutí 40
5 Návrh řešení 42
5.1 Technologie 42
5.1.1 Crawler4j 42
5.1.2 Elasticsearch 42
5.1.3 Kibana 43
5.1.4 Carrot2 43

5.2 Architektura řešení 44
5.3 Datové zdroje 44
5.3.1 Bitcoin Forum 46
5.4 Shrnutí úkolů analytické části 49
6 Analytická část 50
6.1 Získání dat 50
6.2 Předzpracování 53
6.3 Analýza a interpretace dat 54
6.3.1 Bitcoin 56
6.3.1.1 Hlavní témata 56
6.3.1.2 Vývoj témat v roce 2010 57
6.3.1.3 Vývoj témat v roce 2011 61
6.3.1.1 Vývoj témat v roce 2012 65
6.3.1.2 Vývoj témat v roce 2013 68
6.3.1.3 Vývoj témat v roce 2014 71
6.3.1.4 Vývoj témat v roce 2015 74
6.3.1.5 Vývoj kurzu 75
6.3.1.6 Dílči shrnutí 79
6.3.2 Nxt 80
6.3.2.1 Hlavní témata 81
6.3.2.2 Vývoj témat v roce 2013 81
6.3.2.3 Vývoj témat v roce 2014 82
6.3.2.4 Vývoj témat v roce 2015 85
6.3.2.5 Vývoj kurzu 86
6.3.2.6 Dílčí shrnutí 88
6.3.3 Monero 88
6.3.3.1 Hlavní témata 88
6.3.3.2 Vývoj témat v roce 2014 89
6.3.3.3 Vývoj témat v roce 2015 91
6.3.3.4 Vývoj kurzu 91
6.3.3.5 Dílčí shrnutí 92
6.3.4 Dash 93
6.3.4.1 Hlavní témata 93
6.3.4.2 Vývoj témat v roce 2014 94
6.3.4.3 Vývoj témat v roce 2015 97

6.3.4.4 Vývoj kurzu 97
6.3.4.5 Dílčí shrnutí 99
6.3.5 Counterparty 99
6.3.5.1 Hlavní témata 100
6.3.5.2 Vývoj témat v roce 2014 101
6.3.5.3 Vývoj témat v roce 2015 104
6.3.5.4 Vývoj kurzu 105
6.3.5.5 Dílčí shrnutí 106
6.3.6 Shrnutí poznatků – globální pohled 107
7 Závěr 111
Seznam literatury 113
Seznam obrázků, tabulek a grafů 125
Přílohy 131

Úvod

První kryptoměna Bitcoin vznikla v roce 2009 a postupem času vyvolala zájem u široké skupiny lidí. Během krátké doby od roku 2009 do doby psaní této diplomové práce vzniklo i mnoha dalších kryptoměn, které zájemci aktivně využívají při určitých činnostech, jako jsou obchodování s pravými měnami, nákup v obchodech, které přijímají kryptoměny, a také při nelegálním obchodování. Jako hlavní zdroj považuji především Internet, prostřednictvím kterého jsou určité kryptoměny prezentovány uživatelům. Následně jsou zde uživateli diskutovány ohledně vlastností daných kryptoměn, příležitostech jejich využití a o dalších tématech, která jsou v této práci identifikovaná.

Tato diplomová práce se zabývá souvislostmi vývoje hodnoty kryptoměn s dynamikou uživateli tvořeného obsahu na Internetu. Motivací tohoto téma byl především zájem o analýzu nestrukturovaných dat, kterou považuji v dnešní době za nedílnou součást hodnocení průběhů určitých činností a samozřejmě zájem o oblast kryptoměn. Domnívám se, že se otevřely nové možnosti pro investování, které je rizikové, ale přináší odpovídající potenciální zisky.

Hlavními cíli této práce je analyzovat nestrukturovaná data se zaměřením na vývoj témat jednotlivých kryptoměn v čase (cíl 1) a propojit klíčová témata a souvislosti s vývojem hodnoty jednotlivých kryptoměn (cíl 2).

Pro dosažení hlavních cílů je nutné stanovit další podpůrné cíle. Před započetím analýzy nestrukturovaných dat v oblasti kryptoměn v čase a propojením zjištěných témat s vývojem hodnoty kryptoměn je potřeba nejprve provést teoretické zakotvení analýzy nestrukturovaných dat (cíl 3) a vysvětlit problematiku kryptoměn (cíl 4). Dalšími podpůrnými cíli jsou identifikace internetových zdrojů s vysokým výskytem příspěvků o kryptoměnách (cíl 5) a zmapování existujících kryptoměn (cíl 6). Z důvodu dodržení rozsahu diplomové práce je zmapováno 20 nejdiskutovanějších kryptoměn u zvoleného zdroje, ze kterých je vybráno pět kryptoměn se třemi různými metodami generace bloků pro zjištění odlišností vývoje jejich klíčových témat. Dalším kapacitním omezením práce je stažený objem dat ze zvoleného zdroje, kdy je analytická část založena na vzorku staženého objemu dat. Hlavní cíle této práce jsou aplikované na pět zvolených kryptoměn, jelikož předpokládám, že vývoj klíčových témat u jednotlivých kryptoměn má obdobný průběh. Hodnota jednotlivých kryptoměn v této práci je určena jejím kurzem vůči americkému dolaru.

Globální pohled na strukturu diplomové práce je zobrazen na schématu č. 1, kde jsou jednotlivé kapitoly uvedeny v kulatých závorkách. Práce je rozdělená na tři části, teoretickou, návrh řešení pro analytickou část a samotnou analytickou část.

Teoretická část práce je rozdělená na čtyři kapitoly. První kapitola se zabývá úvodem do problematiky dat (viz Schéma 1 (1)), konfrontací pojmu dat s informacemi a znalostmi a zmapováním a popisem typů dat. Obsahem druhé kapitoly je popis procesu analýzy nestrukturovaných dat, kde jsou identifikované a popsané jednotlivé fáze tohoto procesu (viz Schéma 1 (2)). Ve třetí kapitole je uveden základní přehled metod analýzy dat (viz Schéma 1 (3)), použitelné metody pro analytickou část a následně výběr a popis zvolených metod. Čtvrtá kapitola teoretické části se zabývá problematikou kryptoměn (viz Schéma 1 (4)). Tato kapitola se věnuje pojmu kryptoměna a identifikaci a popisu klíčových prvků spojených s kryptoměnou. Pro přiblížení k oblasti kryptoměn jsou některé prvky uvedeny na příkladu nejstarší kryptoměny Bitcoin.

Druhá část této diplomové práce je návrh řešení pro analytickou část (viz Schéma 1 (5)). V této části jsou popsány vybrané technologie umožňující splnit stanovené cíle, znázorněná architektura řešení a identifikovány datové zdroje pro následnou analýzu. Součástí této části je shrnutí úkolů pro analytickou část.

Analytická část (viz Schéma 1 (6)) je rozdělena na tři základní kapitoly, které odpovídají jednotlivým krokům procesu analýzy nestrukturovaných dat. První kapitola se zabývá zjištěním potřebných atributů a určením způsobu stažení dat (viz Schéma 1 (6.1)). Další kapitola je věnována předzpracování stažených dat (viz Schéma 1 (6.2)), kde je na základě stažených dat představen plugin "head" Elasticsearch a ukázán přehled struktury indexu stažených dat. Třetí kapitola se zabývá analýzou a postupnou interpretací dat (viz Schéma 1 (6.3)). Středem zájmu této kapitoly je zmapování nejdiskutovanějších kryptoměn a identifikace hlavních témat s následným objasněním souvislostí hlavních témat s obsahem zdroje. Také se tato kapitola zabývá analýzou vývoje hlavních témat a následným propojováním klíčových témat s vývojem kurzů zvolených pěti kryptoměn.

Závěrem práce je zjištění splnění stanovených cílů a shrnutí výsledků provedené analýzy (viz Schéma 1 (7)).

Za vlastní přínos práce považuji identifikaci klíčových témat v čase u zvoleného zdroje a nalezení odpovědi na otázku, o čem uživatelé nejvíce diskutují v souvislosti s pěti vybranými kryptoměnami. Dalším přínosem práce je zjištění existence a případný popis reakcí uživatelů na vývoj kurzů zvolených kryptoměn. Za dílčí přínos považuji zmapování nejdiskutovanějších kryptoměn u zvoleného zdroje, jelikož vývoj kurzu závisí kromě jiných důvodů také na zájmu o určitou kryptoměnu. Tento zájem je částečně měřen aktivitou zájemců o tuto kryptoměnu. Dílčím přínosem je praktická ukázka využití zvolené technologie při analýze nestrukturovaných dat. Jako další dílčí přínos vidím způsob analýzy dat, využitý v analytické části.

Práce je určena pro zájemce o analýzu nestrukturovaných dat, kteří díky ní mohou získat teoretické a analytické znalosti o procesu analýzy nestrukturovaných dat. Metody a technologie aplikované v této práci, pak můžou uplatnit ve své vlastní činnosti. Práce je také určena pro začínající i pokročilé zájemce o problematiku kryptoměn, kteří se mohou seznámit s populárními kryptoměnami a udělat si vlastní názor na informační hodnotu využitého v této diplomové práce zdroje při jeho použití pro určité činnosti.

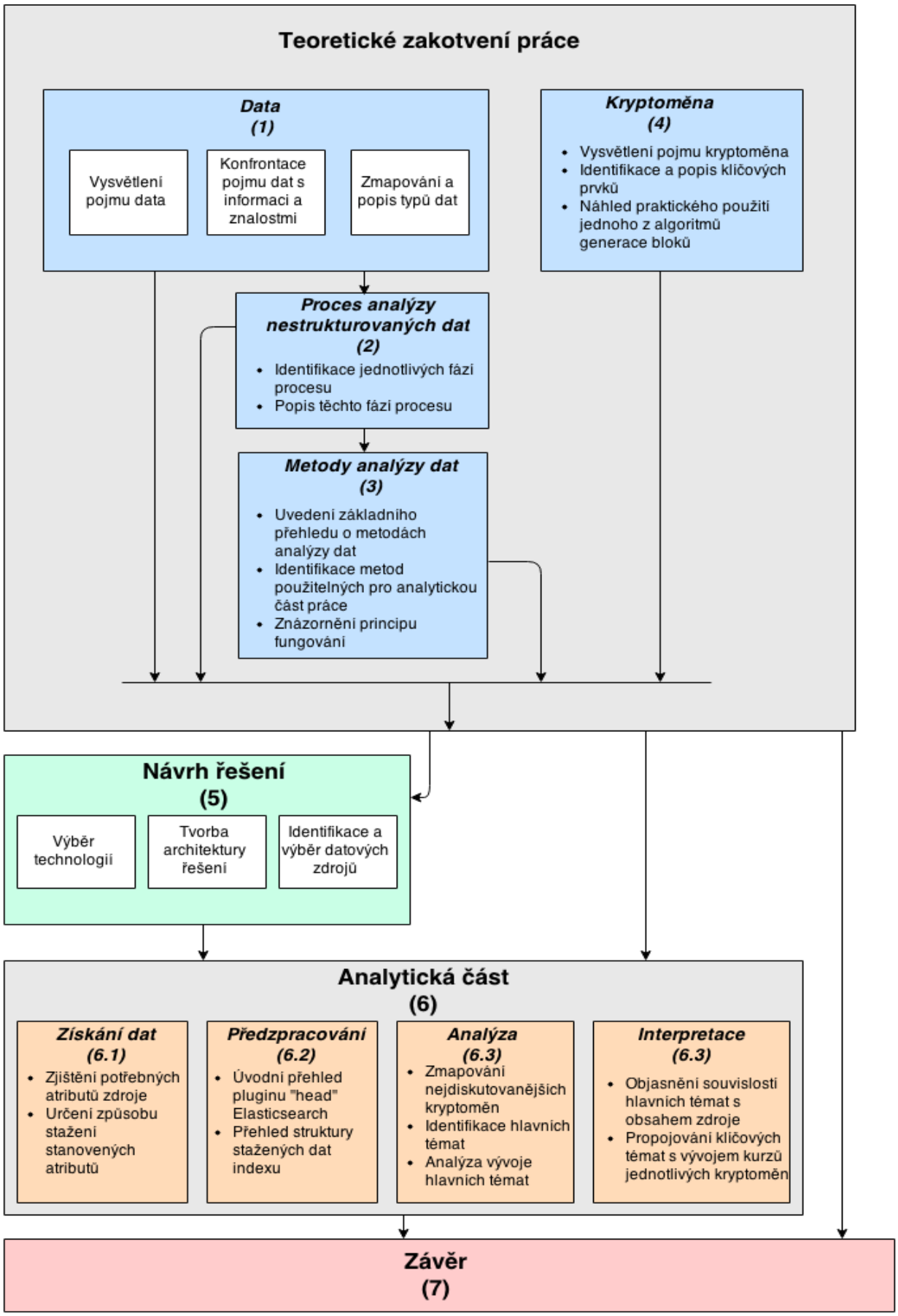


Schéma 1: Globální pohled na strukturu diplomové práce, zdroj vlastní zpracování

Rešerše zdrojů

Zdroje teoretické části této diplomové práce jsou domácí a zahraniční odborné články a knihy, normy, případová studia podniků a informační portály.

V části vysvětlující pojem data a další související s tím pojmy, je zde také popisující proces a metody analýzy dat za významný zdroj z domácí literatury považuji publikace Viléma Sklenáka a kol. "*Data, informace, znalosti a Internet*", která je využitelná pro tuto diplomovou práci tím, že se zabývá mimo jiných témat pojmem data, popisem procesu a metod analýzy dat.

Ze zahraniční literatury významnými zdroji považuji odborné knihy Tuffery Stéphane "*Data mining and statistics for decision making*" a Larose Daniel T "*Discovering knowledge in data: an introduction to data mining*", které se podrobně zabývají procesem a metodami analýzy dat. Ze všech dalších použitých zdrojů vyčleňuji odborný článek Takako Hashimoto, Tesuji Kuboyama, Basabi Chakraborty a Yukari Shirota "*Discovering Topic Transition about the East Japan Great Earthquake in Dynamic Social Media*", který je praktickou ukázkou využití analýzy nestrukturovaných dat.

Primárně pro kapitolu metody analýzy dat základní literaturou jsou odborná kniha autorů Leonarda Kaufmana, Petera J. Rousseeuw "*Finding Groups in Data: An introduction to Cluster Analysis*", odborný článek autorů Milos Ilic, Petar Spalevic a Mladen Veinovic "*Suffix Tree Clustering - Data mining algorithm*" a odborný článek autorů Stanislaw Osiński a Dawid Weiss "*Conceptual Clustering Using Lingo Algorithm: Evaluation on Open Directory Project Data*", které se podrobně zabývají shlukovou analýzou a její algoritmy.

V této části také používám přibližující k určitému tématu příklady ze zúčastněného mnou projektu analýzy bankovních stránek na sociální sítě Facebook na Vysoké škole ekonomické ve spolupráci s Hewlett-Packard, s.r.o.

V teoretické části, která se zabývá problematikou kryptoměn za významný zdroj považuji článek Satoshi Nakamoto "*Bitcoin: A Peer-to-Peer Electronic Cash Systém*", který popisuje koncept elektronického peněžního systému založeného na principech kryptografii první kryptoměny Bitcoin.

Základními zdroji analytické časti, jsou příručka autorů Rafał Kuć a Marek Rogoziński "*Elasticsearch Server. Second Edition*", která vysvětluje a popisuje postupy práce s nástrojem Elasticsearch a Bitcoin Forum, jelikož analýza je uskutečněná na základě dat z tohoto fóra. Dalšími zdroji jsou specializované zpravodajské webové stránky, informační portály, burzovní webové stránky, webové stránky jednotlivých organizací a odborné diskuzní fóra.

1 Data

Data, informace a znalosti v dnešní informační době hrají jednu z nejdůležitějších rolí, která nesporně napomáhá k vývoji organizací. Měli bychom tady rozlišit pojmy data, informace a znalost.

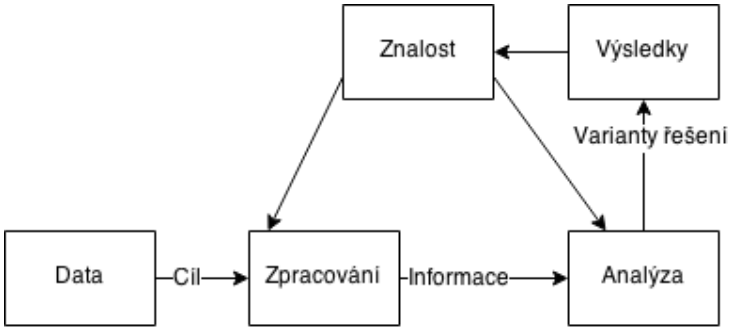
Pod pojmem data se rozumí: "*Opakovaně interpretovatelná formalizovaná podoba informace vhodná pro komunikaci, vyhodnocování nebo zpracování*" (ČSN ISO/IEC 2382-1, 1998). ISO/IEC 15939 (2007) pak uvádí, že data jsou souborem hodnot umístěné v základních a odvozených stupních a/nebo indikátorech. ISO/IEC/IEEE 24765 (2010) definuje data jako reprezentace faktů, konceptů nebo instrukci pro vhodný způsob komunikace, interpretace nebo zpracování člověkem nebo strojem. Sklenák a kol. (2001) pak píše o datech v kontextu počítačové vědy, jako o číslech, textu, zvuku, obrazu anebo jiné formě souborů umožněných ke zpracování počítačem.

S pojmem data je spojen pojem metadata. Obecně se pod pojmem metadata rozumí "*data o datech*" (Sklenák a kol., 2001, s. 333), přesněji "*data sdružená s objekty, která zbavují jejich potenciální uživatele nutnosti předběžné znalosti existence či charakteristik těchto objektů*" (Sklenák a kol., 2001, s. 333). Jako příklad metadat slouží data, která jsou zapsána v záhlaví HTML dokumentu v elementu META, který definuje základní popis obsahu webové stránky, její autory, klíčová slova, povolení vyhledavačům tuto webovou stránku indexovat či nikoliv atd. (Sklenák a kol., 2001).

Informace podle ISO/IEC 2382-1 (1993) jsou data týkající se objektů, jako faktů, situace, věcí, procesů nebo nápadů, včetně konceptů, které v určitém kontextu mají význam. ISO/IEC 10746-2 (2009) pak dodává, že informace představují poznatek, který je vyměnitelný mezi uživateli.

Znalost podle IEEE 1320.2 (1998) je aspekt instančních specifikací, které jsou determinované atributy, vlastnostmi a konstantami, operacemi jenom pro čtení. Sklenák a kol. (2001) definují znalost jako "*vzájemně provázané (měnitelné, rozšiřitelné) struktury souvisejících poznatků*». Z toho lze odvodit, že znalost je druh informace, která je uložená v bázi znalostí a reprezentuje znalosti lidí v určitém oboru, jinými slovy je to intelektuální kapitál. (Tkačenko, 2007).

Z toho vyplývá, že existuje soubor dat, která se pak za určitým cílem dá zpracovávat člověkem anebo strojem, z čehož dostaneme informace. Po analýze informací dostaneme nějaké výsledky, které se pak stávají znalostmi určenými buď pro další zpracování, nebo analýzu (viz Obrázek 1).



Obrázek 1: Souvislost pojmů data, informace a znalostí, zdroj Tkačenko (2007)

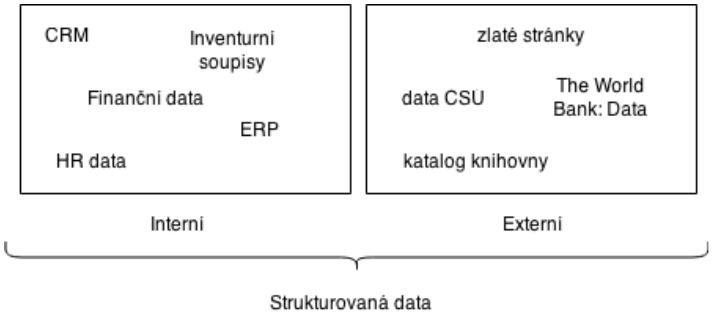
Tato diplomová práce se zabývá analýzou dat, a proto další části budou zaměřené na pojem data, která se z hlediska práce rozlišují na dva typy: strukturovaná a nestrukturovaná data

(Sklenák a kol., 2001). Také s pojmem data je spojen pojem Big Data, který je popsán v samostatné kapitole.

1.1 Strukturovaná data

Strukturovaná data podle Sklenáka a kol. (2001) představují fakta, atributy, objekty apod., které jsou charakterizované nějakými elementy jako například hierarchie elementů v relačních databázových systémech pole → záznam → relace → databáze, pomocí kterých pak vyhledávání dat je snadnější a rychlejší.

Strukturovaná data pak podle dostupnosti lze rozdělit na interní a externí (viz Obrázek 2). Interní strukturovaná data nejsou veřejná a nejvíce pochopitelná organizaci, která je vlastní. Příkladem takových dat mohou být data uložená v podnikových informačních systémech jako CRM, ERP atd., inventurní soupisy s daty o skutečných stavech majetků, finanční data v podobě účetních výkazů, HR data o uchazečích o práci atd. Externí data pak představují zveřejněná data, příkladem mohou být Zlaté stránky, data ČSÚ anebo The World Bank, dále také databáze knihoven lze počítat za externí strukturovaná data.

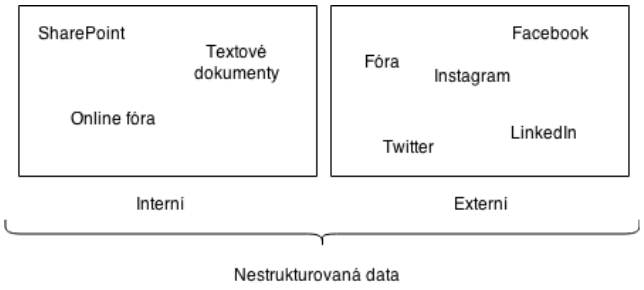


Obrázek 2:Příklady strukturovaných dat, zdroj Nair a Narayanan (2012) a vlastní zpracování

1.2 Nestrukturovaná data

Stewart (2013) definuje nestrukturovaná data jako obsah (angl. content), pro který není předem definovaný žádný data model. Obvykle jsou data vytvořená uživateli a mají uživatelsky orientovaný obsah (people-oriented content), který nelze zařadit do žádné databázové tabulky.

Nestrukturovaná data stejně jako i strukturovaná lze rozlišit na interní a externí (viz Obrázek 3). Interní data představují nestrukturovaný tok dat uvnitř firmy, příkladem jsou SharePoint pro podporu týmové spolupráce, různé textové dokumenty a online firemní fóra atd. Příkladem zveřejněných nestrukturovaných dat mohou být různá diskuzní fóra a sociální sítě: Facebook, Instagram, LinkedIn, Twitter atd.



Obrázek 3: Příklady nestrukturovaných dat, zdroj Nair a Narayanan (2012) a vlastní zpracování

S pojmem strukturovaná data a nestrukturovaná data je spojen pojem Big data, která se dělí na strukturovaná a nestrukturovaná, přičemž podíl nestrukturovaných dat podle Beala činí 90 % z veškerého objemu Big dat.

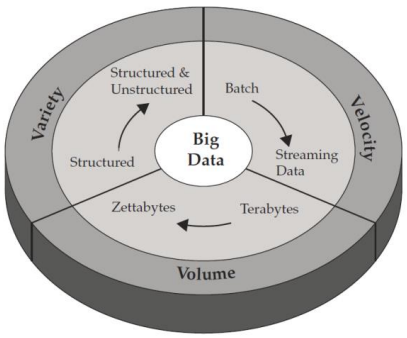
1.3 Big Data

V dnešní době lze najít několik různých definic vysvětlujících pojem Big data. Důkazem toho jsou následující definice:

- "*Big data – je velký bazén dat, která mohou být zachycená, sdělená, agregovaná, uložená a analyzovaná*[1]" (McKinsey Global Institute, 2011, s. 4),
- "*Big data představují data, která převážně překračují řídící kapacity běžných databázových systémů. Objem dat je velký, rychlé generující nebo nehodící struktuře tradičních databázových architektur* [2]" (Sundaresan a Kandavel, 2014, s. 1),
- "*Big data jsou technologií popisující novou generací technologií a architektur, které byly navřené pro ekonomickou extrakci dat z velkého objemů široké variace dat umožňující velmi rychlé získat data, objevit a analyzovat*[3]" (Gantz a Reinsel, 2011, s. 6).

Z výše uvedených definic je vidět, že pod pojmem Big data si lze představovat jak data, tak i technologie, které jsou určené ke zpracování strukturovaných a nestrukturovaných dat. v této diplomové práci jsou Big data popsaná z pohledu dat a popis technologií není cílem této práce.

Koncepce pojmu Big data ale podle většiny autorů je stejná, většina autorů charakterizuje Big data prostřednictvím tzv. „tři V[4]" nebo „čtyři V[5]". Například Lodha, Jain, Kurup. (2014) nebo McKinsey Global Institute (2011) uvádí tři charakteristiky Big dat, která jsou zobrazená na obrázku č. 4. Těmito charakteristikami jsou rozmanitost (variety) v datech, která se dělí na strukturovaná a nestrukturovaná, rychlost (velocity) dávkového a potokového zpracování dat a objem (volume) dat, který je měřen v zetabytech nebo terabytech.



Obrázek 4: Charakteristiky Big dat, zdroj Zikopoulos, Rousseeuw, Pei (2012)

V současné době stále častěji někteří autoři jako například Cichocki (2013) nebo Lukoianova a Rubin (2014) uvádějí čtvrtou charakteristiku Big dat, kterou je věrohodnost získaných dat (veracity).

Důležité je pak rozlišovat, z jakých zdrojů data pochází. Obecně lze tyto zdroje rozdělit podle Lalor (2014) na tři základní druhy:

---

[1] "*Big data—large pools of data that can be captured, communicated, aggregated, stored, and analyzed*" (McKinsey Global Institute, 2011, s. 4).

[2] "*Big data is data that predominantly exceeds the managing capacity of a conventional database systems.The amount of data is huge, produced fast, or doesnt fit the structures of traditional database architectures*" (Sundaresan a Kandavel, 2014, s. 1).

[3] "*Big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis*" (Gantz a Reinsel, 2011, s. 6).

[4] three V's

[5] fourth V

- Uživatelské generované obsahy (Human-sourced information), do tohoto typu zdroje patří sociální sítě, blogy nebo jiné sítě, které byly založené na principu a pomocí technologií Web 2.0 a které dovolují vytvářet a měnit uživatelský generovaný obsah (náplň) (angl. user generated content) (Kaplan a Haenlein, 2010), (Kurniawati, Shanks, Bekmamedova, 2013),
- Procesně zprostředkovatelská data, která jsou tvořená během byznysové činnosti (Process-mediated data). Tato skupina zahrnuje data, která byla vytvořená firmou buď uvnitř vlastní organizace anebo navenek pro dodavatele nebo pro zákazníky,
- Data vygenerovaná strojem (Machine-generated data), to jsou data, která byla vytvořená pouze strojem bez pomoci člověka, příkladem může být GPS systém nebo výstupy z lékařských přístrojů atd.

Další část této kapitoly je zaměřená na podrobnější popis každé z charakteristik a jejich odlišnosti podle typu zdrojů.

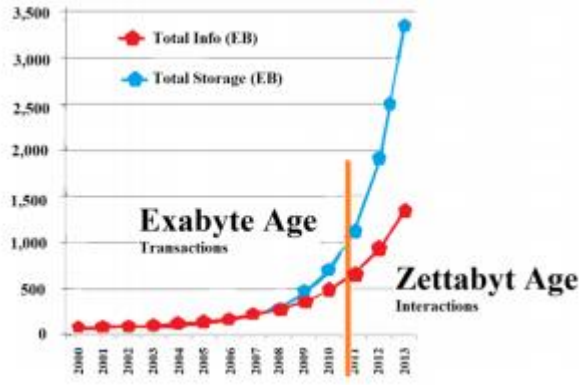
1.3.1 Volume (Objem)

Touto charakteristikou lze definovat Big data jako rychle objemově rostoucí data, která jsou nashromážděná z různých zdrojů jak v režimu off-line uvnitř firmy, tak i v síti Internet. Podle Lalor (2014) je objem Big dat u každého druhu zdroje různý, nejvíce vygenerovaných dat pochází z dat vytvořených strojem, méně, ale stále velký objem dat se tvoří na sociálních sítích anebo jiných sítích, kde je obsah uživatelský generován. v porovnání s těmito zdroji data vznikající v rámci firmy objemově zůstávají pozadu (viz Tabulka 1).

Tabulka 1: Porovnání objemu dat u různých zdrojů Lalor (2014) a vlastní zpracování[6]

Charakteristika/druh zdroje	Uživatelské generované obsahy	Procesně zprostředkovatelská data	Data vygenerovaná strojem
Objem (volume)	2	3	1

Objem uložených dat roste exponenciálně (Zaslavsky, Perera, Georgakopoulos, 2013), v roce 2000 objem dat činil 800 000 PB (Zikopoulos, Rousseeuw, Pei, 2012), v roce 2010 objem dat byl cca jeden ZB, na konci roku 2011 objem dat vzrostl až na 1,8 ZB (Zaslavsky, Perera, Georgakopoulos, 2013) (viz Graf 1), předpokládá se, že do roku 2020 objem dat dosáhne 35 ZB (Zikopoulos, Rousseeuw, Pei, 2012).



Graf 1: Objem vygenerovaných dat, zdroj (Zaslavsky, Perera, Georgakopoulos, 2013)

[6] Čísla v tabulce ukazují na porovnání pořadí velikosti objemu vytvořených dat u různých zdrojů, kde 1 je největší vytvořený objem dat a 3 je nejmenší v porovnání s uvedenými zdroji.

Pro představu dole v tabulce č. 2 je uveden obsah bajtů v jednotlivých jednotkách měření množství dat, označení písmen je podle mezinárodního standardu IEC 60027-2 (2005).

Tabulka 2: Obsah bajtů v jednotlivých jednotkách měření množství dat, zdroj Ram (2004)

Jednotka	Bajty	Bajty
KiB	$1\ 024^1$	1 024
MiB	$1\ 024^2$	1 048 576
GiB	$1\ 024^3$	1 073 741 824
TiB	$1\ 024^4$	1 099 511 627 776
PiB	$1\ 024^5$	1 125 899 906 842 624
EiB	$1\ 024^6$	1 152 921 504 606 846 976
ZiB	$1\ 024^7$	1 180 591 620 717 411 303 424

Z grafu č. 1 je pak vidět, jak rychle objemově rostou Big data, která jsou tvořená uživateli na Internetu, ve firmách a stroji. Tady je pak velký prostor pro analýzu Big dat, co se v těch datech skrývá, co lze z toho vytáhnout a jaký přínos ponesou získané informace.

1.3.2 Velocity (Rychlost)

Pod charakteristikou velocity (rychlost) se rozumí, s jakou frekvencí (v milisekundách, v sekundách, minutách, hodinách, dnech atd.) jsou data vygenerovaná (Zaslavsky, Perera, Georgakopoulos, 2013), uložená a zpracovaná (Zikopoulos, Rousseeuw, Pei, 2012).

Frekvence pak záleží na druhu zdroje, nejrychleji rostou data vygenerovaná strojem, pak u zdrojů, kde jsou data tvořená uživateli, rychlost vzniku dat ale závisí na obsahu, některá témata jsou uživateli diskutovaná více a některá méně. Nejméně rychle v porovnání s předchozími zdroji roste ve firmách (viz Tabulka 3) Lalor (2014).

Tabulka 3:Porovnání rychlosti vytvořených dat u různých zdrojů Lalor (2014) a vlastní zpracování[7]

Charakteristika/druh zdroje	Uživatelské generované obsahy	Procesně zprostředkovatelská data	Data vygenerovaná strojem
Rychlost (velocity)	2	3	1

Zpracování dat probíhá podle potřeb zpracovatele, obecně lze rozlišit tři základní kategorie nutnosti zpracování dat: příležitostně, často a zpracování v reálném čase (Zaslavsky, Perera, Georgakopoulos, 2013). V dnešní době ale stále více dat je třeba zpracovat co nejrychleji buď z důvodu jejich krátké existence, anebo z důvodu nutnosti rychlé identifikace trendu, problému nebo příležitosti. Pak čím rychleji data budou zpracovaná, tím rychleji zpracovatel může dosáhnout přínosných výsledků.

Uvedu příklad charakteristiky rychlosti Big dat a nutnosti rychlého zpracování v návaznosti (Škola, Šperková, Timofeeva, 2014), který se zabýval problematikou využití sociální sítě Facebook v bankovní sféře jako komunikačního prostředku mezi bankami a jejich klienty. Z rychle rostoucího počtu příspěvků jsme měli odhalit využitelné informace pro banky. Jeden ze závěrů byl, že čím rychleji banky odhalí nejaktivnějšího uživatele na Facebooku, jeho postoj k bance a v jaké míře ovlivní názory ostatních uživatelů, tím rychleji mohou najít prostředek k udržení stálých zákazníků, k získání nových, anebo naopak mohou

[7] Čísla v tabulce ukazují na porovnání pořadí rychlosti vytvořených dat u různých zdrojů, kde 1 je nejrychlejší vytvořený objem dat a 3 je nejméně rychlý v porovnání s uvedenými zdroji.

odhalit uživatele, který brzdí rozvoj banky. Z toho plyne, že z rychle rostoucích dat na Faceebooku banky je možné zjistit názory uživatelů, řešit včas jejich problémy a odhalit jiné příležitosti a tím ovlivnit vlastní zisk.

1.3.3 Variety (Rozmanitost)

Big data jsou daná svojí rozmanitostí, data mohou být strukturovaná, nestrukturovaná a semistrukturovaná. Strukturovaná a nestrukturovaná data jsou vysvětlena výše, takže když jenom shrneme - strukturovaná data jsou organizovaná do sémantických entit, jsou uložená do relačních databází, entity v jedné skupině mají stejný popis (atribut), atributy pro všechny entity v jedné skupině mají stejný definovaný formát a předdefinovanou délku (Abiteboul, Buneman, Suciu, 1999). Nestrukturovaná data naopak nejsou zařazená do žádných relačních databází a nemají žádné pravidlo uspořádání. Semistrukturovaná data představují takovou formu dat, která nemá striktní vymezenou strukturu tabulek a relací v relačních databázích, jako mají strukturovaná data, nebo jejich struktura se často mění. Entity v jedné skupině mohou mít různé atributy. Zároveň ale tahle forma dat obsahuje tzv. tagy nebo jiné označení (angl. marker) pro roztřídění sémantických elementů a pro zabezpečení hierarchické struktury polí dat, příkladem je elektronická pošta (Abiteboul, Buneman, Suciu, 1999)

Podle druhu zdroje se liší i rozmanitost dat. Při porovnání tří základních druhů zdroje lze říci, že u dat vygenerovaných stroji je rozmanitost nejvyšší, u uživatelsky generovaných obsahů rozmanitost je v porovnání s ostatními daty střední, nejmenší různorodost dat mají procesně zprostředkovatelská data (viz Tabulka 4).

Tabulka 4:Porovnání rozmanitosti dat u různých zdrojů Lalor (2014) a vlastní zpracování[8]

Charakteristika/druh zdroje	Uživatelské generované obsahy	Procesně zprostředkovatelská data	Data vygenerovaná strojem
Rozmanitost (variety)	2	3	1

1.3.4 Veracity (Věrohodnost)

Rychle objemově rostoucí rozmanitá data mohou, ale nemusí být věrohodná. Podle (Sharma a kol. 2014) ve světě byznysu, lídři zcela nevěří informacím vytaženým z Big dat, ačkoliv pro lepší rozhodnutí informace získané z Big dat využívají.

Věrohodnost u každého ze zdrojů dat se pak liší podle toho, kým nebo čím ta data byla vytvořená (viz Tabulka 5). Věrohodnost dat vygenerovaných strojem záleží na kvalitě využitého hardwaru nebo poskytovatele služeb, pokud data vznikla procesně zprostředkovatelsky, pak věrohodnost záleží na společnosti, která ta data vytvořila. Nejméně snadné odhalení věrohodnosti je u obsahu tvořeném uživateli, literatura říká, že věrohodnost dat zaleží na typu providera, společnosti nebo instituce. Také lze říct, že věrohodnost u tohoto druhu zdroje závisí na druhu sociální sítě nebo na tématu diskuze Lalor (2014).

[8] Čísla v tabulce ukazují na porovnání pořadí rozmanitosti vytvořených dat u různých zdrojů, kde 1 je nejvyšší rozmanitost dat a 3 je nejmenší v porovnání s uvedenými zdroji.

Tabulka 5: Porovnání věrohodnosti u různých zdrojů Lalor (2014) a vlastní zpracování

Charakteristika/druh zdroje	Uživatelské generované obsahy	Procesně zprostředkovatelská data	Data vygenerovaná strojem
Věrohodnost (veracity)	Provider / Společnost a instituce / Druh sociálních sítí / Témata	Společnost	Hardware / Provider

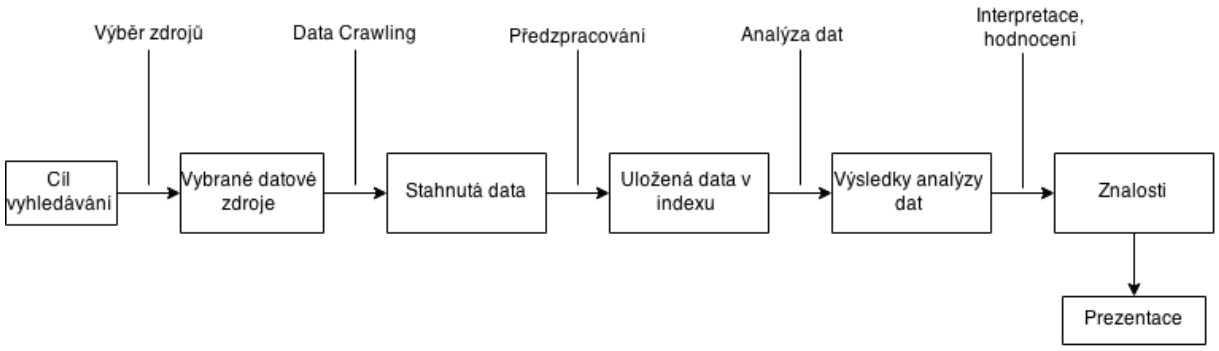
Tato práce se zabývá analýzou nestrukturovaných dat, a proto se další kapitola bude věnovat popisu procesu analýzy nestrukturovaných dat.

2 Proces analýzy nestrukturovaných dat

Z důvodu rychle rostoucího objemu rozmanitých dat je třeba mít určitý postup pro jejich zpracování. Proces analýzy dat lze rozdělit na několik základních fází. Obecně lze proces analýzy dat rozdělit na následující kroky: vyhodnocení, jaká data je třeba stáhnout, a zvolení zdrojů, kde ta data jsou umístěná, následuje pak stáhnutí, dalším krokem je vyčištění dat, pak analýza, interpretace a hodnocení výsledků, posledním krokem je vizualizace a prezentace výsledků (Runkler, 2012). Literatura nabízí různé modifikace těchto kroků, jako například analýza dat prostřednictvím získávání znalostí z databází, metodiku CRISP-DM a jiné možnosti dostupné pro analýzu strukturovaných a nestrukturovaných dat. Analytická část této diplomové práce se zabývá analýzou nestrukturovaných dat, a proto další část této kapitoly se bude zabývat procesem analýzy nestrukturovaných dat.

Proces analýzy nestrukturovaných dat je charakteristický tím, že zvolená data, jak už bylo popsáno v subkapitole č. 1.2, nemají žádný datový model, jelikož analytická část této diplomové práce se zabývá analýzou nestrukturovaných dat tvořených uživateli, pak popis procesu analýzy dat bude zaměřen na uživatelský generovaný obsah.

Pomocí publikací autorů Hashimoto a kol. (2012), Kuć a Rogoziński (2014), Sklenák a kol. (2001) byl sestaven obecný model procesu analýzy nestrukturovaných dat (viz Obrázek 5).



Obrázek 5: Proces analýzy nestrukturovaných dat, zdroj Hashimoto a kol. (2012), Kuć a Rogoziński (2014), Sklenák a kol. (2001) a vlastní zpracování

Cíl vyhledávání

Zpracovatel předem stanoví cíl vyhledávání, občas je ale problémem nestrukturovaných dat skutečnost, že je těžké stanovit, co přesně se bude v těchto datech vyhledávat (Škola, Šperková, Timofeeva, 2014). Cíle vyhledávání v kontextu sociálních sítí mohou být různá, například identifikace témat nebo řešení problémů, které jsou nejběžnější na sociálních sítích ve spojení s určitou událostí (Hashimoto a kol., 2012) nebo vyhledávání nejaktivnějšího uživatele

nebo skupiny uživatelů na sociálních sítích, identifikace zájmů uživatelů nebo hodnocení uživateli určité organizace (Škola, Šperková, Timofeeva, 2014). Důvodem těchto vyhledávání v komerční sféře je především dosažení většího zisku, zvýšení tržní hodnoty organizace a získání většího podílu na trhu, podpůrnými cíli jsou udržení dobrého image organizace, udržení stálých zákazníků a získání nových.

Výběr zdrojů a data crawling

Po stanovení cílů vyhledávání je třeba identifikovat zdroje, ze kterých lze data získat. Data použitá pro dosažení cílů by měla být dostupná ke stažení z těchto zdrojů. v tomto procesu z důvodu existence či dostupnosti dat může dojít k přeformulování hlavních cílů vyhledávání (Sklenák a kol., 2001). Zdroje dat jsou nestrukturovaná data, která nemají žádný předem identifikovaný data model a nepatří do žádné relační databáze. Zpracovatel by měl vybrat vhodné zdroje podle stanoveného cíle a pak tato data ze zvolených zdrojů stáhnout.

Dále vyhledává data podle předem definovaných podmínek a následně je stahuje tzv. web crawler, který prochází HTML kód každé webové stránky a stahuje pouze ta data, které těmto podmínkám odpovídají. Základní algoritmus web crawlera obsahuje následující komponenty:

- komponenta pro ukládání seznamu URL ke stáhnutí v URL frontě (URL frontier),
- komponenta pro určení IP adresy hostitele,
- komponenta pro stáhnutí dokumentů prostřednictvím HTTP protokolu,
- komponenta pro extrakci odkazů z HTML dokumentů a
- komponenta, která zjistí, jestli určitý URL před tím už byl zpracován (Heydon a Najork, 1999).

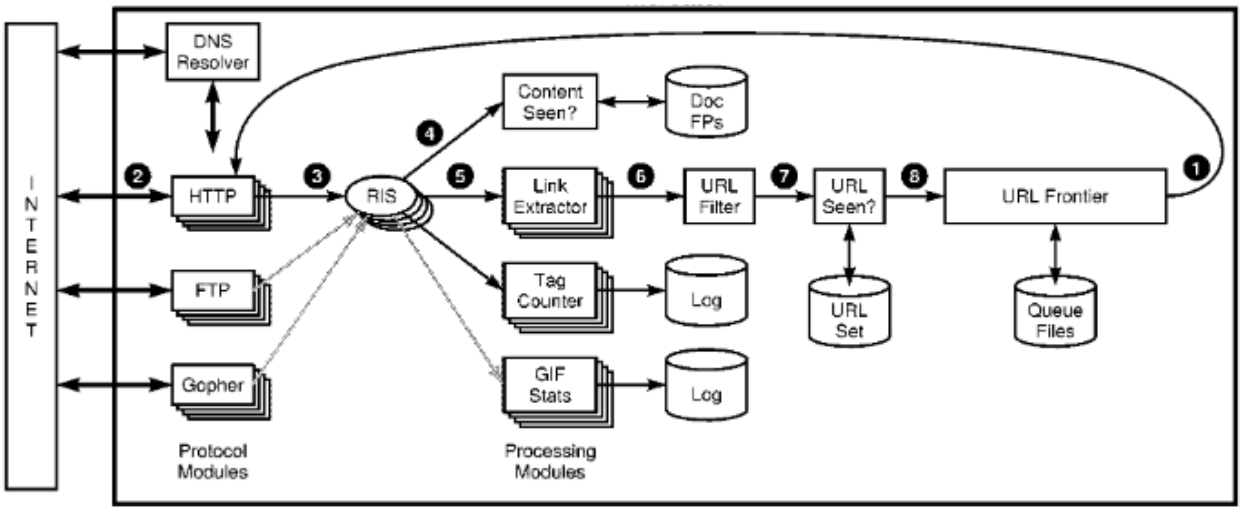
Pro podrobnější vysvětlení principu fungování web crawlera uvedeme architekturu Web crawlera Mercator (viz Obrázek 6). Základní algoritmus web crawlera je následující: vyhledává seznam předem definovaných URL a opakovaně provádí následující kroky.

Na začátku web crawler vytáhne absolutní URL ze seznamu URL v URL frontier používající FIFO princip, kde se zpracovává ten URL, který je ve frontě nejdelší dobu. Každý absolutní URL začíná s “http”, který definuje, jaký síťový protokol by měl být stažený. U web crawlera Mercator tyto síťové protokoly jsou definované v protokolových modulech, které definují základní konfigurace při startování crawlera, v našem případě protokolové moduly obsahují konfigurace pro HTTP, FTP a Gopher.

Vycházející z URL schématu crawler identifikuje vhodný protokol pro stažení dokumentů, kde následně druhým krokem stáhne pomocí tohoto protokolu dokument z internetu a umístí ho v tzv. RewindInputStream (RIS). RIS je I/O abstrakcí, která dovoluje náhodné vstupní potoky dokumentů přečíst několikrát. ¨

Čtvrtým krokem dokument prochází obsahovými testy a zjišťuje se, jestli tento dokument už byl předtím prohlédnut i z jiných URL. Pokud dokument předtím byl prohlédnutý, pak crawler odstraňuje URL z fronty. Každý stažený dokument má MIME typ, například text/html nebo image/gif a na základě tohoto MIME typu crawler stahuje data z HTML, každý odkaz umístěný v HTML kódu pak konvertuje na absolutní URL a odesílá na kontrolu do URL filtrů, kde se určuje, jestli URL má být stažená do fronty.

Pokud URL prošla URL filtry, dalším krokem jsou URL testy, kde se kontroluje, jestli daná URL byla předtím prohlednutá, pokud URL je nová, pak se přidává do URL fronty. Po ukončení zpracování jednotného URL se pak ten proces opakuje do té doby, pokud nebudou stažená data ze všech URL podle předem vymezených podmínek (Heydon a Najork, 1999).



Obrázek 6: Architektura Web Crawlera Mercator, zdroj Heydon a Najork (1999)

Předzpracování

Pod předzpracováním se obecně rozumí příprava stažených dat pro následnou analýzu. Během tohoto procesu se data čistí a transformují z různých zdrojů do jednotné podoby (Runkler, 2012).

V kontextu s analýzou nestrukturovaných dat se lze setkat s dalším pojmem indexování. Stažené crawlerem dokumenty se pak přepisuji do struktury tzv. obracený index (inverted index). Tento index je tabulkovou datovou strukturou, která mapuje pouze hlavní termíny nebo hlavní elementy dokumentů, žádné relační vztahy tady nejsou (Kuć a Rogoziński, 2014). Jako příklad poslouží následující ukázka dvou dokumentů o finančních institucích:

- Česká spořitelna,
- Česká pojišťovna.

Index pak bude vypadat následovně:

Tabulka 6: Index, zdroj Kuć a Rogoziński (2014) a vlastní zpracování

Termín	Počet	Dokument
Česká	2	<1>, <2>
Spořitelna	1	<1>
Pojišťovna	1	<2>

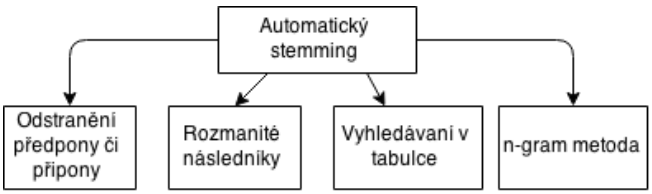
Z tabulky č. 6 je vidět, že index má tabulkovou podobu a obsahuje počet každého slova ve dvou dokumentech dohromady, což umožňuje rychlé vyhledávání založené na vyhledávání podle termínů.

Z dokumentové podoby do indexu transformuje tzv. strojový analytik, který transformuje data podle předem definovaných podmínek. Příkladem je jazykový analytik, který bez žádných podmínek rozkládá do jednotlivých řádků tabulky každý tvar určitého slova (Kuć a Rogoziński, 2014).

Transformace dat a očištění jsou spojené s procesem tzv. stemmingem, který znamená převedení slov do základních forem za účelem snížení nesouladu ve výsledcích dotazování uživateli na nějaké slovo v dokumentech, například pokud uživatel chce vyhledat dokumenty na téma "Jak se naučit vařit" a napíše dotaz pouze "vaření", pak bez stemmingu software může neposkytnout všechny relevantní výsledky. Pokud dotaz bude očištěn pomocí algoritmu stemming, pak slovo "vaření" bude převedeno na "vařit", čímž pak budou vyhledány všechny

relevantní dokumenty. Dále se stemming využívá pro zlepšení fungování systémů vyhledávání informací (angl. Information Retrieval Systems) u různých jazyků, například pro češtinu, francouzštinu atd. Pak se stemming používá také pro snížení velikosti indexů (Sharma, 2012).

Stemming lze provádět ručně, ale také i automaticky. Automatický stemming se pak dělí na čtyři různé metody (viz Obrázek 7).



Obrázek 7: Automatický stemming, zdroj Sharma, (2012)

První metodou je odstranění předpon či přípon za účelem převedení slova do základního tvaru.

Další metodou je rozmanitý následník, který počítá, kolik různých písmen a která jsou umístěná v jednom řádku a porovnává to s hledaným, například v řádku se slovy

telefon, televizor, teleskop

uživatel chce vyhledat rozmanitého následníka u slova *telefax*. Vyhledává se to pak následovně. v uvedeném řádku má písmeno *t* jenom jednoho následníka, a je to písmeno *e*. Dále každé písmeno do druhého *e* má stejný počet následníků. Potom zůstane jenom jedno slovo se stejným následníkem druhého *e*. Nakonec v tomto řádku nejsou stejné následníky, jako u vyhledávaného slova. Tímto způsobem se zjišťují základní tvary slov.

Další metodou stemminigu je vyhledávaní v tabulce. Termíny a jejich základní tvary jsou uložené do tabulek, stemming pak probíhá způsobem vyhledávání slov z těchto tabulek.

Poslední metodou je n-gram metoda, která je reprezentována diagramy, trigramy atd. složenými z písmen jdoucích po sobě. Pak se prostřednictvím určitých koeficientů vypočítává podobnost slov. Například uvedené slova rozložíme do diagramů:

technologie -> te ech chn no ol lo og gi ie

technologický -> te ech chn no ol lo og gi ic ck ký

technolog-> te ech chn no ol lo og

Pak má každé slovo vlastní počet unikátních diagramů, například technologie má devět unikátních diagramů a technologický jedenáct atd. Podobnost těchto slov se dále vypočítává prostřednictvím určitých koeficientů, například pomocí Dicího (angl. Dice) podobnost prvních dvou slov se rovná 80 %[9] (Sharma, 2012).

Analýza dat

Po uložení dat do indexu pokračuje dál analýza dat, úkolem které je "*redukce, organizace, syntéza a sumarizace informací s cílem dát výsledkům význam a provádět úvahy podporované validní evidencí*" (FTVS, 2012). Pro analýzu dat lze uplatnit několik různých metod, přičemž se některé šíří jak pro strukturovaná, tak i pro nestrukturovaná data, některé jsou

[9] S = 2*8*(9+11), kde 2 je dané číslo ze vzorku, 8 je počet podobných diagramů, 9 je počet diagramů v prvním slovu, 11 je počet diagramů v druhém slovu.

určené ale jenom pro analýzu nestrukturovaných dat. Jednotlivými metodami analýzy dat se bude zabývat následující kapitola této diplomové práce.

Interpretace a vyhodnocení

Analýzou dat celý proces nekončí, jelikož výsledky analýzy neposkytují smysluplné závěry pro předem stanovený cíl. Výsledky je třeba vyhodnotit a vhodně interpretovat. Jednotná pravidla pro interpretaci a vyhodnocení neexistují, vždy záleží na cílech vyhledávání, které pokud byly stanovené na začátku procesu získání dat, ovlivní interpretaci. Obecně lze rozlišit tři základní druhy interpretace (viz Obrázek 8):

- Souvislost
  - Vnitřní

    Pod vnitřním porovnáním se rozumí porovnání zjištěných hodnot mezi sebou nebo porovnání zjištěných hodnot s vnitřními předem stanovenými hodnotami.

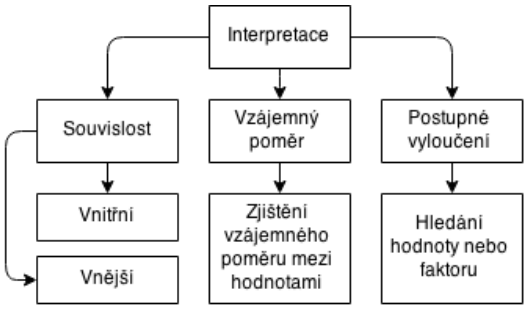
  - Vnější

    Vnější souvislost znamená konfrontaci zjištěných hodnot s vnějšími hodnotami, například s ukazateli trhu, s výsledky jiných průzkumů atd.

- Vzájemný poměr

  Interpretace výsledků analýzy se projevuje v existenci předpokladu, že zjištěné hodnoty jsou ve vzájemném poměru, například věk uživatele a jeho zájmy (Škola, Šperková, Timofeeva, 2014).
- Postupné vyloučení

  Interpretace postupem vyloučení znamená to, že se ze zjištěných hodnot analýzou vyhledává faktor nebo hodnota, které mají vliv na dosažení předem stanovených cílů (Gorshkov a Sheregi, 2012).



Obrázek 8: Interpretace dat, zdroj Gorshkov a Sheregi (2012)a vlastní zpracování

Prezentace

Literatura pak vyčleňuje poslední krok procesu analýzy dat, kterým je prezentování znalostí, jinými slovy prezentace interpretace výsledků analýzy dat. Často se v literatuře lze setkat s pojmem vizualizace, která představuje metody grafického zpracování interpretace výsledků analýzy v podobě grafů, diagramů, animací nebo jiných způsobů grafického zobrazení pro usnadnění porozumění získaným výsledkům (Han, Kamber, Pei 2012).

3 Metody analýzy dat

Na výsledky analýzy Big dat má vliv, jaká metoda analýzy je zvolená, analytická architektura a využité softwarové technologie (Chen a kol., 2014). Obecně lze metody analýzy dat rozlišit na tradiční metody analýzy dat a i jiné metody, které nepatří do tradičních modelů. Cílem této kapitoly je popsat různé metody analýzy dat, které budou využité v analytické části této diplomové práce.

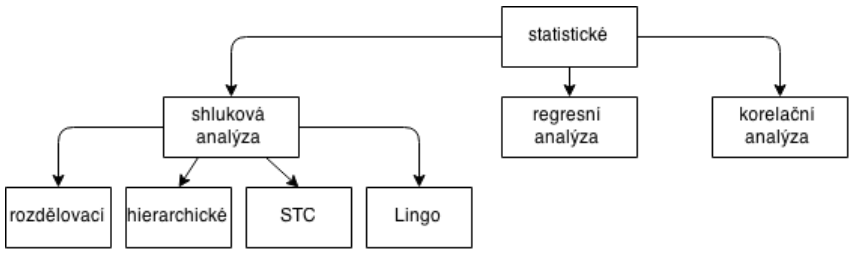
Tradiční metody analýzy dat, které se využívají pro analýzu jak strukturovaných, tak i nestrukturovaných dat, jsou založené na vhodných metodách pro stažení primárních a sekundárních dat (first-hand data and second-hand data), vyčištění užitečných dat z chaoticky uspořádaných dat a následnou analýzu (Chen a kol., 2014). Obecně mezi tradiční metody analýzy dat podle Sklenáka a kol. (2001) patří:

- **Statistické metody** jsou vhodné především pro analýzy numerických dat, využívá se v těchto metodách matematické funkce, vektory nebo podmíněné pravděpodobnosti, jako příklady těchto metod slouží regresní metody, diskriminační analýza, shluková analýza nebo bayesovské metody.
- **Metody umělé inteligence**
    - **Symbolické**, které vyhledávají zajímavé vztahy v datech a zjišťují jejich strukturu. Na rozdíl od statistických metod jsou tyto metody založené na logických usuzováních, a proto jsou snadnější pro uživatele. Představitelé těchto metod jsou rozhodovací stromy, metoda IF-THEN, případové usuzování atd.
    - **Subsymbolické** jsou rozšířením symbolických metod umělé inteligence, reprezentanty jsou neuronové sítě nebo genetické algoritmy.

Metody analýzy dat řeší několik následujících úloh:

- Klasifikace – hledání a zařazení nového objektu do předem stanovaných kategorií, např. diskriminační a regresní analýzy, rozhodovací stromy, *k*-nejbližší soused, neuronové sítě (Řezanková, 2001), (Berry a Linoff, 2004).
- Odhad hodnoty vysvětlené proměnné – určuje hodnotu cílové proměnné, například váhu u neúrovňových sítí, příklady tohoto typu úloh jsou lineární a nelineární regresní analýza, neuronové sítě (Řezanková, 2001).
- Predikce – důležitou roli u této úlohy hraje čas (Sklenák a kol., 2001), jelikož na základě trénovacích historických dat se predikuji budoucí výsledky. Příklady mohou být všechny metody, které patří do úloh klasifikace a odhad (Berry a Linoff, 2004)).
- Segmentace (shlukování) – cílem je seskupit data na základě jejich shody do určité míry, přičemž u této metody není žádná cílová hodnota, která by byla společná pro všechny skupiny. Představitelé této úlohy jsou shluková analýza, genetické algoritmy a neuronové shlukování (Kohonenovy mapy) (Řezanková, 2001).
- Asociační pravidla – cílem této úlohy je vyhledat atributy, které se vyskytují hromadně a pak přiřadit nejčetnější atributy cílové hodnotě, příklady metod této úlohy jsou analýza nákupního košíku (Berry a Linoff, 2004).
- Deskripce – cílem najít strukturu dat nebo vazby mezi nimi, které jsou v těchto datech skryté (Sklenák a kol., 2001).

V analytické části této diplomové práce budou použité pouze statistické metody tradičních metod analýzy dat jako shluková, regresní a korelační analýza (viz Obrázek 9), jelikož tyto metody napomohou k dosažení cílů této práce.



Obrázek 9: Zvolené metody pro analytickou část, zdroj vlastní zpracování

3.1 Statistické metody

Statistické metody jsou základem statistiky, která je definována jako věda a větev matematiky zabývající se sběrem, klasifikací, analýzou a interpretací faktů nebo dat v numerické podobě s využitím teorie pravděpodobnosti (*Statistics*, 2015). Statistické metody pak jsou metody analýzy a reprezentace statistických dat (*Statistics*, 2015).

Statistika nabízí různé metody analýzy dat, které popisují skutečnost, analyzují vztahy mezi daty, kontrolují platnost určitých hypotéz u určitého souboru dat atd. Dále v této diplomové práci budou popsány statistické metody, které budou využité v analytické části, takovými metodami jsou shluková analýza, regresní a korelační analýzy.

3.1.1 Shluková analýza

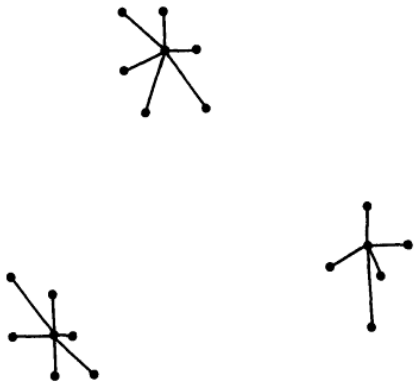
Podle Chena a kol. (2014) za shlukovou analýzu nebo analýzu clusterů je považována statistická metoda, která klasifikuje objekty podle určitých rysů a pak tyto objekty se společnými prvky spojuje do skupin (kategorií nebo clusterů) s různou úrovní podobnosti.

Podle Kaufmana a Rousseeuw (2005) existují různé typy shlukových algoritmů, ale jaký druh je třeba vybrat pro analýzu, závisí na typu a na účelu analyzovaných dat. Literatura uvádí různé druhy algoritmů analýzy clusterů. Tato diplomová práce se zabývá rozdělovacími, STC a Lingo algoritmy shlukové analýzy, které budou použité v analytické části této práce. Také jsou popsané hierarchické algoritmy shlukové analýzy pro základní přehled obecně existujících základních algoritmů shlukové analýzy. Kaufman a Rousseeuw (2005) dělí algoritmy analýzy clusterů na dvě základní skupiny: rozdělovací metody a hierarchické metody. Osiński a Weiss (2002) rozlišují další algoritmy, tzv. STC a Lingo.

3.1.1.1 Rozdělovací metody

Rozdělovací metody (partitionining methods) klasifikují data na k skupiny (angl. k-means), které by měly uspokojit následující požadavky: každá skupina by měla obsahovat objekt a každý objekt by měl patřit ke každé skupině, to znamená, že počet skupin by neměl přesáhnout počet objektů ($k \leq n$) a že dva různé clustery nemohou mít jeden společný objekt. Rozdělovací metody se používají v případech, kdy uživatel potřebuje klasifikovat objekty na k clustery, ve kterém objekty mezi sebou těsně souvisí, a z toho je i vyplývající účel těchto metod, a to najít nejvhodnější cluster s k objekty. Kaufman a Rousseeuw (2005) pak vyčleňují tři rozdělovací metody:

- Rozdělování kolem medoidů (angl. Partitioning Around Medoids), kde je cílem získání k clusteru, metoda si vybere k objekty, které se nazývají representativními objekty, příslušné clustery pak vyhledají každý objekt, který se podobá representativnímu objektu (viz Obrázek 10). v průměru destinace v jednom clusteru mezi representativními objekty a všemi objekty, které se jim podobají, by měla být minimální, v literatuře pak optimální representativní objekty se nazývají medoidy clusteru a metodu rozdělení kolem medoidů nazývají k-medoid technikou.



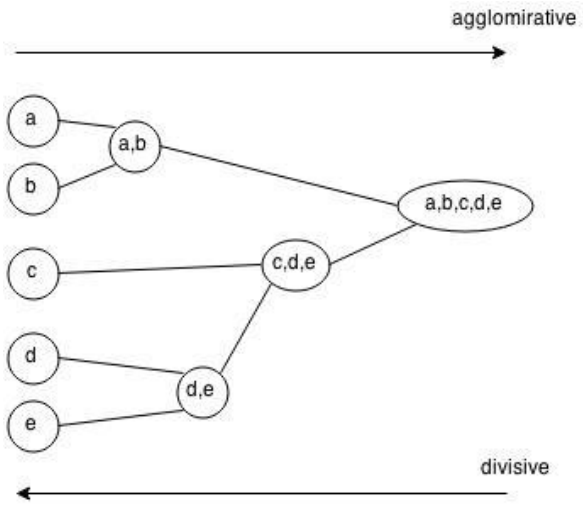
Obrázek 10: Ilustrace rozdělování kolem medoidů, zdroj Kaufman a Rousseeuw (2005)

- Obsáhlé aplikace pro analýzu clusterů (angl. Clustering Large Applications), programy v této skupině rozdělovacích metod využívají podobný algoritmus, jako i předchozí popisovaná metoda rozdělování kolem medoidů, kde se hledají *k* representativní objekty, které jsou centrálně umístěné v definovaném clusteru. Programy využívající předchozí metodu, ale na rozdíl od programů založených na této metodě nemohou analyzovat velké objemy dat, jelikož se velikost veškerých destinací u těchto programů ukládá centrálně, softwary pro analýzu velkého objemu dat ukládají velikost destinací mezi objekty rozdílně.
- Fuzzy Analysis (Měkká analýza). Tato metoda je založena na "měkkých" principech, a proto se vyhýbá "tvrdým" řešením, což znamená, že se tady definuje, do jaké míry nějaký objekt patří do clusteru.

3.1.1.2 Hierarchické metody

Druhou skupinou podle Kaufmana a Rousseeuw (2005) jsou hierarchické metody, které nevytvářejí jednotlivé rozdělení na *k* clustery, ale pracují se všemi objekty *k* dohromady, a proto u těchto metod rozdělení má k=1 (všechny objekty jsou dohromady v jednom stejném clusteru), v případě, kdy k=n, pak každý cluster byl vytvořen z jednoho elementu, k=2,3…n-1 pak zobrazením gradací clusterů.

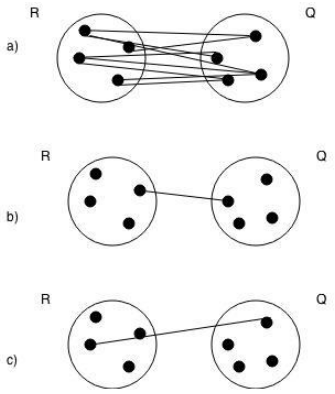
Rozpoznávají se dva druhy hierarchických metod: aglomerativní a divizní, které se liší opačným přístupem analýzy dat (viz Obrázek 11). Aglomerativní technika začíná pracovat v okamžiku, kdy se všechny objekty nacházejí zvlášť (v kroku 0 máme *n* clusterů), v dalších postupech se clustery spojují s účelem nechávání co nejméně clusterů. Divizní technika naopak začíná pracovat v okamžiku, kdy všechny objekty jsou spojené do jednoho clusteru a v následujících krocích se tyto clustery pak rozpadají na dílčí do okamžiku, kdy zůstane *n* clusterů.



Obrázek 11: Rozdíly mezi metodami aglomerativní a divizní, zdroj Kaufman a Rousseeuw (2005)

Tyto druhy hierarchických metod Kaufman a Rousseeuw (2005) pak popisují podrobně u následujících algoritmů:

- Aglomerativní hnízdění (angl. Agglomerative Nesting), tento algoritmus je zobrazen na obrázku č. 11, ve kterém objekty n-1 se nacházejí zvlášť a v dalších krocích se objekty spojují do jednoho clusteru, jeden z algoritmů této skupiny je metoda průměru párových skupin bez vah, kde se sčítá průměr délky mezi clusterem R a clusterem Q d(i, j), kde je i libovolný objekt v clusteru R a j je libovolný objekt v clusteru Q. Na obrázku č. 12a) je zobrazen d(R, Q), kde d je střední hodnota délky mezi objekty v clusterech R a Q. Na obrázku č. 12b) a 12c) je zobrazen další algoritmus této metody: nejbližší soused definovaný d(R, Q), kde d je nejkratší délka mezi clusterem R a Q, a nejdelší soused definovaný d(R, Q), kde d je nejdelší délka mezi clusterem R a Q.



Obrázek 12: Reprezentace definice rozdělení mezi clustery, a) Střední hodnota skupiny, b) Nejbližší soused, c) Nejdelší soused, zdroj Kaufman a Rousseeuw (2005)

Metoda průměrové délky se využívá pro hledání koulových clusterů, nejbližší soused vyhledává clustery s nejbližšími objekty mezi předchozími clustery a buduje protáhlé clustery. Metoda nejdelšího souseda v odlišnosti od předchozích vytváří velmi kompaktní clustery s malým diametrem.

- Divizní analýza (angl. Divisive Analysis) na rozdíl od předchozích popsaných algoritmů analyzuje opačným směrem, protože se předpokládá, že se uživatel na začátku zajímá o hlavní strukturu dat s několika velkými clustery bez detailního rozštěpení a pak už ty clustery začíná rozkládat na dílčí (viz obrázek 11).

- Monoteistické analýza (angl. Monothetic Analysis), tento typ algoritmu patří do typu divizních algoritmů, ale liší se od předchozích tím, že obsahuje data pouze s binárními proměnnými, kde 1 znamená, že nějaký atribut je přítomný, 0 pak znamená, že atribut chybí. Cílem algoritmu je rozdělit clustery na dílčí, kde nějaké clustery budou zahrnovat objekty s proměnnou 1, a jiné clustery se budou skládat s objekty s proměnnou 0. S cílem rozhodnutí, jakou proměnnou zvolit, prvním krokem se sčítá míra asociace mezi proměnnými (tento algoritmus se někdy nazývá asociační analýzou) a pak bude zvolena ta proměnná, která nejvíce representuje určitý objekt.

3.1.1.3 STC

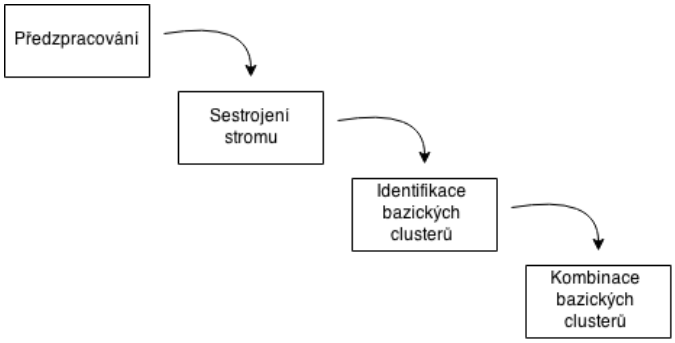
STC (angl. Suffix Tree Clustering) – je lineární shlukový algoritmus, který je založen na identifikaci slov nebo frází společných pro určitou skupinu dokumentů. STC využívá strukturu tzv. stromu z přípon (angl. sufix tree), kde clustery vznikají v uzlech tohoto stromu, který se skládá ze slov a frází výchozího dokumentu (Ilic, Spalevic, Veinovic, 2014).

Sestrojení stromu se provádí podle čtyř základních etap: předzpracování, sestrojení stromu, identifikace bazických clusterů a kombinace bazických clusterů do dalších clusterů (Ilic, Spalevic, Veinovic, 2014).

Na začátku se všechny přijaté dokumenty vystavují prvotnímu zpracování, například očištěni od interpunkce a pomocí algoritmů morfologické analýzy nebo tzv. stemmingu nebo převodu slov do jejich základního tvaru.

Dále se určuje jednotka stromu, mohou to být jak jednotlivá písmena, tak i slova a fráze. U stromů skládajících se ze slov a frází se jednotka nachází na žebru tohoto stromu, každý vrchol obsahuje odpovídající frázi. Lze ji získat po spojení všech slov nebo frází nacházejících se na žebrech od kořene stromu do určitého vrcholu.

Každý vrchol mající potomka obsahuje odkazy na dokumenty, kde jsou umístěná slova nebo fráze odpovídající vrcholu. Tyto dokumenty pak tvoří základní clustery a po kombinaci základních clusterů vznikají soubory clusterů (Kiričenko a Gerasimov, 2001). Celý proces shlukování lze znázornit na obrázku č. 13.

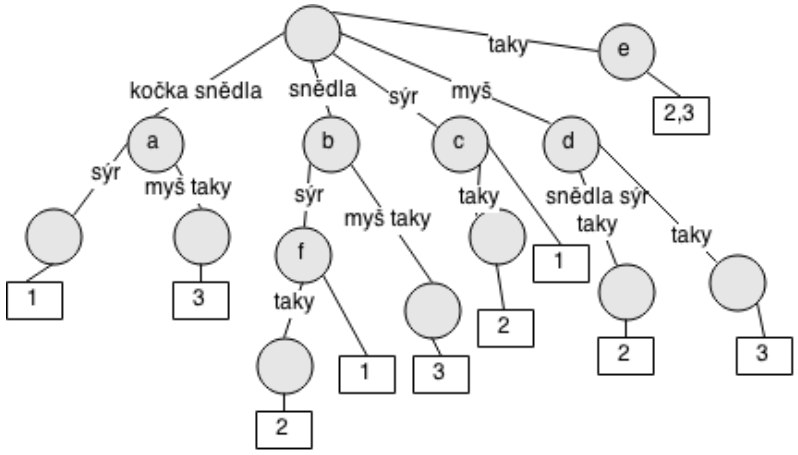


Obrázek 13: Proces shlukování pomocí algoritmu STC, zdroj Ilic, Spalevic, Veinovic (2014) a vlastní zpracování

První fáze očištění dokumentů znamená morfologické rozložení slova a převedení ho do základního tvaru anebo očištěni od určitých stopslov, které nenesou žádný informační význam (Kiričenko a Gerasimov, 2001). Tato etapa nese charakter fáze předzpracování. Spočívá ve vytvoření základního tvaru slova, získání a zpracování dat, která byla popsána v předchozí kapitole.

Druhá fáze je sestrojení stromu, které spočívá v postupném rozložení dokumentů na jednotlivá slova a fráze. Příkladem sestrojení stromu rozložení tří dokumentů s následujícím obsahem na dílčí slova a fráze (viz Obrázek 14):

1. Kočka snědla sýr
2. Myš snědla sýr taky
3. Kočka snědla myš taky



Obrázek 14: Příklad generování stromu přípon, zdroj Ilic, Spalevic, Veinovic (2014) a vlastní zpracování

Z obrázku č. 14 je vidět, že první rozložení dokumentů proběhlo na pět částí, které se pak dělí dál podle toho, v jakém smyslu a pořadí se setkávají v dokumentech. Například fráze "kočka snědla" se setkává v kontextu se "sýrem" a s "myš taky", celkový výraz "kočka snědla sýr" je uveden v prvním dokumentu, a proto na obrázku č. 14 v obdélníku je uvedeno číslo jedna (Ilic, Spalevic, Veinovic, 2014).

Třetí fází je identifikace základních clusterů. v sestrojeném stromu každý uzel v podobě kruhu s písmeny na obrázku č. 14 reprezentuje základní cluster. Potom každý cluster je definován rovnici

$$s(B) = |B| \times f(|P|) \qquad (3.1)$$

kde |B| je počet dokumentů v základním clusteru a |P| je počet slov ve frázích (Ilic, Spalevic, Veinovic, 2014).

Poslední etapou je kombinace clusterů, kde opakující se nebo blízko shodné clustery se spojují do jednoho velkého clusteru. Kombinace probíhá jen v tom případě, pokud budou splněny následující nerovnice:

$$|B_m \cap B_n| \div |B_m| > 0{,}5 \qquad (3.2)$$

$$|B_m \cap B_n| \div |B_n| > 0{,}5 \qquad (3.3)$$

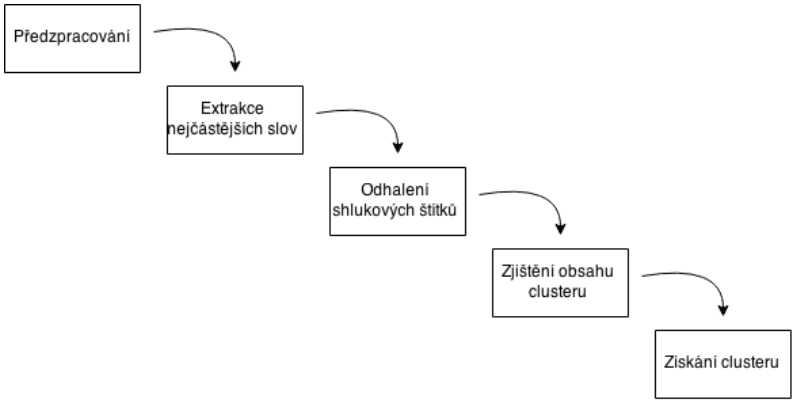
kde B_m a B_n jsou základní clustery, $|B_m|$ a $|B_n|$ rozměry těchto clusterů a $|B_m \cap B_n|$ je počet společných dokumentů mezi clustery B_m a B_n (Kiričenko a Gerasimov, 2001).

Tento algoritmus je charakteristický obsáhlými clustery, vysokou škálovatelnosti a méně vysvětlujícími názvy clusterů (Osiński a Weiss, 2002).

3.1.1.4 Lingo

Lingo je shlukovým algoritmem, který klade důraz na kvalitu tzv. shlukových štítků. Většina shlukových algoritmů na začátku identifikuje obsah a potom na jeho základě je zvolen název clusteru nebo podle Osiński a Weiss (2004) štítek (angl. labels). Tento proces ale není bezproblémový, může nastat situace, kdy obsah clusteru a jeho štítek si

nekorespondují a analytik tomu tudíž nerozumí. Za účelem odstranění tohoto problému byl vymyšlen algoritmus Lingo, který má tento proces obrácený. Na začátku se vytváří pro analytika pochopitelný štítek a pak se do toho štítku sypou určité dokumenty. První etapou je předzpracování, pak následuje vytahování nejčastějších slov nebo frází z dokumentů, které pravděpodobně vysvětlují témata obsažené v dokumentech. Potom se za pomocí sestrojení matice dokumentů prostřednictvím singulárního rozkladu (SVD[10]) zjišťuje skrytá struktura v rozmanitých vyhledávacích tématech. Nakonec se spojují skupinové popisy s vytaženými tématy a ukládají se k nim i relevantní dokumenty (Osiński a Weiss, 2004). Celý proces lze znázornit na obrázku č. 15.



Obrázek 15: Proces shlukování pomocí algoritmu Lingo, zdroj Stefanowski, Osiński, Weiss (2004) a vlastní zpracování

První fáze začíná předzpracováním, které zahrnuje vytvoření základního tvaru slova - popsáno v kapitole č. 2 a odstraněni stopslov, jelikož bez tohoto procesu předzpracování kvalita výsledků bude ovlivněna slovy, která buď nemají informační význam, anebo mají několik forem. První etapou předzpracování je filtrace textů, která odstraňuje HTML tágy a interpunkci. Pak se identifikuje jazyk každého úryvku (angl. snippet), poslední etapou je identifikace základních tvarů slov a odstranění stopslov. v tomto algoritmu stopslova slouží k identifikaci jazyka. Další metody tvorby základního tvaru slova, které se případně využívají. jsou například n-gramy popsané v kapitole č. 2. Pro tvorbu základního tvaru slov různých jazyků se využívají buď vlastní slovníky anebo obecně uznané, například pro identifikaci základních tvarů angličtiny se používá Porterův slovník (Stefanowski, Osiński, Weiss, 2004). Dále pro ukázku je uveden příklad jednoho z pravidel z Porterova slovníku podle (*Snowball*):

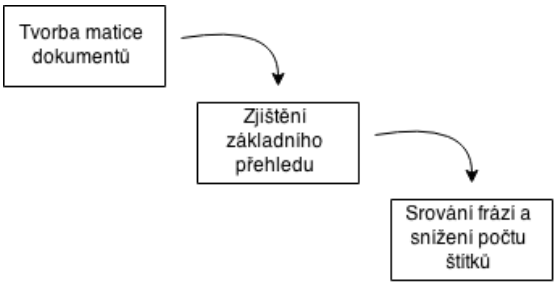
"*SSES -> SS*	*caresses -> carem*
IES -> I	*ponies ->poni*
SS -> SS	*carem -> carem*
S ->	*cats -> cat*"

Druhou etapou je extrakce nejčetnějších slov nebo frází. Podle Stefanowski, Osiński, Weiss (2004) při tvorbě jakéhokoliv dokumentu autor používá nějaká slova a jejich synonyma častěji než ostatní slova. Algoritmus Lingo, zvládá odhalit tyto slova prostřednictvím singulárního rozkladu dokumentů na matici pro identifikaci základních přehledů o dokumentech. Je zde třeba zmínit, že do nejčetnějších slov by neměla být zahrnutá stopslova a interpunkce.

Na druhou etapu navazuje i odhalení shlukových štítků. Tuto etapu lze rozdělit na tři kroky, které jsou částečně přebrané z druhé etapy. Těmito kroky jsou:

[10] SVD - z angličtiny Singular Value Decomposition (český singulární rozklad)

- tvorba matice dokumentů, pro kterou na základě frekvence slov stanovují váhy jednotlivých dokumentů a pomocí vektorového modelu (VSM[11]) se sestavuje singulární rozklad dokumentů,
- zjištění základního přehledu ze singulárního rozkladu
- a srovnání frází a snížení počtu štítků v rámci jednoho vektorového prostoru, kde se zjišťuje shoda základního přehledu s jednotlivými slovy a frázemi. Za finální fázi tohoto kroku se počítá snížení přikrývajících se štítků (viz Obrázek 16).



Obrázek 16: Proces tvorby štítků, zdroj Stefanowski, Osiński, Weiss (2004) a vlastní zpracování

Čtvrtá etapa se týká zjištění obsahu clusteru, kde prostřednictvím VSM do identifikovaných štítků z předchozí etapy se ukládá obsah z původních dokumentů.

Pátou etapou je získání clusteru, který lze obecně vyjádřit následujícím vzorečkem

$$C_{skóre} = štítky_{skóre} \times \|C\| \qquad (3.3)$$

kde $C_{skóre}$ je finální cluster, $štítky_{skóre}$ zjištěné v druhé fázi a $\|C\|$ je počet dokumentů v clusteru (Stefanowski, Osiński, Weiss, 2004).

Tento algoritmus je charakteristický objemově menšími clustery a vyšší informativností názvů clusterů na rozdíl od STC a skupinového shlukování a nízkou škálovatelností na rozdíl od STC algoritmu. Lingo pro zpracování více než zhruba 1000 dokumentů potřebuje více času a větší objem paměti (Osiński a Weiss, 2002).

3.1.2 Regresní analýza

Další použitou statistickou metodou je regresní analýza, která může nabývat lineárních a nelineárních podob (Anděl, 2003). V analytické části této diplomové práce bude použitá lineární regrese. Tato metoda může být použitá pro predikci výstupů daných lineární funkcí této predikce a regresní koeficienty mohou být vyjádřeny porovnáním predikovaných hodnot nebo porovnáním mezi průměry v datech (Tuffery, 2011). Lineární regrese lze rozdělit na dva základní druhy: jednoduchou a vícenásobnou lineární regresi.

Jednoduchá lineární regrese nám umožní spojit vektory Y s X. Vektory X a Y, jsou závislé proměnné Y a nezávislé proměnné X. Jinými slovy vektor Y z lineárního prostoru $Y \in y_1 \ldots y_n$ je vysvětlovanou proměnnou, když vektor X z lineárního prostoru $X \in x_1..x_k$ je vysvětlující vektor, který vysvětluje vektor Y, tím pádem X a Y nejsou v nezávislém vztahu a díky znalostem o vektoru X můžeme najít vektor Y, za podmínky, pokud X=x, pak ten vztah lze vyjádřit následovně: E (Y|X=x), kde E (Y) je střední hodnotou závislého vektoru a tím pádem se vyhledává střední hodnota vektoru Y za pomocí vektoru X. (Tuffery, 2011). Rovnice jednoduché lineární regrese potom podle Tuffery (2011) představená v následném tvaru:

[11] VSM – z angličtiny Vector Space Model (český vektorový model)

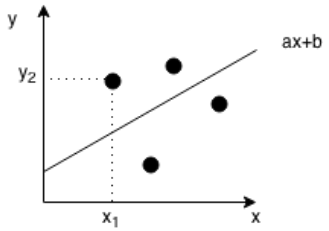
$$E(y_i) = \alpha + \beta x_i \text{, kde i = 1, …n,} \quad (3.4)$$

Po úpravách ekvivalentní rovnice jednoduché lineární regrese vypadá

$$y_i = \alpha + \beta x_i + \varepsilon_i \text{, kde } E(\varepsilon_i) = 0 \text{ pro každé i = 1, …n.} \quad (3.5)$$

Vektor ε_i je tzv. rezidua, které znamená velikost přípustné chyby při odhadu. Vektor ε_i není možné pozorovat, ukazuje na nepřesnosti vektoru Y (Anděl, 2003), které mohou být vyvolané například tím, že je měřený několikrát během různých dob (Tuffery, 2011).

Grafický vzoreček lineární regrese můžeme zobrazit na grafu č. 2, kde x_i je měřený v závislosti na vývoje y_i. Jednotlivé hodnoty jsou zobrazené tečkami, pak průměrná závislost hodnot se rovná přímce $ax+b$. Rozdíl mezi tečkou a přímkou je pak reziduum, které po splnění určitých podmínek lze vyhledat pomocí metody nejmenších čtverců.



Graf 2: Měření y v závislosti na x, zdroj vlastní zpracování

Vícenásobná lineární regrese pak obsahuje více než jeden vysvětlující vektor X, jelikož nějaká data nebo vektor Y nelze vysvětlit pomocí pouze jednoho vektoru. Rovnice vícenásobné lineární regrese pak vypadá následovně podle Tuffery (2011):

$$y = \beta_0 + \beta_1 X_1 + \ ... + \beta_p X_p + \varepsilon \quad (3.6)$$

Pomocí lineární regrese a korelační analýzy popsané v další subkapitole této diplomové práce se budou vyhledávat závislosti určitých hodnot uvedených v analytické části.

3.1.3 Korelační analýza

Spolu s regresní analýzou bude v analytické části využitá analýza korelace, která determinuje míru závislosti mezi pozorovatelnými hodnotami prostřednictvím tzv. korelačního koeficientu. Po výpočtu korelačního koeficientu se pak vyhodnocuje míra závislosti mezi vysvětlujícími proměnnými. Obecně se rozlišují tři základní druhy korelačních koeficientů: párový, parciální a vícenásobný. Párový a parciální koeficienty mohou nabývat hodnot v rozmezí <-1,+1>, v případě korelačního koeficientu <-1, 0) pak proměnné jsou v nepřímé závislosti, což znamená, že pokud roste jedna proměnná, pak druhá klesá. Naopak když korelační koeficient nabývá hodnot (0, +1>, pak proměnné jsou v přímé lineární závislosti, což znamená, že pokud roste jedna proměnná, druhá roste také. Pokud se korelační koeficient rovná nule, pak proměnné nejsou závislé. Párové korelační koeficienty pak vyjadřují vztah mezi dvěma proměnnými, parciální měří závislost mezi proměnnými za předpokladu, že vliv jiných proměnných je konstantní. Vícenásobný korelační koeficient ukazuje míru závislosti jedné proměnné s lineární kombinací více proměnných než jedna (*Korelace*).

Další část práce se zabývá problématikou kryptoměn, jelikož analytická část je postavená na analýze nestrukturovaných dat v oblasti kryptoměn.

4 Kryptoměna

V dnešní době hlavním prostředkem směny jsou peníze ve fyzické nebo digitální podobě, které držitel využívá pro různé účely. V posledních přibližně pěti letech se objevila mezi různými prostředky pro směnu nová forma prostředků a tou je kryptoměna.

Cílem této kapitoly je vysvětlit pojem „kryptoměna“ a popsat její princip fungování. Pro bližší představu principu fungování kryptoměny bude využit popis kryptoměny Bitcoin.

4.1 Definice kryptoměny

Kryptoměna je podle Cocco, Concas, Marchesi (2014) digitální měnou, která je založená na zásadách kryptografie a je alternativní k pravým měnám. Většina kryptoměn je navržená tak, aby se postupně se vytvořila nová měna s cílem zabránit inflačním tlakům, jako je to u pravých měn. Evropská centrální banka ECB (2012) definuje kryptoměnu jako neregulovanou digitální měnu, kterou vydávají a většinou i regulují tvůrci této měny, a která je přijatelná a použitelná mezi členy nějaké specifické virtuální skupiny.

První kryptoměnou a v dnešní době podle tržní kapitalizace také nejrozšířenější je Bitcoin. Pro pochopení principu fungování kryptoměny a jejích odlišných charakteristik od pravých měn bude podrobněji vysvětlen na definici Bitcoin.

4.1.1 Bitcoin

Bitcoin (BTC) (z angl. bit – jednotka informace, coin – mince) (Bitcoin Security, 2015) – je podle Nakamoto (2009) digitální nezávislou a důvěryhodnou měnou nebo elektronickou mincí, která funguje v systému peer-to-peer, což umožňuje provádět anonymní online platby za pomocí kryptografických pravidel z jedné části světa do druhé bez využití zprostředkovatelských služeb bank.

Decentralizace a nezávislost

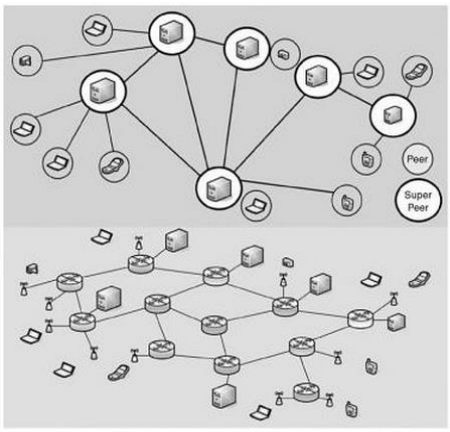
To znamená, že neexistuje žádná centrální instituce, která dodává likviditu na trh a reguluje vývoj této měny, tím pádem žádný subjekt v podobě centrální banky neexistuje a žádné techniky vlivu na kurz, jako řízený floating, currency board, crawling peg (nebo crawling bads) atd. nejsou. Také tady lze říct, že neexistuje žádný centrální subjekt, který by prováděl konvenční nebo nekonvenční měnovou politiku, která se stará především o cenovou stabilitu prostřednictvím v případě kryptoměn regulace měnových kursů. Neexistuje žádný centrální subjekt, který by na systém směny Bitcoinů dohlížel a pomáhal určitým subjektům v případě velké ztráty.

Takový decentralizační systém kryptoměn má své výhody a nevýhody. Výhodou je především to, že emise například Bitcoinů je omezená do výše 21 000 000 BTC (Bitcoin Security, 2015), tím pádem se v době dosažení limitu předpokládá, že BTC se vytvářet nebudou a uživatelé si budou vyměňovat vlastní BTC. v případě centralizačního systému může centrální banka dodat likviditu (v podobě domácí měny) na trh a tím roste inflace tohoto státu. Případně lze takovýmto způsobem likviditu snižovat. Naopak nevýhodou decentralizačního systému je to, že neexistuje žádný centrální subjekt, který tuto směnu reguluje nebo dohlíží, jak bude popsáno dál, směna se provádí anonymně a tím pádem se jí mohou zúčastnit i nelegální obchody. Existují ale i kryptoměny, které žádné limitní omezení v objemu nemají, představitelem je kryptoměna Peercoin.

Peer-to-peer

S pojmem decentralizace je spojený pojem peer-to-peer, který podle Buforda, Yu, Lua (2009) znamená decentralizační síťová architektura klient-klient, ve které komunikace probíhá mezi jednotlivými uživateli sítě bez zprostředkujícího serveru (viz Obrázek 17). Klienti (peers) v P2P aplikaci komunikují s ostatními klienty přes internet nebo přes jinou síť prostřednictvím zpráv díky protokolům, které mají určité specifičnosti. Za prvé jsou tyto protokoly výsledkem aplikační vrstvy množiny síťových protokolů, za druhé každý klient má unikátní identifikační číslo, kterým je klientský ID anebo klientská adresa, za třetí zprávy v architektuře P2P jsou často definovány stejně a za čtvrté, protokol určuje směr určitých typů zpráv.

Jako příklady peer-to-peer architektury slouží původní verze Napster aplikace, Gnutella, FastTrack a BitTorrent.



Obrázek 17: P2P, zdroj Buford, Yu, Lua (2009)

Anonymnost

Platby se provádí přímo, bez nějaké zprostředkující finanční instituce. i když převody prostřednictvím bank lze zrušit, ale zase výjimečně (v případě chyby na straně banky), zrušit transakce v systému směny kryptoměn nelze. Informace o uskutečněné transakci dostává každý klient sítě. Tzv. adresa klienta nemá žádné osobní údaje o klientovi, a proto se nazývají anonymními. Adresa je v textovém formátu a skládá se z 34 symbolů z písmen latinské abecedy a čísel, která nazýváme hash (princip hash bude probrán později). Klient může vytvářet několik adres a tím se mu vygeneruje několik klíčů. Kryptoměny jsou pak uloženy v tzv. peněžence a mohou být zaslány na jinou libovolnou adresu v tomto systému.

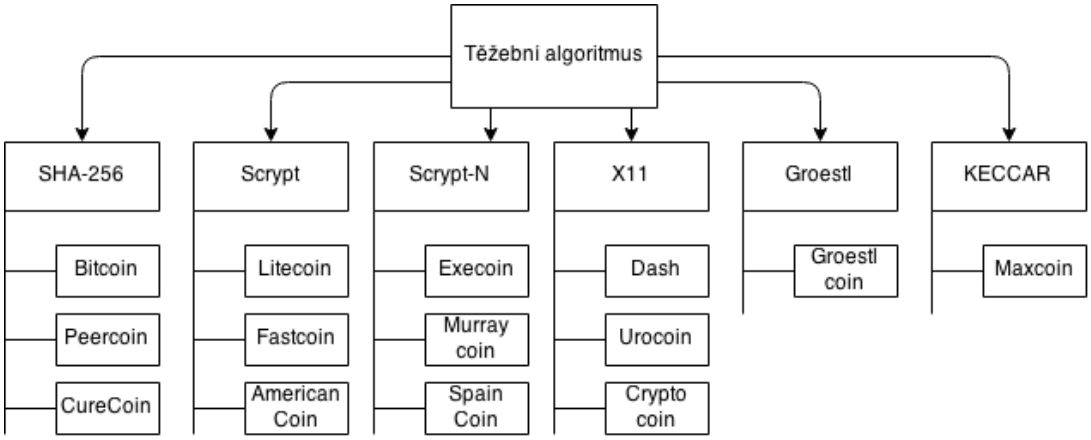
Bezpečnost

Základem kryptoměny je kryptografie, která znamená: "*z řeckého kryptós (skrytý) a lógos (slovo) — je vědní disciplína, zabývající se bezpečnou a tajnou komunikací*" (Bulant, 2001, s. 1). Pomocí technik kryptografie odesilatel zašifruje nějakou zprávu nebo převede ji do nečitelné podoby, do tzv. hash a rozšiřovat ji může pouze "správný" příjemce. Kryptografie pak zabezpečuje možnosti proti utracení peněz z cizího účtu nebo poškození sítě bloků. Kryptografie je založená na třech základních principech:

- integrita – příjemce může zkontrolovat, jestli zpráva nebyla na cestě změněná,
- autenticita – identita komunikačních stran je prokazatelná,
- nepopiratelnost – žádná ze stran nemůže později popřít provedené akce, například odesilatel nemůže tvrdit, že není tvůrcem elektronického podpisu (Bulant, 2001).

Kryptoměny využívají metodu asymetrické kryptografie, která je založená na tvorbě dvou klíčů: veřejného a privátního, přičemž veřejný klíč je znám všem uživatelům sítě a privátní pouze majiteli tohoto klíče. Při odesílání nějaké částky přikládá odesilatel vlastní privátní klíč pro autenticitu a veřejný klíč příjemce, který pak po přijetí částky zhodnotí, jestli částka patří právě jemu.

Pro těžení kryptoměn se využívají různé šifrovací algoritmy. Například Bitcoin používá šifrování SHA-256. Některé alternativní kryptoměny byly odvozené od open source Bitcoinového kódu, některé ale byly vytvořené za pomoci jiných šifrovacích algoritmů než Bitcoin, příkladem je Litecoin, Dogecoin, které byly vytvořené na základě Skrypt šifrovacího algoritmu (*Bitmakler*). Zástupci některých těžebních algoritmů jsou zobrazeni na obrázku č. 18.



Obrázek 18: Vybrané těžební algoritmy, zdroj Bitmakler a vlastní zpracování

Pro zabezpečení hlavní knihy s historií transakcí, kterou u většiny metod nelze změnit bez přepracování celé historii operací, dále pro těžbu nebo tvorbu nových jednotek měny a zabránění duplicitních transakcí se využívají různé kryptografické metody:

- Metoda **Proof of Work** v překladu z angličtiny pak znamená důkaz práce a ukazuje na větší pravděpodobnost potvrzení jednotlivých bloků, které mají větší výkon výpočetní techniky (Nakomoto, 2009). Představitelem je kryptoměna Bitcoin.
- Alternativní metodou Proof of Work je **Proof of Stake**, z anglického je důkaz na kapitálovém podílu. Podle této metody se bloky generují na základě množství kryptoměn a času, jak dlouho se nějaká kryptoměna drží v peněžence u jednoho uživatele. Představitelem této metody je například BlackCoin (Vasin, 2014).
- Využívá se i **kombinace** metod **Proof of work** a **Proof of Stake**, představitelem hybridního systému metod je například NovaCoin (Vasin, 2014).
- Jinou známou metodou je **delegovaný Proof Of Stake** (angl. Delegacy Proof of Stake). Princip této metody je hlasování v režimu reálného času, kde se vybírá delegát na základě jeho reputace. Vybírat může každý majitel určité kryptoměny, počet zvolených delegátů není stanoven, ale často se volí 101 delegátů. Pravomoc delegátů Potvrzuje tvůrce systému kryptoměny. Delegáti pak mají právo tvořit bloky pro následné zapojení do sítě (angl. blockchain) a bránit vstupu cizích osob do procesu. Delegáti mezi sebou před zvolením konkurují způsobem snížení vlastní odměny za vlastní práci. Představitelem této metody je BitShares (Rait, 2015).
- Používá se další metoda s názvem **Proof of Burn**, z anglického je důkaz "spálením", což znamená, že prostřednictvím jedné kryptoměny dochází ke "spálení" nebo ke zničení druhé kryptoměny. Uživatel jedné kryptoměny pošle na neexistující adresu své kryptoměny, které se pak zničí a tomu uživateli se pak vrací jiná kryptoměna, která je založená na této metodě. Kryptoměny založené na této metodě pracují s kryptoměnami,

které jsou založené na metodě Proof of Work (Tristan Winters, 2014). Příkladem této metody je kryptoměna Counterparty.

- **Proof of HDD Capacity**, z anglického je důkaz pevným diskem. Princip této metody je založen na tom, že těžba probíhá na pevném disku uživatele. Těžební program vytvoří tzv. „ploty" s určitým objemem dat, které se pak ukládají na pevný disk uživatele. Čím více místa na pevném disku uživatel má, tím větší je šance rychle vygenerovat blok (*Burst*). Představitelem této metody je kryptoměna Burst.
- **Proof of Importance**, z angličtiny se překládá jako důkaz důležitosti. Princip této metody je založen na vyhodnocení informace o tom, jací uživatelé jsou pro systém důležití. Uživatele, kteří aktivně obchodují v určité kryptoměně založené na této metodě, tím podporují chod systému této kryptoměny a lze je označit též za důležité. Také reputace uživatele závisí na objemu vlastněné kryptoměny. Čím více kryptoměny této metody uživatel má, tím vyšší má reputaci (Beikverdi, 2015) Představitelem této metody je kryptoměna NEM (plný název je New Economy Movement).

Důvěryhodná

Důležitou vlastností prostředků směn je jejich důvěryhodnost, klienti musí věřit v to, co kupují, nebo co přijímají. Nyní takovýto specifický nástroj nemáme. Za důvěryhodný nástroj lze označit jakékoliv aktivum, které je všeobecně přijímáno při placení za zboží a služby nebo při úhradě dluhů. Můžeme tedy vycházet z podstaty těchto aktiv a funkce peněz. Kryptoměny lze dnes využit jako peníze pro platbu za zboží nebo služby. Za prvé, podle Revendy a kol. (2005) peníze mají plnit funkci prostředku směny, jimi by bylo možné, snadné a relativně levné platit za zboží a služby. Druhou funkcí je to, že by peníze měly být jako účetní jednotka a měly by vyjadřovat cenu zboží a služeb, kapitálu, práce, zahraničních měn atd. Poslední funkcí je schopnost zachování hodnoty, peníze by měly uchovat vlastní hodnotu, která se vyvíjí od kupní síly v čase, a ukazovat celkové bohatství jednotlivého klienta.

Z výše uvedeného lze předpokládat, že definice peněz a jejich funkce se částečně hodí i pro kryptoměny, protože i když si lze vyměňovat kryptoměnu za jiné pravé měny nebo platit kryptoměnami za zboží v určitých obchodech, ale zase nejsou všeobecně přijímané jako peníze. Kryptoměna závisí pouze na důvěře uživatelů, bez ní by kryptoměny neexistovaly, a proto lze říct, že kryptoměna je důvěryhodná a plní všechny výše uvedené funkce peněz.

4.2 Peněženka a klíče

Aby uživatel mohl obchodovat s kryptoměnami nebo koupit nějaké zboží za kryptoměny, musí vytvořit tzv. peněženku, ve které se budou ukládat digitální měny. Lze rozlišit tři druhy peněženek podle typu úložiště klíčů a prováděných transakcí:

- stažený na koncové zařízení uživatele (tlustý klient),
- transakce se ukládají na serveru třetí strany a klíče, o kterých řekneme dále, jsou chráněny uživatelem (tenký klient),
- anebo uživatel může využít cloudových služeb, kde historie transakcí a klíče jsou chráněné na serveru třetích stran (web klient) (*Wallet*).

Peněženka je jako virtuální bankovní účet bez specifických vlastností vztažených k bankám a jejím službám, který umožňuje uživatelům provádět interpersonální, obchodní transakce nebo ukládat peníze (Lee, 2014).

Po stažení peněženky na koncové zařízení se vytvoří první veřejný klíč nebo jinými slovy adresa rozměrem 512 bitů. Uživatel pak může provádět různé transakce, kde se nějaké množství Bitcoinů předává mezi Bitcoin-peněženkami, informace o těchto transakcí se zapisuje do sítě

blocků (angl. blockchain), které ukládají historii všech provedených transakcí od vzniku Bitcoinu. Peněženky obsahují vygenerovaný privátní klíč o rozměru 256 bitů pro podepsání transakcí, který je jedinečným identifikátorem peněženky. Při ztracení privátního klíče dojde ke ztracení všech kryptoměn v peněžence. Pak lze tvrdit, že peněženka je souborem privátního a veřejného klíče, kde se privátní klíč zobrazuje pouze vlastníkovi peněženky a veřejný klíč je dostupný všem uživatelům sítě. Pro tvorbu nových klíčů například Bitcoin využívá algoritmus ECDSA s elementy SECP256k1 (Bitcoin, 2015).

Příklady klíčů:

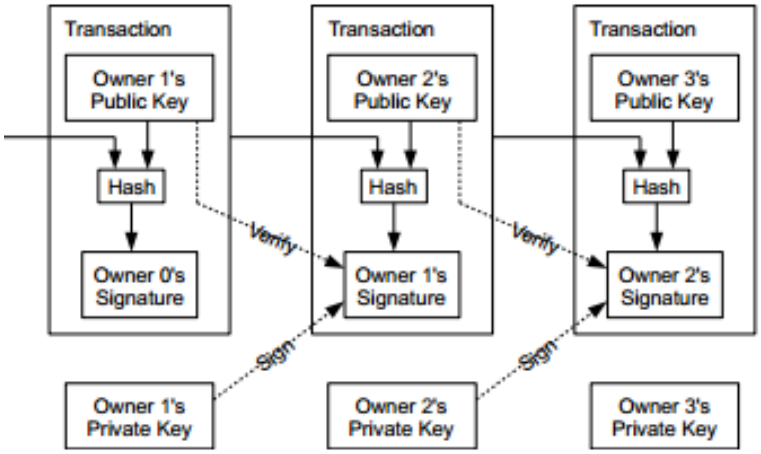
"*1F3sAm6ZtwLAUnj7d38pGFxtP3RVEvtsbV – Veřejný klíč*

L4rK1yDtCWekvXuE6oXD9jCYfFNV2cWRpVuPLBcCU2z8TrisoyY1 – Privátní klíč" (Bitcoin, 2015).

4.3 Transakce

Pro přiblížení čtenáře k problematice kryptoměn, další subkapitoly budou popsány na příkladě kryptoměny Bitcoin s metodou generace bloků PoW. Podle Nakamoto (2009) vlastník Bitcoinů posílá jinému uživateli systému určitou částku Bitcoinů, tím se vytvoří nová transakce, která obsahuje hash a privátní klíč vlastníka předchozí transakce a veřejný klíč uživatele, kterému jsou Bitcoiny zasílané (viz Obrázek 19). Tato informace je odeslána do P2P Bitcoinové sítě, kde jiní uživatelé před tím, než přijali tuto transakci, kontrolují veřejný klíč, jestli částka patří jim, posílanou částku dostává ten příjemce, jemuž patří veřejný klíč. Transakce se pak odesílá do bloku a následně, aby tento blok byl zapsán do sítě bloků, očekává se potvrzení transakce. Díky těžbě, o které bude zmínka později, každých zhruba deset minut se do sítě bloků přidává nový blok.

Z důvodu absence centrálního subjektu může dojít k duplicitním transakcím. S cílem zamezení dvojitých transakcí byla vytvořena jednotná historie, která je dostupná každému uživateli sítě a která obsahuje všechny transakce provedené v historickém pořadí. V síti Bitcoinů se využívá pro jeden z účelů výše uvedená kryptografická metoda Proof of Work.



Obrázek 19: Transakce, zdroj Nakamoto (2009)

4.4 Blok a síť bloků

Blok se skládá z nedávno provedených transkakcí čekajících na potvrzení, obsahuje hash současného a předchozího bloku, po těžbě a po zapojení bloku do sítě bloku přidává se hash i dalšího bloku, pak obsahuje hash, který je kořenem všech transakcí (viz Obrázek 20).

Hashes	
Hash	00000000000000000121ba61c07bb9da501e0ee244d04ead52a34d9a702561eb5
Previous Block	00000000000000000012217d4273b0f1bd55502f0621c7e90ec0bac9dc94dbd0d3
Next Block(s)	
Merkle Root	8d964d88521b6dec95e2491e2d5b79c1b80ce18065964e8ca3c300774815a7c6

Obrázek 20: Hash určitého bloku, zdroj Blochchain Info

Dále lze v bloku vyhledat informaci o čísle bloku v síti bloků, na tento okamžik máme číslo bloku Bitcoinů 348549, který zatím nemá žádný nalezený následující blok. Blok také obsahuje počet transakcí, které byly provedené za posledních zhruba deset minut a které byly potvrzené těžaři. Počet výstupů a celková částka v BTC nám ukazují celkovou částku BTC, která byla převedená v tomto bloku. Také lze najít informaci o poplatcích z transakcí, které se počítají rozdílem mezi vstupem a výstupem (viz Obrázek 21). Blok také obsahuje podrobnější informace o provedených transakcích a to, od koho a komu byla odeslaná určitá částka, číslo transakcí, velikost dat určité transakce a poplatky.

Block #348549

Summary	
Number Of Transactions	368
Output Total	2,085.80985324 BTC
Estimated Transaction Volume	282.52332228 BTC
Transaction Fees	0.05194463 BTC
Height	348549 (Main Chain)
Timestamp	2015-03-21 10:47:54

Obrázek 21: Informace o bloku č. 348549, zdroj Blochchain Info

V tuto chvíli těžaři hledají hash, který se bude hodit ke každé transakci, a v okamžik, jak bude tento hash nalezen, blok se zapisuje do sítě bloků (angl. “blockchain”). V síti bloků lze nalézt informaci o všech provedených transakcích za celou dobu existence Bitcoinů. Síť bloků obsahuje číslo bloků v síti, počet transakcí, celkovou obchodovanou částku u jednotlivých bloků, jakým subjektem ten blok byl vynalezen a velikost dat bloku v kB (viz Obrázek 22).

Height	Age	Transactions	Total Sent	Relayed By	Size (kB)
348549	4 minutes	368	2,085.81 BTC	BW.COM	273.55
348548	12 minutes	112	615.45 BTC	KnCMiner	112.91
348547	2 minutes	128	961.98 BTC	148.251.6.18	96.44
348546	12 minutes	1038	18,832.50 BTC	BTCChina Pool	487.49
348545	31 minutes	72	406.69 BTC	AntPool	45.59
348544	31 minutes	394	2,623.11 BTC	F2Pool	474.29

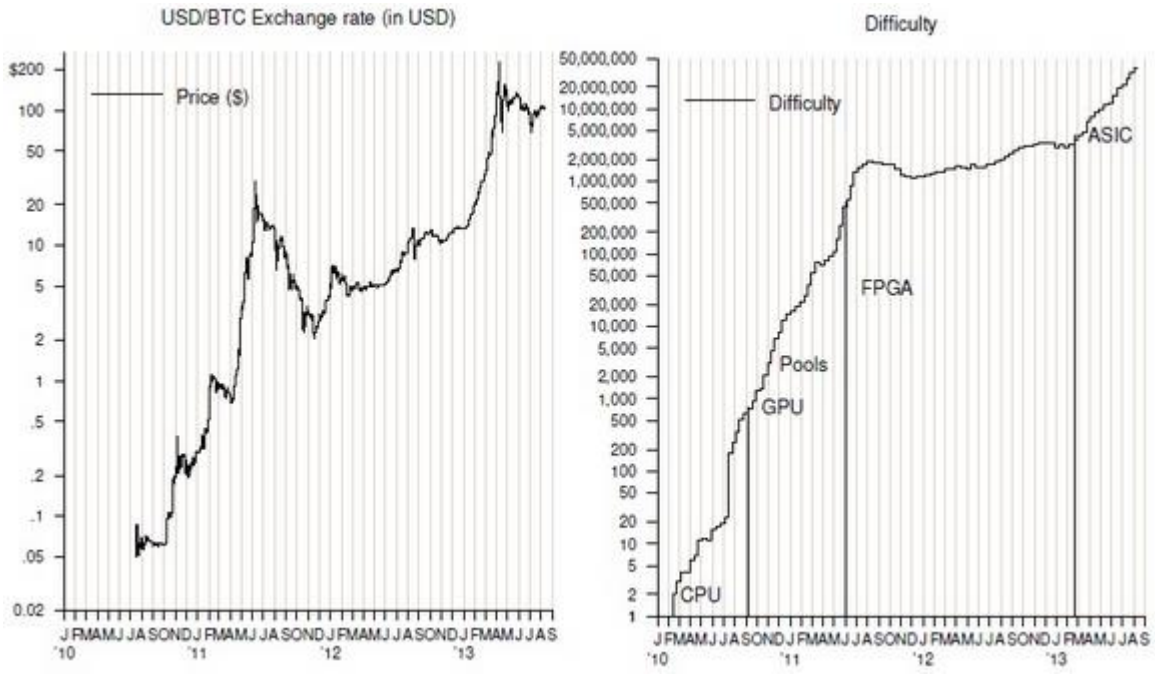
Obrázek 22:Sít bloků, zdroj Blochchain Info

4.5 Těžba

Na rozdíl od pravé měny, která je tvořená centrální bankou nebo jinou pověřenou institucí pro tvorbu elektronických peněz, kryptoměnu nevytváří žádná centrální instituce, vytváří ji tzv. těžaři, přičemž těžařem se může stát libovolný uživatel sítě.

Těžba je rozdělovací systém, který se používá pro potvrzení čekajících transakcí na zařazení do sítě bloků. Těžba zabezpečuje chronologické pořadí transakcí v síti bloků. Aby byla transakce potvrzená, musí být zabalená do bloku, který je striktně zabezpečen kryptografickými požadavky. Tato pravidla neumožňují ani změnit předchozí bloky, ani neumožní odvolat transakci.

Těžba je také proces, ve kterém se pomocí výpočetní techniky provádí složité matematické výpočty s cílem potvrzení transakcí a podpory sítě Bitcoin. Za provedenou práci těžaři dostávají pak odměnu v podobě nových Bitcoinů. Těžba je specifický a konkurenční trh, na kterém se odměna dělí počtem provedených výpočtů (*Vocabulary*). První těžba Bitcoinů byla na dnešní okamžik poměrně lehká, jelikož úroveň náročnosti byla nízká. Uživateli postačoval procesor obyčejného stolního počítače pro zvládnutí matematické úlohy najít blok a vytěžit na ten okamžik 50 BTC. Po nalezení každých 210 000 bloků se odměna zmenšuje dvakrát, předpokládá se, že po nalezení těžby všech Bitcoinů systém bude podporován poplatky od uživatelů za potvrzení transakcí a odeslání do sítě bloků (Bitcoin Security, 2015). Jelikož v době vzniku Bitcoinů postačoval pouze procesor stolního počítače, ostatní uživatelé začali rychle a na rozdíl od dnešní doby snadno těžit. Aby všechny bloky s odměnou v Bitcoiny nebyly nalezené, jinými slovy, aby všechny existující Bitcoiny v síti nebyly vytěžené hned a cena BTC se udržovala na nějaké úrovni, náročnost těžby postupně stoupala. Čím složitější je najít Bitcoin, tím vyšší je jeho cena a tím výkonnější výpočetní techniku těžař potřebuje (viz Obrázek 23). Na levém grafu na obrázku č. 23 je zobrazen vývoj kurzu a na pravém grafu je vývoj náročnosti od počátku roku 2010 do konce roku 2013.



Obrázek 23: Vývoj kurzu USD/BTC a náročnosti těžby, zdroj Taylor (2013)

Z historického hlediska jsou známé čtyři druhy těžební výpočetní techniky:

- CPU – standardní procesor. Jedná se o nejstarší těžební výpočetní techniku a podle výkonu je nejslabší z dalších uvedených.
- GPU – počítačová grafická karta je druhou generací těžební výpočetní techniky. GPU je výkonnější než CPU, je ekvivalentem podle *Heida* zhruba 44.64 CPU.
- FPGA – programovatelné hradlové pole, které má větší výkon než CPU a GPU a hlavně na rozdíl od GPU spotřebovává méně elektřiny. Typická 600 MH/s GPU spotřebovává zhruba 400w elektřiny, když 826 MH/s FPGA spotřebovává zhruba 80w (*Weusecoins*)
- ASIC – integrovaný obvod pro specifickou činnost. To je představitel poslední generace těžební výpočetní techniky a na rozdíl od všech předchozích je nejvýkonnější za použití menšího objemu elektřiny (*Heid*).

V dnešní době pro těžbu Bitcoinů už nepostačuje výkon grafických karet a ani programových hradlových polí, na směnu přišel ASIC. Pro těžbu kryptoměn s menší náročností postačují i menší výkony procesoru, využití ASIC není třeba, jelikož by se nákup drahé výpočetní techniky měl vyplatit z důvodu menší náročnosti výpočtu hash, menších cen a menšího zájmu o jiné kryptoměny a tím i zdůvodnění potřeby méně výkonnější techniky. Dále těžař rozhoduje, jestli bude těžit, a to na základě ceny energie a rychlosti a ceny internetu. Tím pádem s rostoucím zájmem o těžbu Bitcoinů roste náročnost těžby, potom rostou i nároky na výpočetní techniku, na energii a na podmínky poskytovaného internetu. v okamžiku, kdy náklady na těžbu budou vyšší než odměna za těžbu, těžit už ztrácí smysl.

4.6 Shrnutí

V dnešní době se staly velmi populární kryptoměny, které jsou zcela postaveny na důvěře uživatelů, jsou anonymní v rámci adres, kam se peníze posílají, decentralizované a zabezpečené kryptografií. Pro obchodování kryptoměnami je třeba založit peněženku, která je složená z privátního a veřejného klíče. Pro odeslání kryptoměn je třeba znát veřejný klíč přijímajícího. Odeslaná částka a adresa přijímajícího je dostupná všem uživatelům sítě. Přijímající ale po přijetí částky nějakou dobu obchodovat s touto částkou nemůže, jelikož tato transakce je zapsána do bloku a blok čeká na potvrzení. Způsob potvrzení transakce a nalezení

nového bloku závisí na určité metodě generace bloků. v případě Proof-of-Work těžaři v tuto chvíli vyhledávají správné chronologické pořadí všech nepotvrzených transakcí prostřednictvím výpočtu složitých matematických úloh. Rychlost výpočtu záleží na nastaveném systému náročnosti úlohy, která je daná počtem těžařů a výkonností jejich výpočetní techniky, náročnost pak nastavená tak, aby úloha byla řešená za určitou dobu. Ten těžař, který odhadne správné chronologické pořadí, dostává odměnu, která se pak dává do oběhu. Tím je nalezen nový blok, který se zapíše do sítě bloků ve správném chronologickém pořadí, transakce jsou potvrzené, přijímající může používat přijatou částku.

5 Návrh řešení

Tato kapitola se zabývá návrhem řešení analytické části této diplomové práce. Cílem této kapitoly je vybrat a popsat analytický nástroj pro získání, zpracování a analýzu dat. Dalšími cíli jsou zpracování architektury řešení, vybrat vhodné datové zdroje v souladu se základními cíli této diplomové práce a shrnout úkoly pro analytickou část.

5.1 Technologie

Pro splnění základního cíle této diplomové práce, přesněji pro analýzu nestrukturovaných dat se zaměřením na vývoj témat jednotlivých kryptoměn v čase bylo zapotřebí vybrat vhodný nástroj, který umí data ze zvoleného zdroje získat, zpracovat a pak zanalyzovat. Pro tento účel byly zvolené nástroje Crawler4j, Elasticsearch verze 1.5.0., Kibana 4.0.2 a Carrot2. Pro regresní analýzu využívám MS Excel.

5.1.1 Crawler4j

Pro účely získání dat jsem zvolila open source konektor s názvem Crawler4j. Crawler4j je Java knihovnou s Apache License 2.0, která poskytuje jednoduché rozhraní pro procházení webových stránek a je schopná stáhnout 200 stránek za sekundu s Quad core PC kabelovým připojením (*Java-Source.net*), (*Yasserg*). Pro účely této práce bude zapotřebí stáhnout základní crawler (angl. Basic Crawler) s tzv. kontrolérem (angl. Controller). Upravení kódu crawleru bude popsáno v této práci dál.

5.1.2 Elasticsearch

Elasticsearch je distribuovaným open source (s Apache 2.0 licenci) vyhledávačem a analytickým nástrojem, který nabízí větší škálovatelnost, rychlost, spolehlivost, easy management ve vyhledávání a analýze strukturovaných a nestrukturovaných dat a který umožňuje data vyhledávat a analyzovat v reálném čase. Elasticsearch byl vyvíjen na základě Apache Lucene v programovacím jazyce Java (*Elastic*), což umožňuje tento nástroj nainstalovat na několika operačních systémech. Nástroj využívá API a většina operací může být provedená za použití RESTful API, kde samotné dokumenty ve formátu JSON jsou zasílány prostřednictvím protokolu HTTP přes webové rozhraní nebo jiné instrumenty nebo knihovny, které podporují odesílání dotazů a obdržení odpovědí přes HTTP, například CURL. v okamžik indexování JSON dokumentů, Elasticsearch automaticky rozpozná datovou strukturu požadovaného dokumentu, vytvoří index a umožní data pak vyhledávat (*Elastic*).

Distribuovaný systém

Elasticsearch je škálovatelným systémem, který dovoluje pracovat s různými objemy dat. v okamžiku, kdy server, na kterém Elasticsearch běží, je zatížen, je snadné přidat více serverů (angl. nodes), tím vzniklé clustery budou automaticky roztříděné mezi dostupnými servery (angl. nodes). (*Elastic*)

Vysoká dostupnost a conflict management

Clustery, které byly vytvořené v rámci Elasticsearch, jsou pružné. Ve chvíli, kdy se objeví nějaká chyba nebo konflikt u určitých uzlů, automaticky přetřídí data tak, aby nebyla ztracená, ale bezpečně uložená a dostupná.

Fulltextové vyhledávání

Elasticsearch je vyvíjen na základě Apache Lucene, aby dokázal zajistit fulltextové vyhledávání dostupné pro jakýkoliv open source software. Tento pro vývojáře přívětivý dotazovací API podporuje více jazyků vyhledávání, geolokaci, návrhy "mysleli jste na", automatické dokončení a vyhledávání podobných záznamů podle úryvků (*Elastic*).

Analyzer

Důležitým prvkem pro analýzu dat je předzpracování. Zvolený nástroj obsahuje základní tzv. analyzer, který dokáže zpracovat většinu západních jazyků. Analyzer Elasticsearch obsahuje prvky a postupuje v tom samém pořadí jako i body uvedené níže (viz Obrázek 24):



Obrázek 24: Proces analyzeru Elasticsearch, zdroj Elastic a vlastní zpracování

- typický filtr, který očišťuje dokumenty od HTML "tagů" před tím, než následuje rozdělení textu na mezery, které je uvedeno v dalším bodu,
- standardní tzv. tokenizer, který rozděluje vstupní text na mezery a odstraňuje interpunkci, existují i další druhy tokenizerů, například tokenizer pro rozdělení textu prostřednictvím regulérních výrazů atd.
- filtr symbolů (angl. token filter), který pak po předchozím kroku provádí určité činnosti:
  - modifikuje symboly (angl. tokens), například převod slov do malých písmen, který provádí tzv. filtr pro malá písmena (angl. lowercase filter),
  - odstraňuje symboly, tzv. stop filtr (angl. stop filter), který odstraňuje stopslova, v angličtině to jsou například *a, the, and, is* atd.
  - anebo přidává symboly, například synonyma,
  - stemmer token filtr zahrnuje převod slova do základního tvaru, pro angličtinu Elasticsearch se využívá Porterův algoritmus,
  - filtr pro odstranění diakritiky,
  - další možnosti lze nakonfigurovat podle potřeb (Kuć a Rogoziński, 2014).

Do Elasticsearch bude nainstalován plugin head, který je webovým front-end pro prohlížení clusterů instalovaných do Elasticsearch a pro vzájemnou komunikaci s Elasticsearch (*Elasticsearch-head*).

5.1.3 Kibana

Kibana je open source (Apache Licence) platformou pro analýzu a vizualizaci, která byla vyvíjena pro účely spolupráce s Elasticsearch. Kibana se využívá pro vyhledávání, nahlížení, provádění určitých analýz a jinou práci s daty, která byla uložená v Elasticsearch (*Elastic*).

5.1.4 Carrot2

Carrot2 je open source nástrojem pro tvorbu a vyhledávání clusterů, který automatický seskupí dokumenty podle určitých kategorií. Carrot2 je realizován v Java, a proto může být integrován s Java softwary, ale také podporuje C# / .NET API. Spolupráce s jinými jazyky jako například PHP nebo Ruby probíhá prostřednictvím REST (*Carrot2*).

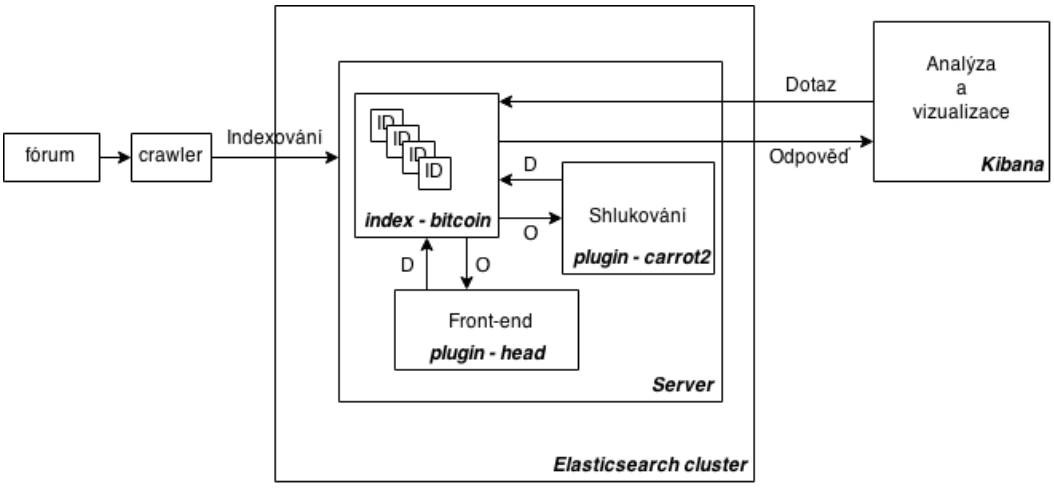
Pomocí Carrot2 lze analyzovat prostřednictvím dvou shlukových algoritmů, kterými jsou STC a Lingo. Také může shlukovat dokumenty 19 různých jazyků, jedním z nichž je angličtina (Osiński a Weiss, 2002).

Carrot2 není vyhledávácím nástrojem, jelikož nemá crawler a neindexuje (Osiński a Weiss, 2002), proto byly zvolené nástroje popsané výše jako Crawler4j a ELasticsearch. Carrot2 může fungovat i přímo v Elasticsearh v podobě pluginu, ale také prostřednictvím tzv. Workbench (Osiński a Weiss, 2002). Nainstalování pluginu Carrot2 do Elasticsearch je řešením pro tuto diplomovou práci.

5.2 Architektura řešení

Architektura řešení je zobrazená na obrázku č. 25. Data z předem určeného diskusního fóra budou stažená na server pomocí crawlera, který projde všechny relevantní URL a z HTML kódu vytáhne potřebná data. v crawleru využitém v této diplomové práce budou nastavené podmínky, aby prošel stránky zvoleného fóra a z nich primárně stáhl všechny příspěvky se jménem autora, s obsahem diskuzí a datem publikování těchto příspěvků, jelikož je třeba zjistit vývoj témat v čase a propojit tato témata s vývojem kryptoměn. Žádné další informace z fóra není třeba stahovat. Dále nestrukturovaná podoba příspěvků bude prostřednictvím Elasticsearch indexovaná. Jelikož všechny příspěvky budou stahována na jeden sever, pak tzv. Elasticsearch cluster bude obsahovat jenom jeden server. Data budou stažená do indexu s názvem „bitcoin“. Tento index bude obsahovat dokumenty s jediným identifikátorem, který se opakovat nemůže, jelikož všechny dokumenty budou ve stejném JSON formátu.

Pro analýzu a vizualizaci dat využiji nástroj Kibana, která data z Elasticsearch vytáhne a umožní využít různé metody analýzy dat a ukázat výsledky graficky. Pro shlukovou analýzu využiji plugin Carrot2 a pro informativní účely o indexu a nastavení shlukování v pluginu Carrot2 použiji plugin Head, které fungují obdobně jako i Kibana na základě dotazu do indexu na indexovaná data a odpovědi indexu na položený dotaz v podobě poskytnutí potřebných dat.



Obrázek 25: Architektura řešení, zdroj vlastní zpracování[12]

5.3 Datové zdroje

Pro dožení cílů jsem si vybrala relevantní, fungující, veřejně dostupné zdroje, které patří do skupiny sociálních sítí v podobě internetových nebo diskusních fór. Internetovým fórem podle Van Dulmena a kol. (2008) je webová stránka, která spojuje určité lidi do jedné specifické skupiny, kde si členové této skupiny podle Kitzingera (1994) vyměňují vlastní názory a zkušenosti na určitou věc. Diskuse je pak "*rozpravou, výměnou a obhajobou názorů,*

[12] D = dotaz, o = odpověď.

debatou" (*Diskuze*). Spojené diskusní fórum pak znamená seskupení lidí podle společného "zájmu", kteří poskytují vlastní názory a aktivně reagují na názory ostatních uživatelů.

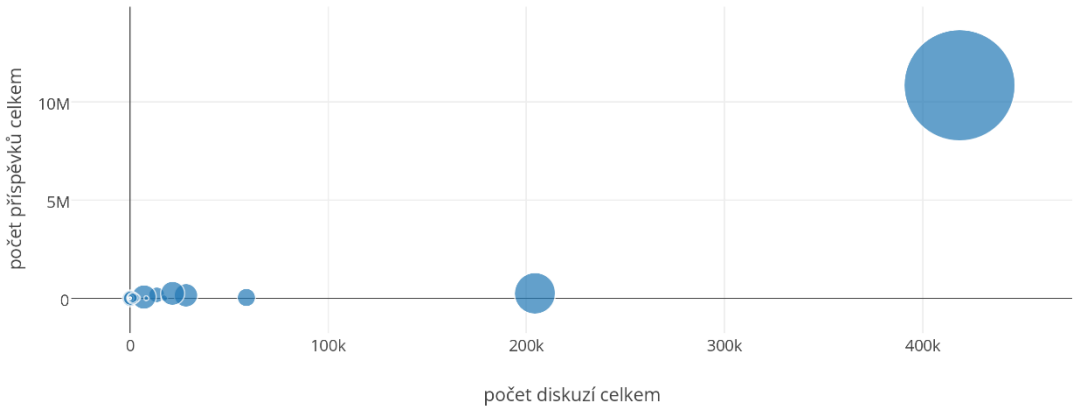
V dnešní době na internetu existuje spousta diskusních fór, kde uživatelé propagují novou kryptoměnu, píšou dotazy ohledně principu fungování kryptoměn, ptají se na vhodnou výpočetní techniku pro těžbu, diskutují ohledně vývoje kryptoměn a aj. Je třeba tady ale upozornit, že mnohá z těchto fór nejsou funkční, obsahují málo uživatelů nebo uživatelé na těchto fórech nejsou aktivní.

Zdroje jsou zvolené na základě diskuzí o kryptoměnách v angličtině, na základě počtu registrovaných uživatelů, obsahu fóra a počtu příspěvků uživatelů s porovnáním s počtem vytvořených diskuzí.

Na základě Guzman (2014) bylo vybraných 47 fór s kryptoměnami buď obecnými anebo se specifickou kryptoměnou, kde se mluví i o jiných kryptoměnách. Fóra, která byla v době psaní této diplomové práce funkční a která bezpodmínečně obsahovala současně více než nulu registrovaných uživatelů, počtu příspěvků a počtu vytvořených diskuzí. Zároveň zde bylo možno dohledat potřebné informace bez nutnosti registrace.

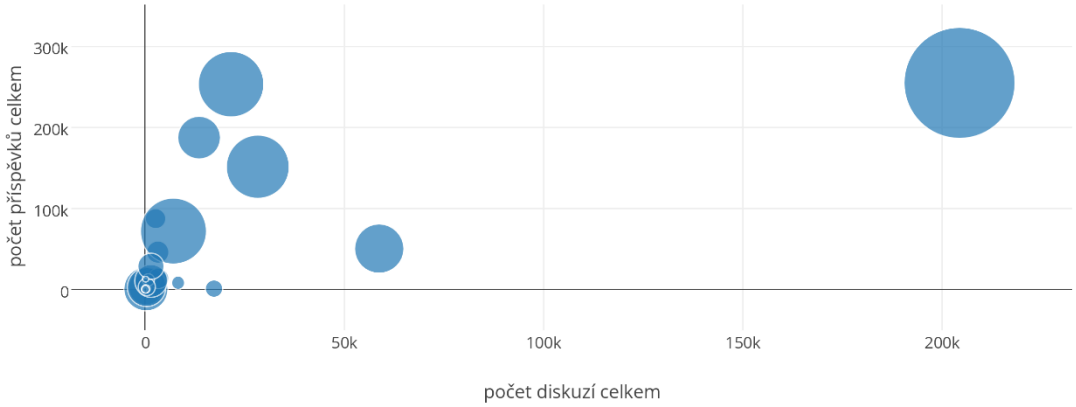
Na základě dohledatelných informací z jednotlivých fór o počtech registrovaných uživatelů, o celkovém počtu příspěvků a vytvořených diskuzí jsem sestrojila následující grafy s cílem vybrat potřebný zdroj pro budoucí stažení dat (viz Obrázek 26). Na ose *y* je počet příspěvků celkem v milionech, na ose *x* je počet vytvořených diskuzí v tisících, jednotlivé kruhy jsou jednotlivá fóra, velikost kruhu se odvíjí od počtu registrovaných uživatelů na daném fóru.

Z Obrázku č. 26 je vidět, že kruh, který se nachází v horním pravém rohu a který představuje fórum Bitcoin Forum, je hodně daleko od ostatních a je značně větší než ostatní představitelé fór. Z toho lze usoudit, že toto fórum obsahuje značně větší počet příspěvků, registrovaných uživatelů a počet vytvořených diskuzí. Přičemž při srovnání počtu vytvořených diskuzí s počtem příspěvků a počtem registrovaných uživatelů na jedno téma vychází v průměru cca 25 příspěvků a na jednoho registrovaného uživatele vychází v průměru cca 22 příspěvky. Na základě těchto informací můžu pak udělat závěr, že se na tomto fóru vytvářejí diskuze, které jsou aktivní, a že Bitcoin Forum je nejrozšířenější fórum mezi uživateli diskuzí o kryptoměnách ze zvolených zdrojů. Tento zdroj bude využitelný pro další analýzu.



Obrázek 26: Zmapování fór podle počtu příspěvků, počtu diskuzí a počtu zaregistrovaných uživatelů, zdroj vlastní zpracování

Zhodnotím i další zdroje. Dalším významným zdrojem je Marinecoin Forum (viz Obrázek 27), který je zobrazen v pravém horním rohu. Tento zdroj ale nebude pro moji analýzu využitelný, jelikož i když v průměru jeden registrovaný uživatel má cca tři příspěvky, ale jak je vidět z obrázku č. 27, celkový počet příspěvků a celkový počet diskuzí se pohybuje mezi skoro stejnými veličinami, kde na jednu diskuzi v průměru vychází cca jeden příspěvek.



Obrázek 27: Zmapování fór bez Bitcoin Forum podle počtu příspěvků, počtu diskuzí a počtu zaregistrovaných uživatelů, zdroj vlastní zpracování

Pro podrobnější zhodnocení Marinecoin Forum zjistím obsah na jeho webové stránce. Jedna ze zájmových sekcí fóra této diplomové práce je "Diskuze" (angl. "Discussion"). Ostatní sekce neobsahují na rozdíl od sekce "Diskuze" tak značný rozdíl mezi počtem příspěvků a vytvořených diskuzí. Sekce "Diskuze" obsahuje v době psaní této části diplomové práce 18 687 příspěvků a 6 705 diskuzí, což napovídá o podílu v průměru dva příspěvky na jednu diskuzi. Podívala jsem se na tuto sekci ještě podrobněji a zjistila jsem, že jedna diskuze s počtem příspěvků více než 2 000 nemá přínosnou hodnotu pro analýzu, je to téma přivítání na fóra, jehož obsahem nejsou diskuze o kryptoměnách. Další téma s obsahem více než 2 000 příspěvků je diskuze, kterou lze nalézt na předchozím zvoleném fóru Bitcoin Forum atd. Reálnou hodnotu pro analýzu mají zhruba čtyři témata ze sekce "Diskuze", která ale celkem obsahují cca 6 000 příspěvků ohledně jedné kryptoměny. z jedné strany v porovnání s Bitcoin Forum je to velmi malý objem dat pro analýzu pouze tohoto fóra. Na druhou stranu po spojení všech dat z tohoto fóra a z Bitcoin Forum dojde ke zkreslení analýzy nejdiskutovanějších měn, jelikož bez změny počtu příspěvků jiných kryptoměn zvýšíme počet příspěvků, kde se mluví o Marinecoin o 6 000. Tím z tohoto důvodu mi tento zdroj pro analýzu nebo použitelný.

Další zdroje s největším počtem příspěvků mají podíl na Bitcoin Forum zhruba 2 % na celkovém počtu příspěvků, přičemž kryptoměny, pro které byla vytvořená fóra, jsou diskutována i na Bitcoin Forum. Proto pro důkladnou analýzu byl vybrán jeden zdroj Bitcoin Forum s největším podílem uživatelů, příspěvků a vytvořených diskuzí mezi zvolenými fóra o kryptoměnách v anglickém jazyce.

5.3.1 Bitcoin Forum

Ze všech navřených anglických fór Guzman (2014) se fórum Bitcoin Forum těší největšímu zájmu uživatelů.

Bitcoin Forum je bezplatné fórum, které umožňuje uživatelům diskutovat na různá témata týkající se kryptoměn. Fórum bylo založeno v roce 2009 Satoshi Nakomoto a nyní ho spravují Michael Marquardt, Martti Malmi a Stefan Thomas (*Bitcoin Forum*). Fórum bylo založeno prostřednictvím softwarového balíčku Simple Machines Forum, který byl napsán v jazyce PHP a používá MySQL (Simple Machines, 2013)

Pro účely analýzy jsem sestrojila schéma sledované části struktury fóra (viz Příloha 1). Fórum se dělí na pět základních sekcí: Bitcoin, ekonomika, jiné, alternativní kryptoměny a lokalita. Každá ze sekcí se pak dělí na další dílčí sekce, kde uživatelé mohou snadněji najít určité užší téma a diskutovat na něj. Poslední sekce "Lokalita" v Příloze č. 1 se dělí ještě na 19 oddílů podle různých jazyků (upřesním později, o které jazyky se jedná). Fórum není přeloženo do těchto jazyků, ale uživatelé mluvící jiným jazykem než angličtinou mají možnost diskutovat ve vlastním jazyce ve zvláštně vytvořené sekci na Bitcointalk. Správci těchto sekcí potom vytvářejí vlastní strukturu hlavní stránky sekce a jejích dílčích sekcí.

Upřesnění pojmů z Přílohy č. 1:

- MultiBit – je mezinárodní bezpečnou open source Bitcoin peněženkou, kterou lze nainstalovat na Windows, MacOs a na Linux (Multibit, 2011).
- Electrum – je bezpečnou Bitcoin open source klient-server peněženkou (*Electrum*).
- BitcoinJ – Bitcoinj je Java knihovou pro práci s Bitcoin protokolem, ale může být použitá i v jiných jazycích kompatibilních s Javou, například Python nebo JavaScript (*Bitcoinj*).
- Armory – je bezpečnou open source peněženkou vyvíjenou na základě jazyku Python a dostupnou na platformách Windows, MacOs a na Linux (Armory, 2015).
- Pooly – je skupinou několika těžařů. v dnešní době představiteli poolu pro Bitcoin jsou větší společnosti, které mají prostředky na zabezpečení těžby, například F2Pool nebo AntPool atd. Menší těžaři, kteří nemají prostředky na nákup drahé výpočetní techniky nebo na zabezpečení těžby, se připojují do poolů a po nalezení bloku mají podíl na odměně.

V době psaní této části diplomové práce fórum obsahovalo 473 702 registrovaných uživatelů, v sekci diskuse o Bitcoinech bylo 34 416 vytvořených diskuzí a 579 977 příspěvků, v oddílu těžba bylo 20 685 diskutovaných diskuzí a 959 724 příspěvků, v části diskutování o ekonomice bylo 15 475 diskuzí a 726 356 příspěvků atd. (viz Tabulka 7).

Pro upřesnění uvedu tady krátký komentář ke každému oddílu diskusního fóra Bircointalk, který je uveden v tabulce č. 7 pro pochopení, jaká data toto fórum obsahuje.

Diskuze o Bitcoinu. Oddíl fóra, který obsahuje diskuze o systému Bitcoin, zprávách, inovacích a různých jiných specifických tématech, která se týkají kryptoměny Bitcoin (*Bitcoin Forum*). Z podrobnějšího prohlédnutí tohoto oddílu bylo zjištěno, že v něm uživatelé diskutují i o jiných existujících kryptoměnách.

Diskutování o technické stránce Bitcoinu. Diskuze o technických věcech systému Bitcoin a jiných systémů.

Těžba. v tomto oddílu uživatelé primárně píšou všechno, co je spojeno s těžbou Bitcoinu.

Technická podpora. Další oddíl, kde se uživatelé ptají na technickou stránku různých systémů kryptoměn.

Rozvoj projektu. Uživatelé tady řeší organizační otázky spojené s Bitcoinem a jinými souvisejícími projekty.

Ekonomika. Tento oddíl se věnuje otázkám ohledně ekonomiky států, porovnání kryptoměn s jinými měnami nebo drahocennými kameny a jiným problémům, které se týkají ekonomické stránky kryptoměn.

Trh. Diskuze ohledně tržních systémů.

Obchodování. Tento oddíl poskytuje možnost uživatelům diskutovat ohledně podnikání s Bitcoinem a psát vlastní best practices v podnikání s Bitcoinem.

Diskuze o Bitcoin Forum. Obecné otázky ohledně spravování fóra.

Politika a společnost. Tento oddíl se věnuje diskuzi ohledně politických situací ve světě.

Začátečníci a pomoc. v této části je začátečníkům umožněno pokládat otázky na systémy kryptoměn nebo obchodování s těmito kryptoměnami.

Jiná témata. Tady uživatelé mohou psát jiná témata, která mohou být zajímavá pro uživatele systémů kryptoměn.

Archiv. Obsahuje témata, která nejsou aktuální.

Diskuze o altcoinech, inzeráty, těžba a trh s altcoiny. Tyto části fóra by měly obsahovat témata, která se týkají altcoinů (*Bitcoin Forum*) nebo alternativních kryptoměn Bitcoinů. z pozorování bylo zjištěno, že uživatelé tuto strukturu striktně nedodržují a o altcoinech mluví i v jiných oddílech fóra.

Tabulka 7: Struktura sledované části fóra Bitcoin Forum a počet jeho témat a příspěvků 25. 3. 2015, zdroj Bitcoin Forum a vlastní zpracování

	Počet témat	Počet příspěvků
Diskuze o Bitcoinu	34 416	579 977
Diskutování o technické stránce Bitcoinu	9 485	121 055
Těžba	20 686	959 796
Technická podpora	5 347	40 654
Rozvoj projektu	6 935	100 448
Ekonomika	15 475	726 356
Trh	148 268	2 295 148
Obchodování	9 350	136 993
Diskuze o Bitcoin Forum	4 640	71 781
Politika a společnost	5 290	149 988
Začátečníci a pomoc	43 215	506 716
Jiná témata	13 837	315 803
Archiv	8 121	86 567
Diskuze o altcoinech	33 742	758 073
Inzeráty (altcoiny)	6 035	2 450 088
Těžba (altcoiny)	6 115	165 904
Trh (altcoiny)	5 820	114 552

Dále tato fóra obsahují oddíly i pro uživatele jiných jazyků, než je angličtina, které mají vlastní strukturu Bitcoin Forum trochu modifikovanou, ale diskusní témata jsou podobná. v grafu č. 3 je uveden seznam jazyků, ve kterých uživatelé diskutují ohledně kryptoměn. Nejvíce diskutují rusky mluvící uživatelé, kde počet příspěvků dosahuje skoro 350 000, pak

španělsky a německy mluvící uživatelé. Nejméně diskutují o kryptoměnách na vytvořeném fóru Bitcointalk.org ve svém rodném jazyce korejsky, polsky, hebrejsky mluvící země.



Graf 3: Počet příspěvků u Bitcoin Forum podle jednotlivých jazyků, zdroj Bitcoin Forum a vlastní zpracování

Na základě tabulky č. 7 a grafu č. 3 lze pak usoudit, že v dnešní době jsou kryptoměny mezi lidmi rozšířené, přičemž uživatelé o nich rádi diskutují na společných diskusních fórech, což nám umožní provést věrohodnou analýzu pomocí velkého objemu nestrukturovaných dat.

5.4 Shrnutí úkolů analytické části

Na základě stanovených cílů touto diplomovou práci, při využití vybraných technologií a vybraných zdrojů shrneme úkoly pro analytickou část:

- Stáhnout, instalovat a připravit zvolené nástroje Elasticsearch, Carrot2 a Kibanu pro zpracování analytické části.
- Stáhnout skript zvoleného open source crawleru.
- Nastavit crawler podle potřeb této diplomové práce.
- Stáhnout data ze zvoleného zdroje.
- Na základě dat zmapovat kryptoměny, o kterých se nejvíce mluví.
- Prozkoumat vývoj témat v čase.
- Propojit klíčová témata a souvislosti s vývojem hodnoty jednotlivých kryptoměn.
- Výsledky analýzy podpořit vizualizací.

6 Analytická část

Prostřednictvím analytické části budou dosažené následující předem definované tři cíle:

- Zmapování existujících kryptoměn.
- Analýza nestrukturovaných dat se zaměřením na vývoj témat jednotlivých kryptoměn v čase.
- Propojení klíčových témat a souvislostí s vývojem hodnoty jednotlivých kryptoměn.

V době psaní této části diplomové práce (25. 03. 2015) Bitcoin Forum obsahoval 10 837 360 příspěvků. Z omezených kapacitních důvodů bylo staženo 5 071 051 příspěvků, a proto analytická část je založená na vzorku s více než pětimilionovým objemem příspěvků. Na základě tohoto vzorku budou vytvořené určité analýzy a stanovené předpoklady určité události.

Při psaní této práce existovalo 598 kryptoměn, z nichž bylo 545 běžně obchodovaných s celkovou tržní kapitalizací kolem 4 100 000 000 $ (3. 4. 2015) (*Crypto-Currency Market Capitalizations*). Dá se tvrdit, že část kryptoměn není aktivně obchodovaná a jsou populární pouze u malé zájmové skupiny. Zájmem této práce je zmapovat nejdiskutovanější kryptoměny, a proto prostřednictvím softwaru bude zmapováno 20 nejdiskutovanějších krypotměn.

Hlavní témata diskuzí identifikuji u pěti z těchto kryptoměn a propojím to s vývojem jejich kurzů. Krátký popis ostatních nejdiskutovanějších kryptoměn uvádím v tabulkách v Příloze č. 2. Vybrala jsem si pět kryptoměn, ze kterých tři se liší metodou generace bloků. Tato volba je odůvodněná tím, aby bylo umožněno vidět rozdíl ve vývoji témat u kryptoměn s různými metodami generace bloků. Zvolenými kryptoměnami jsou Bitcoin, Dash, Monero, Nxt a Counterparty (XCP).

6.1 Získání dat

Na základě určitých usuzování uvedených v kapitole s názvem „Návrh řešení" byl zvolen zdroj Bitcoin Forum. Data z veřejně dostupného fóra jsou stažené do Elasticsearch prostřednictvím open source crawleru, který byl opraven podle potřeb práce.

Proces stažení dat je nastaven tak, že po spuštění crawler projde všechny stránky začínající na *https://bitcointalk.org/*. v crawleru jsou nastavené určité podmínky, podle kterých crawler rozpozná, co přesně bude stahovat z tohoto fóra. Pokud stránka odpovídá podmínkám, proces končí stáhnutím příspěvku na určité stránce, pokud ale stránka neobsahuje definované crawlerem prvky, pak crawler půjde na další stránku (viz Diagram 1).

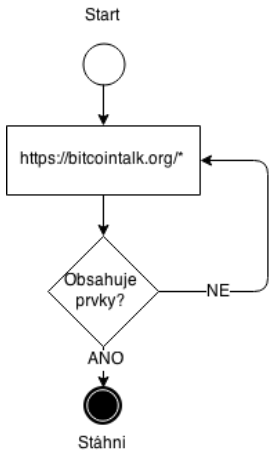
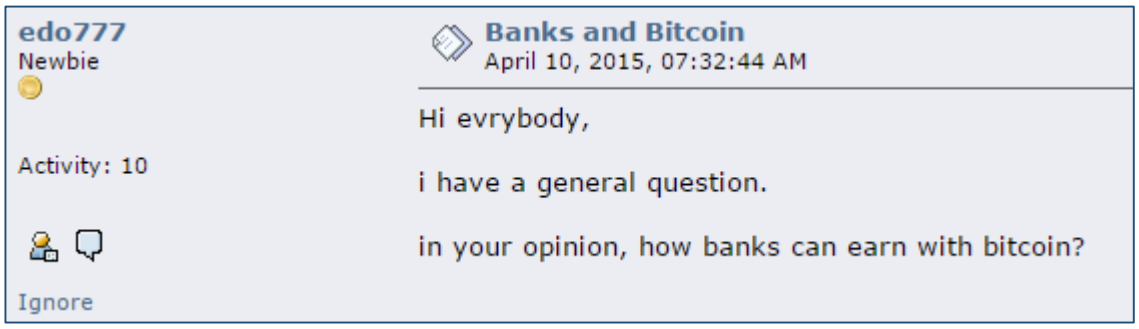


Diagram 1: Proces získání dat crawlerem, zdroj vlastní zpracování

Struktura příspěvků na Bitcoin Forum vypadá následovně (viz Obrázek 28):

- v levé části na obrázku č. 28:
  - jméno uživatele, který publikoval příspěvek,
  - jeho status na fóru, jestli je nováček na fóru, vývojář určité kryptoměny, člen fóra atd.,
  - uživatelská aktivnost na fóru, která neukazuje počet příspěvků, které byly publikované tímto uživatelem,
  - dále existuje možnost zaslat privátní zprávu, podívat se na jeho profil a ignorovat,
- v pravé části obrázku č. 28:
  - název dané diskuze,
  - datum publikování příspěvku a
  - obsah příspěvku.



Obrázek 28: Příspěvek s Bitcoin Forum, zdroj Bitcoin Forum

V HTML kódu celý příspěvek s obsažením jména uživatele, názvem diskuze a datem publikování příspěvku je umístěn v tabulce. Na základě HTML kódu crawler stáhne z určité stránky následující data:

- příspěvek,
- téma diskuze,
- jméno uživatele,
- datum publikování příspěvku.

Příspěvky v HTML kódu u fóru jsou obsažené v tágu:

```
<div class="post"> Příspěvek </div>
```

Obrázek 29: Umístění příspěvku v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování

V crawleru pro stažení příspěvků je pak uveden následující kód,

```
Elements messageEl = item.getElementsByAttributeValue("class", "post");
String message = messageEl.get(0).text();
```

Obrázek 30: Definování v crawleru místa pro stažení příspěvku, zdroj Škola (2015) a vlastní zpracování

který říká: stáhni všechny příspěvky, které jsou obsažené v atributu "class" s hodnotou "post" a začni stahovat od prvního tágu s tímto atributem a jeho hodnotou na určité stránce.

Podobným způsobem je nastaven kód pro stažení jmen uživatelů, témat diskuzí a dat publikování příspěvků. Jméno je obsaženo v HTML kódu v tagu tabulky s atributem "class" a jeho hodnotou „post" v jedné z jejích buněk. Tady je uveden tento tag ve zjednodušené podobě za vyloučením definovaných rozměrů tabulky a jiných hodnot definujících, jak by měla tabulka vypadat:

```
<td class="poster_info">
  <b>
    <a href="Odkaz na profil uživatele"> Jméno uživatele
    </a>
  </b>
```

Obrázek 31: Umístění jména uživatelé v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování

Pro stažení názvu diskuzí v crawleru, který je umístěn HTML tagu hlavičky dokumentů "title", je uveden následující kód:

```
for (Element item2 : itemsD ){
Elements nameEl = item2.getElementsByTag("title");
String name = nameEl.get(0).getAllElements().text();}
```

Obrázek 32: Definování v crawleru místa pro stažení názvu diskuze, zdroj Škola (2015) a vlastní zpracování

Tento kód napovídá, aby crawler pro určitý dokument stáhl obsah uvedený v tagu "title".

V crawleru pak pro stažení jména z tohoto tágu je určen následující kód,

```
Elements userEl = item.getElementsByAttributeValue("class", "poster_info");
String user = userEl.get(0).getAllElements().text();
```

Obrázek 33: Definování v crawleru místa pro stažení jména uživatele, zdroj Škola (2015) a vlastní zpracování

který podobně jako i u stažení příspěvků napovídá: stáhni všechna jména, která jsou obsažená v atributu "class" s hodnotou "poster_info", přičemž začni od prvního tágu s tímto atributem a jeho hodnotou.

Datum je obsaženo v HTML kódu v tágu s atributem "class" a jeho hodnotou "smalltext":

```
<div class="smalltext"> Datum </div>
```

Obrázek 34: Umístění data publikování příspěvku v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování

V crawleru pak kód vypadá následovně:

```
DateFormat formatter = new SimpleDateFormat("MMMM dd, yyyy, HH:mm:ss aaa");
Elements dateTimeEl = item.getElementsByAttributeValue("class", "smalltext");
String dT = dateTimeEl.get(2).getElementsByClass("smalltext").get(0).text().replace("\\s+", " ").replace("\u00a0"," ");
Date dt = new Date();
if (dT.contains("Today")){
 Calendar cal = Calendar.getInstance();
 cal.add(Calendar.DATE, 0);
 dt = cal.getTime();
}
```

Obrázek 35: Definování v crawleru formátu dat pro stažení, zdroj Škola (2015) a vlastní zpracování

Podle Java Tutorial v crawleru bylo definováno, v jakém formátu jsou data na fóru uvedená: *plný název měsíce, číslo dne, rok, hodina publikování příspěvku hodina : minuta : sekunda počítání denních hodin v anglosaském pojetí* (viz Obrázek 35). Dále kde je v HTML kódu datum umístěno. Datum je obsaženo v atributu "class" s hodnotou "smalltex", přičemž crawler bude stahovat od třetího atributu "class" s hodnotou "smalltext", jelikož předchozí atributy s touto hodnotou obsahují data, která nepotřebuji anebo datum

skutečného dne, který pro účely analýzy dat žádnou vypovídací schopnost nemá. Dále pokud se v datu na stránce u zvoleného atributu mohou vyskytnout jiné znaky než mezery, pak tento znak bude zaměněn na běžnou mezeru. Pokud příspěvek byl publikován v ten samý den, kdy i probíhalo stažení, na fóru je datum publikování označeno jako "Today" (z angl. dneska), pak bylo zapotřebí definovat, jaké má skutečně datum "Today".

Celý kód upraveného open source crawleru je uveden v Příloze č. 3 na CD-ROM.

Následným krokem je stažení dat a dále pokračuje krok předzpracování dat.

6.2 Předzpracování

Po stažení dat crawlerem do Elasticsearch pak prostřednictvím pluginu „head" lze zjistit určité informace o vytvořeném indexu.

Po každém novém spuštění Elasticsearch serveru se přiděluje název. Jelikož práce je postavená na jednom serveru v podobě osobního počítače autora, tzv. node (z angl. uzel) u Elasticsearch je jeden a nese název v okamžiku spuštění "Witchfire" (viz Obrázek 36 levá část).

Na pravé části obrázku č. 36 je zobrazen tzv. index s názvem "bitcoin", který z důvodu stažení velkého objemu dat byl rozdělen na pět tzv. shardů (z angl. kusy), kde každý shard je zvláštním Apache Lucene index (Kuć a Rogoziński, 2014). Na obrázku č. 36 jsou tyto shardy zobrazené ve čtverečcích a zelená barva říká, že je určitý shard aktivní anebo funguje bez chyb.

Pod názvem indexu "bitcoin" lze nalézt velikost všech obsažených dat a počet dokumentů, které byly stažené crawlerem. v této práci se jeden dokument počítá za jeden příspěvek z Bitcoin Fora.

Podrobnější informace o serveru, o indexu a o shardech jsou uvedené ve složce "Info" (z angl. informace)ve složce "Actions" (z angl. činnost) jsou uvedené určité aktivity, které lze například s indexem provést, lze ho odstranit, obnovit, testovat atd.



Obrázek 36: vytvořený index v Elasticsearch s názvem "bitcoin", zdroj Plugin head a vlastní zpracování

V složce "Info" také lze nalézt informaci o metadatách indexu, kde na začátku je uvedená informace o analyzeru, u tohoto indexu je nastavený standardní tokenizer, filtry symbolů (anlg. filter tokens): filtr pro převod do malých písmen, odstranění stop slov a stemmer filtr. Dále je zde uvedená informace o vytvořených polích. Informace v shardech je roztříděná podle tzv.

mappingu (z angl. zobrazení). Každý dokument má vlastní mapping, pokud předem mapping nebyl definován, Elasticsearch vytvoří mapping automaticky sám (Kuć a Rogoziński, 2014).

Část mappingu indexu "bitcoin" je zobrazen na obrázku č. 37, kde jsou zobrazeny názvy jednotlivých polí, jejich typ, a jestli byl analyzován prostřednictvím tzv. analyzeru.


```
"mappings": {
    "discussion": {
        "properties": {
            "page.raw": {
                "index": "not_analyzed",
                "type": "string"
            },
            "url.raw": {
                "index": "not_analyzed",
                "type": "string"
            },
            "level": {
                "type": "long"
            },
            "created": {
                "format": "yyyy-MM-dd HH:mm:ss",
                "type": "date"
            }
        }
    }
}
```


Obrázek 37: Část mapping indexu "bitcoin", zdroj Plugin head a vlastní zpracování

Jakmile data jsou předzpracovaná, přestoupíme na analýzu těchto dat. Během zpracování analýzy budou získané určité výstupy, které budou okomentované. Další kapitola se zabývá analýzou a interpretací dat získaných na Bitcoin Forum.

6.3 Analýza a interpretace dat

Tato kapitola je zaměřená na analýzu nestrukturovaných dat prostřednictvím zvolených softwarů.

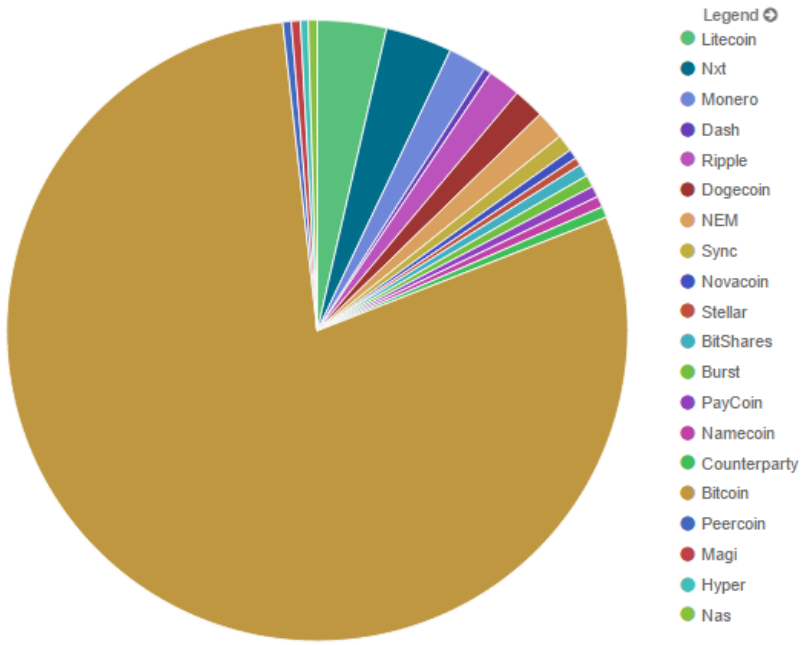
V této části je třeba zmínit, že ve výsledcích clusterů se vyskytují záznamy, které nenesou žádnou hodnotu pro analýzu dat, a proto je komentovat nebudu. Také výsledky shlukové analýzy uvádím v jazyku stažených dat, přesněji v angličtině, jelikož některá slova specifická pro obor kryptoměn se nedají správně přeložit jedním slovem do češtiny. Výsledky clusterů pak komentuji v češtině s uvedením vhodného překladu jednotlivých termínů.

Při identifikaci hlavních témat u jednotlivých kryptoměn na začátku vytvářím cluster spojený s určitou kryptoměnou na celkový objem dat, pak pro podrobnější sledování vývoje témat v čase uvádím témata podle jednotlivých roků podrobněji. Jelikož cluster na celkový objem dat bude obsahovat všechna slova využitá v kontextu s určitou kryptoměnou i do doby spuštění této kryptoměny, výsledky podrobnější identifikace témat se mohou v určité míře lišit.

Na základě stažených dat bylo identifikováno 20 nejvíce diskutovaných kryptoměn (viz Graf 4). Nejvíce diskutují uživatelé na Bitcoin Forum o kryptoměně Bitcoin a jak je vidět z grafu č. 4, podíl příspěvků o Bitcoinech z celkového počtu 20 nejdiskutovanějších kryptoměn je více než 80 %. A na celkové tržní kapitalizaci na celém trhu kryptoměn mají skoro 97% podíl (*Crypto-Currency Market Capitalizations*). Důvody jeho popularity mezi uživateli fóra mohou být následující:

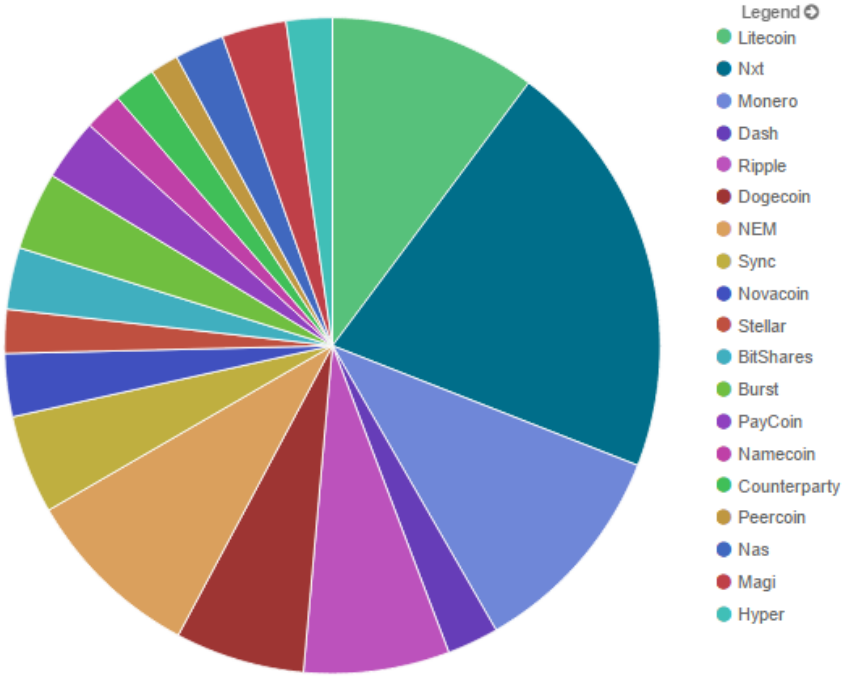
- je to první a nejstarší obchodovaná měna,
- fórum bylo primárně vytvořeno pro Bitcoin, jenom postupem času uživatelé začali psát i o jiných kryptoměnách,

- kurz BTC/USD = 257, 08 $ (5. 4. 2015). Z historie je známý příklad prudkého vzestupu Bitcoinu (viz Graf 9), což vyvolalo zájem velkého počtu lidí, a to pravděpodobně diskutovali na fóru.



Graf 4: 20 nejdiskutovanějších kryptoměn na fóru Bitcoin, zdroj vlastní zpracování

Pro lepší přehlednost jsem vyloučila ze seznamu 20 nejdiskutovanějších kryptoměn Bitcoin (viz Graf 5). Na druhém místě nejvíce diskutovaných kryptoměn je Nxt, pak Monero, Litecoin, NEM, Ripple atd.



Graf 5: 19 nejdiskutovanějších kryptoměn na fóru Bitcoin po vyloučení kryptoměny Bitcoin, zdroj vlastní zpracování

Další část této práce bude zaměřená na analýzu dat pěti kryptoměn z 20 nejdiskutovanějších. Také bude zdůvodněno, čím jsou jednotlivé kryptoměny pro uživateli zajímavé.

6.3.1 Bitcoin

Kryptoměna Bitcoin (BTC), jak už bylo uvedeno v kapitole „Kryptoměna“, je nejstarší kryptoměnou ze všech existujících - byla spuštěná v roce 2009. Maximální množství mincí je 21 000 000 BTC. Průměrná doba generace nových bloků nebo schválení transakcí je deset minut, těžební algoritmus je SHA-256, metoda generace a bezpečnosti bloků je Proof of Work. Pro generaci nových bloků je nastavená určitá úroveň náročnosti, která se mění po nalezení každých 2016 bloků. Odměna těžařovi, který nalezl určitý blok, je 25 BTC. Mince se prostřednictvím odměňování vypouští do obchodování. Odměna za nový blok se snižuje dvakrát po nalezenií každých 210 000 bloků. v době psaní této části diplomové práce byl počet naleznutých 349 314 (viz Tabulka 8).

Tabulka 8: Obecná informace o BTC, zdroj Bitmakler, Blockchain Info, Crypto-Currency Market Capitalizations a vlastní zpracování

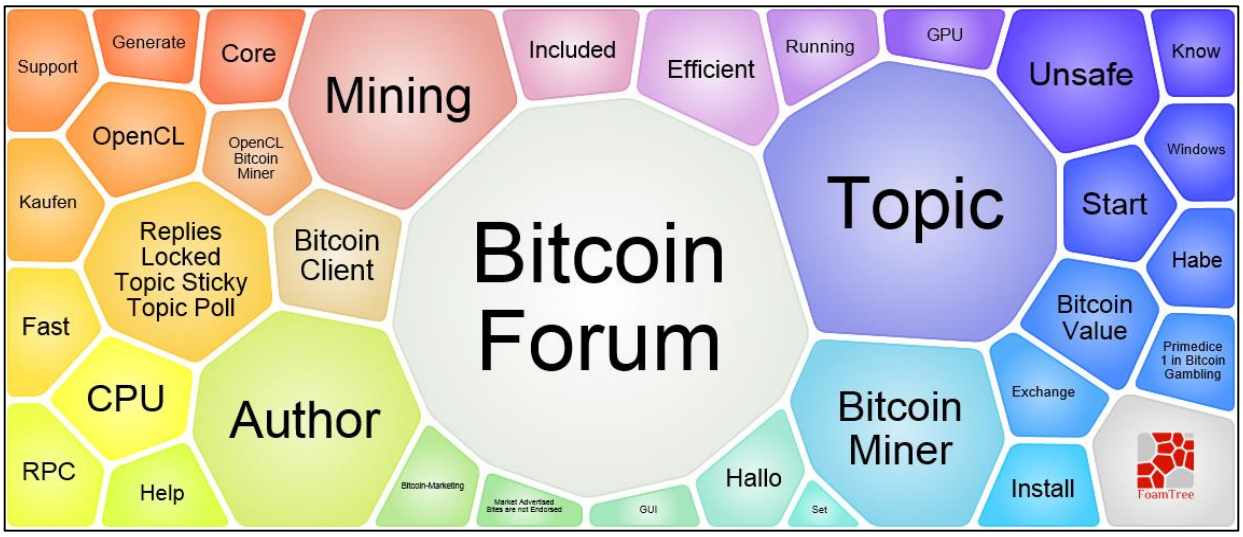
Doba spuštění	04. 02. 2009
Průměrná doba generace bloků	10 minut
Metoda generace a bezpečnosti bloků	Proof of Work
Maximální množství mince	21 000 000 BTC
Těžební algoritmus	SHA-256
Změna náročnosti	každých 2016 bloků
Odměna za nalezený blok (26.03.2015 11:00)	25 BTC
Snížení odměny (kdy?)	po nalezení každých 210 000 bloků
Snížení odměny (o kolik?)	Dvakrát
Počet nalezených bloků (26.03.2015 11:10)	349 314

6.3.1.1 Hlavní témata

Bitcoin podle výše provedené analýzy je nejdiskutovanější kryptoměnou ze všech diskutovaných kryptoměn na Bitcoin Forum.

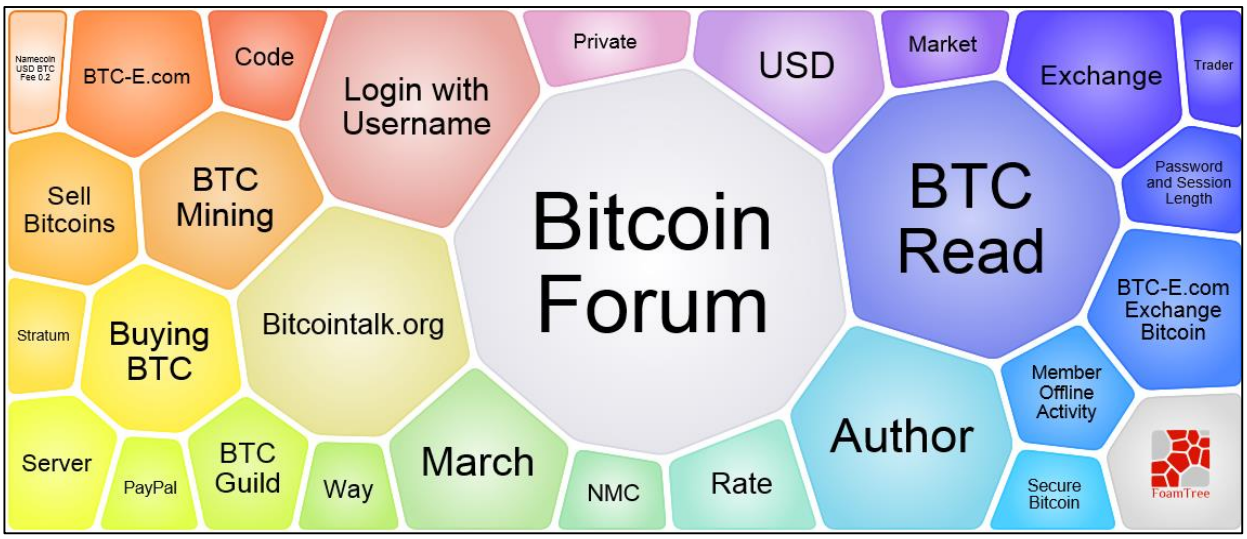
Na základě clusteru tvořeného prostřednictvím algoritmu “Lingo” byla zjištěna slova a témata, která jsou nejvíce spojená se slovem “Bitcoin” (viz Obrázek 38) a “BTC” (viz Obrázek 39) od doby spuštění kryptoměny Bitcoin.

Z obrázku č. 38 je vidět, v jakém kontextu se nejvíce mluví o Bitcoinu. z tohoto clusteru lze usoudit, že na Bitcoin Foru se nejvíce o Bitcoinu mluví v kontextu těžby, softwaru Bitcoin Core, generace bloků, podpory Bitcoin klientů, těžby prostřednictvím CPU a GPU, také prostřednictvím tzv. poolu RPC miner. Na fóru se očividně často ptají na koncept Bitcoinu anebo na pomoc při řešení určitých problémů. Mluví se často o tzv. OpenCL Bitcoin těžařovi, který je založen na OpenCL frameworku pro výpočet hash. v kontextu s Bitcoinem se používají přídavná jmena, jako “nebezpečný, rychlý a výkonný”, což po zkoumání obsahu fóra slova rychlý a výkonný se nejvíce mluví v kontextu s hardware nebo software, se kterými jsou také spojená podstatná jména jako instalování a provozování, přičemž nejvíce uživatelé diskutují o operačním systému Windows. v případě tvorby clusteru prostřednictvím algoritmu “STC” se pak ve výsledku vyskytuje i operační systém Linux. Dále se často mluví o tom, jak začít s Bitcoinem, pak také uživatelé fóra diskutují o marketingové stránce Bitcoinu, kde a jak vyměnit Bitcoin, o hodnotě Bitcoinu a o hazardu.



Obrázek 38: Cluster "Bitcoin", zdroj vlastní zpracování

V kontextu se zkratkou Bitcoinu BTC jsou spojená další témata (viz Obrázek 39). Na fóru uživatelé projednávají kód Bitcoinu, protokol Stratum. Diskutují ohledně burzy BTC - E a platebního systému PayPal. Nejvíce se v kontextu s BTC mluví o měně USD a o kryptoměně NMC, která je zkratkou Namecoin. Na Bitcoin Forum se diskutuje o kurzu BTC, o jeho nákupu či prodeji. Dále v této práci bude uvedeno podrobněji, jestli uživatelé doporučují prodávat nebo kupovat BTC nebo jestli se ohledně nákupu a prodeji mluví v jiném kontextu než doporučení. Také se uživatelé ptají na otázky týkající se hesel, jednají ohledně bezpečnosti Bitcoinu, ohledně obchodování a trhu. Prostřednictvím algoritmu "STC" bylo zjištěno další slovní spojení, které se často používá uživateli fóra v kontextu s BTC. Tímto slovním spojením je investovat do BTC.



Obrázek 39: Cluster "BTC", zdroj vlastní zpracování

6.3.1.2 Vývoj témat v roce 2010

V této části bude uveden vývoj témat v čase, o kterých se nejčastěji mluví v kontextu s Bitcoinem nebo jeho zkratkou BTC. i když Bitcoin byl spuštěn v roce 2009, rozsáhlé diskuze o něm na Bitcoin Forum začínají od roku 2010, a proto vývoj témat budu popisovat od roku 2010.

V každém roce uživatelé mluví o operačních systémech Windows a Linux, které z důvodu neopakovatelnosti uvádět ve výsledcích nebudu.

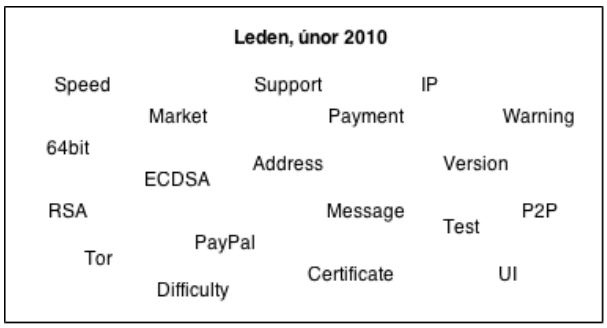
V roce 2010 jsou v kontextu s Bitcoinem nebo BTC spojená slova, která jsou zobrazená na obrázku č. 40. Nejvíce uživatelé fóra diskutují o adresách, jak to vytvořit, jak používat, ptají se na adresy jiných uživatelů nebo řeší nějaké problémy s tím spojené. Dále hodně mluví o náročnosti nalezení nových bloků. Pak byl v roce 2010 publikován vývoj nových verzí softwaru Bitcoin a jeho aktualizací, které uživatelé zkoušejí a určitým způsobem je hodnotí. v kontextu s Bitcoinem pak používají slova jako měna, komodita, peníze, také virtuální, decentralizovaný, P2P, anonymní anebo podvod (angl. scam). Uživatelé také diskutují o těžbě Bitcoinu (v roce 2010 angl. minting) a využívajícím algoritmu těžby Proof of Work (zkratka PoW),jaké technologie mohou použít pro těžbu. v roce 2010 se nejvíce v kontextu s Bitcoinem mluví o CPU, méně o GPU. Uživatelé se ptají také na různé servisy, kde mohou Bitcoin používat nebo vyměnit na nějakou jinou existující měnu.Nejvíce diskutovalo o směně BTC na USD a naopak. Projednávají otázky ohledně hash, IP adresách, posílání zpráv (angl. message) spolu s posíláním mincí a uložení mincí (angl. backup). Uživatelé také projednávají generaci nových bloků, sítě bloků (angl. blockchain), rychlost schvalování transakcí, celkový počet mincí. Dále se mluví o businessstránce, o vytvoření článků na Wikipedii, o hrách přijímajících Bitcoin, o propagaci BTC a o fóru Bitcoin Forum. Pak diskutují o různých problémech spojených s otevřením browserů. Darují BTC, mluví o cenách transakcí, nákupu či prodeji BTC, spotřebě elektřiny, vztahu bank a státu k Bitcoinu. Ptají se na různé otázky u Satoshi Nakamoto.



Obrázek 40: Diskuze spojené s Bitcoinem v roce 2010, zdroj Jason Davies a vlastní zpracování

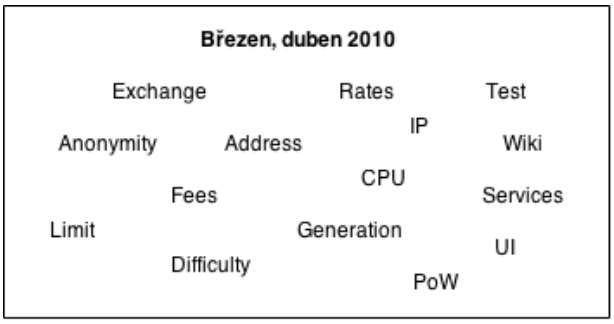
Dále uvádím podrobněji vývoj témat v kontextu s Bitcoin nebo jeho zkratkou BTC v jednotlivých měsících.

V lednu roku 2010 (viz Obrázek 41) uživatelé projednávají taková témata, jako jsou tvorba trhu Bitcoinu, pak diskutují o technické stránce Bitcoinu: o RSA[13], ECDSA[14], také řeší, jestli lze software Bitcoin pustit na 64 bitových systémech. v únoru roku 2010 se přidávají další témata diskuze, uživatelé diskutují o rychlosti generace bloků, mluví o vývoji webové aplikace nebo web UI (User Interface) pro pohodlné obchodování s mincí nebo o nepřátelském UI současných verzí. Mluví o systému Bitcoin jako o Tor[15]. Řeší problémy s browsery, které pro otevření https://bitcoin.org/ potřebovaly certifikáty autority. Uživatelé také v únoru roku 2010 začínají mluvit o náročnosti generaci nových bloků (viz Obrázek 41).



Obrázek 41: Diskuze spojené s Bitcoinem v lednu a únoru roku 2010, zdroj vlastní zpracování

V březnu roku 2010 (viz Obrázek 42) lze pozorovat témata, která jsou méně spojená s technickou stránkou Bitcoinu. Uživatelé diskutují kromě výše uvedeného také o poplatcích za provedené transakce, o budoucí těžbě Bitcoinů po dosažení limitní částky 21 000 000 BTC. Také se ptají na vhodné poskytovatele služeb pro směnu BTC na jinou měnu a jaký má současný kurz. V dubnu se přidává diskuze o přidání tématu Bitcoin na Wikipedii. Uživatelé diskutují o algoritmu těžby Bitcoinu Proof of Work, pak se ptají na vhodné poskytovatele platebních služeb (viz Obrázek 42).



Obrázek 42: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2010, zdroj vlastní zpracování

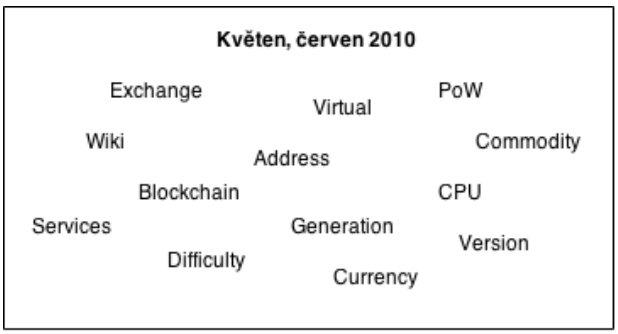
V květnu (viz Obrázek 43) uživatelé diskutují o tom, jak je možné obchodovat s virtuálními měnami, projednávají, jestli Bitcoin je měna nebo komodita, diskutují o síti bloků (angl. blockchain). v červnu uživatelé mluví o možnostech vývoje mobilní verze softwaru Bitcoin a diskutují o skutečné verzi BTC (viz Obrázek 43).

---

[13] "*Algoritmus RSA publikovali v roce 1978 Ronald Rivest, Adi **S**hamir a Leonard **A**dleman. Jedná se o asymetrickou šifru, která je založena na Eulerově větě a která je použitelná jak pro šifrování, tak pro podepisování dokumentů*" (*Mička*)

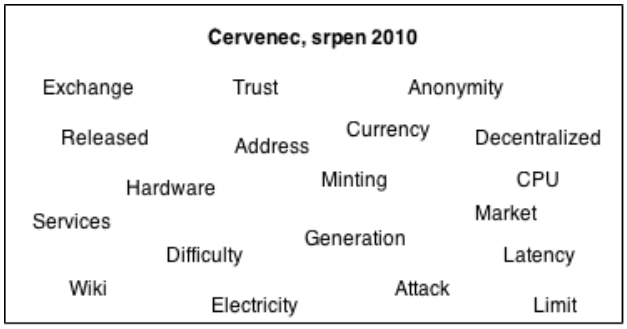
[14] ECDSA (The Elliptic Curve Digital Signature Algorithm) – algoritmus, který je analogem DSA (Digital Signature Algorithm) a používající eliptické křivky. Tento algoritmus využívají pro digitální podepisování (Johnson, Menezes, Vanstone, 2001).

[15] Síť Tor je sdružením několika serverů, které byly dobrovolně zapojené do této sítě. Tato síť umožňuje zvýšit uživatelskou anonymitu a bezpečnost při pohybu na Internetu (*Tor*).



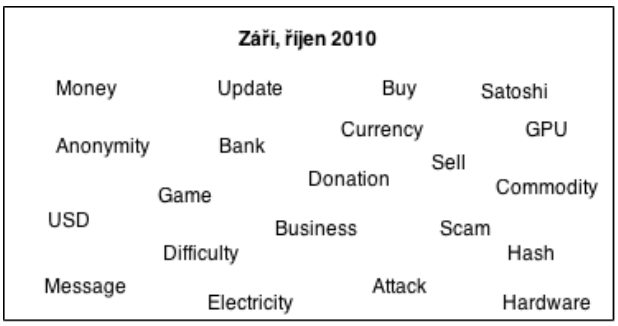
Obrázek 43: Diskuze spojené s Bitcoinem v květnu a červnu roku 2010, zdroj vlastní zpracování

V červenci a srpnu (viz Obrázek 44) se publikují nové verze softwaru Bitcoin, o kterých uživatelé aktivně diskutují. v kontextu s Bitcoin pak probíhají diskuze o důvěryhodnosti, anonymitě, decentralizaci a latenci. Bitcoin se začíná vystavovat útokům hackerů. Uživatelé diskutují o těžbě Bitcoinu, o použití různých hardwarů, o vzniku nových trhů a o spotřebě elektřiny.



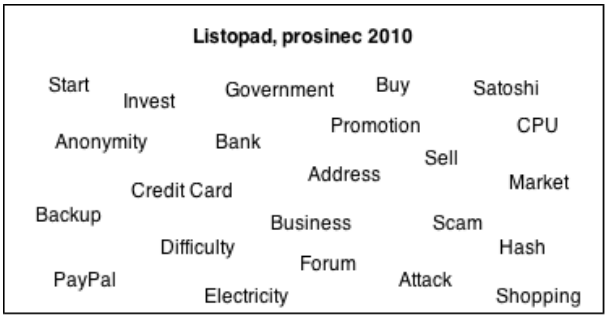
Obrázek 44: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2010, zdroj vlastní zpracování

V září roku 2010 (viz Obrázek 45) uživatelé kromě výše uvedených mluví o aktualizovaných verzích softwaru Bitcoin, o možnosti nákupu či prodeje mincí, přičemž doporučení o tom, kdy kupovat nebo prodávat uživatelé nedávají, diskutují jenom o tom, kde mohou koupit anebo prodat Bitcoiny, a také nabízejí BTC k prodeji či nákupu na Bitcoin Forum. Mluví o nových hrách, kde by uživatelé mohli tratit své mince. V říjnu uživatelé více naznačují o svých penězích, které nechali vyměnit na BTC. Také mluví o dolarech, se kterými BTC může být obchodován. Diskutují o vztahu bank ke kryptoměně a o odlišnostech bankovního systému a kryptoměn. Za provedenou práci nebo na jiné projekty uživatelé posílají darovací částky vývojářům. Stále více mluví o růstu náročnosti a spotřebě elektřiny. Začínají psát o těžbě prostřednictvím GPU, jelikož náročnost vzrůstá. Diskutují o možnostech Bitcoinu v businessu, pak také píšou o Bitcoinu jako o podvodu. Dále se uživatelé obracejí na vývojáře Satoshi Nakamoto s řešením nějakých problémů (viz Obrázek 45).



Obrázek 45: Diskuze spojené s Bitcoinem v září a říjnu roku 2010, zdroj vlastní zpracování

V listopadu (viz Obrázek 46) kromě výše uvedených témat začínají uživatelé více diskutovat o regulaci Bitcoinu ze strany státu nebo o tom, že Bitcoin byl vytvořen státem. Ptají se na možnosti nákupu mincí prostřednictvím kreditních karet a diskutují o příležitostech investování. Někteří uživatelé také mluví o možnostech tvorby nové kryptoměny, jelikož si myslí, že Bitcoin má určité nedostatky. Stále více uživatelé píšou o neefektivnosti těžby prostřednictvím CPU a tím pádem je vidět tendenci odchodu do GPU. Pak vzniká možnost v některých obchodech nakupovat online, což vyvolalo také aktivní diskuzi mezi uživateli (viz Obrázek 46).



Obrázek 46: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2010, zdroj vlastní zpracování

6.3.1.3 Vývoj témat v roce 2011

V roce 2011 se některá témata diskuze změnila. Celý souhrn témat je zobrazen na obrázku č. 47. Stále se ale diskutuje o náročnosti nalezení nových bloků, o Bitcoinu se mluví jako o kryptoměně. Více uživatelé mluví o propagaci Bitcoinu, o jeho hodnotě, ceně a trhu. Také projednávají výši poplatků za transakce, bezpečnost systému, spekulace, darování. Na fóru vznikly diskuze o daních, o Ponziho schématu, Mt. Gox[16], TradeHill[17]. Uživatelé také píšou o ukradených mincích z jejich peněženek a více diskutují o atakovaných burzách a peněženkách hackerů. Na fóru se začíná šířit myšlenka úpadku kryptoměny Bitcoin.

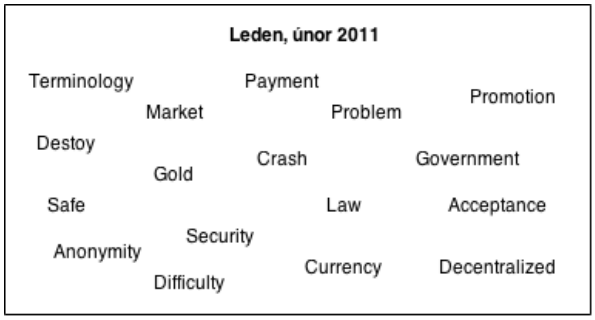
[16] Burza pro obchodování s Bitcoin, která byla založená v Tokiu (Trautman, 2014).
[17]Burza pro obchodování s Bitcoin (Bradbury, 2013).



Obrázek 47: Diskuze spojené s Bitcoinem v roce 2011, zdroj Jason Davies a vlastní zpracování

Dále uvádím podrobněji vývoj témat v roce 2011 podle měsíců.

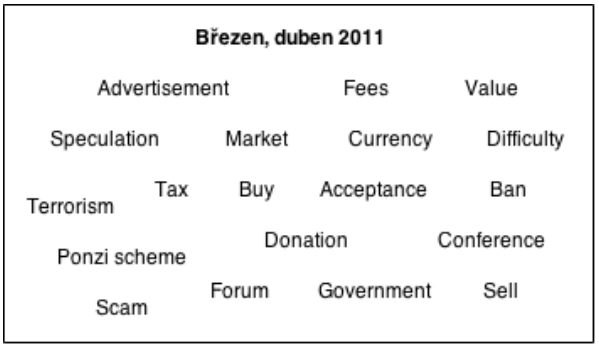
V lednu roku 2011 (viz Obrázek 48) se stále více lidí zajímá o kryptoměnu Bitcoin, kteří se ptají na vytvoření nějakého terminologického základu spojeného s touto kryptoměnou. Na fóru stále diskutují o možnostech obchodování s Bitcoin. Také se uživatelé ptají a píšou, kde lze vlastněné BTC utratit. Pořád panuje téma propagace Bitcoinu pro přilákání více zájemců pro obchodování s BTC. Dále se v lednu řeší určité problémy uživatelů spojené například s instalováním peněženky. V únoru roku 2011 (viz Obrázek 48) uživatelé diskutují o tom, jestli lze Bitcoin zničit, příkladem je William Pitock, který podle uživatelů Bitcoin Forum chtěl zničit Bitcoin. Diskutují také o názoru státu na Bitcoin a o zákonu. Pak uživatelé porovnávají Bitcoin se zlatem. Témata decentralizace, bezpečnosti, náročnosti a Bitcoinu jako měny zůstávají stále populární.



Obrázek 48: Diskuze spojené s Bitcoinem v lednu a únoru roku 2011, zdroj vlastní zpracování

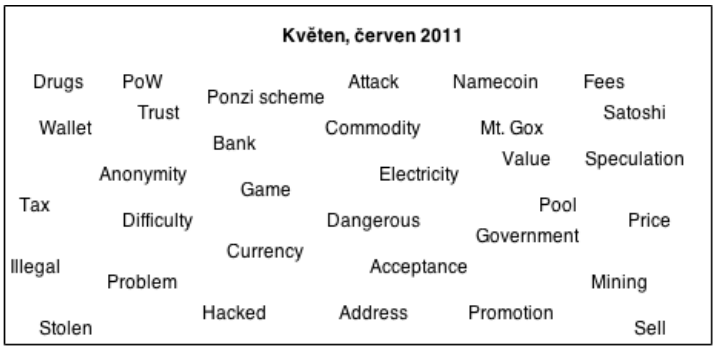
V březnu a dubnu (viz Obrázek 49) uživatelé více diskutují o poplatcích při obchodování s Bitcoinem, stále píšou o nutnosti propagace, o samotném fóru Bitcoin Forum, mluví o hodnotě Bitcoinu a o různých podvodech, o spekulaci s BTC, o činnosti státu ve vztahu k BTC, o existujícím trhu a jeho rozšíření a o darování BTC na různé účely, například pro oběti japonského tsunami. Vznikají nová témata týkající se daní, které mohou být naložené na Bitcoin, píšou o možnostech uplatnění Bitcoinu pro teroristické účely, o konferenci věnované

BTC a porovnávají systém Bitcoinu s Ponziho schématem, které je založeno na tom, že investoři svěří své peněžní prostředky fondu pro budoucí jejich zhodnocení, avšak fond je dále neinvestuje a pouze proplácí zhodnocení některým investorům, kteří investovali dřív z prostředků investovaných novými investory (Hirsch, 2009). Také se uživatelé v dubnu primárně ptají, jestli je pravý čas na nákup BTC, či není. v tomto měsíci ale lze pozorovat doporučení i na prodej BTC.



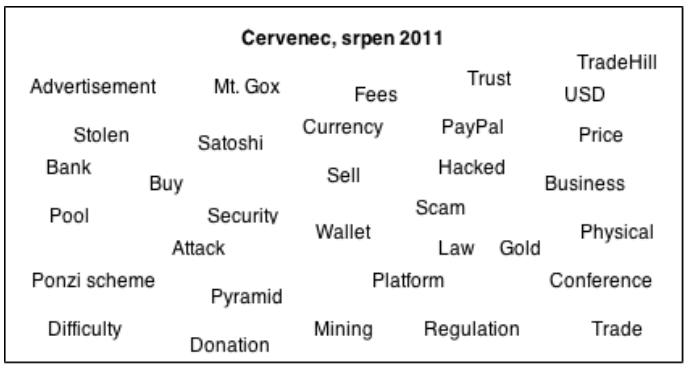
Obrázek 49: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2011, zdroj vlastní zpracování

V květnu (viz Obrázek 50) uživatelé v kontextu s Bitcoin mluví dalších nových tématech, jako jsou drogy (angl. drugs) a nelegální obchody, jelikož obchodování s Bitcoinem je anonymní, pak je prostor pro nelegální obchody. Uživatelé mluví také o možném zdanění BTC. v květnu a v červnu zůstávají populární témata Bitcoinu jako virtuální měny, PoW, spotřeby elektřiny, která pořád roste, využití Bitcoinu ve hrách, vztah banků k BTC, diskutování o peněženkách, atakování a porovnávání BTC s Ponziho schématem atd. (viz Obrázek 50). v červnu se (viz Obrázek 50) pak kromě výše uvedených vyskytují doporučení o prodeji BTC. Uživatelé v kontextu s Bitcoin mluví o další kryptoměně Namecoin, která je uvedená v Příloze č. 2. Také začínají více mluvit o burze Mt. Gox, která v této době utrpěla útok hackery a BTC byly ukradeny (McMillian, 2014). Na fóru se pak v této době aktivně diskutovalo i o ukradených BTC z účtů nebo z peněženky. v této době vznikají i témata spojená s pooly (z angl. bazény), kde se těžaři sdružují do jedné skupiny a těží kryptoměnu společně, pak se odměna za nalezený blok rozděluje mezi členy této skupiny. v této době se mění i diskuze ohledně Satoshi Nakamoto, uživatelé se začínají ptát na to, kdo Satoshi Nakamoto ve skutečnosti byl a proč neodepisuje na fóru. Také se jménem satoshi uživatelé spojují drobné BTC.



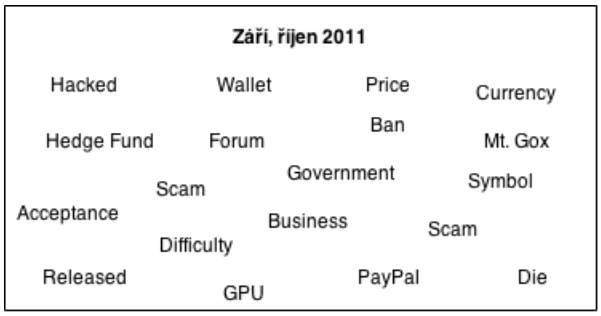
Obrázek 50: Diskuze spojené s Bitcoinem v květnu a červnu roku 2011, zdroj vlastní zpracování

V červenci (viz Obrázek 51) uživatelé kromě výše zmíněných témat porovnávají systém Bitcoin s pyramidou, ptají se na platformy umožňující obchodování s BTC. v srpnu (viz Obrázek 51) kromě výše zmíněných se ještě přidávají témata o burze TradeHill a o možnostech regulace BTC.



Obrázek 51: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2011, zdroj vlastní zpracování

V září a v říjnu roku 2011 (viz Obrázek 52) uživatelé diskutují o možnostech založení Hedge fondů, začínají psát o krachu kryptoměny BTC, mluví o výkonnosti GPU a diskutují o symbolech, které lze přidělit Bitcoinu.



Obrázek 52: Diskuze spojené s Bitcoinem v září a říjnu roku 2011, zdroj vlastní zpracování

V listopadu a v prosinci (viz Obrázek 53) uživatelé začínají aktivně diskutovat ohledně Silk Road[18], která umožní provádět nelegální obchody. Na fóru stále více píšou i o opozici Bitcoin. Také v listopadu se prováděl tzv. giveaway (z angl. jako zadarmo), který znamená, že buď vývojář kryptoměny anebo jiný uživatel rozdává mince zadarmo. S tímto termínem se často lze setkat u čerstvě vzniklých kryptoměn, u nichž chce tvůrce měny takovým způsobem zvýšit zájem ostatních uuživatelů. Také se ale lze setkat s giveaway i u existujících měn, u kterých klesá kurz.



Obrázek 53: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2011, zdroj vlastní zpracování

[18] Je to platforma pro anonymní obchodování, která je založená na síti Tor (*Silk Road Drugs*).

6.3.1.1 Vývoj témat v roce 2012

V dalších měsících každého roku uživatelé pořád diskutují o adresách, transakcích, decentralizaci, anonymitě, náročnosti nalezení bloků, sítě bloků (angl. blockchain), o peněženkách, obchodování, generaci nových bloků, hashi, poolech, bezpečnosti a o Bitcoinu jako o měně. Abych se neopakovala, tato slova z výsledků vyloučím. Ze stejného důvodu ale nevyloučím slova spojená s krádeží a přijetím BTC různými obchody nebo službami, aby bylo vidět tendenci rozšíření BTC a zároveň nebezpečnost obchodování s touto kryptoměnou.

V roce 2012 uživatelé nejvíce diskutují na témata, která jsou zobrazena na obrázku č. 54. Populární se stává diskutování o technologii ASIC, z toho je vidět tendenci, jak uživatelé začínají ASIC potřebovat, jelikož náročnost nalezení bloků neustále roste, výkonnost GPU už nepostačuje. Také je vidět, že uživatelé pořád mluví o státu a bankách a o jejich vztahu ke kryptoměně Bitcoin. v roce 2012 na rozdíl od předchozího roku uživatelé více diskutují o burze Mt. Gox než o TradeHill. Populární mezi uživateli se stává téma Bitcoinica[19], o které se bude hovořit v pozdější subkapitole. Také uživatelé píšou a ptají se, kde lze BTC uplatnit. Peněženky nebo účty uživatelů trpí u burz neustále podle obsahu fóra v roce 2012 útokům hackerů. Uživatelé mluví o daních, o poplatcích za transakce, diskutují o budoucnosti BTC a různých službách, o poolech, o ziskovosti BTC a porovnávají systém BTC s Ponziho schématem,o kterém byla zmínka v předchozí subkapitole. Uživatelé se stále snaží odhadnout, kdo byl ve skutečnosti Satoshi Nakamoto, jestli je ještě živ a proč nenavštěvuje fórum.

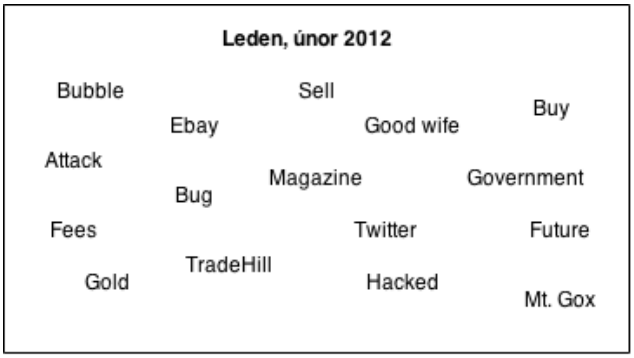


Obrázek 54: Diskuze spojené s Bitcoinem v roce 2012, zdroj Jason Davies a vlastní zpracování

[19] Online servise, který umožňoval provádět CFD (contract-for-difference) při obchodování s BTC (*Bitcoinica*). CFD jsou kontrakty, které se uzavírají mezi kupujícím a prodávajícím a prostřednictvím toho kontraktu určitá strana pak musí zaplatit rozdíl mezi současnou cenou aktiva a domluvenou v tomto kontraktu (Contract for Differences – CFD, 2015).

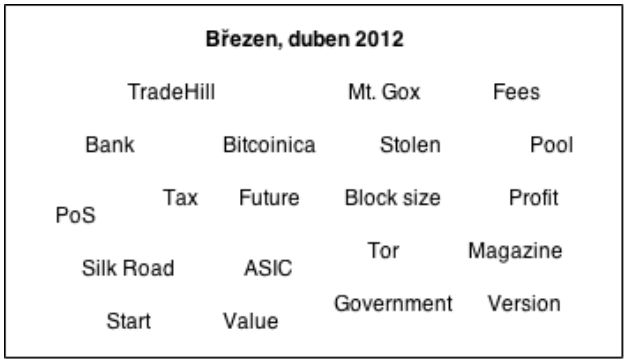
V prvních dvou měsících z předchozích let uživatelé stále diskutují o útocích hackerů, o burzách Mt. Gox a TradeHill a o poplatcích za provedené transakce (viz Obrázek 55).

V lednu roku 2012 uživatelé mluvili o tzv. bublině (angl. bubble), přičemž lze naznačit, že v této době BTC zažíval růst. Pak uživatelé diskutují o možnostech nákupu či prodeji, o přijetí BTC Ebay. Také v tomto měsíci uživatelé píšou o seriálu Good Wife, který v jednom ze svých scénářů použil historii o obchodování s BTC. Dále vychází noviny o BTC, které vyvolaly mezi uživateli velký zájem (viz Obrázek 55). v únoru se pak přidávají diskuze o porovnávání zlata a BTC, o různých chybách v peněženkách anebo systému Bitcoin (viz Obrázek 55).



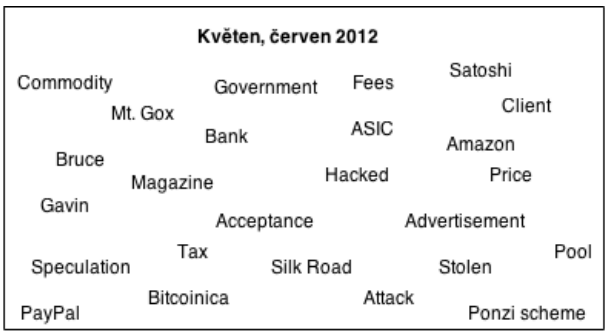
Obrázek 55: Diskuze spojené s Bitcoinem v lednu a únoru roku 2012, zdroj vlastní zpracování

V březnu a květnu (viz Obrázek 56) uživatelé pořád diskutují o burzách TradeHill a Mt. Gox, o vztazích bank a státu k BTC, o něčích ukradených BTC, o poolech a o různých verzích BTC. v těchto měsících uživatelé aktivně začínají diskutovat Bitcoinicu, přičemž Bitcoinica v tento okamžik utrpěla velkou ztrátu z důvodu krádeží útočníky (*Bitcoinica*). Uživatelé pak diskutují o tom, jakou hodnotu má BTC a jestli ji vůbec má. Mezi uživateli se pak stává populární algoritmus těžby Proof of Stake, uživatelé porovnávají PoW a PoS a ptají se na možnosti změny u Bitcoin z PoW na PoS. Dále v této periodě začínají aktivně diskutovat o ASIC a počítat ziskovost těžby BTC. Také mluví i o Silk Road, přičemž podle obsahu dat na fóru bylo zjištěno, že velká část obchodů s BTC probíhá na vrub Silk Road.



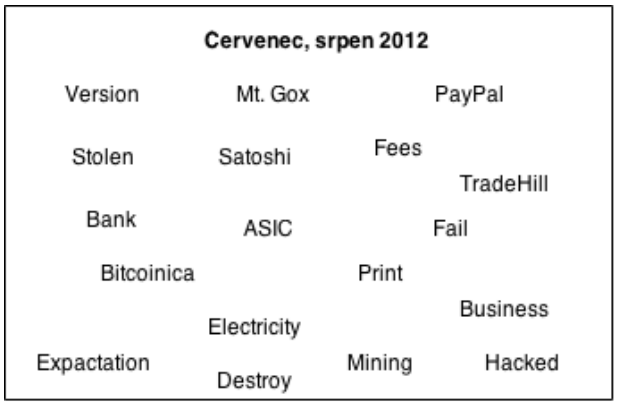
Obrázek 56: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2012, zdroj vlastní zpracování

V dalších dvou měsících roku 2012 (viz Obrázek 57) vznikají další nová témata, uživatelé diskutují a kladou otázky uživatelům se jmény na fóru Gavin a Bruce, kteří pořád aktualizují verzi softwaru Bitcoin a kteří aktivně diskutují o Bitcoin v televizi anebo v novinách. Také uživatelé píšou o tom, že by bylo vhodné, kdyby Amazon přijímal BTC. Zůstávají populární i jiná témata zobrazená na obrázku č. 57 a která byla popsána výše.



Obrázek 57: Diskuze spojené s Bitcoinem v květnu a červnu roku 2012, zdroj vlastní zpracování

V červenci a srpnu (viz Obrázek 58) uživatelé diskutují o vlastních očekáváních ohledně kurzu BTC, panuje myšlenka, že BTC neuspěje. Další témata, která jsou zobrazená na obrázku č. 58, z předchozích měsíců zůstávají mezi uživateli stále populární.

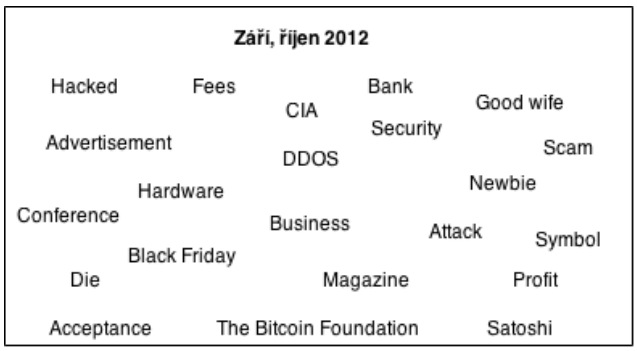


Obrázek 58: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2012, zdroj vlastní zpracování

V září a říjnu (viz Obrázek 59) uživatelé určitou informaci považují za podvod (angl. scam), panuje myšlenka, že BTC umírá, ale zároveň stále více lidí se začíná zajímat o BTC, ptají se, jak s tím začít a kladou další otázky spojené s BTC. Mezi uživateli se pak prosazuje termín DDoS[20], jelikož vznikají problémy tohoto typu. v té době vzniká také nezisková organizace The Bitcoin Foundation, která poskytuje finanční zdroje pro vývoj a infrastrukturu pro Bitcoin (Bitcoin Foundation, 2015). Tuto skutečnost uživatelé projednávali na Bitcoin Forum. Uživatelé také diskutují o tzv. Černém pátku (angl. Black Friday), který probíhá na konci listopadu, během něhož velké obchody zahajují výprodeje (The Black Friday, 2015). Dále mluví o CIA[21] v kontextu, jestli není Satoshi Nakamoto zaměstnancem této organizaci anebo se projednávají možné účely tvorby Bitcoin touto organizací.

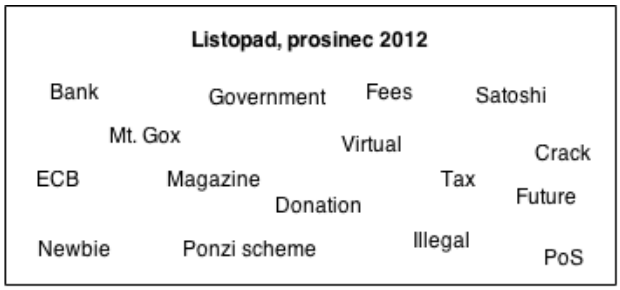
[20] DDoS (Distributed Denial of Service) je to útok, během kterého se skupina lidí domluví na prohlížení nějaké stránky a posílání určitých požadavků na server, tím pádem server bude přetížen a ostatním uživatelům se stránka nenačte (Wift, 2012).

[21] CIA (Central Intelligence Agency), ústřední zpravodajská služba v USA (*Central Intelligence Agency*).



Obrázek 59: Diskuze spojené s Bitcoinem v září a říjnu roku 2012, zdroj vlastní zpracování

V listopadu a prosinci (viz Obrázek 60) vzniká mezi uživateli nové téma diskuze, a to vztah Evropy a Evropské centrální banky (ECB) k Bitcoin. Ostatní témata byla popsána výše.



Obrázek 60: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2012, zdroj vlastní zpracování

6.3.1.2 Vývoj témat v roce 2013

V roce 2013 uživatelé stále diskutují na témata, která jsem vyloučila z výsledků v roce 2012, a proto ze stejného důvodu je vyloučím také z výsledků roku 2013. K vyloučení přidávám také stát (angl. governmnet), Satoshi, rozměr bloků (angl. block size), sítě bloků (angl. blockchain) a banky, o kterých uživatelé mluví každý měsíc v roce 2013.

V roce 2013 zůstává některá témata z předchozích let stále populární (viz Obrázek 61). Takovými tématy jsou útoky, PayPal, aktualizování verzí, konference, krádež BTC, Mt. Gox a ASIC. Vznikají nová témata diskuze, která byla vyvolaná určitými událostmi, například uživatelé začali diskutovat o Dwolla[22], o Maxu Keiserovi[23], o FBI[24], ATM[25] a o NSA[26].

[22] Dwolla je organizace pro elektronické obchodování poskytující služby online plateb prostřednictvím různých koncových zařízení, včetně mobilních (Dwolla, 2015).

[23] Max Keiser je známý moderátor v televizi (Max Keiser, 2015).

[24] FBI (Federal Bureau of Investigation) je Federální úřad pro vyšetřování (*FBI*).

[25] ATM (Automated Teller Machine) je zařízení umožňující vyměnit BTC a hotovost (Automated Teller Machine – ATM, 2015).

[26] NSA (The National Security Agency/Central Security Service) je Národní bezpečnostní agentura/Centrální bezpečnostní službou, která je státní organizací zodpovídající za kryptografii (NSA, 2009).



Obrázek 61: Diskuze spojené s Bitcoinem v roce 2013, zdroj Jason Davies a vlastní zpracování

V lednu a únoru roku 2013 (viz Obrázek 62) uživatelé diskutují o činnosti uživatele Gavin, který neustále aktualizuje verze Bitcoin. Mluví o tzv. nezpracovaném mléku (angl. raw milk), který se v některých okresech považuje za nelegální produkt (*Real raw milk facts*). Uživatelé diskutují o konceptu BTC, o jeho ceně, přičemž podle grafu č. 9 si lze všimnout, že cena BTC začíná růst. Stále ale mluví o tom, že BTC může krachnout, ptají se na možnosti tvorby kreditních karet a diskutují o výhrách v určitých hrách spojených s BTC. Stává se populárním i téma tzv. fork, které znamená opravující činnost, prováděnou z důvodu aktualizování verzí a následnou nekompatibilitou mezi programy, která vyvolává chyby a která v důsledku může způsobit chyby i v síti bloků (Kirk, 2014).



Obrázek 62: Diskuze spojené s Bitcoinem v lednu a únoru roku 2013, zdroj vlastní zpracování

Na konci roku 2012 byl nalezen 210 000. blok, který způsobil snížení odměny v dvakrát za nalezený blok (*Blockchain Info*). v březnu a v dubnu v roce 2013 pak tuto skutečnost aktivně projednávají. Uživatelé mluví v těchto měsících také o IRS[27], jelikož v předchozích letech, jak bylo popsáno v předchozí subkapitole, uživatelé mluvili o zdanění BTC. o Bitcoin se začal zajímat americký moderátor Max Keiser (Max Keiser, 2015), jeho vysílání pak vyvolala aktivní diskuzi na Bitcoin Forum. Uživatelé také projednávají využití BTC v terorismu, využití

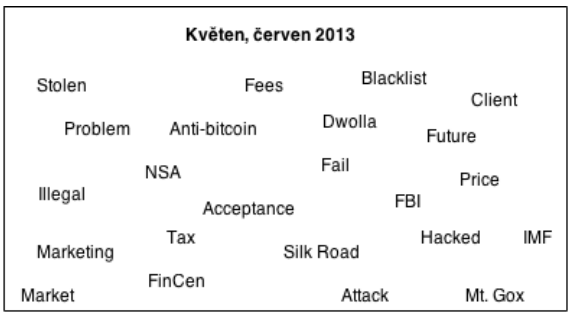
[27] IRS (The Internal Revenue Service) je státní orgán zodpovídající za sběr daní a kontroluje dodržení daňových zákonů (*IRS*).

ATM, o možnosti změny pravidel Bitcoinu, DDoS, o odměně za naleznutí nového bloku atd. (viz Obrázek 63).



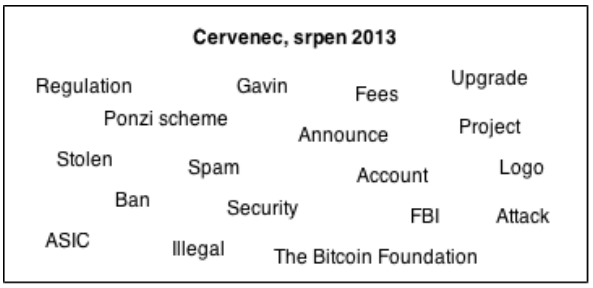
Obrázek 63: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2013, zdroj vlastní zpracování

V květnu a červnu roku 2013 (viz Obrázek 64) stále více uživatelé mluví o státních orgánech, například o FinCen[28], FBI, NSA a IMF[29]. v tomto roce FBI uzavřelo webovou stránku Silk Road a zavřelo Ross William Ulbricht (*Mt.Gox*), což vyvolalo diskuzi na Bitcoin Forum. Také byla diskutována událost, kdy Dwolla zrušil všechny transakce posílané z účtu Mt. Gox (*Mt.Gox*). v těchto měsících se na fóru prosazuje opozice Bitcoin. Ostatní témata z obrázku č. 64 byla popsána výše.



Obrázek 64: Diskuze spojené s Bitcoinem v květnu a červnu roku 2013, zdroj vlastní zpracování

V červenci a srpnu roku 2013 (viz Obrázek 65) uživatelé pokračují v diskuzi na témata, která jsem popsala výše. Zajímavým faktem je podle obsahu fóra během této periody, že americké banky uzavíraly účty svým zákazníků, kteří měli nějaký vztah k Bitcoinu.

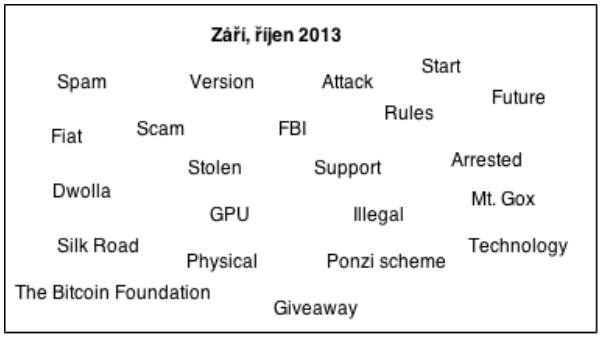


Obrázek 65: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2013, zdroj vlastní zpracování

[28] FinCen (The Financial Crimes Enforcement Network) je americký státní orgán, který sbírá a analyzuje informace o finančních transakcích (*FinCEN*).

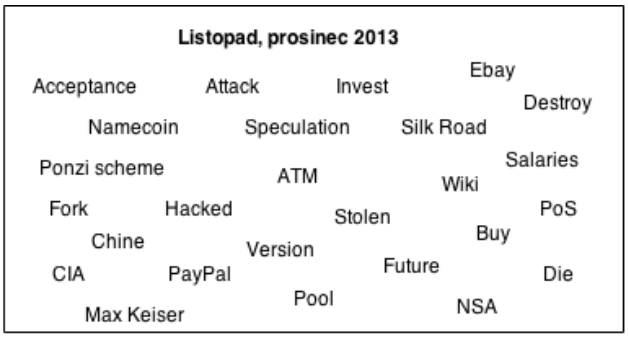
[29] IMF (International Monetary Fund) je mezinárodní měnový fond, jehož primárním cílem je podpora stability měnového systému, přesněji systému měnových kurzů a mezinárodních platebních operací (*International Monetary Fund).*

V září a říjnu roku 2013 se prováděl giveaway, uživatelé aktivně diskutovali o pravých měnách, psali o tom, že GPU už nemá postačující výkonnost a je nepoužitelný, dále také projednávali fyzickou podobu BTC atd. (viz Obrázek 66).



Obrázek 66: Diskuze spojené s Bitcoinem v září a říjnu roku 2013, zdroj vlastní zpracování

V listopadu a v prosinci uživatelé nejvíce diskutují na téma uvedená na obrázku č. 67, většina už byla popsána výše. Na konci roku 2013 uživatelé projednávají legalizaci či zákaz Bitcoinu v Číně. Také diskutují o tom,, jestli nakupovat či prodávat BTC, mluví také o tom, že v některých státech si zaměstnanci některých firem mohou vybrat, jestli dostat plat v pravé měně, anebo v kryptoměně BTC.



Obrázek 67: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2013, zdroj vlastní zpracování

6.3.1.3 Vývoj témat v roce 2014

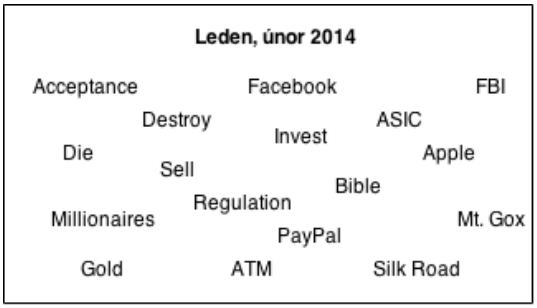
Dále ve výsledcích roků 2014 a 2015 vylučuji útoky (angl. attack a hacked) a poplatky za provedené transakce (angl. fees).

Souhrn témat v roce 2014 je zobrazen na obrázku č. 68. Uživatelé pořád diskutují o přijetí BTC jinými institucemi, mluví o možnosti investování, o událostech spojených se Silk Road, srovnávají BTC s Ponziho schématem, diskutují o službách PayPal, píšou o možnosti bohatnout anebo o lidech, kteří zbohatli prostřednictvím obchodování s BTC. Stále mluví o burze Mt. Gox, píšou o krádežích, o daních a o vztahu se státními institucemi. Další témata jsou zobrazená na obrázku č. 68.



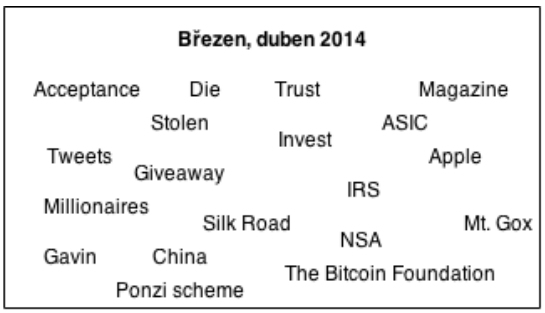
Obrázek 68: Diskuze spojené s Bitcoinem v roce 2014, zdroj Jason Davies a vlastní zpracování

Při prohlédnutí vývoje témat podle měsíců vidíme, že v lednu a únoru roku 2014 uživatelé diskutovali o přijetí BTC různými organizacemi nebo státy. Stále si je určitá skupina uživatelů jistá, že BTC nemá budoucnost a zkrachuje. Určitá skupina uživatelů chce zničit BTC, podle obsahu fóra má stát také zájem o zničení Bitcoinu. Na Bitcoin Forum pak vznikají příspěvky o Bibli v kontextu s Bitcoin. Sociální síť Facebook začíná testovat BTC platby za poskytování prostoru na reklamu. Uživatelé píšou i o firmě Apple, která v tu dobu byla proti BTC a vlastníci zařízení od této firmy měli problémy při obchodování s touto kryptoměnou. Ostatní témata byly okomentovaná dříve v této diplomové práci (viz Obrázek 69).



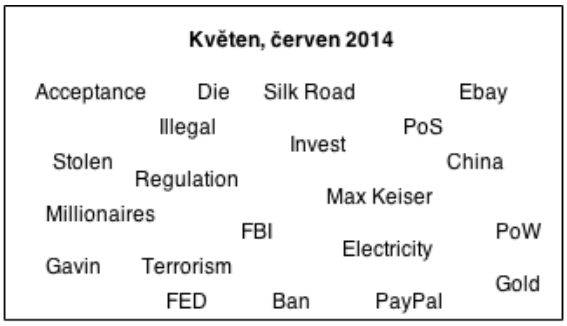
Obrázek 69: Diskuze spojené s Bitcoinem v lednu a únoru roku 2014, zdroj vlastní zpracování

V březnu a dubnu roku 2014 stále uživatelé diskutují na předchozí témata (viz Obrázek 70), přidává se diskuze na témata Číny, která v té době zakazovala obchodování s BTC (Urbánek, 2013). Tato skutečnost ovlivnila kurz BTC, který kvůli tomu také ještě na konci roku 2013 začal padat. Uživatelé diskutují o příspěvcích spojených s BTC, které byly publikované na sociální síti Twitter. Prosazuje se myšlenka o tom, že těžba ASIC zničí BTC. Ostatní témata byla okomentovaná výše.



Obrázek 70: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2014, zdroj vlastní zpracování

V květnu a červnu uživatelé píšou o vztahu Ebay k BTC, o jednání FBI a FED[30], přičemž se prosazuje myšlenka, že FED chce Bitcoin regulovat. Dále diskutují o dalších zákazech obchodování BTC v různých státech, například v Číně nebo v Rusku. Ostatní témata byla popsána výše (viz Obrázek 71).



Obrázek 71: Diskuze spojené s Bitcoinem v květnu a červnu roku 2014, zdroj vlastní zpracování

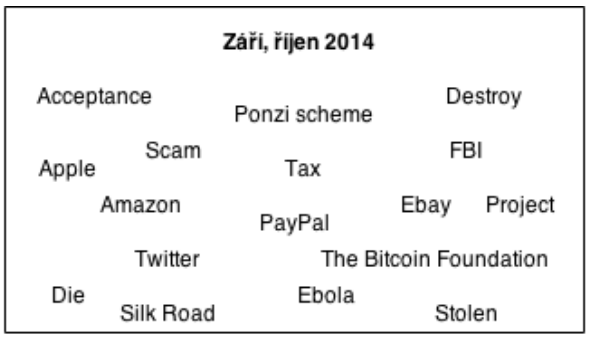
Uživatelé diskutují o dalších zákazech obchodování s BTC, v červenci například píšou o zákazu BTC v Bolívii. Prosazuje se myšlenka, že BTC je projektem CIA. Také uživatelé mluví o tom, že NSA chce zničit Bitcoin. Ostatní témata byla popsána výše (viz Obrázek 72).



Obrázek 72: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2014, zdroj vlastní zpracování

V září a říjnu roku 2014 uživatelé diskutují kromě témat uvedených na obrázku č. 73 také o možnosti přijetí BTC Amazonem, o možnosti napsání různých dopisů Amazonu anebo petici. Také se objevuje diskuze o Ebole, uživatelé chtěli darovat určitou částku BTC obětem této nemoci. Ostatní témata byla popsána výše.

[30] FED (Federal Reserve System) je Federální rezervní systém nebo americkou centrální bankou (FED, 2015).



Obrázek 73: Diskuze spojené s Bitcoinem v září a říjnu roku 2014, zdroj vlastní zpracování

V posledních měsících roku 2014 uživatelé diskutují o možnosti zavedení kreditních karet pro účty s Bitcoiny. Stále diskutují na téma přijetí BTC jinými organizacemi, o hrách, kde lze ztratit BTC, o drogách a Silk Road. Také porovnávají zlato s Bitcoinem, BTC srovnávají s Ponziho schématem, píšou o důvěryhodnosti BTC nebo služeb poskytujících možnost držení a obchodování s BTC atd. (viz Obrázek 74).



Obrázek 74: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2014, zdroj vlastní zpracování

6.3.1.4 Vývoj témat v roce 2015

Jelikož data jsou stáhnutá do konce března roku 2015, poslední rok 2015 obsahuje pouze první tři měsíce.

Diskutovaná témata jsou zobrazená na obrázku č. 75, přičemž je vidět, že stále je aktuální téma krádeží a více diskutovanou se stává téma investování. Také se mluví o tom, co nebo kdo může BTC zničit. Z nových témat lze vyčlenit diskuze o tvorbě Bitcoin banků a o možném úpadku BTC. Ostatní témata jsou popsána v předchozích subkapitolách.



Obrázek 75: Diskuze spojené s Bitcoinem v roce 2015, zdroj Jason Davies a vlastní zpracování

6.3.1.5 Vývoj kurzu

Z výše uvedené identifikaci hlavních témat je vidět, že na základě obsahu tvořeného na fóru nelze predikovat vývoj kurzu. Uživatelé na fóru reagují na různé události spojené s Bitcoinem. Při rozhodování o nákupu či prodeji zájemce může sledovat příspěvky na Bitcoin Forum, kde uživatelé publikují a diskutují na události, které byly před tím zveřejněny v určitých externích zdrojích.

Vývoj kurzu libovolné kryptoměny závisí na poptávce a nabídce po ní. Jelikož cílem této práce je propojit hlavní témata s vývojem kurzu prostřednictvím výše zjištěných diskuzí, budu předpokládat faktory ovlivňující kurz a také to, jestli uživatelé reagují na změnu kurzu a jak. Na základě obsahu tvořeného na Bitcoiun Forum se dá tvrdit, že vývoj kurzu BTC/USD ovlivní několik následujících faktorů:

Změna náročnosti

Podle výše uvedené identifikaci témat je vidět, že uživatelé v průběhu času reagují na změnu náročnosti a diskutují na témata spojená s příležitostí pro těžbu. Vysoká náročnost může odradit těžaře od těžby nových mincí, pokud jejich těžba nebude odměněna částkou, která jim vynahradí náklady.

Podle diskuzí na Bitcoin Forum uživatelé píšou i o zvýšené spotřebě elektřiny z důvodu růstu náročnosti nalezení bloků a i o následné potřebě pořízení výkonnějšího zařízení. v průběhu analyzovaných let je vidět, jak se mění tendence v pořízení výkonnějších zařízení. Také je vidět tendenci sdružování těžařů do poolů, které umožní nalézt nové bloky rychleji. Proto se dá předpokládat, že změna spotřeby elektřiny a výkonnosti zařízení je v přímé korelaci se změnou náročnosti.

Odměna těžařům a poplatky za transakce v USD

Při nalezení nového bloku těžaři dostávají odměnu. v době psaní této diplomové práce byla odměna 25 BTC za nový blok. Do listopadu roku 2012 tato odměna činila 50 BTC (*Bitcoin Forum*). Také těžaři dostávají za své služby podpory systému odměnu ve výši poplatků za provedené transakce, které by měly vynahrazovat náklady spojené s těžbou. Poplatky spojené s transakcemi ovlivní také chování zájemců, jelikož s vysokými poplatky nebudou chtít využívat tuto kryptoměnu a tím bude klesat zájem o ni a následně i její cena.

Podle identifikace hlavních témat je vidět, že jedno z nejdiskutovanějších témat je diskuze o transakcích a jejich poplatcích, takže lze tvrdit, že tohle ovlivní zájem uživatelů o kryptoměnu.

Počet transakcí

Zájem vlastníků kryptoměny lze sledovat na vývoji počtu provedených transakcí.

Regrese

Prostřednictvím zveřejněných dat na Blockchain Info byla sestavená regresní rovnice P=f(D, T, F),

- kde P = kurz BTC/USD,
- D = náročnost,
- T= počet transakcí,
- F = poplatky za transakce a odměna za naleznuté bloky v USD.

Index determinace sestaveného regresního modelu je 0,98, což napovídá o přijatelné kvalitě modelu. Model platí dlouhodobě, krátkodobé výkyvy nevysvětluje.

Regresní rovnice byla sestavená na základě dat od 1. 1. 2011 a do 31. 12. 2014, na základě dat od 1. 1. 2015 a do 31. 3. 2015 byla rovnice ověřená.

$$P = -6{,}68178 + 0{,}000000002 * D + 0{,}00019 * T + 0{,}00002 * F \qquad (6.1)$$

Reziduem pak ostatní faktory, které ovlivňují vývoj kurzu a které lze částečně identifikovat podle diskuzí tvořených na Bitcoin Forum.

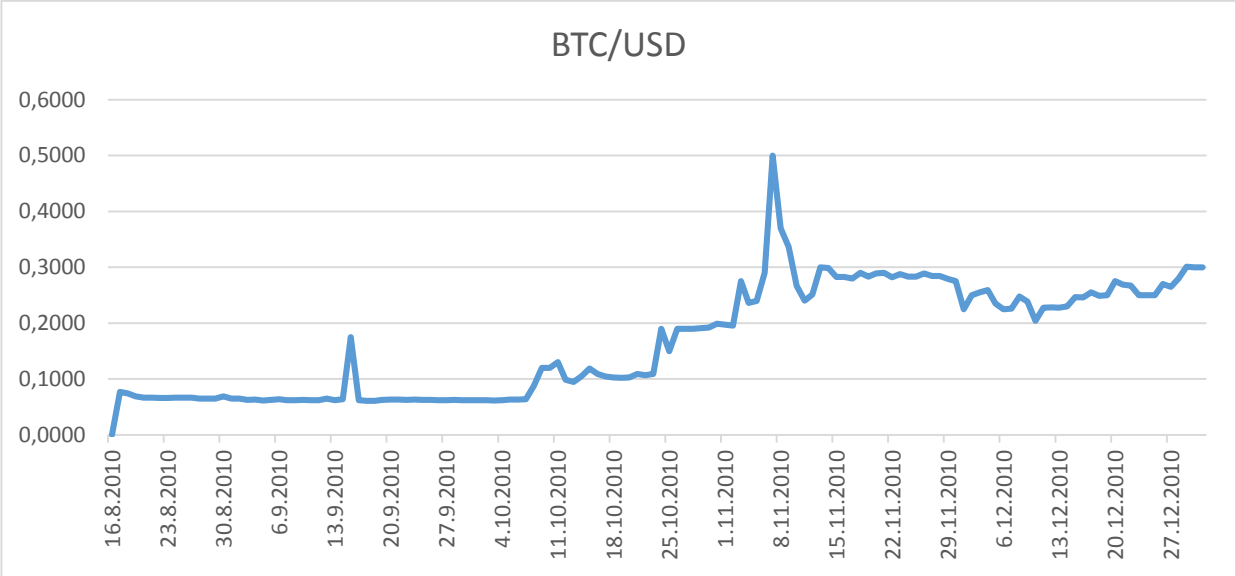
Je třeba tady také zmínit, že počet transakcí a výše poplatků za provedené transakce jsou také ovlivněné ostatními faktory, které uvádím v této práci dále.

Ostatní faktory

Dalšími faktory, které ovlivní kurz podle obsahu Bitcoin Forum, jsou ekonomická situace zemí[31], jednání států ohledně BTC, marketingové akce, vývoj nelegálních obchodů a činnosti prováděné státem pro zákaz těchto obchodů, přijetí BTC jinými obchody a službami, jinými slovy, kde by uživatelé mohli tratit své BTC, a pak důvěra v systém. Podle cílů této diplomové práce uvedu vývoj kurzu v jednotlivých letech a porovnám s obsahem tvořeným na Bitcoin Forum.

Jelikož podle Blockchain Info kurz BTC/USD začal pomalu měnit svou hodnotu od srpna roku 2010, začnu porovnávat kurz a hlavní témata od této doby.

Od srpna roku 2010 se kurz Bitcoinu začal vůči americkému dolaru růst (viz Graf 6). Podle výše uvedené identifikace hlavních témat je vidět, že se uživatelé zajímají o kolik je Bitcoin důvěryhodný, anonymní a latentní. Také Bitcoin se začíná vystavovat útokům hackerů, což potvrzuje rostoucí zájem o kryptoměny. Od října kurz BTC/USD začal růst a podle identifikace hlavních témat je vidět růst zájmu uživatelů o kryptoměnu BTC, který se projevuje kromě ostatních v tom, že uživatelé píšou o businessu, o tom, že BTC nechali vyměnit za USD a o tom, kde by mohli své BTC uplatnit.



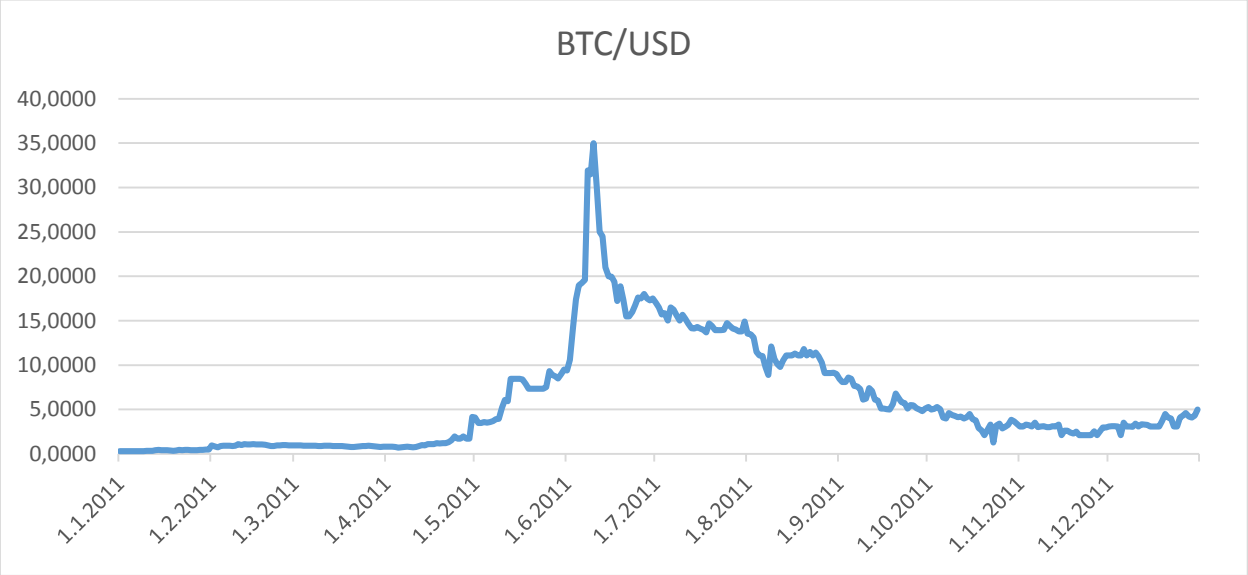
Graf 6: Vývoj kurzu BTC/USD v roce 2010, zdroj Blockchain Info a vlastní zpracování

V roce 2011 uživatelé, jak bylo zjištěno výše, pokračují od roku 2010 BTC propagovat, čímž kromě jiných faktorů vyvolali větší zájem o kryptoměnu. Dále jak je vidět na grafu č.

[31] Analýza ekonomické situace zemí není zájmem této diplomové práce, a proto jí nebude věnována pozornost.

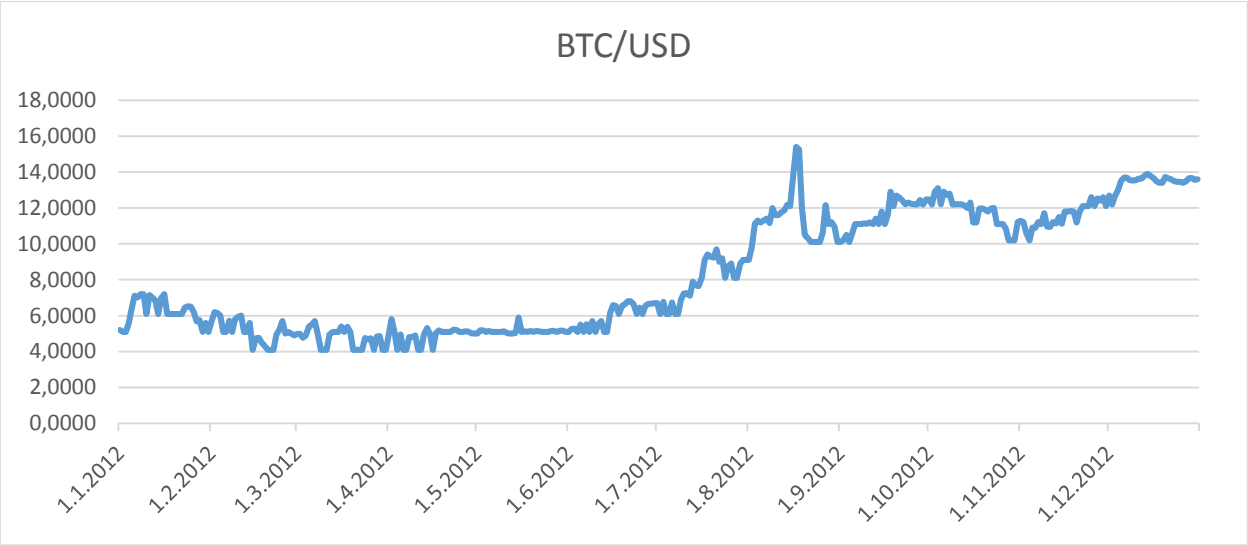
7, v červnu kurz BTC/USD šel prudce nahoru a uprostřed června pak padal, podle obsahu tvořeného na Bitcoin Forum je vidět, že uživatelé na toto padnutí reagovali tím, že doporučovali BTC prodávat. Také v této době byl útok na burzu Mt. Gox, což mohl způsobit pád kurzu a zase uživatelé na tuto skutečnost reagovali na fóru.

Na pokles kurzu pak uživatelé reagují příspěvky o krachu BTC, také vzrůstá počet uživatelů, kteří negativně píšou o BTC, což může ovlivnit i na názory ostatních uživatelů vůči BTC. Reakci na padnutí kurzu u uživatelů také bylo provedení nějakých akcí na povzbuzení kurzu, například uskutečnění tzv. giveaway.



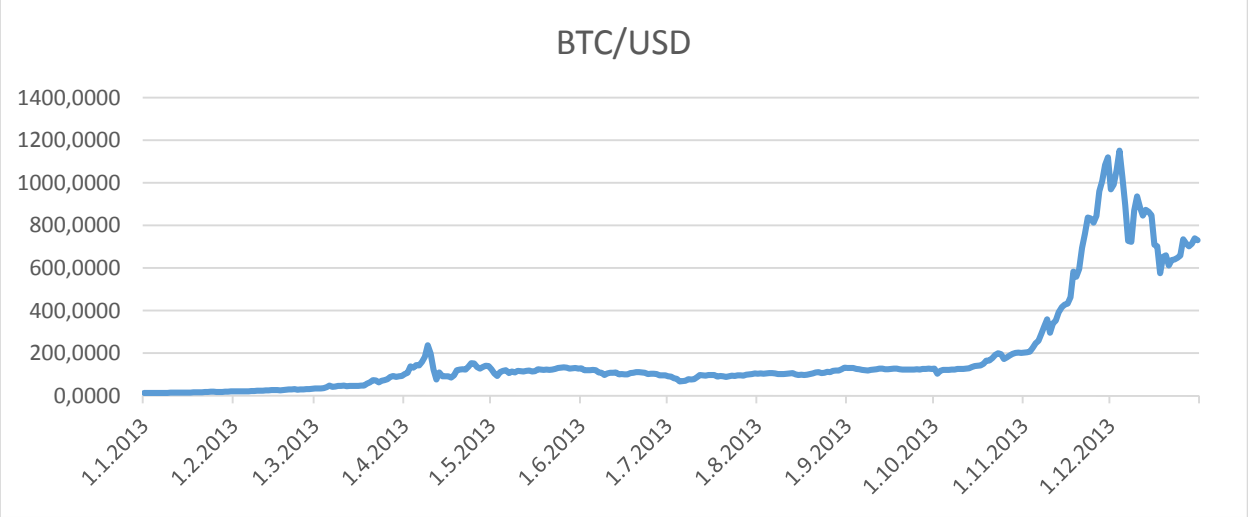
Graf 7: Vývoj kurzu BTC/USD v roce 2011, zdroj Blockchain Info a vlastní zpracování

V roce 2012 lze sledovat růst kurzu, přičemž uživatelé na něj reagovali na Bitcoin Forum, psali o růstu kurzu nebo diskutovali o bublině. Uživatelé také jsou docela aktivní při doporučení na nákup či prodej BTC. Tady ale musím říct, že na základě této informace nejde stanovit, jestli lidi budou více nakupovat či prodávat, jelikož názor na nákup či prodej mezi zájemci je rozdílný, a proto určitá část lidí může nakupovat a jiná část prodávat. Postupný nárůst kurzu, jak je vidět z grafu č. 8, se uskutečnil od půlky června, na což uživatelé reagovali v červenci příspěvky ohledně vlastních očekávání na kurz. Také koloval názor o neúspěchu kurzu, který podle porovnání s grafu č. 8 byl opodstatněný, jelikož na konci srpna kurz padal. Následné kolísaní kurzu mimo jiné faktory mohlo by být vyvoláno podle identifikovaných témat útoky typu DDoS, vzniku The Bitcoin Foundation, Černému pátku a vztahu bank a států k BTC.



Graf 8: Vývoj kurzu BTC/USD v roce 2012, zdroj Blockchain Info a vlastní zpracování

Jak je vidět na grafu č. 9, kurz šel nahoru, přičemž z obsahu tvořeného na Bitcoin Forum se dá předpokládat, že stále více lidí se zajímá o Bitcoin, jelikož o Bitcoinu už začínají mluvit známí moderátoři, staví ATM přijímající Bitcoiny a postupně se zvyšuje počet obchodů nebo služeb přijímajících BTC. Na konci listopadu, jak je vidět na grafu č. 9, šel kurz prudce nahoru, uživatelé pak na tuto skutečnost reagovali doporučením více nakupovat či prodávat, což podle ekonomických zákonů odpovídá růstu kurzu.

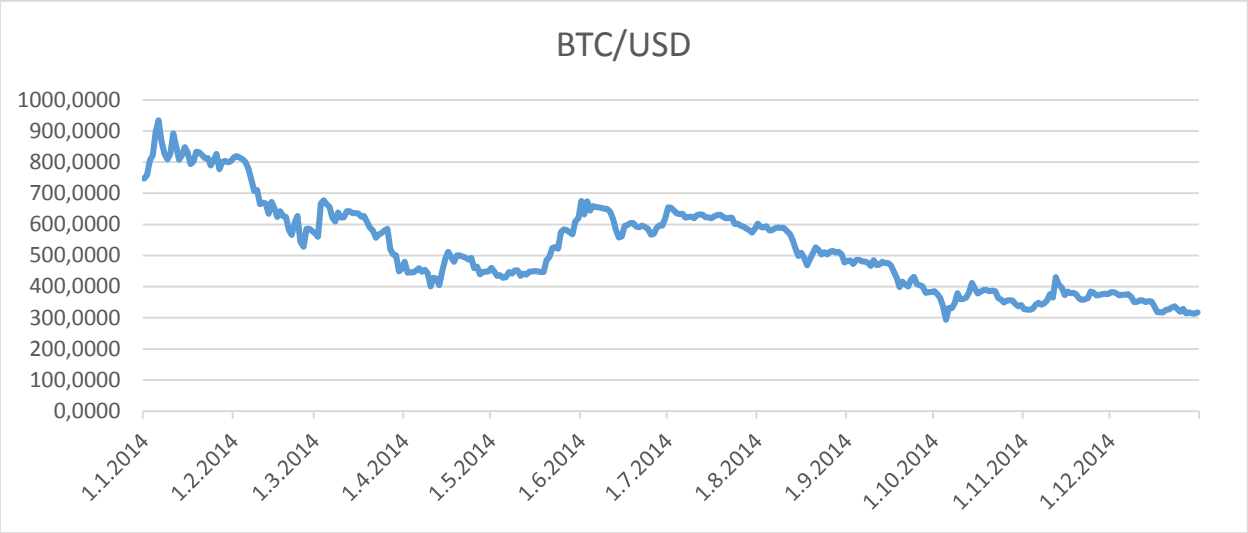


Graf 9: Vývoj kurzu BTC/USD v roce 2013, zdroj Blockchain Info a vlastní zpracování

Na začátku roku 2014 je vidět pokles kurzu BTC/USD (viz Graf 10), na který uživatelé fóra reagují diskutováním o různých událostech čerpaných z externích zdrojů. Za hlavní příčinu poklesu kurzu považují zákaz obchodování s BTC v různých státech jako v Číně a v Rusku, někteří uživatelé prosazují názor o pádu BTC nebo jeho zničení. v lednu a únoru také na fóru převládá nálada k prodeji BTC, přičemž na grafu č. 10 je vidět, že vlastnici BTC se snaží prodat, čímž padá cena Bitcoinu.

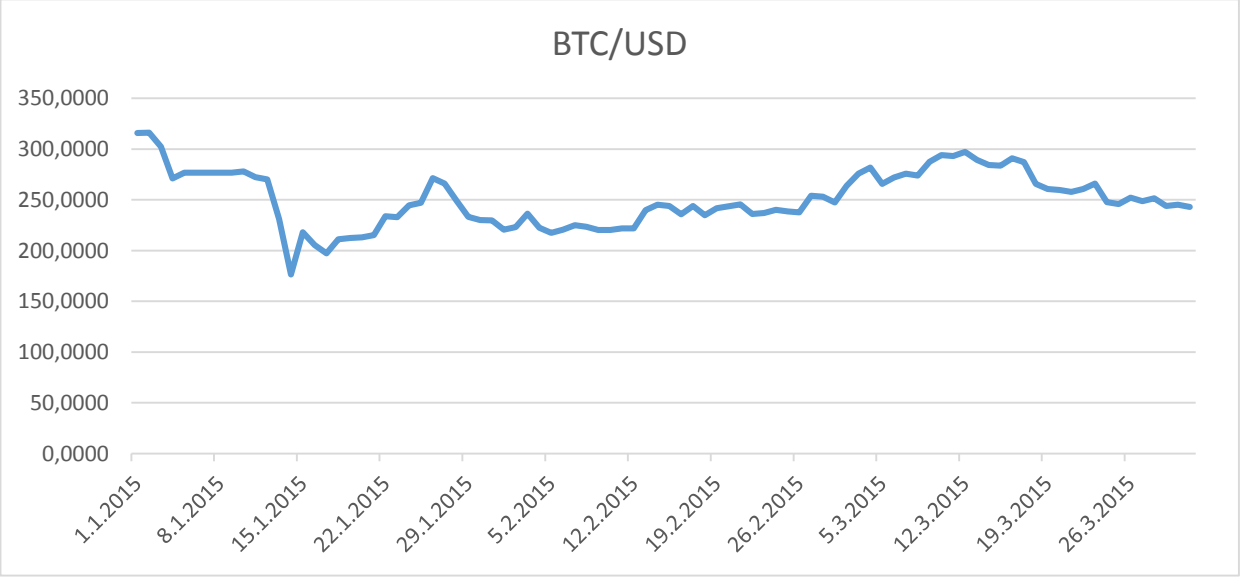
Důvěryhodnost v tuto kryptoměnu u uživatelů nepadá do nulové hodnoty, přičemž jak je vidět na grafu č. 10, kurz na konci května šel nahoru, což napovídá o aktivním nakupování kryptoměny. Tuto skutečnost lze spojit s obsahem tvořeným na Bitcoin Forum tak, že uživatelé, i když kurz padal, se stále ptali na příležitosti investování do této kryptoměny, což napovídá o neutuchajícím zájm uživatelů o BTC.

Následující tendence poklesu kurzu podle obsahu Bitcoin Forum je vyvolána opatřeními států a bank proti obchodům uskutečněných prostřednictvím kryptoměny Bitcoin.



Graf 10: Vývoj kurzu BTC/USD v roce 2014, zdroj Blockchain Info a vlastní zpracování

Na začátku roku 2015 je kurz poměrně stabilní (viz Graf 11). Podle identifikovaných témat lze soudit, že státy a banky mají negativní postoj k Bitcoinu a snaží se potlačovat obchody s tím spojené. Uživatelé ale nejsou zcela ovlivněni touto náladou a pokračují v obchodování s Bitcoinem. Na fóru na začátku roku 2015 psali o lidech, kteří pomocí BTC zbohatli, a o příležitostech znovu zbohatnout. Ptají se na pozitiva či negativa investování do této kryptoměny. a proto kurz nepadá na nulovou hodnotu.



Graf 11: Vývoj kurzu BTC/USD v roce 2015, zdroj Blockchain Info a vlastní zpracování

6.3.1.6 Dílči shrnutí

Z výše provedené analýzy a interpretace dat je vidět tendenci v postupných změnách v tématech.

V době vzniku uživatelé projednávali převáženě technická témata a ptali se na koncept Bitcoinu.

Následně se uživatelé začínají zajímat o výši poplatků za provedené transakce, o budoucí těžbu, o možnosti obchodování, o poskytovatele služeb pro směnu BTC na jinou měnu. Také Bitcoin začíná být vystavován útokům hackerů, což naznačuje zvýšení zájmu o tuto kryptoměnu. Uživatelé také aktivně píšou o tom, že se vyměnily pravé měny na BTC. Panuje téma propagace kryptoměny. Uživatelé diskutují i o jednání států, státních a finančních institucí, což naznačuje, že se Bitcoin prosazuje do života velké skupiny lidí. Zájemci o kryptoměnu BTC také píšou o podvodech, spekulacích a o nelegálních obchodech. Snaží se seskupit do poolů pro zefektivnění těžby. Uživatelé Také diskutují o peněženkách, transakcích a jejich poplatcích, adresách, chybách a o burzách.

V určitou dobu panuje myšlenka na krachu Bitcoinu. Uživatelé píšou na Bitcoin Forum méně aktivně.

Z výsledků analýzy dat bylo zjištěno, že vývoj kurzu BTC/USD je dlouhodobě ovlivněn několika různými faktory, mezi něž patří změna náročnosti, počet transakcí, odměna těžařům a poplatky za transakce. Krátkodobě pak kurz ovlivní ostatní faktory, kterými jsou ekonomické a politické události. Uživatelé na fóru kurz nepredikují, převážně o nákupech či prodejích kryptoměny Bitcoin nepíšou, ale občas na výkyvy kurzu na fóru reagují, a to nejvíce na externí události ekonomického a politické charakteru.

Také prostřednictvím analýzy příspěvků spojených s Bitcoin nebo BTC je vidět tendenci ve změně výběru technologií pro těžbu, zvýšení či pokles diskutování ohledně reklamy, zvýšení či pokles zájmu o kryptoměnu mezi uživateli a také ostatními institucemi.

Využití této komunity pro predikci kurzu BTC/USD není správným řešením, jelikož se uživatelé nezabývají pouze diskutováním o vývoji kurzu, a také není sledované časté doporučení na nákup či prodej. Toto fórum je využitelné při určitých technických problémech, vysvětlení možné budoucnosti kryptoměny a informování o určitých externích událostech.

6.3.2 Nxt

Další známou a mezi zájemci o kryptoměny populární je kryptoměna Nxt. Nxt je jedna z kryptoměn, která je založená na algoritmu těžby mincí Proof of Stake. Tento algoritmus eliminuje jeden z nedostatků kryptoměn založených na Proof of Work v nutnosti nakupování drahých a výkonných technologií, také prostřednictvím PoS eliminuje nedostatek velké spotřeby elektřiny, jak je to u PoW. Další nedostatek, který se ale pomocí PoS neodstraňuje, ale snižuje, je možnost centralizace těžby mincí nebo tzv. 51% útoku (podíl jedním těžařem na celkové síti). Tím pádem obě metody těžby mají sklon k centralizaci.

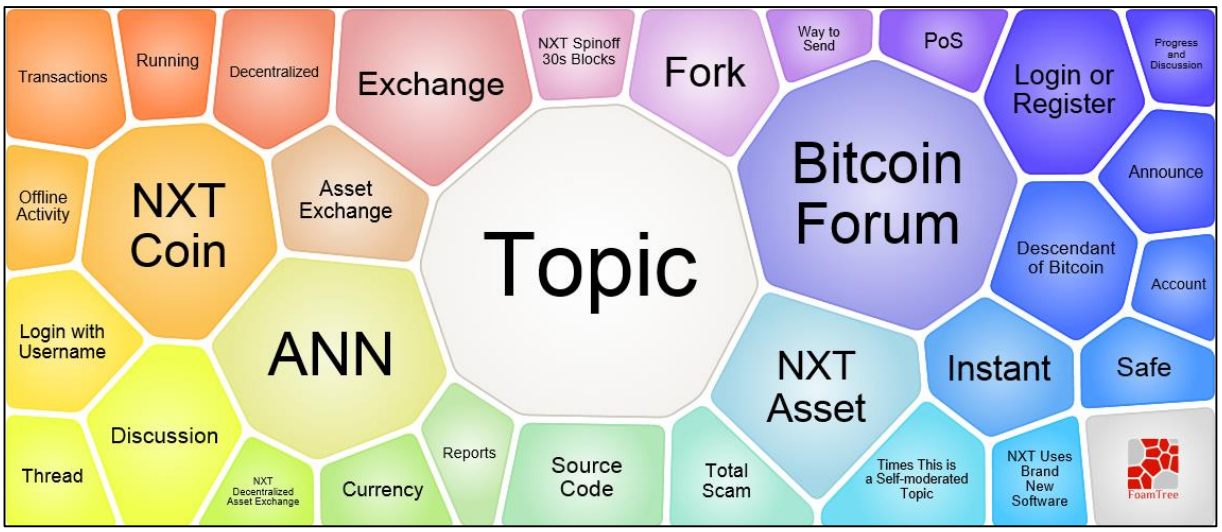
Kryptoměna Nxt byla spuštěná v září roku 2013. Maximální množství, které bylo puštěno do prvního bloku a do ruky uživatelů, je 1 000 000 Nxt, kteří za to zaplatili určitou částku BTC (*Bitcoin Forum*). Tento princip, jak bylo uvedeno v kapitole "Kryptoměna", je založen na době držby kryptoměny Nxt v peněžence jejího vlastníka. Odměnu za nalezený blok pak těžař dostává z komise za provedené transakce nebo za jiné služby ukázané v tabulce č. 9. Těžební algoritmus je SHA-256, průměrná doba generace bloků je jedna minuta (viz Tabulka 9).

Tabulka 9: Obecná informace o Nxt, Nxt (2014), Nxt Blockchain Explorer, Bitmakler a vlastní zpracování, http://www.coinssource.com/crypto-coins/nxt/

Doba spuštění	28. 09. 2013
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof of Stake
Maximální množství mince	1 000 000 000 NXT
Těžební algoritmus	SHA-256
Odměna	Minimální komise je 1 NXT. Komise z transakcí, odeslání zpráv, registrace v tzv. Alias Systém, nákup aktiv na burze nebo v obchodě digitálních věcí (Digital Goods Store).
Počet nalezených bloků (5. 4. 2015 22:03)	396 888

6.3.2.1 Hlavní témata

Prostřednictvím algoritmu "Lingo" byl vytvořen cluster "Nxt" na celkový objem dat (viz Obrázek 76). Z toho clusteru je vidět, s jakými slovy je spojeno "Nxt" na Bitcoin Forum. O Nxt se mluví nejvíce jako o minci, měně a potomku Bitcoinu. i když vývojář tvrdí, že Nxt není větví Bitcoinu a kód Bitcoinu nebyl okopírován (*Bitcoin Forum*), stejně se na fóru šíří myšlenka o tom, že Nxt je větví Bitcoinu. v kontextu s Nxt se mluví o PoS (Proof of Stake), o okamžitých transakcích, o možnosti rozběhnutí, decentralizaci, o bezpečnosti, o způsobech posílání kryptoměny, o účtech, o odvozené kryptoměně od Nxt a o novém softwaru. Také uživatelé diskutují o tom, jestli Nxt není podvodem (angl. scam). Často se vyskytuje slovní spojení "Asset Exchange", to je zabudováno do funkce Nxt softwaru, umožňující směnné operace s kryptoměnou Nxt (Nxt, 2015). Prostřednictvím algoritmu shlukování "STC" jsem dostala další výsledky, o kterých se mluví v kontextu s Nxt. Uživatelé mluví o systémech, které jsou příležitostí pro investování.



Obrázek 76: Cluster "Nxt", zdroj vlastní zpracování

6.3.2.2 Vývoj témat v roce 2013

Jelikož kryptoměna Nxt vznikla v roce 2013, zmapovala jsem hlavní témata od tohoto roku.

V roce vzniku kryptoměny Nxt uživatelé fóra nejvíce diskutovali o tom (viz Obrázek 77), že Nxt je potomkem Bitcoinu, a jestli není pouze jeho kopií. Diskutovali o zdrojovém kódu této kryptoměny. o Nxt pak mluvili jako o podvodu, ptali se na potvrzení transakcí a na základní koncept této kryptoměny. Po spouštění této kryptoměny bylo zapotřebí něco otestovat, o čemž se také psalo na Bitcoin Forum. v kontextu s Nxt se mluvilo o algoritmu PoS, o tzv. giveaway, kde se v době vzniku všechny mince rozpouštěly mezi zájemci. Uživatelé se také ptali na způsob těžby a směny této kryptoměny na jinou kryptoměnu nebo pravou měnu. Také uživatelé diskutovali o logu této kryptoměny a o jejím názvu. Mluvili o připojení zpráv k posílaným mincím, o jiných možných projektech, které lze vytvořit na základě Nxt, o obchodování s Nxt, kde lze koupit či prodat tuto kryptoměnu. Také psali o vytvoření bankovek pro Nxt, o Nxt online peněžence, o možných spekulacích a o ziskovosti těžby. Dále uživatelé mluvili o Ponziho hře, která demonstruje Ponziho schéma (*The Ponzi Game*). Také mluvili o zainteresovaných osobách, o využití nového softwaru pro Nxt, o API a o možnosti vývoje barevných mincí, které sledují vznik původních Nxt mincí (Colored Coins, 2014). Uživatelé dále nabízeli vytvořit zvláštní fórum pouze pro zájemce o Nxt.



Obrázek 77: Diskuze spojené s Nxt v roce 2013, zdroj Jason Davies a vlastní zpracování

6.3.2.3 Vývoj témat v roce 2014

Souhrn hlavních témat diskuzí za rok 2014 je zobrazen na obrázku č. 78. v tomto roce v kontextu s Nxt uživatelé diskutují o transakcích, o způsobech těžby mincí a ptají se na pooly nebo pooly samy nabízejí své služby, mluví o API, o platformě podporující Nxt, píšou o Nxt jako o první kryptoměně IPO[32], často oznamují určité události a vyzývají ke

[32] IPO (Initial Public Offering) je veřejná nabídka akcií nebo také první vstup na burzu. Obvykle se využívá malými a mladými firmami hledajícími kapitál, také se ale využívá i velkými firmami (Initial Public Offering – IPO, 2015).

společné diskuzi. v kontextu s Nxt panuje názor o podvodu, ale zase bezpečné měny. v diskuzích se lze setkat s dalšími slovy, jako Nxt zadarmo (angl. free), testování softwaru, píší o tzv. diskuzi (angl. thread) v kontextu s oficiálním představením Nxt a jeho určitých charakteristik. Diskutují o budoucnosti obchodování s měnou Nxt, o trhu Nxt, uživatelé mluví o marketingu Nxt, pak doporučují, jestli prodat nebo nakoupit měnu, mezi členy komunity Nxt na Bitcoin Forum se prosazuje názor o krachu kryptoměny Nxt.



Obrázek 78: Diskuze spojené s Nxt v roce 2014, zdroj Jason Davies a vlastní zpracování

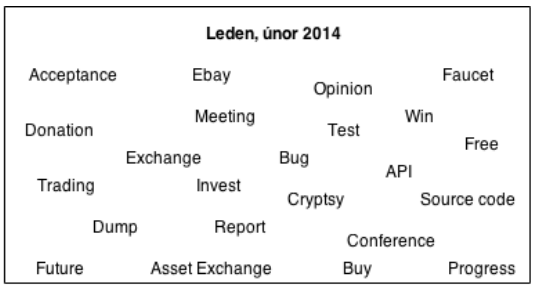
Dále uvádím vývoj témat v roce 2014 podrobněji.

V lednu a únoru roku 2014 (viz Obrázek 79) uživatelé psali o službách nebo obchodech, kde přijímají kryptoměnu Nxt, jedním z nejdiskutovanějších obchodů, kde přijímají Nxt, byl Ebay. Uživatelé se pak také dělili o názory ohledně této kryptoměny a o jejích možných zlepšeních. Někteří uživatelé vyvíjeli tzv. faucet, který je webovou stránkou představující kryptoměnu. Obvykle se to dělá pro úvod nových uživatelů do kryptoměn, také se často faucety dělají pro giveaway, přičemž by zájemce o kryptoměnu měl pro obdržení bezplatných mincí něco udělat, třeba každý den navštěvovat web (*Cryptocurrency Faucets*). Také psali o různých výhrách.

Dále uživatelé psali o burzách Nxt. Jedna z nejdiskutovanějších je Cryptsy, pak mezi uživateli je také populární BTER. Mluvili o darování někomu nějaké částky kryptoměny, o budoucnosti kryptoměny, o organizaci určitých schůzek nebo konferencí.

V těchto měsících lze pozorovat i nástroj uživatelů na nákup či prodej kryptoměny. Uživatelé více píšou o nákupu kryptoměny a také se také ptají na příležitost k investování.

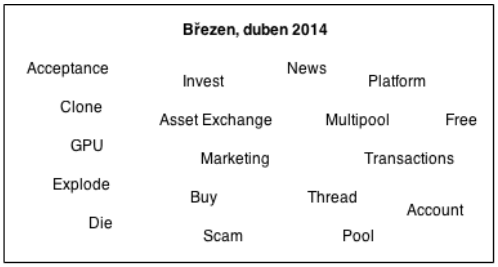
Uživatelé komunikují o zdrojovém kódu, reportují například opravy a píšou o API. Ostatní vyskytující slova nebo slovní spojení byla vysvětlovaná výše.



Obrázek 79: Diskuze spojené s Nxt v lednu a únoru roku 2014, zdroj vlastní zpracování

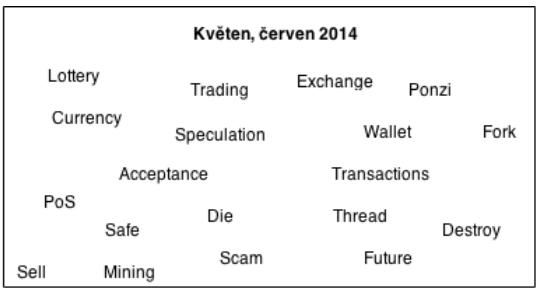
V březnu a dubnu se přidávají diskuze (viz Obrázek 80) na témata kopií Nxt o platformě Nxt a o multipoolech, kde několik poolů se sdružuje pro jediné společné těžení. Ptají se na těžbu prostřednictvím GPU. Uživatelé pak píšou i o marketingu, domnívají se, že marketing není postačující, dělí se o různé zprávy spojené s kryptoměnou Nxt. o kryptoměně Nxt nebo o určité informaci s ní spojené panuje názor o podvodu. Někteří uživatelé jsou si pak jistí krachem této kryptoměny.

V těchto měsících převládá názor na nákup kryptoměny. Dále se uživatelé ptají, jestli cena Nxt ještě vybuchne.



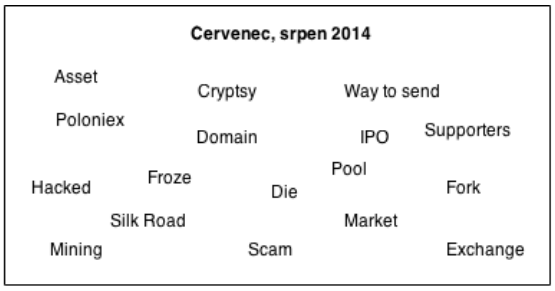
Obrázek 80: Diskuze spojené s Nxt v březnu a dubnu roku 2014, zdroj vlastní zpracování

V květnu a červnu roku 2014 (viz Obrázek 81) vznikají různé loterie. Uživatelé píšou o hře vysvětlující Ponziho schéma, mluví o příležitosti pro spekulace, o možnostech zničení kryptoměny. Panuje názor o prodeji Nxt. o ostatních tématech diskuze už bylo napsáno výše v této práci.



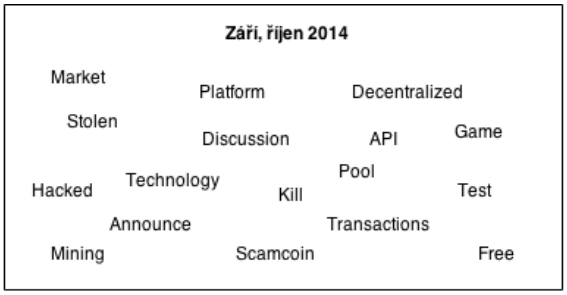
Obrázek 81: Diskuze spojené s Nxt v květnu a červnu roku 2014, zdroj vlastní zpracování

V červenci a srpnu kromě výše zmíněných a zobrazených na obrázku č. 82 uživatelé aktivně diskutovali o tom, že BTER zablokovala kryptoměny na účtech uživatelů a chtěli tyto mince vrátit. Také psali o tom, že BTER byla vystavená útokům hackerů. Nově diskutovali o burze Poloniex, o anonymním obchodování Silk Road, někteří uživatelé nabízeli určité domény.



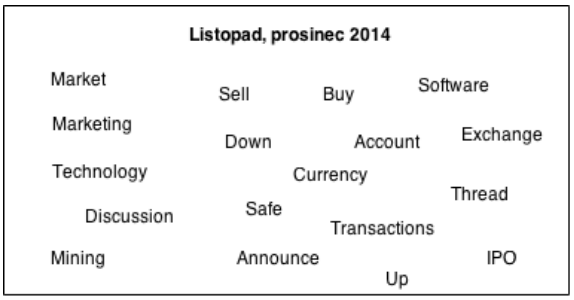
Obrázek 82: Diskuze spojené s Nxt v červenci a srpnu roku 2014, zdroj vlastní zpracování

V září a říjnu z obrázku č. 83 se přidává k diskuzím i téma krádeží, uživatelé mluví o možnostech zničení kryptoměny, oznamují určité události, projednávají technologie. Ostatní témata byla popsána výše.



Obrázek 83: Diskuze spojené s Nxt v září a říjnu roku 2014, zdroj vlastní zpracování

V listopadu a prosinci (viz Obrázek 84) uživatelé chválí technologii Nxt, ale zároveň nejsou spokojeni s marketingem, diskutují na téma, jestli Nxt půjde nahoru či dolů, jestli mohou prodávat či kupovat kryptoměnu. Ostatní témata už byla vysvětlená výše.



Obrázek 84: Diskuze spojené s Nxt v listopadu a prosinci roku 2014, zdroj vlastní zpracování

6.3.2.4 Vývoj témat v roce 2015

V roce 2015, jak je vidět z obrázku č. 85, zůstávají populární témata ohledně funkce tzv. „Asset Exchange“, uživatelé často využívají termín měna (angl. currency), pak se stále ptají na příležitost investování do této Nxt. Stále uživatelé píšou o účtech u Nxt, o burze Cryptsy, ptají se, jestli Nxt je další protokol Bitcoinu. Zůstávají populární mezi uživateli témata z minulých let, jako Nxt je kopií BTC, je podvod, je měna, která byla vytvořená pro účely rychlého zbohatnutí (angl. scamcoin), diskutují o transakcích, o přijetí Nxt různými obchody nebo poskytovateli služeb, píšou o zdrojovém kódu, o směně na USD a o dalších tématech týkajících se budoucnosti Nxt.



Obrázek 85: Diskuze spojené s Nxt v roce 2015, zdroj Jason Davies a vlastní zpracování

6.3.2.5 Vývoj kurzu

Na základě identifikovaných témat se dá tvrdit, že některé faktory ovlivňující vývoj kurzu pro Bitcoin neovlivní kurz Nxt, jelikož kryptoměna Nxt je založena na jiném algoritmu generace bloků. v této části se propojují hlavní témata s vývojem kurzu kryptoměny Nxt.

Na základě identifikovaných témat po spuštění v roce 2013 se uživatelé ptali na základní otázky ohledně Nxt, také bylo pozorováno, že se na fóru uskutečňoval tzv. giveaway. Při porovnání s kurzem se dá tvrdit, že v době seznámení s novou kryptoměnou uživatelé aktivně s touto kryptoměnou neobchodují a ptají se na její koncept. Také při vhodném okamžiku pořizují kryptoměnu za přijatelnou cenu. Tím pádem kurz NXT/USD v roce 2013 neměl tendenci růst (viz Graf 12).

V prosinci na konci roku 2013 kurz ale začínal prudce růst a pokračoval do začátku ledna roku 2014. Pak naopak prudce klesal a následně zase stoupal. Uživatelé na Bitcoin Forum podle identifikovaných témat reagovali na tuto skutečnost příspěvky o nákupu Nxt, což odpovídá úseku nejnižšího kurzu v lednu a únoru roku 2014. Prudký růst kurzu je vyvolán především zvýšením poptávky po této kryptoměně, která byla podpořena aktivitami na Bitcoin Forum v roce 2013. Růst zájmu o krytoměnu ještě lze sledovat z příspěvků o příležitostech investování do této krytoměny nebo do aktiv, která jsou v seznamu v jejím seznamu.

Dále v únoru roku 2014 lze z grafu č. 12 pozorovat klesající tendenci kurzu NXT/USD. Podle identifikovaných témat lze tvrdit, že pokles kurzu je vyvolán poklesem zájmu lidí o tuto kryptoměnu. Uživatelé se domnívají, že pokles kurzu je také výsledkem nepostačujícího marketingu. Na fóru panuje negativní nálada, uživatelé více píšou o podvodech a o krachu kryptoměny Nxt. Někteří uživatelé komunity také doporučují v této době nakupovat, což odporuje ekonomickým zákonům. Ale podle podrobnějšího obsahu fóra bylo zjištěno, že uživatelé doporučují nakupovat, aby povzbudili růst kryptoměny, jelikož většina drží kryptoměnu anebo prodává, kurz nebude růst.

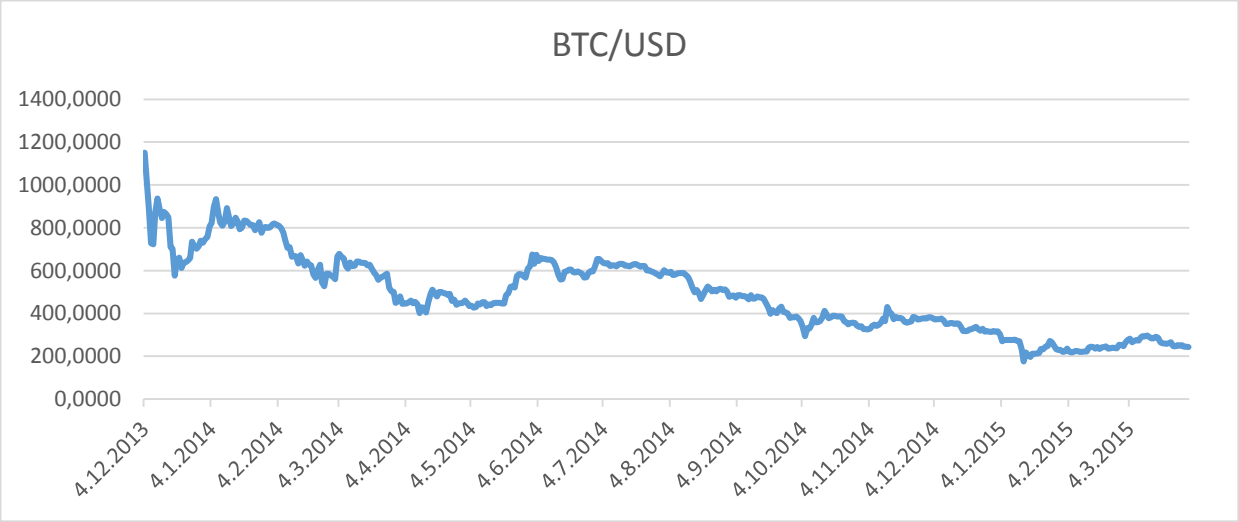
V květnu roku 2014, jak je vidět z grafu č. 12, kurz začíná růst a potom v červnu zase klesat. Na fóru v červnu nejvíce lidí doporučuje prodávat, což neodporuje pravidlům ekonomiky a má za následek pokles kurzu. Uživatelé také diskutovali o spekulacích, což napovídá o zájem uživatelů o Nxt, kteří se ptají nebo píšou o možnostech spekulací anebo dělí se názorem o spekulacích jiných zájemců.

V dalších měsících roku 2014 a 2015 měl kurz tendenci klesat (viz Graf 12). Prostřednictvím obsahu tvořeného na fóru se dá předpokládat, že pokles kurzu byl vyvolán mimo jiné z důvodu blokování kryptoměn na účtech klientů burzou BTER, také lze sledovat pokles aktivity mezi uživateli na Bitcoin Forum, a tím i předpokládaný zájem o kryptoměny. Podle reakcí uživatelů na pád kurzu lze označit jako příčinu poklesu nepostačující marketingové akce. v roce 2015 uživatelé reagují na pokračující pokles kurzu příspěvky s názory o kryptoměně, která byla vytvořena pro rychlé zbohatnutí.



Graf 12: Vývoj kurzu NXT/USD od prosince roku 2013 do konce dubna roku 2015, zdroj CoinGecko (2015)

Kromě reakce uživatelů na změnu kurzu uvádím dále také graf vývoje kurzu BTC/USD (viz Graf 12 a Graf 13) pro porovnání změny kurzů BTC/USD a NXT/USD. a z porovnání těchto dvou grafů se dá předpokládat, že vývoje kurzů BTC/USD a NXT/USD jsou v přímé korelaci, když kurz BTC/USD roste, pak kurz NXT/USD pravděpodobněji bude růst a naopak, pokud kurz BTC/USD bude padat, pak se předpokládá, že kurz NXT/USD bude rovněž klesat.



Graf 13: Vývoj kurzu BTC/USD od prosince roku 2013 do konce března roku 2015, zdroj Blockchain Info a vlastní zpracování

Tím pádem při vývoji kurzu NXT/USD je třeba sledovat vývoj kurzu BTC/USD a vývoj určitých událostí, o kterých komunita Bitcoinu aktivně diskutuje.

6.3.2.6 Dílčí shrnutí

Z výše provedené analýzy a interpretace dat je vidět, že v době vzniku uživatelé mluví více o konceptu kryptoměny Nxt se ptají na základní otázky, které se týkají algoritmu, potvrzení transakcí, obchodování, propagace. Postupem času začínají mluvit i o možných zlepšeních kryptoměny, píšou o způsobech těžby, o spekulacích, o přijetí kryptoměny různými obchody nebo službami. Také uživatelé píšou i o ekonomických událostech spojených s obchodováním s kryptoměnou Nxt. Uživatelé diskutují o peněženkách, transakcích a jejich poplatcích, adresách, chybách a o burzách. Uživatelé aktivně píšou také o investování, krádežích a aktivech, které lze prostřednictvím kryptoměny nakoupit.

Uživatelé této komunity reagují na změnu kurzu NXT/USD určitými příspěvky, které nějakým způsobem vysvětlují výkyvy kurzu.

Při rozhodování ohledně nákupu či prodeje nelze spoléhat u této kryptoměny pouze na diskuze na Bitcoin Forum stejně jako u Bitcoin, jelikož uživatelé nevyvolávají změnu kurzu, ale reagují na interní nebo externí události, které primárně vyvolávají změnu kurzu.

Také prostřednictvím analýzy příspěvků spojených s Nxt je vidět tendenci ke zvýšení či poklesu diskutování ohledně reklamy.

6.3.3 Monero

Monero (zkratka XMR) je kryptoměnou spuštěnou na konci dubna roku 2014. Využívá algoritmus CryptoNight a metodu generace bloků Proof of Work. Maximální množství mincí, které lze vytěžit, je 18 446 744 XMR. Průměrná doba generace bloků je jedna minuta a změna náročnosti probíhá po každém bloku. Odměna za nalezený blok se počítá podle vzorečku, který je znázorněn v tabulce č. 10.

Tabulka 10: Obecná informace o Monero, Bitmakler, Monero Blockchain

Doba spuštění	25. 04. 2014
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof of Work
Maximální množství mince	18 446 744 XMR
Těžební algoritmus	CryptoNight
Změna náročnosti	Každý 1 blok
Odměna za nalezený blok (počítá se podle vzorečku)	$H = \frac{M-A}{2^{20}}/10^{12}$, $M = 2^{64} - 1$, A – počet nalezených bloků, H – odměna za nalezený blok, M – maximální množství mincí.
Počet nalezených bloků (5. 4. 2015 22:40)	508 042

6.3.3.1 Hlavní témata

Při tvorbě clusteru prostřednictvím algoritmu “Lingo” na slovo Monero se vyskytují výsledky, které jsou zobrazeny na obrázku č. 86. z toho je vidět, že v kontextu se slovem Monero se během celé doby existence Bitcoin Forum mluví o příspěvcích (angl. posts), o změnách, o jednotlivých diskuzích spojených s Monero (angl. single), o vysvětleních (angl. explains), o tzv. diskuzích (angl. thread). Uživatelé používají v kontextu s Monero slovo

kryptoměna. Dále diskutují na témata těžby Monero, porovnávají s podvodem a píšou o podpoře při obchodování s kryptoměnou nebo při řešení nějakých technických problémů.

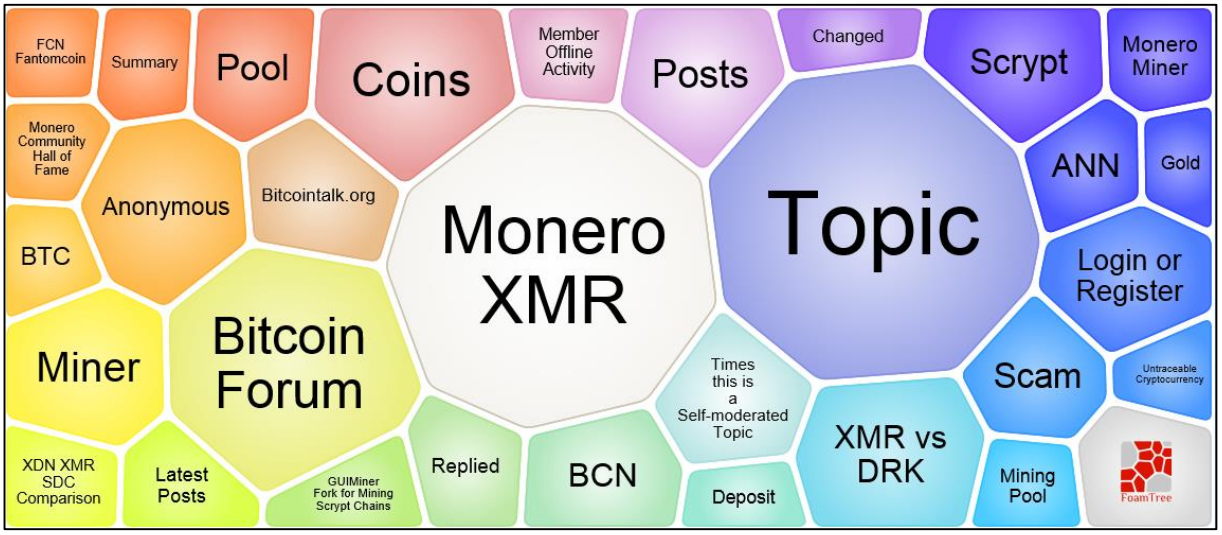
V případě tvorby clusteru prostřednictvím algoritmu "STC" se vyskytují další výsledky jako CPU a GPU, což je typické pro metodu generace bloků PoW.



Obrázek 86: Cluster "Monero", zdroj vlastní zpracování

Jelikož Monero má zkratku XMR, vytvořila jsem cluster na celý objem dat i se zkratkou této kryptoměny. Na obrázku č. 87 je vidět, která slova se nejvíce používají v kontextu s XMR. Uživatelé mluví o poolech, o podvodech, o anonymitě, o těžbě, o komunitě Monero, o zlatě, o změnách, o příspěvcích a o shrnutí. Také se v kontextu s XMR používají zkratky ostatních kryptoměn, jako jsou BCN – Bytecoin, DRK – bývalý Darkcoin, v dnešní době Darkcoin změnil název na Dash se zkratkou DASH a FCN – Fantomcoin.

Prostřednictvím tvorby clusteru s pomocí algoritmu "STC" se pak v kontextu se zkratkou XMR vyskytuje zkratka LTC, která patří kryptoměně Litecoin.



Obrázek 87: Cluster "XMR", zdroj vlastní zpracování

6.3.3.2 Vývoj témat v roce 2014

V roce 2014 v kontextu s Monero nebo jeho zkratkou se nejvíce vyskytovala témata, která jsou zobrazená na obrázku č. 88.

V době spuštění Monero na fóru se psalo o představení této kryptoměny, její charakteristice a odlišnosti od ostatních kryptoměn. Uživatelé potom kladli otázky zaměřené na novou kryptoměnu. Také se psalo o tzv. giveaway. Mezi uživateli bylo v roce 2014 populární téma těžby prostřednictvím různých poolů, také psali o CPU a GPU a řešení problémů s tím spojených, přičemž více se vyskytuj, CPU než GPU, z čehož lze odvodit, že pro těžbu této kryptoměny v roce 2014 postačoval obyčejný procesor stolního počítače. Uživatelé také projednávali logo Monero, publikovali příspěvky o různých peněženkách nebo řešili problémy s tím spojené. Uživatelé této komunity se zajímali o tržní kapitalizaci Monero.

V kontextu s Monero a jeho zkratkou XMR se vyskytují zkratky jiných kryptoměn, které jsou zobrazené na obrázku č. 87 a obrázku č. 88. Uživatelé porovnávají Monero s jinými kryptoměnami, ptají se, jestli nějaká kryptoměna může způsobit pád XMR nebo ovlivnit jeho kurz.

Dalšími tématy roku 2014 byly tvorba fyzické podoby kryptoměny Monero, zabezpečení od DDoS, spekulace s XMR a diskuze o tzv. OTC[33], kde uživatelé přirovnávali OTC k obchodování s XMR.

Uživatelé také mluvili o ceně Monero. v určitých úsecích poklesu Monero uživatelé se ptali, co se děje, také zmiňovali manipulační techniku Pump and Dump[34], psali o prodeji či nákupu kryptoměny, někdo píše o možnostech krachu kryptoměny, jiní píšou o možnostech zničení XMR.

Vyskytují se také příspěvky se slovem troll, které v internetovém slangu znamená provokativní nebo nevhodné zprávy od některých uživatelů diskuze, cílem kterých je přerušit věcnou diskuzi (Troll, 2010).

[33] OTC (Over-the-counter) nebo takové uspořádaní trhu cenných papíru, kde nabídka těchto cenných papírů probíhá bez dohledu určité oficiální instituce, účastnici se na podmínkách domlouvají mezi sebou přímo (Over-The-Counter-OTC, 2015).

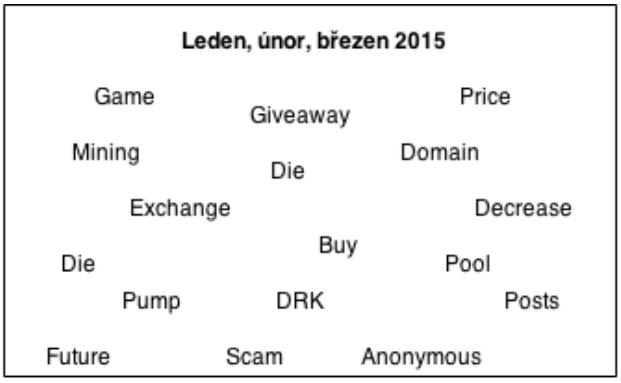
[34] Pump and Dump je manipulační a nelegální technika, ve které investor investoval určitou částku do akcií a pak určitým způsobem se snaží přilákat další investory, jejichž prostřednictvím částka jeho investice vzroste. Následným krokem tohoto investora bude prodej akcií za vyšší cenu, než před tím investoval, čímž vyvolá pád ceny (Pump and Dump, 2015).



Obrázek 88: Diskuze spojené s Monero v roce 2014, zdroj Jason Davies a vlastní zpracování

6.3.3.3 Vývoj témat v roce 2015

Uživatelé stále diskutují o možných hrách, kde přijímají Monero, diskutují o burzách, o poolech nebo pooly nabízí své služby uživatelům fóra. Na začátku roku 2015 se uživatelům nabízejí k nákupu různé domény. Uživatelé stále píšou o kryptoměně Monero jako o podvodu, diskutují o její budoucnosti a komentují vývoj kurzu. Ostatní témata z obrázku č. 89 byla okomentovaná výše a budou propojena s vývojem kurzu v další subkapitole.



Obrázek 89: Diskuze spojené s Monero v roce 2015, zdroj vlastní zpracování

6.3.3.4 Vývoj kurzu

Kryptoměna Monero byla spuštěna na konci dubna roku 2014 a je založena na stejném algoritmu PoW jako i Bitcoin, tím pádem se dá předpokládat, že vývoj kurzu Monero ovlivní stejné faktory jako Bitcoin.

V této subkapitole uvádím srovnání obsahu tvořeného na Bitcoin Forum v kontextu s Monero s vývojem kurzu XMR/USD, jelikož podle obsahu tvořeného na fóru se dají identifikovat ostatní faktory kromě uvedených v subkapitole "Vývoj kurzu" u Bitcoinu.

Po spuštění nové kryptoměny Monero se uživatelé seznamují s její koncepcí. Tím pádem se dá očekávat, že kurz po spuštění bude klesat, což se potvrzuje úsekem na grafu č. 14 v květnu roku 2014. Po vyvolání zájmu uživatelů se dá předpokládat, že kurz bude růst, uživatelé reagují diskuzí o spekulaci a příspěvky spojenými s tržní kapitalizací Monero. Komunita aktivně reagovala na změny kurzu také porovnáním kryptoměny Monero s jinými kryptoměnami. v době klesání kurzů uživatelé reagovali příspěvky ohledně využití techniky Pump and Dump, což znamená, že někteří vyvolali zájem o kryptoměnu u ostatních uživatelů. To vedlo k tomu, že cena Monero stoupla. Následným krokem po dosažení vysokého kurzu je to, tito uživatelé začínají rychle prodávat držené kryptoměny, což vede k jejímu prudkému poklesu.

Další reakcí uživatelů na pokles kurzu byly příspěvky týkající se krachu kryptoměny nebo možného zničení.

V lednu roku 2015 cena XMR pomalu klesala (viz Graf 14), uživatelé psali o možnosti nákupu, motivem těchto příspěvků byl nákup za levnější cenu. Cena ale nešla v lednu nahoru, doporučení malého počtu uživatelů k nákupu není postačující k zdvihnutí celého trhu XMR. v únoru roku 2015 uživatelé komentují pád. v březnu ale kurz začal růst a uživatelé na něj reagují příspěvky o využití manipulační techniky Pump and Dump, přičemž v březnu je první fáze Pump, tím pádem podle příspěvků se dá očekávat v budoucnosti prudký pokles či využití metody Dump.



Graf 14: Vývoj kurzu XMR/USD od května roku 2014 do konce dubna roku 2015, zdroj CoinGecko (2015)

6.3.3.5 Dílčí shrnutí

U vývoje diskuzí této kryptoměny je také jako i u předchozích analyzovaných kryptoměn vidět, že v době spuštění kryptoměny uživatelé aktivně neobchodují s novou kryptoměnou, ale seznamují se s ní. V době zvýšení či poklesu kurzu XMR/USD uživatelé určitým způsobem reagují na tyto výkyvy. Uživatelé diskutují také o peněženkách, transakcích a jejich poplatcích, adresách, chybách a o burzách. Často se vyskytuje názor na využití techniky Pump and Dump. Uživatelé aktivně píšou také o spekulacích a útocích.

V době poklesu aktivity uživatelů nebo častějšího vyskytování názorů na krach kryptoměny kurz také klesal.

Z diskuzí této komunity je vidět také volby určitých technologií pro těžbu.

Uživatelé této komunity stejně jako i u ostatních analyzovaných komunit nevyvolávají změnu kurzu, ale reagují na změnu kurzu psaním o možných příčinách změny kurzu.

Tato diskuze může nových uživatelům přiblížit koncepci kryptoměny, napomoci při určitých technických problémech, vysvětlit možnou budoucnost kryptoměny a možné příčiny výkyvů kurzu.

6.3.4 Dash

Bývalý Darkcoin, nový název kryptoměny, vstoupil v platnost 25. března roku 2015, změna se týkala pouze názvu kryptoměny, žádné vlastnosti Darkoinu se nezměnily (*Bitcoin Forum).*

Kryptoměna byla spuštěna v lednu roku 2014, metoda generace bloků je Proof of Work, těžební algoritmus je X11, maximální množství je 22 000 000 DASH, průměrná doba generace bloků jsou dvě a půl minuty. Náročnost pro výpočet hash se mění každý blok, odměna za nalezený blok v době psaní této části diplomové práce byla 2,8 DASH, přičemž výše odměny záleží na výši náročnosti, která se počítá podle vzorečku uvedeného v tabulce č. 11.

Tabulka 11: Obecné informace o Dash, zdroj Bitmakler, Dash (2015), Dash Cryptocurrency (2015)

Doba spuštění	18. 01. 2014
Průměrná doba generace bloků	2,5 min.
Metoda generace a bezpečnosti bloků	Proof of Work
Maximální množství mincí	22 000 000 DASH
Těžební algoritmus	X11
Změna náročnosti	1 blok
Odměna za nalezený blok (6. 4. 2015 9:42)	2,83971433 DASH
Výpočet odměny	$2222222/\left(\frac{Náročnost + 2600}{9}\right)^2$
Počet nalezených bloků (6.4:2015 9:41)	247 763

Dash odstraňuje další nedostatek Bitcoinu, kterým je transparentnost plateb, při níž každý uživatel může prohlédnout zůstatek účtu libovolného jiného uživatelé, také se podívat na jeho historii transakcí. Dash je více konfidenciální, nedovoluje prohlédnout odesilatele. Princip Dash je následující, klienti uživatelů odesílají tzv. masternodes (z angl. mastery uzlů) Dash, masteři pak tuto částku rozdělují na určité jednotky. v dalším kroku masteři posílají částku dalšímu masterovi anebo zpět do peněženky uživatele s náhodně vygenerovanou adresou. Když uživatel odešle Dash jinému uživateli, přijímající neuvidí původní adresu odesilatele, ale náhodně vygenerovanou adresu (Dashcoin, 2014).

6.3.4.1 Hlavní témata

Prostřednictvím algoritmu shlukování "Lingo" byl vytvořen cluster pro celý objem dat (viz Obrázek 90). Jelikož název kryptoměny je stejný, jako je jeho zkratka, byl vytvořen jeden cluster na celý objem dat.

Uživatelé v kontextu s Dash nejvíce mluví o kryptoměně Darkcoin a její zkratce DRK, jelikož je to starý název kryptoměny Dash, pak také mluví o Bitcoin, SDC - ShadowCash a XMR – Monero.

Dash se vyskytuje se slovy anonymní a altcoin, jelikož je alternativní kryptoměnou Bitcoinu.

Uživatelé pak mluví o tzv. Darksend[35] a InstantX[36], o zařízení algo, o poolech, o těžbě, o transakcích, o peněženkách, o trhu, o podpoře a o možnosti bývalého DRK nebo v okamžik psaní této části diplomové práce Dash nechání ceny růst.



Obrázek 90: Cluster Dash, zdroj vlastní zpracování

6.3.4.2 Vývoj témat v roce 2014

Na začátku této subkapitoly je třeba zmínit, že ze všech analyzovaných kryptoměn je to kryptoměna, u které komunita nejčastěji publikuje na Bitcoin Forum různé grafy a predikce budoucího vývoje kurzu.

Souhrn témat v kontextu s Dash v roce 2014 je zobrazen na obrázku č. 91.

V roce 2014 uživatelé diskutovali o technologii těžby, o využití služeb poolů nebo multipoolů a o úrovni náročnosti. Projednávali použití peněženky a určité problémy s tím spojené. v kontextu s Dash uživatelé diskutují o masteřích (angl. masternodes), ptají se na princip fungování Dash a na role masterů v procesu převodu adres na plnou anonymitu. Populární diskuze i o DarkSend, uživatelé popisují jeho vlastnosti anebo řeší určité problémy s tím spojené. Uživatelé také vytváří tzv. faucety, na fóru se ptají na názory ostatních, jestli vytvořili vhodný faucet anebo čekají na určité doporučení k vyvíjenému jejich faucetu.

Diskuze v roce 2014 se týkají také i porovnání různých kryptoměn s Dash. Uživatelé píšou o obchodování Dash s jinými kryptoměnami. Na fóru se vyskytují nejenom zkratky, ale i celé názvy kryptoměn, uvádím na obrázku č. 91 pouze zkratky kryptoměn, ale diskuze používají celý název i zkratky.

Dále mluví o Bitfinex[37], o Cryptsy, o změně ceny. Tady je třeba zmínit, že komunita kryptoměny Dash aktivně reaguje na změny ceny a při podrobnějším studiu obsahu bylo zjištěno, že uživatelé občas doporučují prodávat nebo nakupovat v době budoucího přelomu kurzu.

35 DarkSend je software, jehož prostřednictvím se odesílají určité částky kryptoměny (*DarkSend Documentation).*

36 InstantX je software, který umožnuje potvrzovat provedené transakce mnohem rychleji než deset minut, jak je to u Bitcoin (Yashu, 2015).

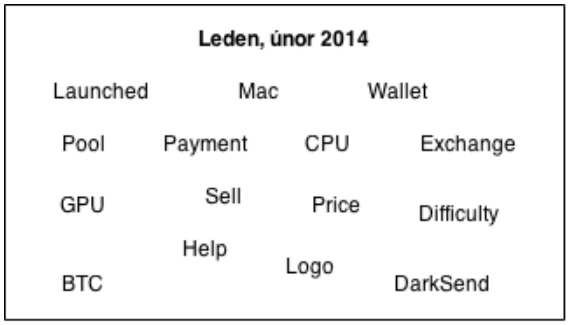
37 Bitfinex se počítá za největší platformu pro obchodování s kryptoměnami (Bitfinex, 2013).

Na Bitcoin Forum v kontextu s Dash se v roce 2014 vyskytují i příspěvky, na které jiní uživatelé reagují odesláním odpovědi se slovem troll. Ostatní témata budou rozebrána v této práci dále.



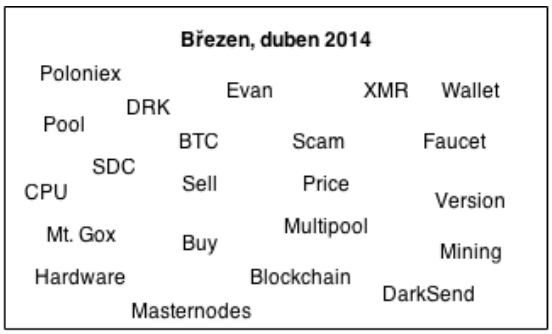
Obrázek 91: Diskuze spojené s Dash v roce 2014, zdroj Jason Davies a vlastní zpracování

Dále uvádím vývoj témat podle dvou měsíců. Dash, bývalý Darkcoin byl spuštěn v lednu roku 2014, a proto dá se očekávat, že v této době se vyskytují příspěvky o jeho spuštění, o obecných otázkách týkajících se peněženek, technologií těžby, na fóru se mluvilo o CPU a GPU. Uživatelé se ptají na pooly, na náročnost na DarkSend. v únoru roku 2014 uživatelé doporučují držet v té době DRK, jelikož s růstem náročnosti a při růstu zájmu cena bude růst (viz Obrázek 92).



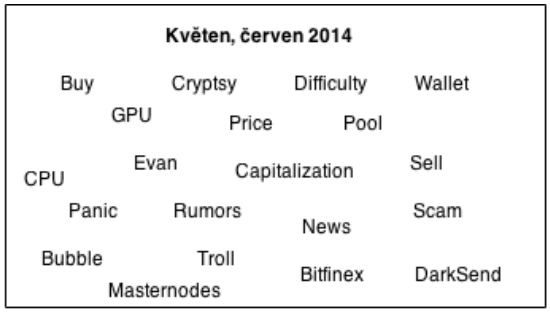
Obrázek 92: Diskuze spojené s Dash v lednu a únoru roku 2014, zdroj vlastní zpracování

V březnu a v dubnu uživatelé (viz Obrázek 93) aktivně diskutují jednání Evana, který je vývojářem Dash tehdejší DRK. Ptají se na webové stránky, kde si lze prohlédnout síť bloků. Uživatelé zase doporučují držet kryptoměnu, ani neprodávat, ani nekupovat. v kontextu s Dash se objevují názory na podvod. Dalším názorem uživatelů je to, že bývalý DRK je verzí BTC.



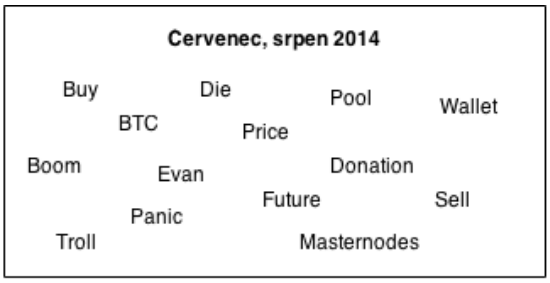
Obrázek 93: Diskuze spojené s Dash v březnu a dubnu roku 2014, zdroj vlastní zpracování

V květnu a červnu (viz Obrázek 94) uživatelé diskutují o růstu náročnosti těžby, o nákupu či prodeji kryptoměny. Uživatelé píšou, aby ostatní nepropadali panice v době pádu kurzu. Spojují růst kurzu s bublinou. Dále uživatelé také oznamovali tržní kapitalizaci ostatních členů komunity, o čemž pak celá komunita nějakou dobu diskutovala. Vyskytoval se názor, že bylo zapotřebí sledovat fámy, které mohou být signálem k nákupu, ale signálem prodeje mohou být věrohodné zprávy. Ostatní témata zobrazená na obrázku č. 94 jsou komentovaná výše.



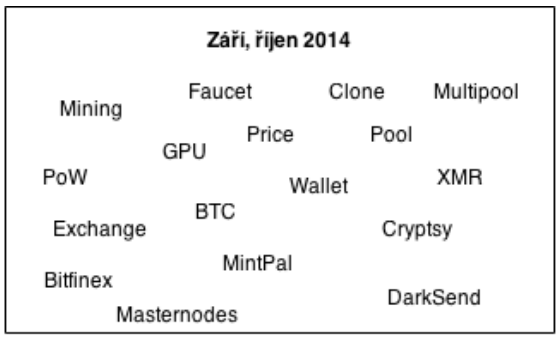
Obrázek 94: Diskuze spojené s Dash v květnu a červnu roku 2014, zdroj vlastní zpracování

V červenci a srpnu uživatelé pokračovali v diskuzi na témata uvedená na obrázku č. 95. Mezi nová témata můžeme zařadit názor o zkrachování Dash. Vyskytovali se názory, že vývoj kurzů BTC/USD a DASH/USD jsou nepřímé korelované, když cena BTC roste, cena DASH padá. Jiní uživatelé ale mají jiný názor, myslí, že BTC a Dash jsou naopak v přímé závislosti. Tuto skutečnost ověřím v následující subkapitole.



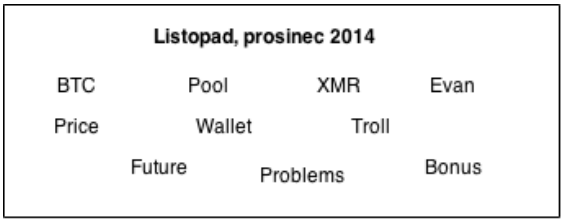
Obrázek 95: Diskuze spojené s Dash v červenci a srpnu roku 2014, zdroj vlastní zpracování

V září a v říjnu (viz Obrázek 96) uživatelé také diskutují, jestli lze tehdejší DRK kopírovat, stále vytváří a komentují facety, mluví o platformě Bitfinex a o burze MinPal, která v této době byla okradena o 3 700 BTC (Gillespie, 2014). Pak také uživatelé diskutují o algoritmu těžby PoW, na které je tehdejší DRK založen. Ostatní témata byla okomentovaná výše.



Obrázek 96: Diskuze spojené s Dash v září a říjnu roku 2014, zdroj vlastní zpracování

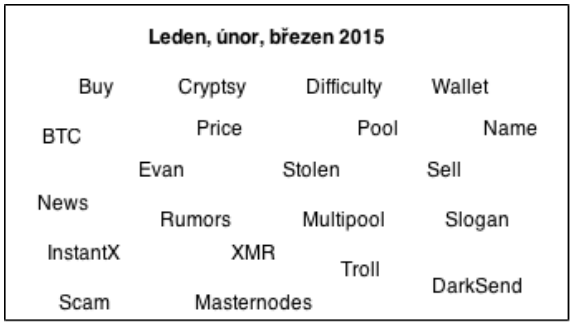
Na konci roku 2014 (viz Obrázek 97) se uživatelům nabízí bonusy za nalezený blok. Uživatelé pak řeší své různé technické problémy. Také diskutují ohledně budoucnosti kryptoměny. Ostatní témata byla popsána výše.



Obrázek 97: Diskuze spojené s Dash v listopadu a prosinci roku 2014, zdroj vlastní zpracování

6.3.4.3 Vývoj témat v roce 2015

Souhrn témat za rok 2015 je zobrazen na obrázku č. 98. Na začátku tohoto roku uživatelé aktivně diskutují o InstantX. Mluví o změně názvu tehdejší kryptoměny DRK a pak o následné změně názvu na Dash, vyjadřují své názory na nový název. Projednávají sloganu Dash. Také diskutovali o možných důvodech současné situace kryptoměny Monero. Ostatní témata byla popsána výše.



Obrázek 98: Diskuze spojené s Dash na začátku roku 2015, zdroj vlastní zpracování

6.3.4.4 Vývoj kurzu

Dash, v minulosti Darkcoin byl spuštěn v lednu roku 2014. Generace bloků Dash je založená na metodě PoW, a proto se dá předpokládat, že vývoj kurzu DASH/USD ovlivní stejné faktory jako i kurz BTC/USD. v této subkapitole propojuji vývoj kurzu s identifikovanými klíčovými tématy z Bitcoin Forum.

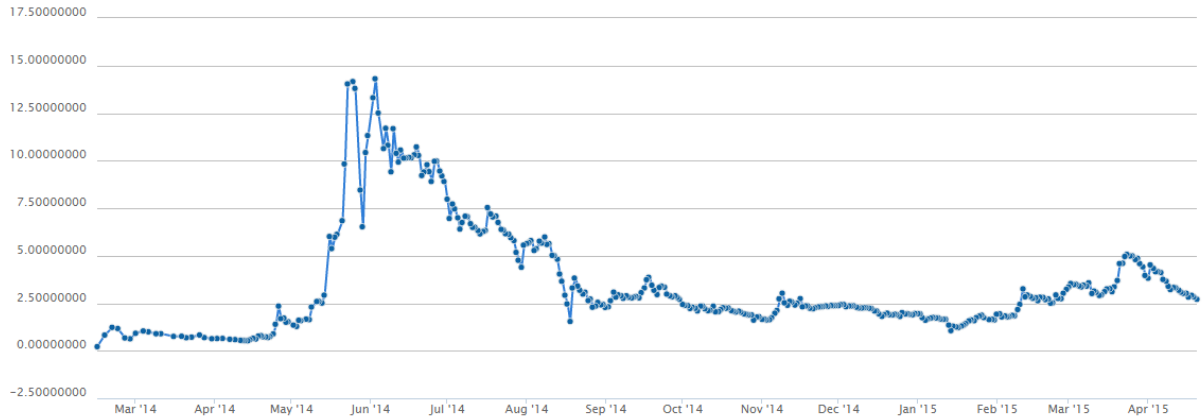
Po spuštění vývoj kurzu DASH/USD nemá tendenci růst, jelikož se uživatelé seznamují s novou kryptoměnou, zjišťují její vlastnosti a zvažují její budoucnost.

V březnu a dubnu roku 2014 je kurz DRK/USD stabilní a podle obsahu tvořeného na fóru se doporučuje kryptoměnu držet, jelikož se neočekává podle názoru uživatelů v této době ani růst a ani pokles kurzu.

Kurz začíná prudce růst až na začátku května roku 2014, přičemž uživatelé na Bitcoin Forum píšou o nákupu kryptoměny na začátku května a na konci. Před nákupem na konci května doporučují prodávat, přičemž při prohlédnutí graf č. 15 je vidět, že se cena DRK hýbá ve stejném směru. v této době, jak bylo naznačeno při identifikaci hlavních témat, uživatelé mluví o bublině, což odpovídá skutečnosti vývoje kurzu podle grafu č. 15.

V době poklesu kurzu DRK/USD uživatelé reagovali příspěvky ohledně krachu DRK. Také se v této době vyskytovaly názory, jak už bylo naznačeno výše, že vývoj kurzu DRK/USD a BTC/USD je ve vztahu nepřímé korelace. Někteří uživatelé ale naopak tvrdí, že co je dobré pro vývoj kurzu BTC/USD, to samé bude fungovat i pro DRK/USD. Při porovnání kurzu DRK/USD na grafu č. 15 a vývoje kurzu BTC/USD grafu č. 16 lze vyvrátit názor nepřímé závislosti kurzů, jelikož podle zjištěné skutečnosti je vývoj kurzů DRK/USD a BTC/USD po dubnu roku 2014 v přímé závislosti. Tím pádem se dá tvrdit, že vývoj kurzu reaguje na události spojené s BTC, a proto je třeba v souvislosti s propojováním hlavních témat s kurzem DRK/USD sledovat i diskuze tvořené v souvislosti s BTC, které byly v této diplomové práci popsané výše.

V ostatních měsících uživatelé o změně kurzu a o možných příčinách pohybu kurzu aktivně nediskutují. Vyskytovaly se příspěvky s dotazy ohledně nákupu či prodeje, ale pouze na základě těchto příspěvků predikovat vývoj kurzu nedá, jelikož názory na nákup či prodej jsou rozdílné a na vývoj kurzu ovlivní i jiné faktory.



Graf 15: Vývoj kurzu DASH/USD od února roku 2014 do dubna roku 2015, zdroj CoinGecko (2015)



Graf 16: Vývoj kurzu BTC/USD od února roku 2014 do března roku 2015, zdroj Blockchain Info a vlastní zpracování

6.3.4.5 Dílčí shrnutí

Po vzniku této kryptoměny vznikala témata typická pro nově vzniklou kryptoměnu. Pak uživatelé začínali více diskutovat o obchodní stránce kryptoměny, což napovídá o zvýšeném zájmu uživatelů o kryptoměnu. Také uživatelé diskutují o peněženkách, transakcích a jejich poplatcích, adresách, chybách a o burzách. Komunita této kryptoměny je jediná, která tvoří různé analýzy pro predikci vývoje kurzu. V určitých úsecích uživatelé doporučují ostatním nakupovat či prodávat. Pak aktivita uživatelů postupně začala klesat.

Informační hodnota diskuzí ohledně kryptoměny Dash může spočívat kromě seznamování a pomoci při určitých problémech také i v predikci budoucího vývoje kurzu. Dá se tady tvrdit, že uživatelé ovlivní vývoj kurzu této kryptoměny vůči americkému dolaru.

Také prostřednictvím analýzy příspěvků spojených s Dash je vidět tendenci ke změně výběru technologií pro těžbu, zvýšení či poklesu zájmu o kryptoměnu mezi uživateli a také jinými institucemi.

6.3.5 Counterparty

Counterparty je bezplatnou platformou, která je doplňkem pro systém Bitcoin (Counterparty, 2014). Tato platforma využívá kryptoměnu XCP, která je založená na odlišném algoritmu generace bloků od všech nejdiskutovanějších zmapovaných kryptoměn v této diplomové práci. XCP využívá metodu Proof of Burn, která „spaluje“ BTC, který generuje bloky prostřednictvím algoritmu Proof of Work. XCP využívá sítě bloků Bitcoin, v níž byly v určité periodě “spálené” BTC. Tato perioda byla od spuštění kryptoměny XCP v lednu roku 2014 a probíhala do začátku února roku 2014. Během ní zájemci posílali na určitou neexistující adresu BTC a dostávali zpět kryptoměnu XCP. Tím se vygenerovalo celkové množství XCP ve výši 2 648 735,92 XCP. BTC byly tímto procesem zcela zničeny a žádný zájemce je nedostal (*Bitcoin Forum*). Po procesu spálení uživatelé začali s kryptoměnou XCP obchodovat.

V době, kdy zájemci odesílali BTC, těžaři BTC přidávali tyto transakce do sítě bloků BTC a za to dostávali určitou odměnu, která se vypočítávala podle vzorečku uvedeného v tabulce č. 12.

V dnešní době se celkové množství XCP zmenšuje nebo „spaluje“ na základě registračních poplatků, ale zároveň celkové množství XCP nikdy nedosáhne nuly, jelikož poplatky pak mohou být upravené (Counterparty, 2014).

Tabulka 12: Obecná informace o Counterparty, Counterparty (2014), Bitmakler, Crypto-Currency Market Capitalizations, Blockscan, Deweller, Bitcoin Forum

Doba spuštění	02. 01. 2014
Metoda generace a bezpečnosti bloků	Proof of Burn
Maximální množství mince	2 648 735,92 XCP
Odměna BTC těžařům za dosazení transakce do blockchain (6. 4. 2015 10:41)	$XCP_{ODMĚNA} = BTC_{SPÁLENÉ} * (1000 * (1 + .5 * (\frac{POSLEDNÍ_{BLOK} - SOUČASNÝ_{BLOK}}{POSLEDNÍ_{BLOK} - PRVNÍ_{BLOK}})$ Poslední blok – poslední blok po spalovacím procesu. První blok – první blok, od kterého spalovací proces začal.
Počet nalezených bloků (6. 4. 2015 11:27)	350 957 (na základě Bitcoin blockchain)
“Spálení”	BTC

6.3.5.1 Hlavní témata

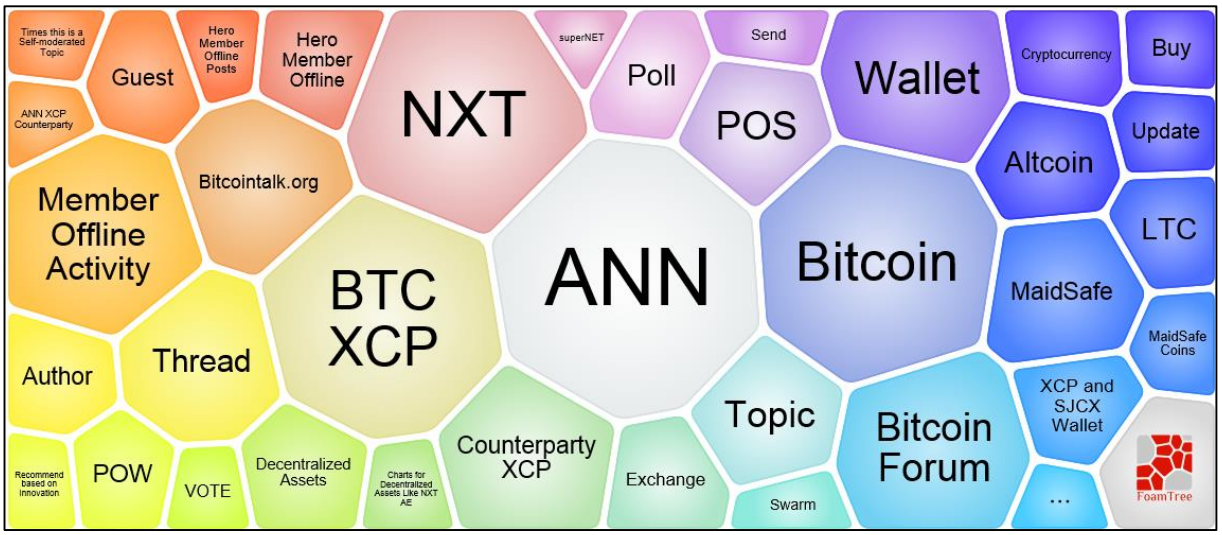
Prostřednictvím algoritmu “Lingo” byl vytvořen cluster “XCP” (viz Obrázek 99). Z něj je vidět, že se nejvíce v kontextu s kryptoměnou XCP mluví s BTC – Bitcoinem, LTC - Litecoin a kryptoměnou Nxt. Také se v kontextu využívají tzv. diskuze (angl. thread), altcoiny nebo alternativní kryptoměna Bitcoinu, pak v kontextu s platformou Counterparty, s peněženkami, s metodami generace bloků jako PoW a PoS. O kryptoměně XCP se v kontextu s PoW mluví jako o kryptoměně, která spaluje kryptoměny, jež generují bloky prostřednictvím metody PoW a v kontextu s PoS se mluví jako o porovnání různých metod a zjištění nejvhodnější metody generace bloků.

XCP se také zúčastňuje různých voleb, například uživatelé volí, jaká kryptoměna přežije nebo jaká kryptoměna je nejvhodnější.

Uživatelé také diskutují v kontextu s XCP o kryptoměně SJCX (Strojcoin X), která je založená na Counterparty.

V kontextu s XCP se vyskytuje také aktivum Swarm[38] a jiné aktiva. Uživatelé mluví o aktualizacích verzí, o nákupu či prodeji kryptoměny a o obchodování s XCP.

[38] Swarm je nezávislou organizaci a inovativním modelem fondu, který umožnuje transformovat rizikový kapitál. (*Swarm*).



Obrázek 99: Cluster "XCP", zdroj vlastní zpracování

6.3.5.2 Vývoj témat v roce 2014

Souhrn témat vyskytujících se v kontextu s XCP je zobrazen na obrázku č. 100. Těžba XCP založená na jiném principu než PoW a PoS, ale na PoB. a proto se místo těžby prostřednictvím CPU, GPU nebo ASIC vyskytuje tzv. "spálení" (angl. burn). Kvůli tomu, že ve spalovacím procesu se spalovaly Bitcoiny, se pak v příspěvcích často vyskytuje BTC. Uživatelé v kontextu s XCP také porovnávají metody PoW a PoS.

Další kryptoměnou, která je založená na PoB, je Mastercoin (zkratka MSC). Uživatelé komunity XCP na Bitcoin Forum porovnávají XCP se MSC, diskutují o jejich vlastnosti, hledají jejich pozitiva a negativa.

Jelikož XCP je v seznamu aktiv Nxt, pak se vyskytuje i Nxt v kontextu s XCP. Uživatelé také porovnávají tyto dvě kryptoměny.

Uživatelé píšou i o dalších aktivech, která lze prostřednictvím XCP nakoupit, jeden z nejdiskutovanějších je Swarm.

Dále uživatelé píšou o burzách, jednou z nejdiskutovanějších je Poloniex. Diskutují o ceně XCP a o prodeji či nákupu této kryptoměny.

Vyskytují se i různé chyby v instalování peněženky nebo jiné, které uživatelé na fóru řeší. v kontextu s XCP se objevují i další témata jako riziko, zprávy, protokol XCP, poplatky za připojení transakcí k síti bloků Bitcoin a jiné, které budou popsány dále.



Obrázek 100: Diskuze spojené s XCP v roce 2014, zdroj Jason Davies a vlastní zpracování

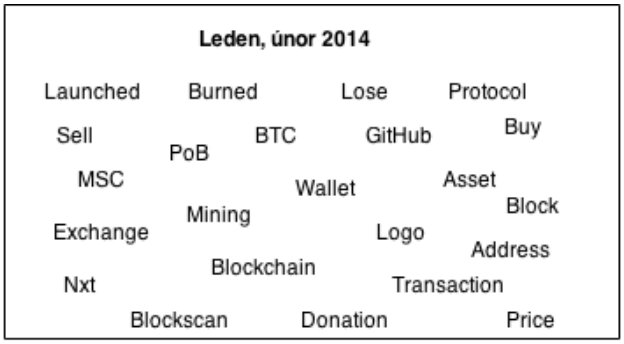
Dále popisuj vývoj témat v roce 2014 podrobněji podle měsíců.

V lednu a únoru po spouštění kryptoměny XCP se uživatelé ptali na princip fungování. Vyskytují se příspěvky o "spálení", o ztrátě BTC, uživatelé se ptali, co se stane s posílanými BTC a jestli vývojáři nedostanou uživatelské BTC. Dále se uživatelé z důvodu nepochopení principu ptali na způsob těžby, diskutovali o adresách peněženek a o samotných peněženkách.

Psali o protokolu, na kterém je založena kryptoměna XCP. Vývojář upozorňoval na možnosti stažení kódu na GitHub. Za práci vývojáře se navrhovalo darovat mu nějakou částku XCP.

Uživatelé se rovněž ptali na tzv. blockscan (z angl. skenování bloků), který kopíruje síť bloků Bitcoinu. Po spalovacím procesu blockscan začal vydávat výsledky s delším časovým úsekem na rozdíl od sítě bloků BTC a to stalo předmětem diskuze komunity XCP.

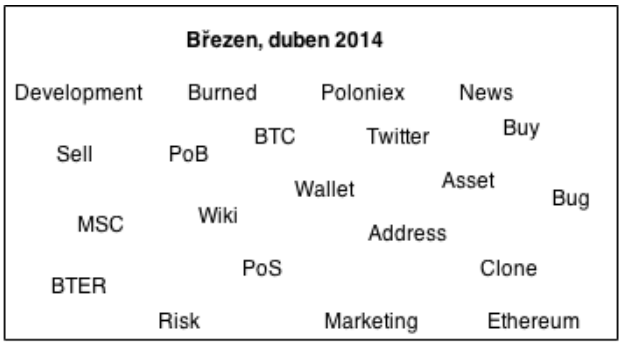
Uživatelé také mluvili o příležitostech nákupu či prodeje XCP. Pak také diskutovali o logu XCP, psali o Nxt a o blocích XCP (viz Obrázek 101).



Obrázek 101: Diskuze spojené s XCP v lednu a únoru roku 2014, zdroj vlastní zpracování

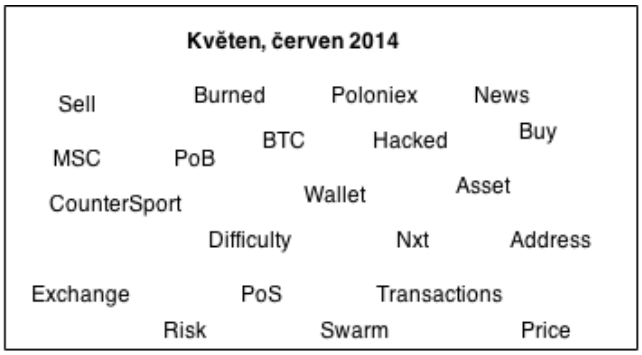
V březnu a dubnu se přidávají další témata diskuze v kontextu s XCP, jako jsou porovnání XCP se MSC a kryptoměnou Ethereum, přidání XCP na Wikipedii, příspěvky na sociální síti Twitter, řešili určité problémy s peněženkou. Také mluvili o určitých zprávách spojených s XCP, o vývoji XCP a jeho propagaci ostatním zájemcům o kryptoměny.

S obchodováním XCP uživatelé spojují riziko. Předmětem diskuzí v těchto měsících se také stala platforma pro směnné operace BTER a burza Poloniex a možnost kopírování této kryptoměny. (viz Obrázek 102). Ostatní témata diskuzí byla popsána výše.



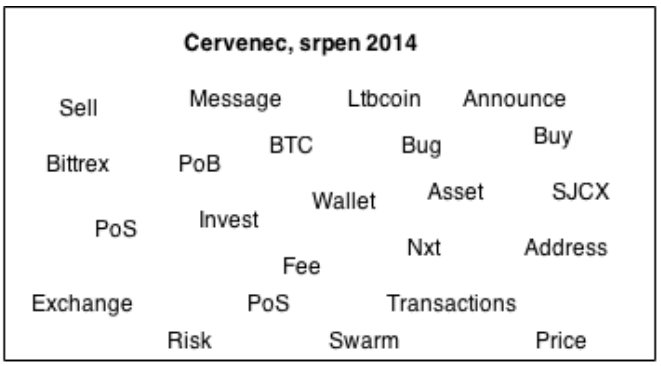
Obrázek 102: Diskuze spojené s XCP v březnu a dubnu roku 2014, zdroj vlastní zpracování

V květnu a červnu (viz Obrázek 103) se přidávají další diskuze spojené s XCP, novým témat je aplikace CounterSport, která je založená na síti bloků BTC a která umožní sledovat sport a vsázet (*Bitcoin Forum*). Dále uživatelé diskutují o náročnosti nalezení nových bloků Bitcoin a o aktivu Swarm. Další témata byla popsána výše.



Obrázek 103: Diskuze spojené s XCP v květnu a červnu roku 2014, zdroj a vlastní zpracování

V červenci a srpnu roku 2014 (viz Obrázek 104) uživatelé mimo výše popsaných témat navíc diskutují o Ltbcoin, který je odměňovacím systémem (*LTBcoin*) a je v seznamu aktiv XCP. Dále také začínají mluvit o kryptoměně SJCX. v těchto měsících diskutují o možnosti vývoje funkce posílání zpráv spolu s odesíláním částek prostřednictvím peněženky. Uživatelé také diskutují o příležitosti investovat do XCP anebo aktiv, která jsou v seznamu XCP. Pak uživatelé oznamují, že by bylo vhodné přidat XCP na burzu pro kryptoměny Bittrex.

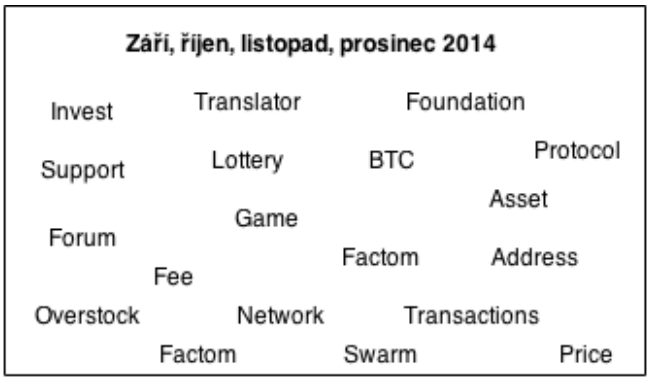


Obrázek 104: Diskuze spojené s XCP v červenci a srpnu roku 2014, zdroj a vlastní zpracování

Další měsíce - září, říjen a listopad jsem spojila do jednoho obrázku č. 105, jelikož značné změny v tématech diskuzí v těchto měsících nebyly pozorované. Uživatelé stále mluví o aktivu Swarm a dalších aktivech, o ceně XCP, o transakcích a jejich poplatcích, o protokolu Counterparty, o příležitostech investovat a o Bitcoinu. Novými tématy jsou diskutování o komunitě XCP s názvem The Counterparty Foundation, o první hře, která přijímá XCP, vývojáři přemýšlí o přeložení stránky do různých jazyků a aktivně hledají překladatele.

Uživatelé také diskutují o tom, jestli Bitcoin Forum je vhodné fórum pro diskutování o různých věcech ohledně XCP a získávají možnost zúčastnit se loterie, kterou aktivně projednávají.

Dále diskutují o tzv. Factom[39], který bude naimplementován pro kryptoměny XCP a MSC. Uživatelé se také ptají na to, jestli Overstock.com přijímá XCP. Předmětem diskuzí je i podpora uživateli kryptoměny XCP.



Obrázek 105: Diskuze spojené s XCP v září, říjnu, listopadu a prosinci roku 2014, zdroj a vlastní zpracování

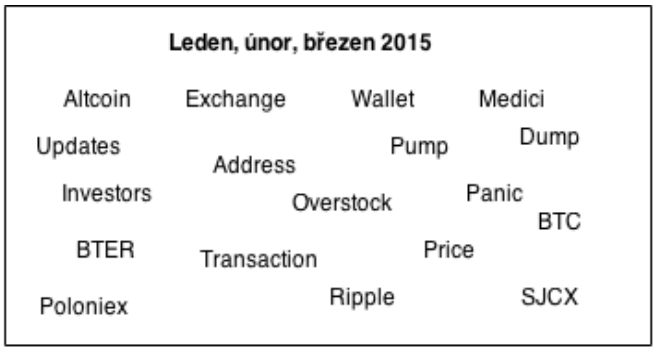
6.3.5.3 Vývoj témat v roce 2015

Na začátku roku 2015 (viz Obrázek 106) uživatelé diskutovali na některá nová témata, jako je kryptoměna XCP, alternativní kryptoměna BTC, porovnávali ceny burz BTER a Poloniex, hovořilio tom, že Overstock.com vyřadila XCP z přijímajících kryptoměn.

[39] Factom je P2P síť uživatelů a federativních serverů, které ovládají decentralizované aplikace (obsahuje také DACs (digital-to-analog converter), který konvertuje digitální data na jiné analogie (ADC and DAC Glossary, 2002) a DAOs (decentralized autonomous organization) je decentralizovanou sítí autonomních agentů (*Swan*)) a protokol, který převádí data na hash pro verifikaci v síti bloků Bitcoin (Wagner, 2014).

Mezi uživateli pak panuje názor, že XCP využila technik Pump and Dump, že na začátku zájemci upoutali pozornost jiných uživatelů o této kryptoměně, zvýšili její hodnotu a potom se jí začali zbavovat, čímž vyvolali její prudký pokles. Z důvodu tohoto prudkého poklesu další vlastníci kryptoměny podléhají panice a také prodávají své kryptoměny, čímž cena XCP jde dolů.

Dalším tématem diskuze v kontextu s XCP kryptoměna Ripple. Ostatní témata byla popsána v předchozích subkapitolách.



Obrázek 106: Diskuze spojené s XCP v roce 2015, zdroj a vlastní zpracování

6.3.5.4 Vývoj kurzu

Kryptoměna XCP byla spuštěna v lednu roku 2014. Generace bloků probíhá prostřednictvím algoritmu PoB, který, jak už bylo naznačeno výše, se liší od PoW a PoS, tím pádem faktory ovlivňující vývoj kurzu XCP/USD jsou jiné.

V této subkapitole porovnám vývoj kurzu s výše identifikovanými tématy. Jako u všech předchozích analyzovaných kryptoměn po spuštění nové kryptoměny se nedá očekávat růst kurzu, jelikož podle identifikovaných diskuzí se uživatelé seznamují s touto novou kryptoměnou. Zajímají se o algoritmus PoB, jelikož neměli před tím s PoB žádné zkušenosti.

Podle identifikovaných diskuzí v březnu se dá tvrdit, že se více uživatelů začíná zajímat o tuto kryptoměnu, jelikož podle příspěvků stále více lidí píše o XCP v různých jiných zdrojích než Bitcoin Forum. Vývoj kurzu XCP/USD tato zjištěná skutečnost ovlivní pozitivně, protože čím více uživatelů se zajímá o kryptoměnu a čím více se o tom píše i na externích zdrojích, aby se o tom dozvěděli i ostatní lidé, tím větší je pravděpodobnost růstu kurzu.

Uživatelé mluví také často o kryptoměně BTC, jelikož nové XCP vznikaly prostřednictvím zničení BTC. A z grafu č. 17 a grafu č. 18 je vidět podobný vývoj kurzů XCP/USD a BTC/USD.

Podle grafu č. 17 je vidět, že kurz XCP/USD od začátku října do přibližně půlky října stoupal, také kurz prudce rostl od listopadu roku 2014. Na tento vývoj uživatelé reagovali příspěvky o příležitosti investovat do XCP a byli také upoutáni vyvíjenou aplikaci umožňující vsázet ve sportu. Probíhaly i různé marketingové akce. Chování vývojářů při snaze přeložit jejich webovou stránku se také dá považovat za pozitivní signál růstu kurzu, jelikož přeložená stránka do několik různých jazyků umožní uživatelům dobře se v ní orientovat a tím bude povzbuzena jejich spokojenost v pochopení obsahu, tím i možný růst zájmu o kryptoměnu XCP.

Graf č. 17 ukazuje, že kurz XCP/USD má tendenci klesat, přičemž za jednu z příčin tohoto poklesu bez ohledu na ekonomické důvody lze podle obsahu tvořeného na Bitcoin Forum považovat vyřazení kryptoměny XCP z přijímajících kryptoměn elektronickým obchodem Overstock. To znamená, že někteří uživatelé, kteří mají zájem nakupovat něco v tomto obchodě, už nemají možnost utratit zde své kryptoměny, a proto je mění na ty, které Overstock přijímá. Za

důležitou příčinu podle identifikovaných témat považuji také využití techniky Pump and Dump, které vysvětluje prudký růst a prudký pokles kurzu XCP/USD na konci roku 2014 a na začátku roku 2015. Potom lze ekonomický vysvětli, že v době prudkého poklesu kurzu ostatní uživatelé začínají také prodávat se záměrem prodat za přiměřenou částku.



Graf 17: Vývoj kurzu XCP/USD od února roku 2014 do konce dubna roku 2015, zdroj CoinGecko (2015)



Graf 18: Vývoj kurzu BTC/USD od března roku 2014 do konce března roku 2015, zdroj Blockchain Info a vlastní zpracování

6.3.5.5 Dílčí shrnutí

Jako u všech předchozích analyzovaných kryptoměn v době seznámení s kryptoměnou uživatelé s ní aktivně neobchodují a tím pádem i kurz nemá tendenci růst. V době, když kurz začíná růst, uživatelé reagují na to porovnáním s jinými kryptoměnami a diskutováním o příležitostech investování. Uživatelé také opět diskutují o peněženkách, transakcích a jejich poplatcích, adresách, chybách a o burzách. U této komunity je sledováno aktivní využití různých marketingových akcí, které podpořily růst zájmů a tím i růst kurzu.

Naopak v době poklesu kurzu uživatelé přestávají aktivně diskutovat ohledně témat, která se týkají kryptoměny. Diskuze postupně začíná hasnout. Zase vzniká názor na využití techniky Pump and Dump.

Tato diskuze může novým uživatelům přiblížit koncepcí kryptoměny, napomoci při určitých technických problémech, vysvětlit možnou budoucnost kryptoměny a možné příčiny výkyvů kurzu.

6.3.6 Shrnutí poznatků – globální pohled

Globální pohled na výsledky analýzy vykazuje společné prvky a umožnuje tak identifikovat obecně fáze vývoje kryptoměn v čase (viz Schéma 2):

- Diskuze spojená se vznikem kryptoměny,
- Diskuze spojená s rozběhnutím a rozšířením kryptoměny mezi uživateli,
- Témata spojená s poklesem kurzu určité kryptoměny a útlumem aktivity uživatelů ohledně diskutování o určité kryptoměně.

Celý průběh jsem zobrazila na schématu č. 2 a dále uvádím popis tohoto schématu podrobněji podle bodů, kterými jsem označila jednotlivé šipky na tomto schématu.

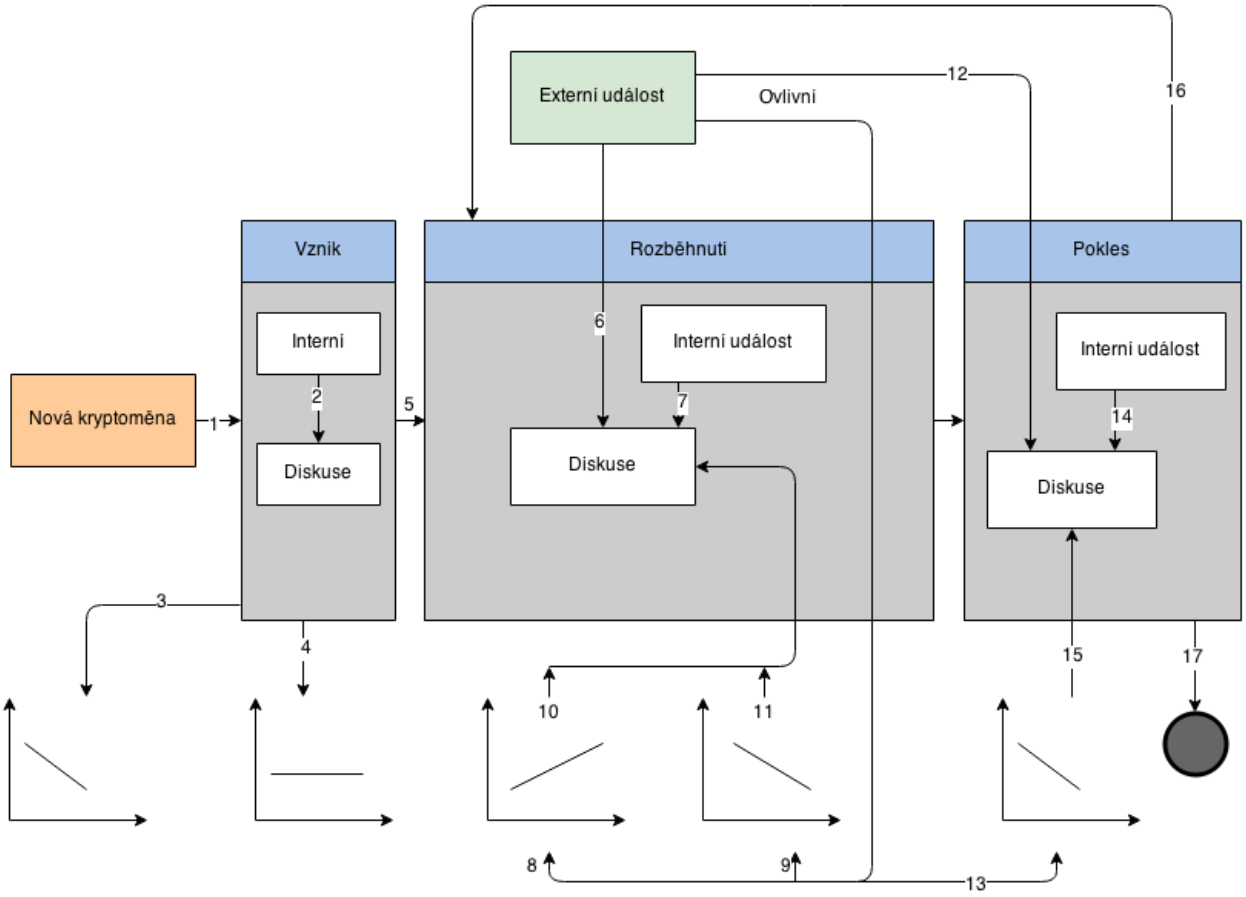


Schéma 2: Globální pohled na vztah mezi vývojem témat v čase a kurzem kryptoměn Bitcoin, Monero, Nxt, Counterparty a částečně Dash, zdroj vlastní zpracování

[1] Vznik nové kryptoměny,

[2] Interní témata, která ovlivní na témata diskusí,

[3] Negativní vliv na kurz vlivem diskusí ve fázi vzniku nové kryptoměny,

[4] Neutrální vliv na kurz vlivem diskusí ve fázi vzniku nové kryptoměny,

[5] Přechod z fází vzniku kryptoměny do fáze rozběhnutí,

[6] Vliv externích událostí na témata diskusí ve fázi rozběhnutí,

[7] Vliv interních událostí na témata diskusí ve fázi rozběhnutí,

[8] Pozitivní vliv externích událostí na vývoj kurzu,

[9] Negativní vliv externích událostí na vývoj kurzu,

[10] Vliv stoupajícího kurzu na vývoj témat diskusí,

[11] Vliv klesajícího kurzu na vývoj témat diskusí,

[12] Vliv externích událostí na témata diskusí ve fázi poklesu,

[13] Negativní vliv externích událostí na vývoj kurzu,

[14] Vliv interních událostí na témata diskusí ve fázi poklesu,

[15] Vliv klesajícího kurzu na vývoj témat diskusí,

[16] Přechod z fáze poklesu zpět do fáze rozběhnutí (oživení),

[17] Deprese.

Vzniká nová kryptoměna, která je na Bitcoin Forum prezentována tvůrci uživatelům fóra [1]. Vznikají určitá témata diskusí (interní), která se tykají koncepce dané kryptoměny [2]. V této fázi jsem identifikovala u pěti analyzovaných kryptoměn následující společná témata:

a. Princip fungování určité kryptoměny,
b. Algoritmus generace nových bloků,
c. Technologie, které jsou využívané pro těžbu kryptoměny,
d. Pooly, které se využívají pro těžbu dané kryptoměny,
e. Výběr peněženky,
f. Vyladění technických problémů, spojených s instalováním peněženky,
g. Dotazy ohledně adres peněženek,
h. Propagace kryptoměny,
i. Giveaway.

Společná identifikovaná témata ovlivňují vývoj kurzu negativně [3]. Kurz určité kryptoměny vůči americkému dolaru klesá v důsledku toho, že uživatelé prostřednictvím tzv. giveaway určité částky kryptoměn obdrželi v době jejího vzniku, ale vlivem malého zájmu o nákup uživatelé prodávají vlastněné nové kryptoměny a tím i klesá kurz.

Kurz určité kryptoměny je vůči americkému dolaru stabilní [4]. Důvodem je držení kryptoměny uživateli s očekáváním růstu kurzu. Vliv těchto témat na vývoj kurzu je stabilní či klesající z důvodu malého zájmů o kryptoměnu. Uživatelé se s kryptoměnou v době vzniku pouze seznamují a tím je způsobena menší aktivita při obchodování.

Témata se postupem času mění a dá se tvrdit, že daná nová kryptoměna dostala do fáze rozběhnutí [5].

V okamžiku rozšíření kryptoměny se o ní začínají zajímat i finanční nebo statní instituce. Mohou se stát určité události externího charakteru, jako [6]:

a. Krádež,
b. Útoky na peněženky a burzy,
c. Politické vlivy na obchodování kryptoměnou,
d. Vliv bank na obchodování kryptoměnou,
e. Vliv určitých burz na obchodování s danou kryptoměnou,
f. Přijetí kryptoměny určitými obchody, burzami nebo službami.

Tato externí událost ovlivňuje průběh diskusí. Uživatelé na tyto události a společně o nich diskutují na Bitcoin Forum.

Stále zůstávají populární i interní společná témata pro všechny analyzované kryptoměny [7], která ovlivňují směr diskusí. U pěti analyzovaných kryptoměn jsem identifikovala následující společná témata v této fázi, která se tykají:

a. Otázek, spojených s koncepcí určité kryptoměny,
b. Řešení technických problému,
c. Výběr poolů pro těžbu kryptoměny,
d. Technologie využité pro těžbu.

Na vývoj kurzu mají vliv externí události, které byly uvedené v bodě č. 6. Tyto události ovlivňují vývoj kurzu určité kryptoměny vůči americkému dolaru buď pozitivně a vývoj kurzu stoupá [8], anebo negativně a vývoj kurzu klesá [9].

Kromě externích událostí jsou předmětem diskusí také rostoucí výkyvy kurzu [10], nebo klesající výkyvy kurzu [11]. V této fázi rozběhnutí kryptoměny, vývoj kurzu ovlivní na vývoj témat v čase. Ne naopak.

Ve fázi neustálého poklesu kurzu a útlum aktivity uživatelů v diskutování ohledně dané kryptoměny mají stále vliv na vývoj témat externí události [12], které byly uvedeny v bodě č. 6. Uživatelé stále reagují na externí události i v době poklesu aktivity nebo neustálého klesání kursu.

Ve fázi rozběhnutí mají na vývoj kurzu vliv externí události, které byly uvedené v bodě č. 6. Ty jej ovlivňují negativně. Kurz má z těchto důvodů tendenci klesat [13]. Na vývoj témat v čase mají samozřejmě vliv i interní události, které byly uvedené v bodě č. 7 [14].

Další vliv na vývoj témat v čase má samotný vývoj kurzu [15]. V této fázi uživatelé reagují na pokles kurzu určitými příspěvky, které se vyskytují i v jiných fázích, ale během fáze poklesu se objevují nejčastěji:

a. Spekulace o využití manipulační techniky Pump and Dump,
b. Diskuze o příležitostech investování,
c. Zvažování budoucnosti,
d. Diskuze o podvodech.

Pokud kryptoměna přežije fázi stálého poklesu kurzu a úpadek aktivity uživatelů na fóru, jelikož jedním ze základu fungování kryptoměn je zájem lidí, vrací se do fáze rozběhnutí [16], kdy se proces opakuje od bodu č. 6. Do této fáze se z pěti analyzovaných kryptoměn dostala pouze Bitcoin.

Pokud kryptoměna nepřežije fázi poklesu, dá se tvrdit, že se uživatelé o tu kryptoměnu zajímat dále nebudou a tím pádem nebudou diskutovat na témata, která jsou s ní spojená. Kurz poklesne na takovou úroveň, která nebude lákavá pro uživatele a samotná kryptoměna se dostane do tzv. deprese [17].

Z toho lze usoudit, že samotná diskuse na Bitcoin Forum ohledně analyzované kryptoměny ovlivní kurz v první fázi, fázi vzniku kryptoměny. V ostatních fázích rozběhnutí a poklesu kryptoměny diskuse nejsou primární příčinou výkyvů kurzů kryptoměn vůči americkému dolaru. Primárním zdrojem se stávají externí události, které pak uživatelé aktivně projednávají na Bitcoin Forum.

Je třeba zmínit, že na rozdíl od Bitcoinu, Nxt, Monero a Counterparty je komunita kryptoměny Dash jediná, kde se uživatelé snaží predikovat vývoj kurzu. V této práci jsem prokázala, že v určitých obdobích uživatelé doporučují nákup, či prodej správně, jelikož kurz se začal pohybovat stejným směrem, o kterém psali uživatelé na Bitcoin Forum. Lze tedy tvrdit, že schéma č. 2 pro kryptoměnu Dash neplatí v bodech č. 10, 11 a 15, kde by šipky měly jít

opačným směrem, jelikož na rozdíl od Bitcoin, Nxt, Monero a Counterparty mají diskuse komunity Dash vliv na vývoj kurzu, ale nikoliv opačně, jak jsem to prokázala u kryptoměn Bitcoin, Nxt, Monero a Counterparty. Tímto závěrem potvrzují existenci vlivu diskusí na vývoj kurzu kryptoměny Dash.

7 Závěr

V diplomové práci jsem se zaměřila na analýzu vývoje hodnoty kryptoměn s dynamikou uživateli tvořeného obsahu na Internetu. Hlavními cíli práce bylo analyzovat nestrukturovaná data se zaměřením na vývoj témat jednotlivých kryptoměn v čase (cíl 1, kapitola 6) a propojit klíčová témata a souvislosti s vývojem hodnoty jednotlivých kryptoměn (cíl 2, kapitola 6). Domnívám se, že se podařilo naplnit cíle stanovené v úvodu práce. Data ze zvoleného zdroje byla zanalyzována v čase a z nich identifikována klíčová témata propojená s vývojem kurzu vybraných pěti nejdiskutovanějších kryptoměn a amerického dolaru.

Celá práce byla rozdělená na tři základní části: na teoretickou, návrh řešení a samotnou analytickou část.

Teoretická část splnila podpůrné cíle. Provedení teoretického zakotvení analýzy nestrukturovaných dat (cíl 3, kapitola 1, 2, 3) a vysvětlení problematiky kryptoměn (cíl 4, kapitola 4).

V části návrhu řešení bylo dosaženo dalšího podpůrného cíle, který spočíval v identifikaci internetových zdrojů s vysokým výskytem příspěvků o kryptoměnách (cíl 5, kapitola 5). Ze všech analyzovaných zdrojů byl zvolen Bitcoin Forum, který v době psaní této diplomové práce obsahoval největší počet příspěvků.

V analytické části bylo dosaženo kromě hlavních cílů také podpůrného cíle, který spočíval ve zmapování 20 nejdiskutovanějších kryptoměn ve zvoleném zdroji (cíl 6, kapitola 6). Z těchto nejdiskutovanějších kryptoměn bylo vybráno pět kryptoměn, z nichž tři se liší algoritmem generace bloků. Zvolenými kryptoměnami byly Bitcoin, Dash, Monero, Nxt a Counterparty. U těchto vybraných kryptoměn pak byla identifikována klíčová témata, která byla porovnána s vývojem jejich kurzu.

Z vystupů analyzy bylo zjištěno, že odlišnosti v metodách generace bloků mají vliv na témata týkající se technologií těžby mincí, vliv odlišností v metodách generace bloků na ostatní témata v této práci nebyl nalezen.

Z výsledků analýzy příspěvků na Bitcoin Forum je zřejmé, že existují určité společné fáze vývoje kryptoměn v čase. V době vzniku nové kryptoměny uživatelé diskutují o její koncepci, vlastnostech, peněženkách a technickém základu. V době rozběhnutí kryptoměny, uživatelé řeší určité technické problémy, diskutují o různých zprávách z externích zdrojů nebo o existujících poolech. Také se přidávají do diskuzí noví zájemci a ptají se na základní otázky spojené s principem fungování kryptoměny. Některé komunity diskutují o vývoji kurzu kryptoměny a snaží se ho určitým způsobem predikovat. Z provedené analýzy v této práci je patrné, že z pěti analyzovaných kryptoměn o vývoji kurzu nejvíce diskutuje komunita zastupující kryptoměnu. Další fází je úpadek aktivity uživatelů a prosazení názoru o selhání určité kryptoměny. Následuje obnovení aktivity uživatelů a opakování procesu znovu od druhé etapy. Z analyzovaných kryptoměn se pouze komunita Bitcoinu dostala do fáze obnovení. Ostatní komunity zvolených kryptoměn prošly pouze fázemi mezi rozběhnutím a úpadkem. Nebo následuje fáze deprese kryptoměny.

Po propojení této struktury a vývoje kurzu jednotlivých kryptoměn vůči americkému dolaru bylo zjištěno, že v době vzniku kryptoměny má její kurz tendenci klesat či být stabilní, jelikož se uživatelé seznamují s touto kryptoměnou. V této fázi jí aktivně nekupují. Ve střední fázi se po seznámení s kryptoměnou vývoj kurzu mění v souvislosti s určitými faktory, na některé uživatelé aktivně reagují svými příspěvky na Bitcoin Forum. Bylo zjištěno, že komunita Bitcoin reaguje na události spojené s politickým jednáním vůči Bitcoinu, na některé činitele ekonomického charakteru, události související s propagováním nebo přijetím Bitcoinu různými obchody, burzami nebo službami, různé krádeže, útoky a podvody a samozřejmě a na samotný

vývoj kurzu. Ostatní analyzované komunity kryptoměn Nxt, Dash, Monero a Counterparty aktivně reagují na stejné události, kromě události spojené s politickým jednáním vůči jednotlivým kryptoměnam, které u těchto komunit na fóru chybí. U reakcí na vývoj kurzu těchto komunit kromě komunity Bitcoinu bylo identifikováno další společné téma využití techniky Pump and Dump, která vede na začátku k prudkému růstu kurzu a následně k rychlému poklesu kurzu. V poslední fázi poklesu aktivity uživatelů a prosazení názoru o neúspěchu kryptoměny na fóru má vývoj kurzu tendenci klesat.

Prostřednictvím uskutečněné analýzy je také sledována tendence vývoje jednotlivých témat, na základě kterých je vidět zvýšení či pokles zájmů o kryptoměnu. Práce ukázala tendenci ke změně výběru určité technologie pro těžbu kryptoměn uživateli, v diskuzích o reklamě, investování, externích událostech působící na vývoj kurzu kryptoměny a krádežích či útocích.

Dalším výsledkem provedené analýzy je skutečnost, že vývoj kurzu kryptoměn Nxt, Dash a Counterparty vůči americkému dolaru je v přímé závislosti s vývojem kurzu Bitcoin vůči americkému dolaru, a proto při vysvětlení změn ve vývoji kurzu alternativních kryptoměn vůči americkému dolaru je třeba spoléhat i na faktory získané z Bitcoin Forum, ovlivňující výkyvy kurzu Bitcoin vůči americkému dolaru.

Komunity kryptoměn Bitcoin, Nxt, Monero, Dash a Counterparty přibližují problematiku těchto kryptoměn novým a stalým uživatelům fóra, napomáhají v řešení určitých technických problémů, poskytují informace o důvodech výkyvů kurzu určité z analyzovaných kryptoměn vůči americkému dolaru a informování o určitých událostech spojených s danou kryptoměnou

Přínos mé práce vidím v provedené analýze diskuzního fóra Bitcoin Forum, ve kterém jsem identifikovala vývoj témat ohledně pěti vybraných kryptoměn v čase a zjistila reakci uživatelů na vývoj kurzu těchto kryptoměn vůči americkému dolaru. Za dílčí přínos považuji praktickou ukázku využití vybrané technologie při analýze nestrukturovaných dat. Za další dílčí přínos vidím způsobu zpracování analýzy nestrukturovaných dat.

V budoucnu je možné rozšíření analýzy o navazující témata směřující především do oblasti vytváření kryptoměn za účelem získání rychlého zisku určitých zájmových skupin. Jelikož v mé diplomové práci byly sledovány názory o využití manipulační techniky Pump and Dump a další podvody uskutečněné prostřednictvím alternativních kryptoměn za účelem zbohatnutí určité skupiny lidí, je tady prostor na další analýzu nestrukturovaných dat uskutečněnou nad altcoinami.

Seznam literatury

Normy

- ČSN ISO/IEC 2382-1, 1998. *Informační technologie - Slovník - Část 1: Základní termíny*. Praha: Český normalizační institute. s. 59. Třídící znak 369001.
- IEC 60027-2, 2005. Letter symbols to be used in electrical technology - Part 2: Telecommunications and electronics. Geneva: IEC, s. 145.
- IEEE 1320.2, 1998. *Standard for Conceptual Modeling Language – Syntax and Semantics for IDEF1X97*. New York: IEEE Computer Society.
- ISO/IEC 2382-1, 1993. *Information technology -- Vocabulary -- Part 1: Fundamental terms.* Geneva: ISO's Central Secretariat. s. 32 s.
- ISO/IEC 15939, 2007. *Systems and software engineering -- Measurement process*. Geneva: ISO's Central Secretariat. s. 38.
- ISO/IEC 10746-2, 2009. *Information technology -- Open Distributed Processing -- Reference Model: Foundations.* Geneva: ISO's Central Secretariat. s. 22.
- ISO/IEC/IEEE 24765, 2010. *Systems and software engineering – Vocabulary*. Geneva: ISO's Central Secretariat. s. 410

Knihy

- ABITEBOUL, Serge, BUNEMAN, Peter, SUCIU, Dan, 1999. *Data on the Web: From Relations to Semistructured Data and XML*. Vyd. 3. San Francisco: Morgan Kaufmann, xiii, s. 257. ISBN- 978-1558606227.
- ANDĚL, Jiří, 2003. *Statistické metody.* Vyd. 3. Praha: Matfyzpress, s. 299. ISBN 80-867-3208-8.
- BERRY, Michael J a Gordon LINOFF, 2004. *Data mining techniques for marketing, sales, and customer relationship management*. 2nd ed. Indianapolis, Ind.: Wiley, xxv, 643 s. ISBN 0471470643.
- BUFORD, John F, YU, Heather, LUA, Eng Keong, 2009. *P2P networking and applications: an introduction to cluster analysis*. Vyd. 2. Burlington: Morgan Kaufmann, xxi, s. 415. The Morgan Kaufmann series in networking. ISBN 978-0-12-374214-8.
- CHEN, Min, MAO, Shiwen, ZHANG, Yin, LEUNG, Victor C.M., 2014. *Big Data. Related Technologies, Challenges and Future Prospects*. New York: Springer, s. 89. ISBN 978-3-319-06244-0.
- GORSHKOV, Michail a Franc SCHEREGI, 2012. *Прикладная социология. Методология и методы*. Moskva: FGANY, s. 404. ISBN 978-5-906001-05-4.
- HAN, Jiawei, KAMBER, Micheline, PEI, Jian, 2012. *Data Mining. Concepts and Techniques*. Waltham: Morgan Kaufmann, s. 703. ISBN 978-0-12-381479-1.
- KAUFMAN, Leonard a Peter J. ROUSSEEUW, 2005. *Finding Groups in Data: An Introduction to Cluster Analysis*. 3rd ed. Hoboken, N.J.: Wiley, xiv, s. 342. ISBN 0-471-73-578-7.
- KUĆ, Rafał, a Marek, ROGOZIŃSKI, 2014. *Elasticsearch Server. Second Edition*. Birmingham: Packt Publishing, s. 400. ISBN 978-1-78398-052-9.

- REVENDA, Zbyněk, MANDEL, Martin, KODERA, Jan, MUSÍLEK, Petr, DVOŘÁK, Petr, BRADA, Jaroslav, 2005. *Peněžní ekonomie a bankovnictví.* Vyd. 4. Praha: Management Press, s. 627. ISBN 80-726-1132-1.
- RUNKLER, Thomas A., 2012. *Data Analytics. Models and Algorithms for Intelligent Data Analysis*. Vyd. 3. New York: Springer Vieweg, ix, s. 137. ISBN 978-3-8348-2588-9.
- SKLENÁK, Vilém, BERKA Petr, RAUCH Jan, STROSSA Petr, SVÁTEK Vojtěch, 2001. *Data, informace, znalosti a Internet*. Praha: C.H. Beck, xvii, s. 507. ISBN 80-7179-409-0.
- TUFFERY, Stéphane, 2011. *Data mining and statistics for decision making*. Hoboken, NJ.: Wiley, xxiv, s. 689. Wiley series in computational statistics. ISBN 978-047-0688-298.
- ZIKOPOULOS, Paul, ROUSSEEUW, Peter J, PEI, Jian, 2012. *Understanding big data: analytics for enterprise class Hadoop and streaming data*. Vyd. 3. New York: McGraw-Hill, xxxi, 141 p. Wiley series in computational statistics.ISBN 978-0-07-179053-6.

Články

- COCCO, Luisanna, CONCAS, Giulio, MARCHESI, Michele, 2014. Using an Artificial Financial Market for studying a Cryptocurrency Market. s. 13.
- HASHIMOTO, Takako, KUBOYAMA, Tesuji, CHAKRABOTY, Basabi, SHIROTA, Yukari, 2012. Discovering Topic Transition about the East Japan Great Earthquake in Dynamic Social Media. IEEE Global Humanitarian Technology Conference, s. 6.
- JOHNSON, Don, MENEZES, Alfred, VANSTONE, Scott, 2001. The Elliptic Curve Digital Signature Algorithm (ECDSA). SpringerLink. ISSN 1615-5262.
- KAPLAN, Andreas M. a Michael, HAENLEIN, 2010. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, s. 10.
- KITZINGER, Jenny, 1994. The methodology of Focus Groups: the importance of interaction between research participants. Sociology of Health & Illness, č. 16-1. ISSN 0141-9889.
- LUKOIANOVA, Tatiana a Victoria L., RUBIN, 2014. Veracity Roadmap: Is Big Data Objective, Truthful and Credible? Advances In Classification Research Online, 24(1).
- SHARMA, Deepika, 2012. Stemming Algorithms: A Comparative Study and their Analysis. International Journal of Applied Information Systems, č. 4-3, s. 6. ISSN 2249-0868.
- SHARMA, Sugam, TIM, Udoyara S, WONG, Johnny, GADIA, Shashi, SHARMA, Subhash, 2014. A brief review on leading big data models. Data Science Journal, č. 13, s. 20.
- SUNDARESAN, Balasubramaniyan a Dinesh KANDAVEL, 2014. Big Languages for Big data. A study and comparison of current trend in data processing techniques for Big data. s. 12.

Elektronické zdroje

- ADC and DAC Glossary, 2002. *Maxim integrated* [online]. [cit. 2015-04-19]. Dostupné z: http://www.maximintegrated.com/en/app-notes/index.mvp/id/641.
- AIMCOIN FORUM. *Aimcoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: https://www.aimcoin.com/.
- APPLEBYTE. *AppleByte* [online]. [cit. 2015-03-08]. Dostupné z: http://applebyteme.invisionzone.com/index.php?.
- ARMORY, 2015. *Armory Secure Wallet* [online]. [cit. 2015-03-08]. Dostupné z: https://bitcoinarmory.com/.
- Automated Teller Machine – ATM, 2015. *Investopedia* [online]. [cit. 2015-04-05]. Dostupné z: http://www.investopedia.com/terms/a/atm.asp.
- BEAL, Vangie. Unstructured data. In: *Webopedia* [online]. [cit. 2015-01-10]. Dostupné z: http://www.webopedia.com/TERM/U/unstructured_data.html.
- BEIKVERDI, Alireza, 2015. Proof-of-Importance: How NEM is Going to Add Reputations to the Blockchain. In: *The Cointelegraph* [online]. [cit. 2015-03-24]. Dostupné z: http://cointelegraph.com/news/113698/proof-of-importance-nem-is-going-to-add-reputations-to-the-blockchain.
- Bitcoin, 2015. In: *Coinbitcoin.info* [online]. [cit. 2015-02-26]. Dostupné z: http://coinbitcoin.info/.
- BITCOIN FORUM. *Bitcoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: https://bitcointalk.org/.
- BITCOIN FOUNDATION, 2015. *Bitcoin Foundation* [online]. [cit. 2015-04-04]. Dostupné z: http://bitcoinfoundation.org/.
- BITCOIN GARDEN FORUM. *Bitcoin Garden Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://bitcoingarden.tk/forum/.
- BITCOIN SECURITY, 2015. Русскоязычный информационный сайт о криптовалюте Bitcoin: Что такое Bitcoin и криптовалюты? In: *Bitcoin Security* [online]. [cit. 2015-03-18]. Dostupné z: http://bits.media/chto-takoe-Bitcoin/.
- Bitcoinica. In: *Bitcoin wiki* [online]. [cit. 2015-04-04]. Dostupné z: https://en.bitcoin.it/wiki/Bitcoinica.
- BITCOINJ. *Bitcoinj* [online]. [cit. 2015-03-08]. Dostupné z: https://bitcoinj.github.io/.
- BITCOINTALK-N-SHOP. *Bitcointalk-n-shop* [online]. [cit. 2015-03-08]. Dostupné z: https://www.bitcointalknshop.com/.
- BITFINEX. *Bitfinex: invest in the future* [online]. [cit. 2015-04-17]. Dostupné z: https://www.bitfinex.com/.
- BITMAKLER. In: *Bitmakler.com* [online]. [cit. 2015-02-23]. Dostupné z: http://en.bitmakler.com/.
- BITSHARES. *Bitshares* [online]. [cit. 2015-03-08]. Dostupné z: https://bitsharestalk.org/.
- BITSHARES BLOCKS. BitShares bloks [online]. [cit. 2015-03-08]. Dostupné z: https://bitsharesblocks.com/.

- BLOCKCHAIN INFO. *Blockchaininfo* [online]. [cit. 2015-03-01]. Dostupné z: https://blockchain.info/.
- BLOCKSCAN. *Blockscan* [online]. [cit. 2015-04-19]. Dostupné z: https://www.blockscan.com/tx?p=1.
- BOOLBERRY COMMUNITY. *Boolberry Community* [online]. [cit. 2015-03-08]. Dostupné z: http://boolberry.com/forum/.
- BRADBURY, Danny, 2013. Tradehill open to holding more auctions as it grapples with regulation. In: *CoinDesk* [online]. [cit. 2015-04-03]. Dostupné z: http://www.coindesk.com/tradehill-open-holding-auctions-grapples-regulation/.
- BULANT, Michal, 2001. Kryptografie a její aplikace. In: *CiteSeerX* [online]. [cit. 2015-02-23]. Dostupné z: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.214.8521&rep=rep1&type=pdf.
- BURST. *Burst* [online]. [cit. 2015-02-24]. Dostupné z: http://burstcoin.info/.
- BURSTCOIN.EU. Burstcoin.eu [online]. [cit. 2015-02-24]. Dostupné z: http://burstcoin.eu/.
- CARROT2. *Carrot2* [online]. [cit. 2015-03-08]. Dostupné z: http://project.carrot2.org/index.html.
- CENTRAL INTELLIGENCE AGENCY. *Central Intelligence Agency* [online]. [cit. 2015-04-05]. Dostupné z: https://www.cia.gov/index.html.
- CICHOCKI, Andrzej, 2013. Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions. In: *Cornell University Library* [online]. [cit. 2015-03-14]. Dostupné z: http://arxiv.org/pdf/1403.2048.pdf.
- COINGECKO, 2015. *CoinGecko* [online]. [cit. 2015-04-16]. Dostupné z: https://www.coingecko.com/en.
- COINZEN. *Coinzen* [online]. [cit. 2015-03-08]. Dostupné z: http://coinzen.org/.
- Colored Coins, 2014. *Nxt Wiki* [online]. [cit. 2015-04-15]. Dostupné z: http://www.thenxtwiki.org/wiki/Colored_Coins.
- COLOSSUSCOINTALK.ORG. *ColossusCoinTalk.org* [online]. [cit. 2015-03-08]. Dostupné z: http://colossuscointalk.org/.
- Contract for Differences – CFD, 2015. *Investopedia* [online]. [cit. 2015-04-04]. Dostupné z: http://www.investopedia.com/terms/c/contractfordifferences.asp.
- COUNTERPARTY, 2014. *Counterparty* [online]. [cit. 2015-04-19]. Dostupné z: http://counterparty.io/.
- CRYPTITALK. *Cryptitalk* [online]. [cit. 2015-03-08]. Dostupné z: http://cryptitalk.com/.
- CRYPTO-CURRENCY MARKET CAPITALIZATIONS. *Crypto-Currency Market Capitalizations* [online]. [cit. 2015-03-08]. Dostupné z: http://coinmarketcap.com/.
- CRYPTOCOINTALK. *Cryptocointalk* [online]. [cit. 2015-03-08]. Dostupné z: https://cryptocointalk.com/.

- Cryptocurrency Faucets. In: *Reddit* [online]. [cit. 2015-04-15]. Dostupné z: http://www.reddit.com/r/CryptoCurrency/comments/1vhy5q/cryptocoin_faucets_a_list_of_all_knownworking_ones/.
- CRYPTOMOMS, 2014. *Cryptomoms* [online]. [cit. 2015-03-08]. Dostupné z: http://cryptomoms.com/forum/.
- CURECOIN. *Curecoin* [online]. [cit. 2015-03-08]. Dostupné z: http://www.curecoin.us/.
- DarkSend Documentation. *Bravenewcoin* [online]. [cit. 2015-04-17]. Dostupné z: http://bravenewcoin.com/assets/Uploads/DarkSendDocumentation.pdf.
- Dash, 2015. In: *BitInfoCharts* [online]. [cit. 2015-04-16]. Dostupné z: https://bitinfocharts.com/darkcoin/.
- Dash Cryptocurrency, 2015. *CoinWarz* [online]. [cit. 2015-04-16]. Dostupné z: http://www.coinwarz.com/cryptocurrency/coins/dash.
- DASHCOIN, 2014. *Dashcoin* [online]. [cit. 2015-04-17]. Dostupné z: http://dashcoin.net/.
- DASHTALK. *DashTalk* [online]. [cit. 2015-03-08]. Dostupné z: https://dashtalk.org/.
- DEWELLER. CounterpartyXCP: Documentation. *GitHub* [online]. [cit. 2015-04-19]. Dostupné z: https://github.com/CounterpartyXCP/Documentation/blob/master/Developers/protocol_specification.md.
- DIAMOND, 2014. *Diamond* [online]. [cit. 2015-03-08]. Dostupné z: https://bit.diamonds/community/.
- DIGITALKCOIN, 2015. *Digitalcoin* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.digitalcoin.co/.
- DIMECOIN OFFICIAL FORUMS. *Dimecoin Official Forums* [online]. [cit. 2015-03-08]. Dostupné z: http://dimecoin.freeforums.net/.
- Diskuze. In: *Slovník cizích slov* [online]. [cit. 2015-03-08]. Dostupné z: http://slovnik-cizich-slov.abz.cz/web.php/slovo/diskuze-diskuse-dyskuse.
- DOGECHAIN. *Dogechain: Dogecoin blockchain explorer* [online]. [cit. 2015-04-20]. Dostupné z: https://dogechain.info/.
- DOLLENTAS, Nigel, 2014. Hyper: Coin Overview. *Bitcoinist.net* [online]. [cit. 2015-04-20]. Dostupné z: http://bitcoinist.net/hyper-coin-overview/.
- DWOLLA, 2015. *Dwolla* [online]. [cit. 2015-04-05]. Dostupné z: https://www.dwolla.com/.
- ECB, 2012. Virtual Currency Schemes. In: *ECB* [online]. [cit. 2015-02-18]. Dostupné z: http://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf.
- ECOINER, 2015. *Ecoiner* [online]. [cit. 2015-03-08]. Dostupné z: http://ecoiner.org/.
- ELASTIC. *Elastic* [online]. [cit. 2015-03-07]. Dostupné z: https://www.elastic.co/products/elasticsearch.
- ELASTICSEARCH-HEAD. In: *Github* [online]. [cit. 2015-03-07]. Dostupné z: http://mobz.github.io/elasticsearch-head/.

- ELECTRUM. *Electrum Bitcoin Wallet* [online]. [cit. 2015-03-08]. Dostupné z: https://electrum.org/#home.
- ENERGYCOIN. *Energycoin* [online]. [cit. 2015-03-08]. Dostupné z: http://energycoin.pw/forum/.
- FASTCOIN FORUM. *Fastcoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://www.fastcointalk.org/.
- FBI. *Federal Bureau of Investigation* [online]. [cit. 2015-04-05]. Dostupné z: http://www.fbi.gov/.
- FEATHERCOIN. *Feasthercoin* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.feathercoin.com/.
- FED, 2015. *Board of Governors of the Federal Reserve System* [online]. [cit. 2015-04-08]. Dostupné z: http://www.federalreserve.gov/.
- FINCEN. FinCEN [online]. [cit. 2015-04-08]. Dostupné z: http://www.fincen.gov/.
- FREICOIN COMMUNITY. *PhPBB* [online]. [cit. 2015-03-08]. Dostupné z: http://freicoin.freeforums.org/index.php.
- FTVS, 2012. Analýza a interpretace dat. In: *Fakulta tělesné výchovy a sportu Univerzity Karlovy* [online]. [cit. 2015-01-25]. Dostupné z: http://www.ftvs.cuni.cz/hendl/metodologie/analyza.htm.
- GANTZ, John a David REINSEL, 2011. Extracting Value from Chaos. In: *EMC*^{2} [online]. [cit. 2015-01-11]. Dostupné z: http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf.
- GILLESPIE, Clay Michael, 2014. Mintpal's 3700 stolen Bitcoin likely in hands of alleged scammer Alex Green. *Cryptocoins news* [online]. [cit. 2015-04-17]. Dostupné z: https://www.cryptocoinsnews.com/mintpals-3700-stolen-bitcoin-likely-in-hands-of-alleged-scammer-alex-green/.
- GOLDCOIN FORUM. *GoldForum* [online]. [cit. 2015-03-08]. Dostupné z: https://www.gldtalk.org/.
- GROESTLCOIN. *g*^{2} [online]. [cit. 2015-03-08]. Dostupné z: http://www.groestlcoin.org/forum/.
- GUUGLL, 2015. XEM & New Economy Movement. *Guugll* [online]. [cit. 2015-04-20]. Dostupné z: http://www.guugll.eu/xem-new-economy-movement/.
- GUZMAN, Aaron, 2014. In: *Regular Writing* [online]. [cit. 2015-03-08]. Dostupné z: http://www.regularwriting.com/2014/02/cryptocurrency-forums/.
- HEID, Alex. Analysis of the Cryptocurrency Marketplace. In: *Hackmiami* [online]. [cit. 2015-03-06]. Dostupné z: http://www.hackmiami.org/whitepapers/HackMiami-Analysis_of_the_Cryptocurrency_Marketplace.pdf.
- HEYDON, Allan a Marc NAJORK, 1999. Mercator: A scalable, extensible Web crawler. In: *NLP Group* [online]. [cit. 2015-02-11]. Dostupné z: http://nlp.uned.es/WebMining/Tema2.Crawling/heydon1999.pdf.
- HIRSCH, Jorge, 2009. BCS theory of superconductivity: it is time to question its validity. In: *UC San Diego Department of Physics* [online]. [cit. 2015-04-03]. Dostupné z: http://physics.ucsd.edu/~jorge/abstracts/bcs.pdf.

- HUNTERCOIN FORUM. *Huntercoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://forum.huntercoin.org/.
- How BitShares Works. In: BitShares [online]. [cit. 2015-03-08]. Dostupné z: https://bitshares.org/how-it-works.
- HYPER BLOCKCHAIN EXPLORER. *CryptoID* [online]. [cit. 2015-04-20]. Dostupné z: https://chainz.cryptoid.info/hyper/.
- HYPER FORUM. *Hyper Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://hypercrypto.com/forum/.
- ILIC, Milos, SPALEVIC, Petar, VEINOVIC, Mladen, 2014. Suffix Tree Clustering – Data mining algorithm. In: *ERK 2014* [online]. [cit. 2015-02-05]. Dostupné z: http://erk.fe.uni-lj.si/2014/ilic%28suffix_tree%29p.pdf.
- INFINITECOINTALK. *IPO* [online]. [cit. 2015-03-08]. Dostupné z: http://infinitecointalk.org/.
- Initial Public Offering – IPO, 2015. In: *Investopedia* [online]. [cit. 2015-04-15]. Dostupné z: http://www.investopedia.com/terms/i/ipo.asp.
- INTERNATIONAL MONETARY FUND. *International Monetary Fund* [online]. [cit. 2015-04-08]. Dostupné z: http://www.imf.org/external/index.htm.
- IPCTALK.ORG. *IPC* [online]. [cit. 2015-03-08]. Dostupné z: http://www.ipctalk.org/.
- IRS. *IRS* [online]. [cit. 2015-04-08]. Dostupné z: http://www.irs.gov/.
- JASON DAVIES. *Jason Davies* [online]. [cit. 2015-03-18]. Dostupné z: https://www.jasondavies.com/.
- JAVA-SOURCE.NET. Crawler4j. In: *Java-Source.net* [online]. [cit. 2015-03-07]. Dostupné z: http://java-source.net/open-source/crawlers/crawler4j.
- JAVA TUTORIAL. *Java Tutorial* [online]. [cit. 2015-03-08]. Dostupné z: http://www.java2s.com/Tutorial/Java/CatalogJava.htm.
- KIRIČENKO, K. a M. GERASIMOV, 2001. Обзор методов кластеризации текстовой информации. In: *Диалог* [online]. [cit. 2015-02-15]. Dostupné z: http://www.dialog-21.ru/Archive/2001/volume2/2_26.htm.
- KIRK, David, 2014. Cryptocurrency: What is a fork? In: *Tech-Recipes* [online]. [cit. 2015-04-07]. Dostupné z: http://www.tech-recipes.com/rx/48517/cryptocurrency-what-is-a-fork/.
- KITTEHCOIN COMMUNITY CENTER. *KittehCoin Community Center* [online]. [cit. 2015-03-08]. Dostupné z: http://kittehcoin.info/forums/.
- Korelace. In: *Trilobyte* [online]. [cit. 2015-02-15]. Dostupné z: http://www.trilobyte.cz/downloadfree/qcemanual/correl.pdf.
- KURNIAWATI, Kurniawati, SHANKS, Graeme, BEKMAMEDOVA, Nargiza, 2013. The Business Impact of Social Media Analytics. In: *AIS Electronic Library* [online]. [cit. 2015-02-12]. Dostupné z: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1271&context=ecis2013_cr.
- LALOR, Thérèse, 2014. How Big is Big Data? In: *UNECE Statistics Wikis* [online]. [cit. 2015-01-15]. Dostupné z: http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=99484307.

- LEE, Jasen, 2014. Bitcoin: The realities of virtual currency. In: *Desert News* [online]. 2014 [cit. 2015-02-26]. Dostupné z: http://www.deseretnews.com/article/865600447/Bitcoin-The-realities-of-virtual-currency.html?pg=all.
- LITECOIN BLOCK EXPLORER. *Litecoin block exporer* [online]. 2013-2015 [cit. 2015-04-20]. Dostupné z: http://ltc.block-explorer.com/.
- LITECOINTALK. *LitecoinTalk* [online]. [cit. 2015-03-08]. Dostupné z: https://litecointalk.org/.
- LODHA, Riya, JAIN, Harshil, KURUP, Lakshmi, 2014. Big Data Challenges: Data Analysis Perspective. In: *Inpressco* [online]. [cit. 2015-01-08]. Dostupné z: http://inpressco.com/wp-content/uploads/2014/09/Paper373286-3289.pdf.
- LTBCOIN. *LTBcoin* [online]. [cit. 2015-04-19]. Dostupné z: http://ltbcoin.com/.
- MAGI BLOCKCHAIN EXPLORER. *CryptoID* [online]. [cit. 2015-04-20]. Dostupné z: https://chainz.cryptoid.info/xmg/.
- MAGI CHEAT SHEET. In: CoinMagi [online]. [cit. 2015-04-20]. Dostupné z: http://coinmagi.org/files/misc/magi-cheat-sheet-v1.pdf.
- MARINECOIN FORUM. *Marinecoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://forum.marinecoin.org/.
- MAX KEISER, 2015. *Max Keiser Financial War Reports* [online]. [cit. 2015-04-05]. Dostupné z: http://www.maxkeiser.com/.
- MCKINSEY GLOBAL INSTITUTE, 2011. Big data: The next fontier for innovation, competition, and productivity. In: *Nist Big Data* [online]. [cit. 2015-01-08]. Dostupné z: http://bigdatawg.nist.gov/MGI_big_data_full_report.pdf.
- MCMILLIAN, Robert, 2014. The inside story of Mt. Gox, Bitcoin's $460 million disaster. In: *Wired* [online]. [cit. 2015-04-04]. Dostupné z: http://www.wired.com/2014/03/bitcoin-exchange.
- MEGACOIN. *Megacoin* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.megacoin.co.nz/.
- MEMORYCOIN SQUARE. *Memorycoin Square* [online]. [cit. 2015-03-08]. Dostupné z: https://www.mmc-square.com/index.php.
- MIČKA, Pavel. Algoritmus RSA. *Algoritmy.net příručka vývojáře* [online]. [cit. 2015-04-03]. Dostupné z: http://www.algoritmy.net/article/4033/RSA.
- MINCOINFORUM.COM. *Mincoinforum.com* [online]. [cit. 2015-03-08]. Dostupné z: http://mincoinforum.com/.
- MONERO BLOCKCHAIN. In: *ChainRadar* [online]. [cit. 2015-04-16]. Dostupné z: http://chainradar.com/xmr/blocks.
- MT.GOX. In: *Bitcoin wiki* [online]. [cit. 2015-04-08]. Dostupné z: https://en.bitcoin.it/wiki/Mt._Gox.
- MULTIBIT, 2011. *Multibit* [online]. [cit. 2015-03-08]. Dostupné z: https://multibit.org/.
- NAIR, Ramesh, NARAYANAN, Andy, 2012. Benefitting from Big Data: Leveraging Unstructured Data Capabilities for Competitive Advantage. In: *Booz & Company Inc*. [online]. [cit. 2015-01-05]. Dostupné z:

http://www.strategyand.pwc.com/media/uploads/Strategyand_Benefitting-from-Big-Data.pdf.

- NAKAMOTO, Satoshi, 2009. Bitcoin: A Peer-to-Peer Electronic Cash System. In: *Bitcoin* [online]. [cit. 2015-02-22]. Dostupné z: https://bitcoin.org/bitcoin.pdf.
- NAMECOIN. *Namecoin* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.namecoin.info/.
- Namecoin Cryptocurrency, 2015. *CoinWarz* [online]. [cit. 2015-04-16]. Dostupné z: http://www.coinwarz.com/cryptocurrency/coins/namecoin.
- NAS. In: *Coinist* [online]. [cit. 2015-04-19]. Dostupné z: https://coinist.co/cryptocurrencies/nas.
- NAS, 2014. In: *Coins Source* [online]. [cit. 2015-04-19]. Dostupné z: http://www.coinssource.com/crypto-coins/nas.
- NEM. *Nem* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.ournem.com/index.php.
- Nem, 2014. *CryptoCoin updates* [online]. [cit. 2015-04-20]. Dostupné z: http://cryptocoinupdates.com/cryptocurrency-nem-features-100-original-codebase-fair-distribution-and-comprehensive-ecosystem/.
- New Economy Movement, 2015. *Nem* [online]. [cit. 2015-04-20]. Dostupné z: http://blog.nem.io/overview/.
- Novacoin Cryptocurrency, 2015. In: *CoinWarz* [online]. [cit. 2015-04-20]. Dostupné z: http://www.coinwarz.com/cryptocurrency/coins/novacoin.
- NOVACOIN EXPLORER. Novacoin Explorer [online]. [cit. 2015-04-20]. Dostupné z: https://explorer.novaco.in/.
- NSA, 2009. *National Security Agency/Central Security Service* [online]. [cit. 2015-04-05]. Dostupné z: https://www.nsa.gov/index.shtml.
- NXT, 2014. *Nxt* [online]. [cit. 2015-04-15]. Dostupné z: http://nxt.org/.
- NXT BLOCKCHAIN EXPLORER. *Nxt Blockchain Explorer* [online]. [cit. 2015-04-15]. Dostupné z: http://www.nxtexplorer.com/nxt/nxt.cgi?action=10.
- OSIŃSKI, Stanisław a Dawid, WEISS, 2002. Carrot^2. User and Developer Manual for version 3.9.0-Snapshot. In: *Carrot2* [online]. [cit. 2015-01-30]. Dostupné z: http://download.carrot2.org/head/manual/index.html#section.getting-started.
- OSIŃSKI, Stanisław a Dawid, WEISS, 2004. Conceptual Clustering Using Lingo Algorithm: Evaluation on Open Directory Project Data. In: *Instytut Informatyki Politechnika Poznańska* [online]. [cit. 2015-02-21]. Dostupné z: http://www.cs.put.poznan.pl/dweiss/site/publications/download/iipwm-osinski-weiss-2004-lingoeval.pdf.
- Over-The-Counter-OTC, 2015. In: *Investopedia* [online]. [cit. 2015-04-16]. Dostupné z: http://www.investopedia.com/terms/o/otc.asp.
- PAYCOIN. Paycoin [online]. [cit. 2015-04-16]. Dostupné z: https://paycoin.com/.
- PLANKTON. *Plankton* [online]. [cit. 2015-03-08]. Dostupné z: http://whalepod.org/forum/.

- PLUGIN HEAD. *Plugin head Elasticsearch* [online]. [cit. 2015-03-18]. Dostupné z: http://localhost:9200/_plugin/head/.
- Pump and Dump, 2015. In: *Investopedia* [online]. [cit. 2015-04-16]. Dostupné z: http://www.investopedia.com/terms/p/pumpanddump.asp.
- RAIT, Maks, 2015. BitShares 101: Руководство о том, как получить прибыть с Bitcoin следующего поколения, криптовалют и децентрализованного бизнеса. In: *Gitbook* [online]. [cit. 2015-03-17]. Dostupné z: http://testzcrypto.gitbooks.io/bitshares101/content/DPOS.html.
- RAM, Stefan L., 2004. Dostupné z: http://userpage.fu-berlin.de/~ram/pub/pub_jf47ht81Ht/code_kbyte_en.
- REAL RAW MILK FACTS. Why is raw milk illegal in some states? In: *Real raw milk facts* [online]. [cit. 2015-04-05]. Dostupné z: http://www.realrawmilkfacts.com/raw-milk-regulations.
- REDDCOIN FORUM. *Reddcoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: http://www.reddcoinforum.com/.
- ŘEZANKOVÁ, Hana, 2001. Postupy používané při analýze dat. In: *Banka dat a modelů ekonomiky ČR* [online]. [cit. 2015-01-26]. Dostupné z: http://badame.vse.cz/clanky/analyza-dat.php.
- SCIFICOINTALK, 2013. *SciFiCoinTalk* [online]. [cit. 2015-03-08]. Dostupné z: http://scificointalk.com/index.php.
- SEXCOINFORUM. *Sexcoinforum* [online]. [cit. 2015-03-08]. Dostupné z: http://www.sexcoinforum.org/.
- SHIBECOIN TALK. *Shibecoin Talk* [online]. [cit. 2015-03-08]. Dostupné z: http://forums.shibecoin.io/.
- SILK ROAD DRUGS. *Silk Road Drugs* [online]. [cit. 2015-04-04]. Dostupné z: http://silkroaddrugs.org/.
- SIMPLE MACHINES, 2013. *Simple Machines* [online]. [cit. 2015-03-08]. Dostupné z: http://www.simplemachines.org/.
- SNOWBALL. The Porter stemming algorithm. In: *Snowball* [online]. [cit. 2015-02-22]. Dostupné z: http://snowball.tartarus.org/algorithms/porter/stemmer.html.
- SOLARCOIN. *Solarcoin* [online]. [cit. 2015-03-08]. Dostupné z: http://solarcoin.org/forum/index.php.
- Statistics, 2015. In: *The free dictionary* [online]. [cit. 2015-01-26]. Dostupné z: http://www.thefreedictionary.com/statistics.
- STEFANOWSKI, Jerzy, OSIŃSKI, Stanisław, WEISS, Dawid, 2004. Lingo: Search Results Clustering Algorithm Based on Singular Value Decomposition. In: *Instytut Informatyki Politechnika Poznańska* [online]. [cit. 2015-02-21]. Dostupné z: http://www.cs.put.poznan.pl/dweiss/site/publications/download/iipwm-osinski-weiss-stefanowski-2004-lingo.pdf.
- STELLAR. In: *Coinist* [online]. [cit. 2015-04-19]. Dostupné z: https://coinist.co/cryptocurrencies/stellar.
- STELLARTALK. *StellarTalk* [online]. [cit. 2015-03-08]. Dostupné z: https://stellartalk.org/.

- STEWART, Darin, 2013. Big Content: The Unstructured Side of Big Data. In: *Gartner* [online]. [cit. 2015-01-10]. Dostupné z: http://blogs.gartner.com/darin-stewart/2013/05/01/big-content-the-unstructured-side-of-big-data/.
- SWAN, Melanie. Blockchain Thinking: The Brain as a DAC (Decentralized Autonomous Organization). *MS Futures Group* [online]. [cit. 2015-04-19]. Dostupné z: http://www.melanieswan.com/documents/BlockchainThinking_SWAN.pdf.
- SWARM. *Swarm* [online]. [cit. 2015-04-19]. Dostupné z: https://medium.com/@Swarm/state-of-the-swarm-7835b413f3d4.
- SYNC, 2014. In: *Coins Source* [online]. [cit. 2015-04-19]. Dostupné z: http://www.coinssource.com/crypto-coins/sync/.
- SYNC BLOCKCHAIN EXPLORER. *CryptoID* [online]. [cit. 2015-04-20]. Dostupné z: https://chainz.cryptoid.info/sync/.
- ŠKOLA, Petr, 2015. *Crawler* [program na CD-ROM].
- ŠKOLA, Petr, ŠPERKOVÁ, Lucie, TIMOFEEVA, Mariya, 2014. *Výsledky analýzy dat ze sociálních sítí – bankovnictví* [online prezentace]. [cit. 2015-04-20]. Dostupné z: https://onedrive.live.com/view.aspx?resid=A1E16B9BF2A2D7FE!2269&app=PowerPoint&authkey=!AF9IGKMiAgCdjiM.
- TAYLOR, Michael B, 2013. Bitcoin and the Age of Bespoke Silicon. In: *UC San Diego* [online]. [cit. 2015-03-06]. Dostupné z: http://cseweb.ucsd.edu/~mbtaylor/papers/bitcoin_taylor_cases_2013.pdf.
- TEACOIN (TEA). *Teacoin* (TEA) [online]. [cit. 2015-03-08]. Dostupné z: http://teacoin.forumotion.com/.
- TKAČENKO, Vladimír, 2007. Теоретические основы экономической информатики: Данные, информация и знания. In: *Обучение в интернет* [online]. [cit. 2015-01-05]. Dostupné z: http://www.lessons-tva.info/edu/e-inf1/inf1-1-2.html.
- THE BLACK FRIDAY, 2015. *The Black Friday* [online]. [cit. 2015-04-04]. Dostupné z: http://www.theblackfriday.com/what-is-black-friday.shtml.
- THE FAIRCOIN FORUM. *The FairCoin Forum* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.fair-coin.org/.
- THE IXCOIN FORUM. *iXC* [online]. [cit. 2015-03-08]. Dostupné z: http://ixc.boards.net/.
- The Ponzi Game, 2015. In: *Windows* [online]. [cit. 2015-04-15]. Dostupné z: http://apps.microsoft.com/windows/en-us/app/the-ponzi-game/5de563cd-c3c5-428e-bb4f-9388004f3bf1.
- TOR. *Tor* [online]. [cit. 2015-04-03]. Dostupné z: https://www.torproject.org/.
- TRAUTMAN, Lawrence, 2014. Virtual Currencies; Bitcoin & What Now After Liberty Reserve, Solk Road, and Mt. Gox? In: *Journal of Law & Technology* [online]. [cit. 2015-04-03]. Dostupné z: http://jolt.richmond.edu/v20i4/article13.pdf.
- TRISTAN WINTERS, 2014. Proof of Burn and the Counterparty Approach. In: *Bitscan* [online]. [cit. 2015-02-24]. Dostupné z: https://bitscan.com/articles/weekly-round-up/developments/proof-of-burn-and-the-counterparty-approach.

- Troll, 2010. In: *Dictionary.com* [online]. [cit. 2015-04-16]. Dostupné z: http://dictionary.reference.com/browse/troll.
- ULTRACOIN. *Ultracoin* [online]. [cit. 2015-03-08]. Dostupné z: http://ultracointalk.org/.
- URBÁNEK, Vladimír, 2013. Bitcoin strmě padá - Čína omezí jeho obchodování, Greenspan ho považuje za bublinu. In: *Kurzycz* [online]. [cit. 2015-04-08]. Dostupné z: http://www.kurzy.cz/zpravy/359792-bitcoin-strme-pada-cina-omezi-jeho-obchodovani-greenspan-ho-povazuje-za-bublinu/.
- VASIN, Pavel, 2014. BlackCoin's Proof-of-Stake Protocol v2. In: *BlackCoin* [online]. [cit. 2015-02-25]. Dostupné z: http://blackcoin.co/blackcoin-pos-protocol-v2-whitepaper.pdf.
- VAN DULMEN, Sandra, SLUIJS, Emmy, VAN DIJK, Liset, DE RIDDER, Denise, HEERDINK, Rob, BENSING, Jozien, 2008. The International Expert Forum on Patient Adherence. Furthering patient adherence: A position paper of the international expert forum on patient adherence based on an internet forum discussion. In: *BMC Health Services Research* [online]. [cit. 2015-03-08]. Dostupné z: http://www.biomedcentral.com/1472-6963/8/47/#B5.
- VOCABULARY. In: *Bitcoin* [online]. [cit. 2015-03-05]. Dostupné z: https://bitcoin.org/en/vocabulary.
- WAGNER, Andrew, 2014. The Facts on Factom. *Cryptobiz Magazine* [online]. [cit. 2015-04-20]. Dostupné z: http://cryptobizmagazine.com/the-facts-on-factom/.
- WALLET. Биткоин кошелек (wallet). In: *Новости криптовалют* [online]. [cit. 2015-03-24]. Dostupné z: http://cryptomap.ru/bitkoin-koshelek-wallet/.
- WEUSECOINS. What is Bitcoin: What is Bitcoin Mining? In: *Weusecoins* [online]. [cit. 2015-03-06]. Dostupné z: https://www.weusecoins.com/en/mining-guide/.
- WIFT, 2012. Co to je DDoS útok a jak se dělá? In: *diit.cz* [online]. [cit. 2015-04-04]. Dostupné z: http://diit.cz/clanek/co-to-je-ddos-utok-a-jak-se-dela.
- YASHU, Gola, 2015. Darkcoin Releases InstantX; Value Continues to Ascend. *Newsbtc* [online]. [cit. 2015-04-17]. Dostupné z: http://www.newsbtc.com/2015/02/11/darkcoin-releases-instantx-value-continues-ascend/.
- YASSERG. Crawler4j [program na CD-ROM]. In: *Github* [online]. [cit. 2015-03-07]. Dostupné z: https://github.com/yasserg/crawler4j.
- ZASLAVSKY, Arkady, PERERA, Charith, GEORGAKOPOULOS, Dimitrios, 2013. Sensing as a Service and Big Data. In: *ICT Centre* [online]. [cit. 2015-01-18]. Dostupné z: http://arxiv.org/ftp/arxiv/papers/1301/1301.0159.pdf.
- ZEITCOIN. *Zeitcoin Community Forum* [online]. [cit. 2015-03-08]. Dostupné z: https://forum.zeit-coin.net/.

Seznam obrázků, tabulek a grafů

Seznam obrázků

Obrázek 1: Souvislost pojmů data, informace a znalostí, zdroj Tkačenko (2007) 13

Obrázek 2:Příklady strukturovaných dat, zdroj Nair a Narayanan (2012) a vlastní zpracování ... 14

Obrázek 3: Příklady nestrukturovaných dat, zdroj Nair a Narayanan (2012) a vlastní zpracování 14

Obrázek 4: Charakteristiky Big dat, zdroj Zikopoulos, Rousseeuw, Pei (2012) 15

Obrázek 5: Proces analýzy nestrukturovaných dat, zdroj Hashimoto a kol. (2012), Kuć a Rogoziński (2014), Sklenák a kol. (2001) a vlastní zpracování 19

Obrázek 6: Architektura Web Crawlera Mercator, zdroj Heydon a Najork (1999) 21

Obrázek 7: Automatický stemming, zdroj Sharma, (2012) 22

Obrázek 8: Interpretace dat, zdroj Gorshkov a Sheregi (2012)a vlastní zpracování 23

Obrázek 9: Zvolené metody pro analytickou část, zdroj vlastní zpracování 25

Obrázek 10: Ilustrace rozdělování kolem medoidů, zdroj Kaufman a Rousseeuw (2005) 26

Obrázek 11: Rozdíly mezi metodami aglomerativní a divizní, zdroj Kaufman a Rousseeuw (2005) 27

Obrázek 12: Reprezentace definice rozdělení mezi clustery, a) Střední hodnota skupiny, b) Nejbližší soused, c) Nejdelší soused, zdroj Kaufman a Rousseeuw (2005) 27

Obrázek 13: Proces shlukování pomocí algoritmu STC, zdroj Ilic, Spalevic, Veinovic (2014) a vlastní zpracování 28

Obrázek 14: Příklad generování stromu přípon, zdroj Ilic, Spalevic, Veinovic (2014) a vlastní zpracování 29

Obrázek 15: Proces shlukování pomocí algoritmu Lingo, zdroj Stefanowski, Osiński, Weiss (2004) a vlastní zpracování 30

Obrázek 16: Proces tvorby štítků, zdroj Stefanowski, Osiński, Weiss (2004) a vlastní zpracování 31

Obrázek 17: P2P, zdroj Buford, Yu, Lua (2009) 34

Obrázek 18: Vybrané těžební algoritmy, zdroj Bitmakler a vlastní zpracování 35

Obrázek 19: Transakce, zdroj Nakamoto (2009) 37

Obrázek 20: Hash určitého bloku, zdroj Blochchain Info 38

Obrázek 21: Informace o bloku č. 348549, zdroj Blochchain Info 38

Obrázek 22:Sít bloků, zdroj Blochchain Info 39

Obrázek 23: Vývoj kurzu USD/BTC a náročnosti těžby, zdroj Taylor (2013) 40

Obrázek 24: Proces analyzeru Elasticsearch, zdroj Elastic a vlastní zpracování 43

Obrázek 25: Architektura řešení, zdroj vlastní zpracování 44

Obrázek 26: Zmapování fór podle počtu příspěvků, počtu diskuzí a počtu zaregistrovaných uživatelů, zdroj vlastní zpracování 45

Obrázek 27: Zmapování fór bez Bitcoin Forum podle počtu příspěvků, počtu diskuzí a počtu zaregistrovaných uživatelů, zdroj vlastní zpracování 46

Obrázek 28: Příspěvek s Bitcoin Forum, zdroj Bitcoin Forum 51

Obrázek 29: Umístění příspěvku v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování 51

Obrázek 30: Definování v crawleru místa pro stažení příspěvku, zdroj Škola (2015) a vlastní zpracování 51

Obrázek 31: Umístění jména uživatelé v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování 52

Obrázek 32: Definování v crawleru místa pro stažení názvu diskuze, zdroj Škola (2015) a vlastní zpracování 52

Obrázek 33: Definování v crawleru místa pro stažení jména uživatele, zdroj Škola (2015) a vlastní zpracování 52

Obrázek 34: Umístění data publikování příspěvku v HTML kódu, zdroj Bitcoin Forum a vlastní zpracování 52

Obrázek 35: Definování v crawleru formátu dat pro stažení, zdroj Škola (2015) a vlastní zpracování 52

Obrázek 36: vytvořený index v Elasticsearch s názvem "bitcoin", zdroj Plugin head a vlastní zpracování 53

Obrázek 37: Část mapping indexu "bitcoin", zdroj Plugin head a vlastní zpracování 54

Obrázek 38: Cluster "Bitcoin", zdroj vlastní zpracování 57

Obrázek 39: Cluster "BTC", zdroj vlastní zpracování 57

Obrázek 40: Diskuze spojené s Bitcoinem v roce 2010, zdroj Jason Davies a vlastní zpracování 58

Obrázek 41: Diskuze spojené s Bitcoinem v lednu a únoru roku 2010, zdroj vlastní zpracování 59

Obrázek 42: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2010, zdroj vlastní zpracování 59

Obrázek 43: Diskuze spojené s Bitcoinem v květnu a červnu roku 2010, zdroj vlastní zpracování 60

Obrázek 44: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2010, zdroj vlastní zpracování 60

Obrázek 45: Diskuze spojené s Bitcoinem v září a říjnu roku 2010, zdroj vlastní zpracování 61

Obrázek 46: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2010, zdroj vlastní zpracování 61

Obrázek 47: Diskuze spojené s Bitcoinem v roce 2011, zdroj Jason Davies a vlastní zpracování 62

Obrázek 48: Diskuze spojené s Bitcoinem v lednu a únoru roku 2011, zdroj vlastní zpracování 62

Obrázek 49: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2011, zdroj vlastní zpracování 63

Obrázek 50: Diskuze spojené s Bitcoinem v květnu a červnu roku 2011, zdroj vlastní zpracování 63
Obrázek 51: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2011, zdroj vlastní zpracování 64
Obrázek 52: Diskuze spojené s Bitcoinem v září a říjnu roku 2011, zdroj vlastní zpracování.....64
Obrázek 53: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2011, zdroj vlastní zpracování.....64
Obrázek 54: Diskuze spojené s Bitcoinem v roce 2012, zdroj Jason Davies a vlastní zpracování 65
Obrázek 55: Diskuze spojené s Bitcoinem v lednu a únoru roku 2012, zdroj vlastní zpracování 66
Obrázek 56: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2012, zdroj vlastní zpracování 66
Obrázek 57: Diskuze spojené s Bitcoinem v květnu a červnu roku 2012, zdroj vlastní zpracování 67
Obrázek 58: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2012, zdroj vlastní zpracování 67
Obrázek 59: Diskuze spojené s Bitcoinem v září a říjnu roku 2012, zdroj vlastní zpracování.....68
Obrázek 60: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2012, zdroj vlastní zpracování.....68
Obrázek 61: Diskuze spojené s Bitcoinem v roce 2013, zdroj Jason Davies a vlastní zpracování 69
Obrázek 62: Diskuze spojené s Bitcoinem v lednu a únoru roku 2013, zdroj vlastní zpracování 69
Obrázek 63: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2013, zdroj vlastní zpracování 70
Obrázek 64: Diskuze spojené s Bitcoinem v květnu a červnu roku 2013, zdroj vlastní zpracování 70
Obrázek 65: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2013, zdroj vlastní zpracování 70
Obrázek 66: Diskuze spojené s Bitcoinem v září a říjnu roku 2013, zdroj vlastní zpracování.....71
Obrázek 67: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2013, zdroj vlastní zpracování.....71
Obrázek 68: Diskuze spojené s Bitcoinem v roce 2014, zdroj Jason Davies a vlastní zpracování 72
Obrázek 69: Diskuze spojené s Bitcoinem v lednu a únoru roku 2014, zdroj vlastní zpracování 72
Obrázek 70: Diskuze spojené s Bitcoinem v březnu a dubnu roku 2014, zdroj vlastní zpracování 73
Obrázek 71: Diskuze spojené s Bitcoinem v květnu a červnu roku 2014, zdroj vlastní zpracování 73
Obrázek 72: Diskuze spojené s Bitcoinem v červenci a srpnu roku 2014, zdroj vlastní zpracování 73
Obrázek 73: Diskuze spojené s Bitcoinem v září a říjnu roku 2014, zdroj vlastní zpracování.....74

Obrázek 74: Diskuze spojené s Bitcoinem v listopadu a prosinci roku 2014, zdroj vlastní zpracování.....74
Obrázek 75: Diskuze spojené s Bitcoinem v roce 2015, zdroj Jason Davies a vlastní zpracování 74
Obrázek 76: Cluster “Nxt”, zdroj vlastní zpracování.....81
Obrázek 77: Diskuze spojené s Nxt v roce 2013, zdroj Jason Davies a vlastní zpracování.....82
Obrázek 78: Diskuze spojené s Nxt v roce 2014, zdroj Jason Davies a vlastní zpracování.....83
Obrázek 79: Diskuze spojené s Nxt v lednu a únoru roku 2014, zdroj vlastní zpracování.....84
Obrázek 80: Diskuze spojené s Nxt v březnu a dubnu roku 2014, zdroj vlastní zpracování 84
Obrázek 81: Diskuze spojené s Nxt v květnu a červnu roku 2014, zdroj vlastní zpracování 84
Obrázek 82: Diskuze spojené s Nxt v červenci a srpnu roku 2014, zdroj vlastní zpracování.....85
Obrázek 83: Diskuze spojené s Nxt v září a říjnu roku 2014, zdroj vlastní zpracování.....85
Obrázek 84: Diskuze spojené s Nxt v listopadu a prosinci roku 2014, zdroj vlastní zpracování..85
Obrázek 85: Diskuze spojené s Nxt v roce 2015, zdroj Jason Davies a vlastní zpracování.....86
Obrázek 86: Cluster “Monero”, zdroj vlastní zpracování 89
Obrázek 87: Cluster “XMR”, zdroj vlastní zpracování.....89
Obrázek 88: Diskuze spojené s Monero v roce 2014, zdroj Jason Davies a vlastní zpracování...91
Obrázek 89: Diskuze spojené s Monero v roce 2015, zdroj vlastní zpracování.....91
Obrázek 90: Cluster Dash, zdroj vlastní zpracování 94
Obrázek 91: Diskuze spojené s Dash v roce 2014, zdroj Jason Davies a vlastní zpracování 95
Obrázek 92: Diskuze spojené s Dash v lednu a únoru roku 2014, zdroj vlastní zpracování.....95
Obrázek 93: Diskuze spojené s Dash v březnu a dubnu roku 2014, zdroj vlastní zpracování.....96
Obrázek 94: Diskuze spojené s Dash v květnu a červnu roku 2014, zdroj vlastní zpracování 96
Obrázek 95: Diskuze spojené s Dash v červenci a srpnu roku 2014, zdroj vlastní zpracování 96
Obrázek 96: Diskuze spojené s Dash v září a říjnu roku 2014, zdroj vlastní zpracování 97
Obrázek 97: Diskuze spojené s Dash v listopadu a prosinci roku 2014, zdroj vlastní zpracování97
Obrázek 98: Diskuze spojené s Dash na začátku roku 2015, zdroj vlastní zpracování.....97
Obrázek 99: Cluster “XCP”, zdroj vlastní zpracování 101
Obrázek 100: Diskuze spojené s XCP v roce 2014, zdroj Jason Davies a vlastní zpracování....102
Obrázek 101: Diskuze spojené s XCP v lednu a únoru roku 2014, zdroj vlastní zpracování 102
Obrázek 102: Diskuze spojené s XCP v březnu a dubnu roku 2014, zdroj vlastní zpracování ..103
Obrázek 103: Diskuze spojené s XCP v květnu a červnu roku 2014, zdroj a vlastní zpracování 103
Obrázek 104: Diskuze spojené s XCP v červenci a srpnu roku 2014, zdroj a vlastní zpracování 104
Obrázek 105: Diskuze spojené s XCP v září, říjnu, listopadu a prosinci roku 2014, zdroj a vlastní zpracování.....104

Obrázek 106: Diskuze spojené s XCP v roce 2015, zdroj a vlastní zpracování.....105

Seznam grafů

Graf 1: Objem vygenerovaných dat, zdroj (Zaslavsky, Perera, Georgakopoulos, 2013).....16

Graf 2: Měření y v závislosti na x, zdroj vlastní zpracování.....32

Graf 3: Počet příspěvků u Bitcoin Forum podle jednotlivých jazyků, zdroj Bitcoin Forum a vlastní zpracování 49

Graf 4: 20 nejdiskutovanějších kryptoměn na fóru Bitcoin, zdroj vlastní zpracování.....55

Graf 5: 19 nejdiskutovanějších kryptoměn na fóru Bitcoin po vyloučení kryptoměny Bitcoin, zdroj vlastní zpracování.....55

Graf 6: Vývoj kurzu BTC/USD v roce 2010, zdroj Blockchain Info a vlastní zpracování.....76

Graf 7: Vývoj kurzu BTC/USD v roce 2011, zdroj Blockchain Info a vlastní zpracování.....77

Graf 8: Vývoj kurzu BTC/USD v roce 2012, zdroj Blockchain Info a vlastní zpracování.....78

Graf 9: Vývoj kurzu BTC/USD v roce 2013, zdroj Blockchain Info a vlastní zpracování.....78

Graf 10: Vývoj kurzu BTC/USD v roce 2014, zdroj Blockchain Info a vlastní zpracování.....79

Graf 11: Vývoj kurzu BTC/USD v roce 2015, zdroj Blockchain Info a vlastní zpracování.....79

Graf 12: Vývoj kurzu NXT/USD od prosince roku 2013 do konce dubna roku 2015, zdroj CoinGecko (2015) 87

Graf 13: Vývoj kurzu BTC/USD od prosince roku 2013 do konce března roku 2015, zdroj Blockchain Info a vlastní zpracování 87

Graf 14: Vývoj kurzu XMR/USD od května roku 2014 do konce dubna roku 2015, zdroj CoinGecko (2015) 92

Graf 15: Vývoj kurzu DASH/USD od února roku 2014 do dubna roku 2015, zdroj CoinGecko (2015) 98

Graf 16: Vývoj kurzu BTC/USD od února roku 2014 do března roku 2015, zdroj Blockchain Info a vlastní zpracování.....99

Graf 17: Vývoj kurzu XCP/USD od února roku 2014 do konce dubna roku 2015, zdroj CoinGecko (2015) 106

Graf 18: Vývoj kurzu BTC/USD od března roku 2014 do konce března roku 2015, zdroj Blockchain Info a vlastní zpracování 106

Seznam tabulek

Tabulka 1: Porovnání objemu dat u různých zdrojů Lalor (2014) a vlastní zpracování 16

Tabulka 2: Obsah bajtů v jednotlivých jednotkách měření množství dat, zdroj Ram (2004) 17

Tabulka 3:Porovnání rychlosti vytvořených dat u různých zdrojů Lalor (2014) a vlastní zpracování.....17

Tabulka 4:Porovnání rozmanitosti dat u různých zdrojů Lalor (2014) a vlastní zpracování 18

Tabulka 5: Porovnání věrohodnosti u různých zdrojů Lalor (2014) a vlastní zpracování 19

Tabulka 6: Index, zdroj Kuć a Rogoziński (2014) a vlastní zpracování 21

Tabulka 7: Struktura sledované části fóra Bitcoin Forum a počet jeho témat a příspěvků 25. 3. 2015, zdroj Bitcoin Forum a vlastní zpracování.....48

Tabulka 8: Obecná informace o BTC, zdroj Bitmakler, Blockchain Info, Crypto-Currency Market Capitalizations a vlastní zpracování.....56

Tabulka 9: Obecná informace o Nxt, Nxt (2014), Nxt Blockchain Explorer, Bitmakler a vlastní zpracování, http://www.coinssource.com/crypto-coins/nxt/ 81

Tabulka 10: Obecná informace o Monero, Bitmakler, Monero Blockchain 88

Tabulka 11: Obecné informace o Dash, zdroj Bitmakler, Dash (2015), Dash Cryptocurrency (2015) 93

Tabulka 12: Obecná informace o Counterparty, Counterparty (2014), Bitmakler, Crypto-Currency Market Capitalizations, Blockscan, Deweller, Bitcoin Forum 100

Seznam diagramů

Diagram 1: Proces získání dat crawlerem, zdroj vlastní zpracování.....50

Seznam schémat

Schéma 1: Globální pohled na strukturu diplomové práce, zdroj vlastní zpracování 11

Schéma 2: Globální pohled na vztah mezi vývojem témat v čase a kurzem kryptoměn Bitcoin, Monero, Nxt, Counterparty a částečně Dash, zdroj vlastní zpracování 107

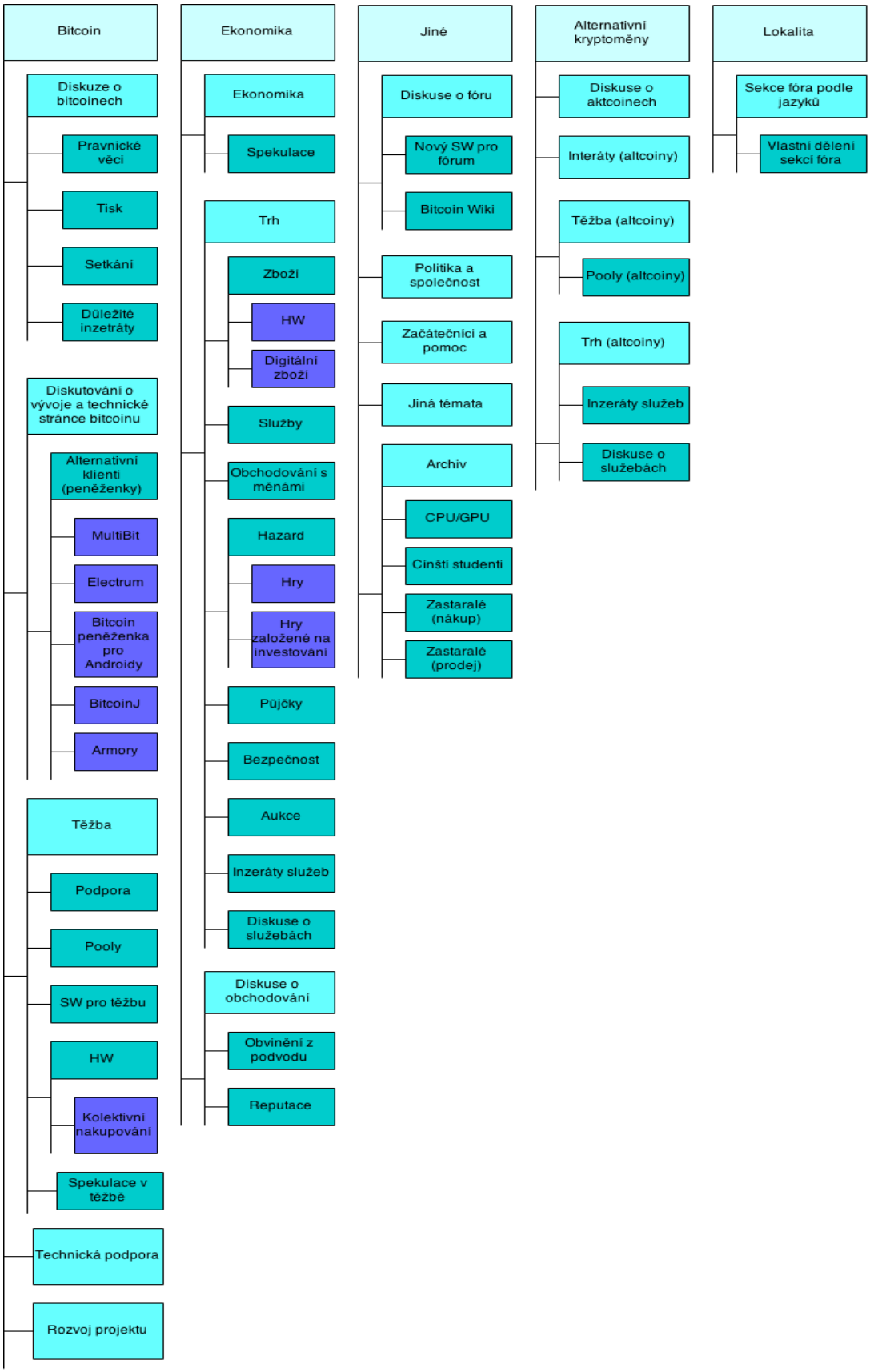
Přílohy

Příloha 1 - Struktura sledované části fóra Bitcoin Forum

Příloha 2 - Ostatní nejdiskutovanější kryptoměny na Bitcoin Forum

Příloha 3 – Crawler na CD-ROM

Příloha 1: Struktura sledované části fóra Bitcoin Forum



Zdroj Bitcoin Forum a vlastní zpracování

Příloha 2: Ostatní nejdiskutovanější kryptoměny na Bitcoin Forum

Tabulka 2. 1: Obecná informace o Litecoin, zdroj Litecoin Block Explorer, Bitcoin Security (2015)

Název kryptoměny a zkratka	Litecoin (LTC)
Doba spuštění	07. 10. 2011
Průměrná doba generace bloků	2,5 minuty
Metoda generace a bezpečnosti bloků	Proof of work
Maximální množství mince	84 000 000 LTC
Těžební algoritmus	Scrypt
Změna náročnosti	každé 2016 bloků
Odměna za nalezený blok (26.03.2015 11:00)	50 LTC
Snížení odměny (kdy?)	po naleznutí každých 840 000 bloků
Snížení odměny (o kolik?)	v dvakrát
Počet nalezených bloků (26.03.2015 11:10)	752 037

Tabulka 2. 2: Obecná informace o Ripple, zdroj Bitmakler

Název kryptoměny a zkratka	Ripple (XRP)
Doba spuštění	01. 05. 2013
Průměrná doba generace bloků	10 sekund
Metoda generace a bezpečnosti bloků	Proof of work
Maximální množství mince	100 000 000 000 XRP
Těžební algoritmus	ECDSA

Tabulka 2. 3: Obecná informace o Dogecoin, zdroj Dogechain, Bitmakler

Název kryptoměny a zkratka	Dogecoin (DOGE)
Doba spuštění	06. 12. 2013
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof of Work
Maximální množství mince	100 000 000 000 DOGE
Těžební algoritmus	Scrypt
Změna náročnosti	Každých 240 bloků
Odměna za nalezený blok (5.4.2015 23:00)	10 000 DOGE
Snížení odměny	od 27. 01. 2015 je fixní
Počet nalezených bloků (5.4.2015 23:11)	653 722

Tabulka 2. 4: Obecná informace o Litecoin, zdroj Nem (2014), New Economy Movement (2015), Guugll (2015)

Název kryptoměny a zkratka	New Economy Movement (NEM)
Doba spuštění	31. 03. 2015
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof of Importance
Maximální množství mince	8 999 999 999 XEM
Odměna	Zaleží na částce naleznutého bloku
Počet nalezených bloků (6.4.2015 15:19)	10 475

Tabulka 2. 5: Obecná informace o Sync, zdroj Sync Blockchain Explorer, Sync (2014)

Název kryptoměny a zkratka	Sync (SYNC)
Doba spuštění	13. 05. 2014
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof-of-Stake
Maximální množství mince	1 000
Těžební algoritmus	X11
Odměna za nalezený blok (6.4.2015 14:15)	0.14 SYNC
Počet nalezených bloků (6.4.2015 14:17)	507 277

Tabulka 2. 6: Obecná informace o BitShares, zdroj How BitShares Works, BitShares blocks

Název kryptoměny a zkratka	BitShares (BTS)
Doba spuštění	20. 07. 2014
Průměrná doba generace bloků	10 sekund
Metoda generace a bezpečnosti bloků	Delegated Proof of Stake
Maximální množství mince	2 000 000 000 BTS
Maximální odměna za nalezený blok (6.4.2015 12:43)	50 BTS
Průměrná odměna za nalezený blok (6.4.2015 12:43)	15, 46 BTS
Změna odměny za nalezený blok	Zaleží na nabídkách delegátů
Počet nalezených bloků (6.4.2015 12:45)	cca 2 212 203

Tabulka 2. 7: Obecná informace o Burst, zdroj Bitmakler, Coingecko (2015), Burstcoin.eu, Burst

Název kryptoměny a zkratka	Burst (BURST)
Doba spuštění	10. 08. 2014
Průměrná doba generace bloků	4 minuty
Metoda generace a bezpečnosti bloků	Proof of HDD Capacity
Maximální množství mince	2 158 812 800 BURST
Odměna za nalezený blok (6.4.2015 11:52)	cca 6 630 BURST
První odměna	10 000 BURST
Snížení odměny	O 5 % každý měsíc
Počet nalezených bloků (5.4.2015 23:32)	84 884

Tabulka 2. 8: Obecná informace o PayCoin, zdroj Bitmakler, Paycoin

Název kryptoměny a zkratka	PayCoin (PYC)
Doba spuštění	14. 11. 2014
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof of Work/Proof of Stake
Maximální množství mince	12 500 000 XPY
Těžební algoritmus	SHA-256
Odměna za nalezený blok (6.4.2015 11:43)	49 XPY
Změna náročnosti	1 blok
Počet nalezených bloků (6.4.2015 9:52)	127 952

Tabulka 2. 9: Obecná informace o Namecoin, zdroj Namecoin Cryptocurrency (2015)

Název kryptoměny a zkratka	Namecoin (NMC)
Doba spuštění	18. 04. 2011
Průměrná doba generace bloků	10 minut
Metoda generace a bezpečnosti bloků	Proof of Work
Maximální množství mince	21 000 000 NMC
Těžební algoritmus	SHA-256
Změna náročnosti	Každých 2 016 bloků
Odměna za nalezený blok (6.4.2015 10:00)	25 NMC
Snížení odměny	Po naleznutí 210 000 bloků
Počet nalezených bloků (6.4.2015 10:01)	225 235

Tabulka 2. 10: Obecná informace o Novacoin, zdroj Novacoin Cryptocurrency (2015), Novacoin Explorer

Název kryptoměny a zkratka	Novacoin (NVC)
Doba spuštění	09. 02. 2013
Průměrná doba generace bloků	10 minut
Metoda generace a bezpečnosti bloků	Proof of Work/Proof of Stake
Maximální množství mince	2 000 000 000 NVC
Těžební algoritmus	Scrypt
Odměna za nalezený blok (6.4.2015 11:35)	7.1522430 NVC
Změna náročnosti	1 blok
Počet nalezených bloků (6.4.2015 11:37)	167 881

Tabulka 2. 11: Obecná informace o Magi, zdroj Magi Blockchain Explorer, Magi Cheat Scheet

Název kryptoměny a zkratka	Magi (XMG)
Doba spuštění	~12. 08. 2014
Průměrná doba generace bloků	POW – 3 minuty, POS -90 sekund
Metoda generace a bezpečnosti bloků	Proof of Work/Proof of Stake
Maximální množství mince	25 000 000 XMG
Těžební algoritmus	SHA256d
Odměna za nalezený blok	Variabilní, zaleží na náročnosti, nejnižší je 5 XMG
Počet nalezených bloků (6.4.2015 17:00)	256 387

Tabulka 2. 12: Obecná informace o Nas, zdroj Nas (2014), Nas

Název kryptoměny a zkratka	Nas (NAS)
Doba spuštění	31. 03. 2014
Průměrná doba generace bloků	1 minuta
Metoda generace a bezpečnosti bloků	Proof-of-Stake
Maximální množství mince	10 000 000 000 NAS

Tabulka 2. 13: Obecná informace o Peercoin, zdroj Bitmakler

Název kryptoměny a zkratka	Peercoin (PPC)
Doba spuštění	12. 08. 2012
Průměrná doba generace bloků	10 minut
Metoda generace a bezpečnosti bloků	Proof of Work/Proof of Stake
Maximální množství mince	Bez limitní PPC
Těžební algoritmus	SHA-256
Odměna za nalezený blok (6.4.2015 13:07)	83.723683228498 PPC
Změna náročnosti	1 blok
Počet nalezených bloků (6.4.2015 13:09)	167 474

Tabulka 2. 14: Obecná informace o Stellar, zdroj Bitmakler, Stellar

Název kryptoměny a zkratka	Stellar (STR)
Doba spuštění	01. 08. 2014
Průměrná doba generace bloků	5 sekund
Metoda generace a bezpečnosti bloků	Proof of work
Maximální množství mince	100 076 294 480 STR
Růst objemu STR	O 1 % každý rok

Tabulka 2. 15: Obecná informace o Hyper, zdroj Hyper Blockchain Explorer, Dollentas (2014)

Název kryptoměny a zkratka	Hyper (HYPER)
Doba spuštění	25. 05. 2014
Průměrná doba generace bloků	30 sekund
Metoda generace a bezpečnosti bloků	Proof of Stake
Těžební algoritmus	Scrypt
Odměna za nalezený blok (6.4.2015 14:15)	72.5 HYPER
Změna náročnosti	10 minutes
Počet nalezených bloků (6.4.2015 18:23)	~1 043 153