VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

HABILITAČNÍ PRÁCE

Brno, 2019 Mgr. KAREL SLAVÍČEK, Ph.D.

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ
DEPARTMENT OF TELECOMMUNICATIONS

MATEMATICKÉ METODY DETEKCE ÚNIKU DAT MEDICÍNSKÉHO KOMUNIKAČNÍHO SYSTÉMU REDIMED

HABILITAČNÍ PRÁCE
HABILITATION THESIS

AUTOR PRÁCE
AUTHOR

Mgr. Karel Slavíček, Ph.D.

BRNO 2019

Obsah

Seznam obrázků

Seznam tabulek

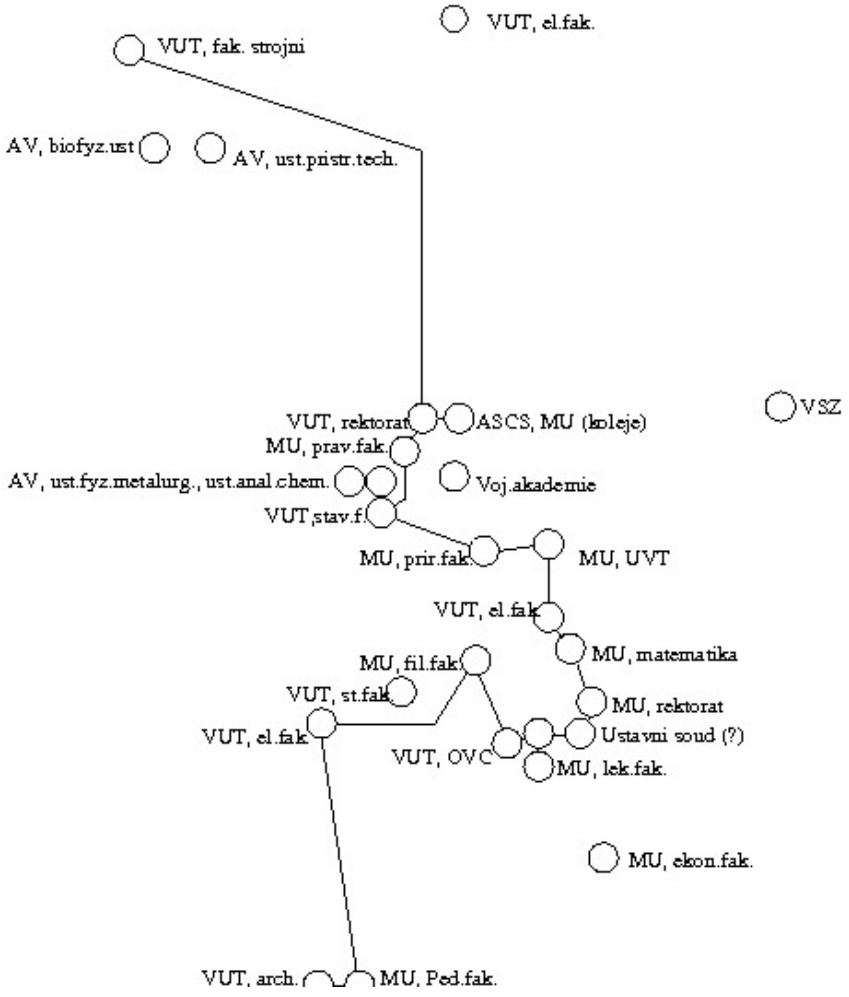
Seznam symbolů, veličin a zkratek

PPP	Point to Point Protocol
ATM	Asynchronout Transfer Mode
ICT	Information and Communication Technology
PACS	Picture Archiving and Communications System
DICOM	Digital Imaging and COmmunications in Medicine
NIS	Nemocniční Informační Systém
RIS	Radiologický Informační Systém
HL7	Health Level Seven International
TAP	Ve skutečnosti není zkratkou, ale užívaným aktronymem pro hardwarové zařízení pro kopírování provozu zpravidla na fyzické vrstvě pro účely monitorování a analýzy.
UTP	Unshielded Twisted Pair
CT	Computer Tomography
DOS	Denial-Of-Service
DDOS	Distributed Denial-Of-Service
GDPR	General Data Protection Regulation
UPS	Uninterruptible Power Supply
NEMA	National Electrical Manufacturers Association
IEEE	Institute of Electrical and Electronics Engineers
ANSI	American National Standards Institute
IOD	Information Objects Definition
UID	Unique IDentifier
DIMSE	DICOM Message Service Elements
SOP	Service Object Pair
WADO	Web Access to DICOM Objects

1 Úvod

Oblast budování a rozvoje vysokorychlostních datových sítí prodělala v uplynulých dvou dekádách bouřlivý rozvoj. Ruku v ruce se změnami kvantitativními, kdy přenosová rychlost vzrostla o několik řádů, dochází i ke změnám kvalitativním, zejména z pohledu spolehlivosti a dostupnosti služeb. S tím souvisí i vznik a rozvoj nových aplikací a uplatnění datových sítí v oborech, které je původně nevyužívaly, například v lékařství.

Rychlost vývoje je možné dokumentovat na příkladech brněnské akademické počítačové sítě a sítě národního výzkumu CESNET na jejich rozvoji a provozu jsem měl tu čest se podílet.



Obr. 1.1: Začátky budování brněnské akademické počítačové sítě - stav v roce 1993. Převzato z [99]

Na obrázku 1.1 je vidět stav brněnské akademické počítačové sítě v roce 1993, kdy začala instalace prvních optických kabelů. Převládající technologií v té době

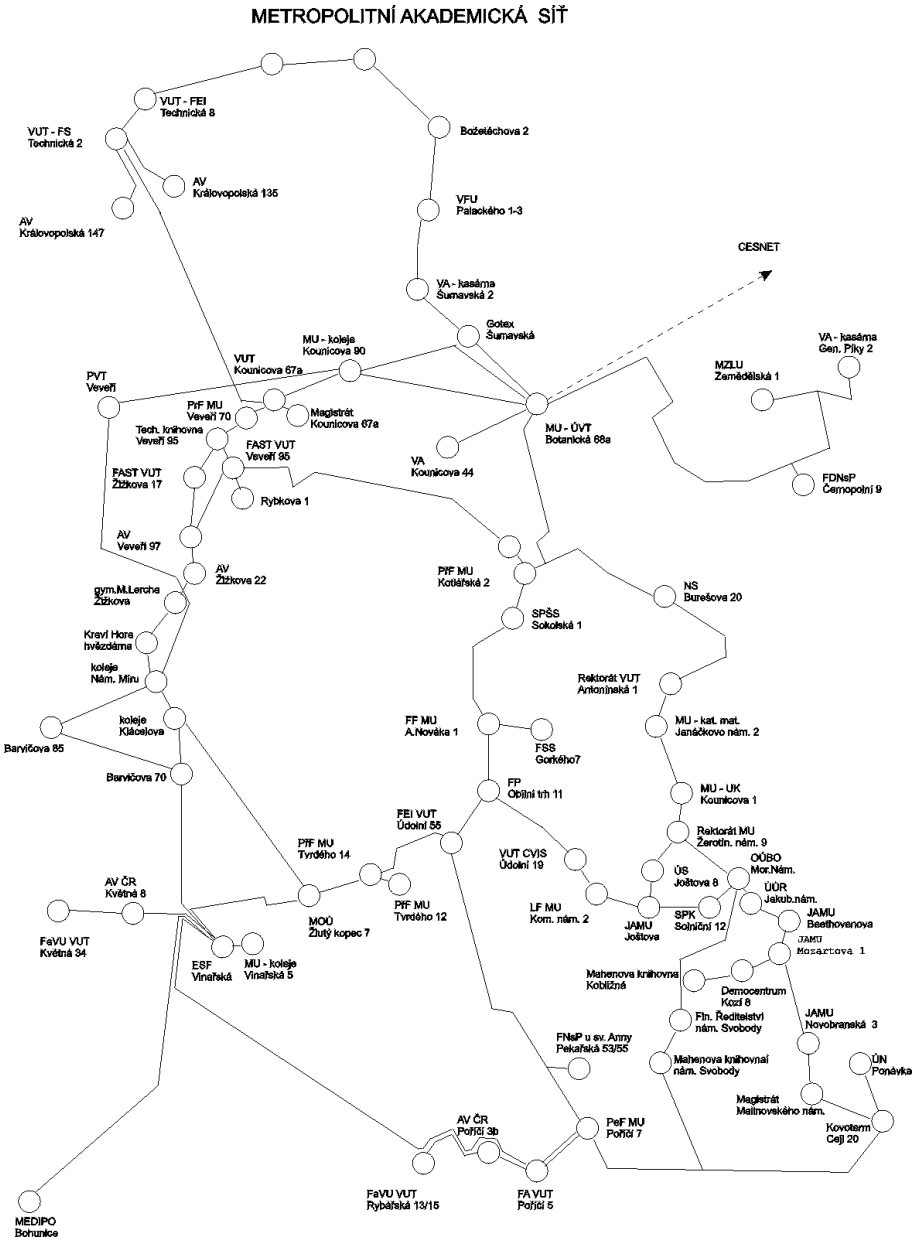
byly pevné metalické okruhy pronajaté od tehdejšího Českého Telecomu, na kterých se používaly převážně synchronní modely o rychlosti 19,2 kb/s a na hlavních trasách 64 kb/s. Jako linkový protokol byl použit protokol PPP. V roce 1997 již bylo jádro sítě provozováno na optických vláknech na tehdy aktuální technologii ATM.

Od roku 2002 se topologie optokabelové sítě brněnských vysokých škol již prakticky nemění, dochází pouze k posilování exponovaných tras a přirozené obměně technologie. Od Ethernetu o přenosové rychlosti 10 Mb/s přešla brněnská akademické počítačová síť přes technologii ATM, Ethernet s přenosobou rychlostí 1 Gb/s a 10 Gb/s až k současné rychlosti 40 Gb/s, která je použita na páteřních trasách a na připojení do sítě národního výzkumu CESNET. Aktuální stav topologie optikabelové sítě je zachycen na obrázku 1.2.

Současně s rozvojem topologie sítě a zvyšováním přenosové rychlosti docházelo ke změnám v oblasti zabezpečení sítí. V počátku budování brněnské akademické počítačové sítě se bezpečnosti nevěnovala pozornost. Tento stav byl vcelku pochopitelný, nejdříve bylo nutné vybudovat základy sítě a získat potřebné znalosti a zkušenosti. Totéž platí i pro vývoj systémů dohledu nad stavem sítě. Vývoj brněnské akademické počítačové sítě je dokumentován v řadě odborných i popularizačních článků [27], [41], [40], [26], [35].

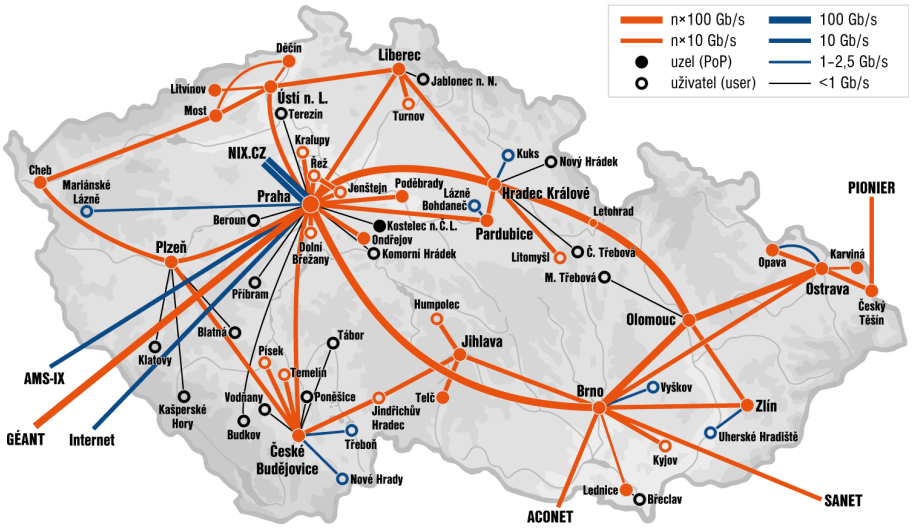
Podobným způsobem se vyvíjela i síť národního výzkumu CESNET. CESNET vždy patřil k průkopníkům budování datových komunikačních sítí v České republice. Na rozdíl od brněnské akademické počítačové sítě byla síť CESNET od začátku postavena na pronajatých datových okruzích a později na pronajatých optických vláknech, které jsou osazeny technologií ve správě CESNETu. V poslední dekádě jsou pronajatá optická vlákna osazena technologií DWDM. Vývoj sítě národního výzkumu CESNET je dokumentován v pravidelných ročních zprávách dostupných ne webu [14] a [15]. Výzukm a postupné nasazování aktuálních technologií je kromě toho zachycen i v řadě publikací [65], [85], [66], [84], [83],[72], [82], [81]. Rovněž síť CESNET prošla technologickým vývojem od modemových linek, přes technologii ATM a POS (Packet Over SONET) až po Ethernet s přenosovou rychlostí 10 GB/s a 100 Gb/s. Aktuálně je jádro páteřní sítě CESNET postaveno na technologii Ethernet s přenosovou rychlostí 100 Gb/s, účastnické přípojky jsou řešeny na technologii Ethernet o přenosových rychlostech 1 Gb/s - 40 Gb/s. V plánu je zvyšování rychlosti v souladu s technickým vývojem. Aktuálně jsou pro nejvýkonnější zařízení předních světových výrobců k dispozici prototypy rozhraní s přenosovou rychlostí 1 Tb/s.

Tak, jak se vyvíjely výkonové parametry a spolehlivost datových sítí a výpočetních systémů, docházelo k postupnému pronikání informačních a komunikačních technologií do všech oborů lidské činnosti. Jedna z posledních oblastí, kde se začaly informační a komunikační technologe využívat, je oblast medicíny. Je to dáno nejen přirozeně vysokými nároky lékařů na výkon a spolehlivost ICT systému, ale do jisté



Obr. 1.2: Aktuální topologie optických kabelů brněnské akademické počítačové sítě. Převzato z [25]

míry i zdravě konzervativním přístupem lékařské veřejnosti. Ve chvíli, kdy výpočetní a komunikační systémy dosáhly parametrů potřebných pro jejich nasazení v oblasti medicíny se situace změnila a dnes si již například radiologii nedokážeme bez ICT systémů představit. K rozvoji využití ICT v lékařství významně přispěl i projekt MeDiMed.



Obr. 1.3: Aktuální topologie sítě CESNET. Převzato z [16]

1.1 Metopolitní PACS systém MeDiMed

Na konci devadesátých let minulého století dosáhl pokrok v ICT technologiích takové úrovně, že je bylo možné začít využívat i pro přenos medicínských obrazových informací z lékařských diagnostických přístrojů, jako např. počítačový tomograf, magnetická rezonance, pozitronový emisní tomograf a další. Přechod na digitální formu zpracování obrazové dokumentace proběhl relativně rychle a dnes se již původní filmový materiál prakticky nepoužívá.

Používání digitálních zobrazovacích systémů a metod a využívání technologie PACS (Picture Archiving and Communication System) má řadu technických i ekonomických výhod. Umožňuje zvýšit přesnost diagnózy, umožňuje rychlejší přístup k obrazovým datům pacienta a nižší potřebu opakovaných vyšetření). PACS systémy jsou zpravidla používány paralelně s nemocničním informačním systémem (NIS), který slouží pro běžnou evidenci pacientů, jejich diagnóz a průběhu léčby. Důvodem je velký objem dat (řádově desítky MB až desítky GB), které produkují lékařské diagnostické přístroje. Jak systémy NIS, tak i systémy PACS jsou dostatečně standardizovány, stávající standardy však počítají s využíváním systému pouze uvnitř jedné zdravotnické instituce.

Pro systémy NIS je nejběžnějším standardem HL7 [11], pro systémy PACS je dnes jediným používaným standardem DICOM [13]. Struktura standardu DICOM je připomenuta v příloze A. Existence standardu DICOM významně přispěla k rozvoji digitalizace medicínských obrazových dat. Průkopníkem tohoto řešení v našem regionu byl Metopolitní PACS systém MeDiMed.

Použitní lokálních PACS systémů pro digitální zpracování medicínských obrazo-

vých dat přináší oproti používání filmového mateiálu celou řadu výhod, stále však využívá možnosti dostupných ICT technologií jen v omezené míře. Na základě výsledků a zkušeností nabytých při řešení projektů na zpracování medicínských obrazových dat bylo postupně vybudováno regionální centrum pro zpracování medicínských obrazových informací (MeDiMed). Přechod z lokálního řešení na metropolitní či spíše regionální řešení umožnilo dosáhnout vyššího komfortu pro lékaře a zlepšení úrovně péče o pacienta. Hlavní přínosy jsou:

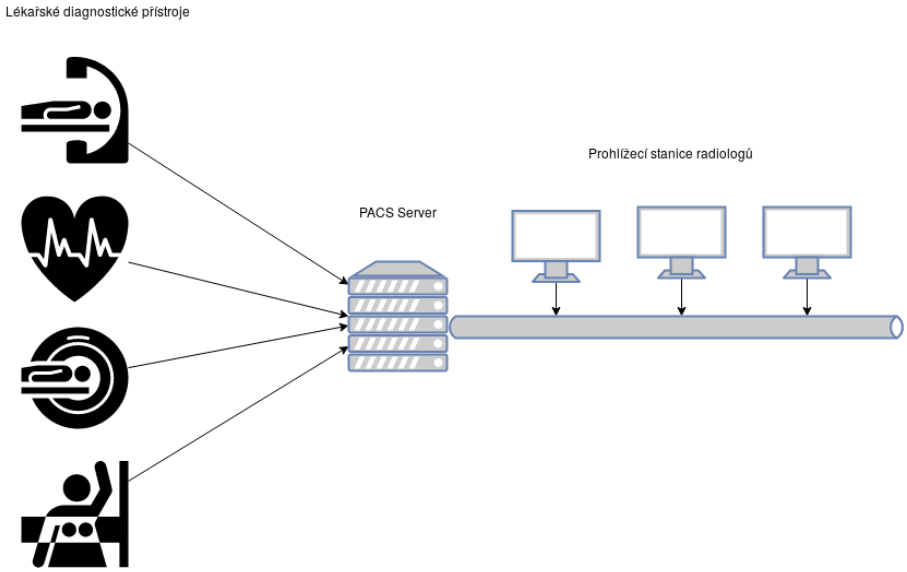
- Výrazné zkrácení doby vyhledávání snímků. Rychlý přístup k obrazové dokumentaci je možný v rámci celého regionálního řešení.
- On-line posouzení obrazové informace více odborníky z různých zdravotnických zařízení, tzv. druhého čtení snímku, které se využívá v případě nejasných příznaků nebo vzácných onemocnění.
- Možnost on-line využití cizího specialisty v případě nepřítomnosti vlastního odborníka v urgentních případech.
- Možnost sdílení top specialistů na vyhodnocování snímků více zdravotnickými zařízeními.
- Snížení počtu požadavků na opakovaná vyšetření.

Vznik systému MeDiMed byl podmíněn existencí kvalitní optokabelové sítě brněnských vysokých škol, neboť běžné internetové přípojky v době vzniku tohoto řešení neposkytovaly dostatečnou přenosovou kapacitu pro přenos velkého objemu dat, které generují lékařské diagnostické přístroje. Pro připojení brněnských nemocnic do tohoto systému proto byly využity vyhrazené optické vlákna a vznikla tak dedikovaná síť pro potřeby lékařské diagnostiky. Vybudování regionálního centra pro podporu zpracování medicínských obrazových informací MeDiMed bylo podpořeno řadou projektů, na jejich řešení jsem měl možnost se podílet. Pokroky budování systému MeDiMed byly popsány v řadě publikací [30], [28], [31], [31], [29], [33], [76],[32], [49], [75], [77], [34], [48], [78], [79], aktuální informace bývají zveřejňovány na webových stránkách projektu [64].

Obecná struktura PACS systému je znázorněna na obrázku 1.4. Jednotlivé diagnostické přístroje, tzv. modality ukládají obrazová data do PACS serveru, odkud tyto data stahují prohlížecí diagnostické stanice radiologů.

Systém MeDiMed je využíván řadou regionálních nemocnic. Umožňuje připojeným nemocnicím využívat pro zpracování svých obrazových systémy mimo areál nemocnice bez nutnosti mít vlastní vyškolený personál. Každý zdravotnický subjekt, který toto řešení využívá, má svůj vlastní server i diskový prostor, kde se nachází jeho obrazová data.

Z důvodů zajištění vysoké odolnosti a dostupnosti sytému jsou data uloženy ve dvou samostatných a geograficky oddělených systémech. Podrobnější schema celého řešení MeDiMed je na obrázku 1.5. Bezpečnost přenosu je zajištěna vyhrazenými



Obr. 1.4: Struktura systému PACS.

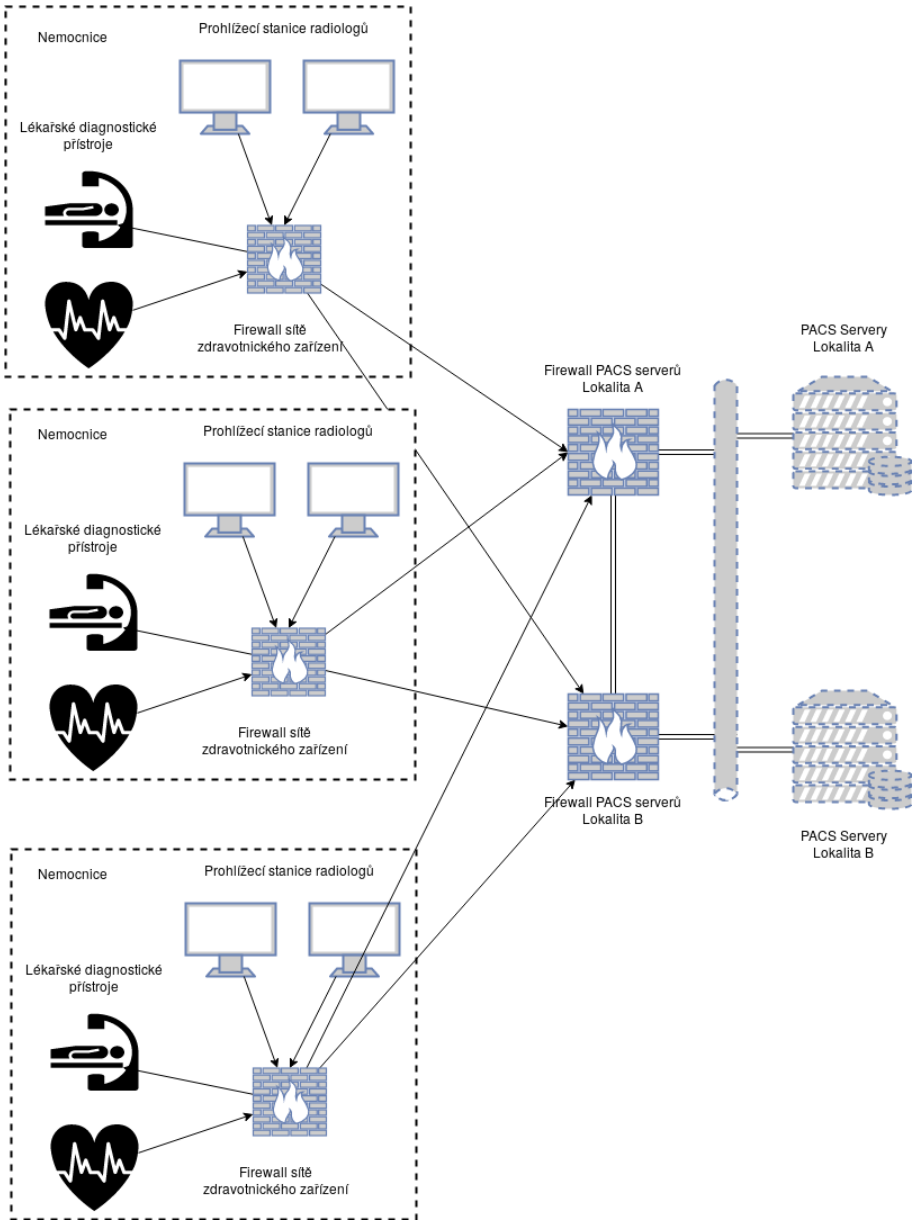
vlákny, případně využitím kryptograficky zabezpečených tunelů.

1.2 Radiologické komunikační centrum Redimed

S rozvojem digitalizace zpracování medicínských obrazových dat rostla i potřeba výměny snímků mezi zdravotnickými institucemi a to nejen v případě velkých nemocnic s rozsáhllým přístrojovýcm vybavením, tak i v případě malých organizací a soukromých lékařských praxí. S nárůstem počtu uživatelů již nebylo možné všechny účastníky připojit vyhrazeným optickým vláknem a v případě institucí s menší potřebou komunikace by to ani nebylo účelné. Z toho důvodu vzniklo v rámci systému MeDiMed speciální radiologické komunikační centrum Redimed.

Komunikační systém Redimed je určený pro elektronickou výměnu medicínské obrazové dokumentace dle standardu DICOM případně dalších souborů mezi zdravotnickými institucemi navzájem. Zdravotnickou institucí zde může být kromě nemocnic a poliklinik i domácí pracoviště radiologů, menší privátní centra a praktičtí lékaři, případně akademická pracoviště lékařských fakult a to nejen v České republice. Jedná se o čistě softwarové řešení, které je použitelné jak pro přímou spolupráci dvou nemocnic případně nemocnice a soukromého radiologa, tak i pro spolupráci v rámci roszáhlých sítí zdravotnických profesionálů. Struktura systému Redimed je znázorněna na obrázku 1.6.

Systém Redimed si velmi rychle získal oblibu mezi zdravotnickými zařízeními všech velikostí, o čemž svědčí objemy přenášených studií. Pro lepší představu jsou tyto kvantitativní parametry uvedeny v tabulce 1.1 a grafech 1.7 a 1.8. Z grafu je

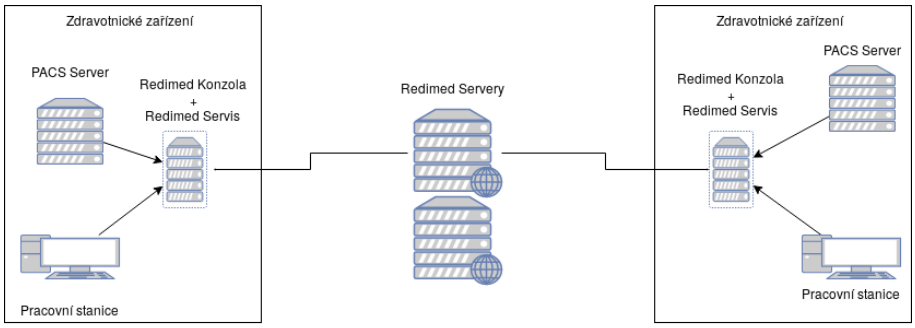


Obr. 1.5: Schema řešení PACS systémů projektu MeDiMed.

patrný přibližně lineární nárůst objemu provozu.

Systém Redimed aktuálně využívá více než 570 institucí a počet uživatelů stále roste.

Přenos principieně citlivých medicínských informací mezi nezávislými zdravotnickými zařízeními a univerzálnost řešení systému Redimed, která je potřeba pro snadné zapojování jednotlivých účastníků do projektu, otevírá i nové možnosti úniku dat o pacientech. Naštěstí dosud nedošlo ke zneužití systému Redimed tímto způsobem, nicméně je zapotřebí mít k dispozici nástroje pro detekci potenciálně nežádoucích přenosů.



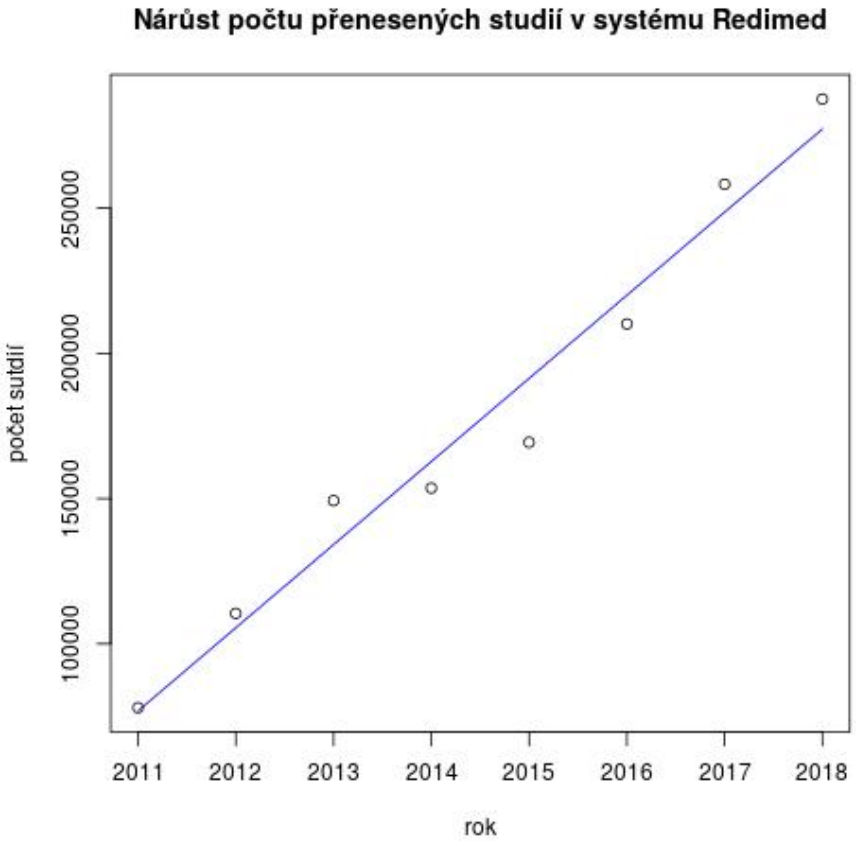
Obr. 1.6: Struktura komunikačního systému Redimed.

Tab. 1.1: Kvantitativní parametry využívání systému výměny medicínských obrazových informací Redimed.

Rok	**Počet přenesených studií**	**Objem přenesených dat**
2010	27.732	1.2 TB
2011	77.969	3.0 TB
2012	110.451	4.5 TB
2013	149.268	5.8 TB
2014	153.607	6.8 TB
2015	169.344	7.9 TB
2016	210.126	9.7 TB
2017	258.240	11.1 TB
2018	287.573	13.0 TB
2019	213.205	10.1 TB

1.3 Bezpečnostní aspekty zpracování medicínských obrazových dat

Současně s rozvojem komunikační a výpočetní technologie narůstá její sepětí s každodenním životem společnosti a tím se bohužel do digitálního prostředí přesouvá i nežádoucí činnost. Prakticky neustále dochází k nejrůznějším útokům na počítačové systémy, obsahující cenná a důležitá data. Různé nežádoucí či přímo podvodné aktivity využívají relativní anonymity elektronického prostředí. Nejinak je tomu i v



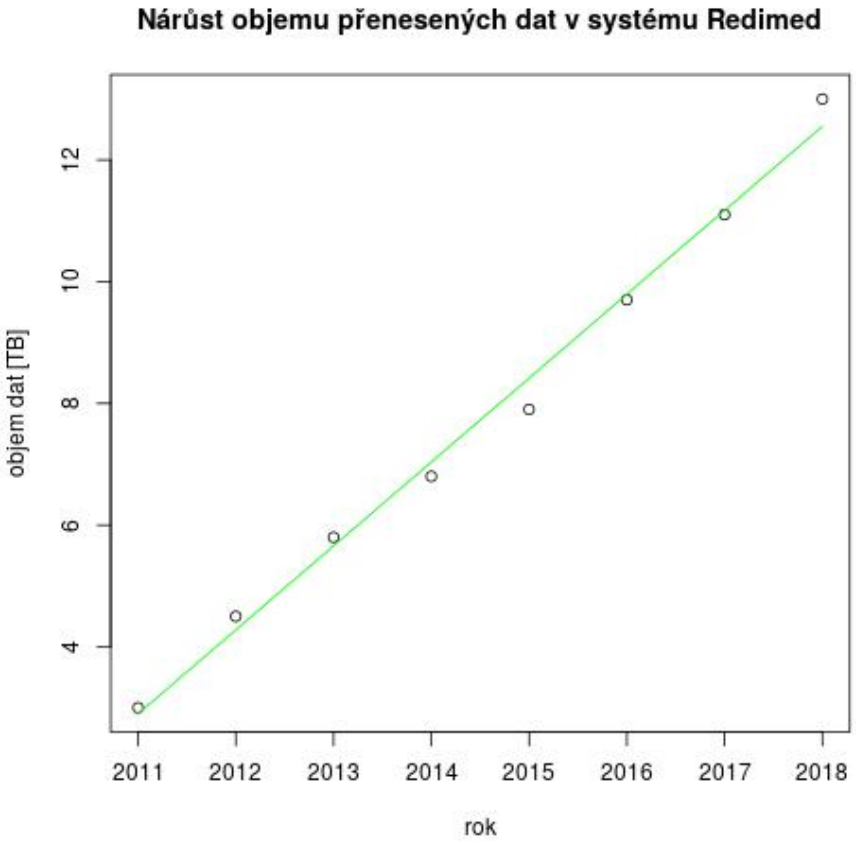
Obr. 1.7: Počet studií přenesených systémem Redimed.

případě systémů, které zpracovávají medicínská obrazová data.

Počítače připojené k celosvětové síti Internet jsou prakticky neustále bombardovány pokusy o prolomení svého zabezpečení ve snaze získat přístup k datům v nich uloženým a tyto počítače ovládnout a využít k dalším útokům. Některé studie uvádí, že pokud k Internetu připojíme nový počítač, aniž by byl zabezpečen alespoň instalací nejnovější bezpečnostních aktualizací, bude útočníkem ovládnut v průměru za 4 minuty.

Medicínská data představují zajímavý cíl potenionálch útoků. Mohou obsahovat velmi citlivé údaje o pacientech. Tato data můžou být předmětem zájmu například pojišťovacích společností, kdy znalost zdravotního stavu člověka může vést ke změně chování pojišťovny při uzavírání např. životního pojištění, či snaze o vypovězení nebo změnu stávajících smluv životního pojištění, pokud se pošťovna dozví o závažném onemocnění klienta.

V případě nemocí psychologického nebo venerologického typu může únik informací ze zdravotní dokumentace pacienta způsobit tomuto pacientovi řady nepříjemností a jistý typ vyloučení ze společnosti. Rovněž informace o existenci např. implan-



Obr. 1.8: Objem dat přenesený systémem Redimed.

tátů používaných v plastické chirurgii můžou v případě veřejně známých osobností působit značné nepříjemnosti.

Vedle těchto snadno představitelných problémů způsobených únikem dat může být velmi problematické i "obyčejné"porušování autorského zákona. Poked např. výzkumný pracovník ve zdravotnické organizaci roku pečlivě shromažďuje a dokumentuje zajímavé případy konkrétního typu onemocnění ‚jistě nebude projevovat přílišné nadšení, pokud mu některý z kolegů obrazová data zkopíruje a sám opublikuje.

Zatím nedošlo k útoku na systémy zpracování medicínských obrazových dat, což můžeme přičítat dílem zodpovědnému chování uživatelů a dílem relativně menší zajímavosti těchto dat ve srovnání s daty vládních institucí či bank. Přesto je nutné se na takovou možnost připravit a mít k dispozici automatizované nástroje, které na případnou nežádoucí aktivitu upozorní.

Navíc s nárůstem počtu uživatelů systému Redimed lze očekávat, že poroste i nebezpečí případného útoku nebo pokusu o únik informací z tohoto systému. Rovněž nástup využívání miniaturní elektroniky moderne označované jako IoT jistě v dohledné době zasáhne i oblast lékařství a přibude tak více dat, která jsou zajímavá

pro potencionální útočníky.

Obecně lze možné útoky na ICT systém s cílem získat neoprávněný přístup k datům rozdělit na dvě kategorie:

- Útok na ICT infrastrukturu. Do této kategorie spadá odposlech na fyzickém přenosovém médiu, napadení aktivních síťových prvků, které nativně poskytují možnost kopírování přenášených dat za účelem diagnostiky případných problémů infrastruktury, dále napadení operačního systému případně programového vybavení serverů uchovávajících a zpracovávajích data.
- Útok na uživatelské úrovni. Touto kategorií rozumíme neoprávněné kopírování dat s využitím přístupových údajů uživatele, který má k těmto datům přístup. V tomto případě není podstatné, zda tak činí uživatel sám, či zda došlo k prozrazení nebo zcizení jeho přihlašovacích údajů.

Tato práce se zabývá možnostni detekce útoku na uživatelské úrovni.

2 Cíle práce

Útoky na ICT infrastrukturu jsou předmětem celé řady výzkumných projektů a prací. Většina typů útoků je dobře prozkoumána a řady odborníků se zabývají možnostmi prevence, detekce a obrany proti těmto útokům. Pro přehlednost uvádím základní přehled možných typů útoků na ICT infrastrukturu rozšířený o některé praktické poznatky v kapitole 3. Dosud málo probádanou oblastí jsou možnosti automatické detekce úniku dat inicializovaných samotnými uživateli systému.

S rozvojem radiologického komunikačního systému Redimed roste i nebezpečí útoku jak na systém Redeimed samotný, tak jeho prostřednictvím na celou ICT infrastrukturu projektu MeDiMed. Dokud se používal pouze systém zpracování medicínských obrazových informací MeDiMed, jednalo se o vyhrazenou privátní síť, kde pouze vzdálené lokality, do kterých nebylo možné získat samostatná optická vlákna, byly připojeny kryptograficky zabezpečeným tunelem (protokol IPSEC). Naproti tomu systém Redimed je určen na masovou výměnu snímků mezi velkým množstvím zdravotnických zařízení. Proto s jeho rozvojen roste i nebezbečí útoků na uživatelské úrovni.

Možnosti dektece takto inicializovaného úniku dat jsou omezené. Pokud např. lékař, který má přístup do nemocničního systému PACS odešle jednu konkrétní studii na svůj soukromý účet systému ReDiMed (nebo účet sprátelené zdravotnické instituce či lékaře), není algoritmicky rozhodnutelné, zda tak učinil oprávněně či nikoli, aniž by to musela posuzovat nějaká další autorita. Pokud by však takto odesílal velké soubory studií, neměla by taková skutečnost uniknout pozornosti vhodných automatizovaných nástrojů pro zpracování logů událostí systému.

Cílem práce je analyzovat možnosti úniku medicínských obrazových dat inicializovancýh uživatelem PACS systému pomocí matematického zpracování logů událostí. Je jisté, že ne všechny útoky je možné tímto způsobem dostatečně včas odhalit, ať již proto, že se jedná o útok malého rozsahu, nebo proto, že je útok příliš sofistikovaný. Přesto však má smysl automatizované nástroje pro detekci nežádoucích datových toků vyvíjet a v praxi nasazovat.

3 Možnosti úniku a odposlechu dat a jejich detekce a prevence

Útočník, který hodlá získat neoprávněný přístup k citlivým údajům, může principiálně napadnout ICT systém na několika úrovních:

- Útok na komunikační infrastrukturu
  - Útok na fyzickou vrstvu sítě
  - Útok na akticní prvky sítě
- Útoky na výpočetní systém
  - Operační systémy a hypervizory
  - Základní SW vybavení - web server, ssh a pod.
- Útoky na informační systém (aplikaci)
  - Prolomení hesla oprávněného uživatele
  - Záměrná činnost oprávněného uživatele

Útoky na komunikační infrastrukturu a výpočetní systém budou podrobněji diskutovány v následující kapitole. Útoky na vlastní informační systém či aplikaci budou dále rozebírány v kapitole 6, které předchází připomenutí nezbytného matematického aparátu v kapitole 4. Důvod tohoto způsobu uspořádání textu bude zřejmý z kapitoly 6.

3.1 Útoky na přenosovou infrastrukturu

Útok na přenosovou infrastrukturu je jednou z možností, jak neprávněně získat citlivá data. Pokud se útočníkovi podaří instalovat nezbytné technické vybavení, může dlouhodobě získávat data aniž by někde vznikal provozní záznam, nebo jakákoli jiná stopa, podle které by bylo možné dohledat, k jakým datům útočník získal přístup. Zásadní nevýhodou je, že útočník nemůže ovlivnit, jaká data získá. Získat může pouze data, která jsou po síti přenášena na popud oprávněného uživatele. Další nevýhodou je, že data budou k dispozici v podobě datových rámců použitého komunikačního protokolu a je proto nutný určitý výpočetní výkon pro získání vlastních dat.

Jedinou obecně použitelnou metodou prevence úniku citlivých dat pomocí útoku na přenosovou infrastrukturu je použití dostatečně silného kryptografického zabezpečení přenosu a to po celé trase přenosu. Použití kryptografických metod vyžaduje příslušný výpočetní výkon jak na straně odesilatele, tak i na straně příjemce zprávy. Navíc některé aplikační protokoly s použitím kryptografického zabezpečení přenosu dat nepočítají a je proto nutné řešit zabezpečení pomocí externího zařízení nebo dalšího software (je-li to možné), což samozřejmě poskytuje o něco menší míru

zabezpečení, neboť alespoň v části komunikační infrastruktury se přenášení data nezabezpečeným způsobem, viz obrázek 3.1.

Obr. 3.1: Schema komunikačního kanálu z pohledu možnosti kryptografického zabezpečení.

Útok na komunikační infrastrukturu je v zásadě možné provést na dvou úrovních:

- Útok na fyzické přenosové medium
- Útok na aktivní síťové prvky

V obou případech se jedná v zásadě o odposlech.

3.1.1 Útok na fyzické přenosové medium

V závislosti na použitém typu fyzického přenosového media je možné instalovat vhodný monitorovací prvek, který umožní neoprávněný příjem přenášených dat. Vedle bezdrátových sítí, kde možnosti neoprávněného přijímání dat jistě není třeba diskutovat, se dnes nejčastěji pro přenos dat používá protokol ethernet v nejrůznějších rychlostech a to na fyzickém přenosovém mediu typu UTP, nebo na optických vláknech.

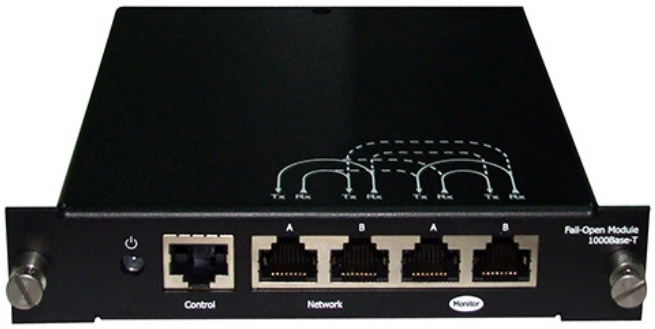
Pro monitorování provozu dat na UTP kabelech dnes existují aktivní prvky, tzv. TAPy, které umožní zrcadlit přenášený datový provoz na speciální rozhraní viz. obrázek 3.2. V tomto případě se jedná o aktivní zařízení, které však při výpadku napájení funguje v transparentním režimu, tj. nezpůsobí přeručení provozu.

Monitorování provozu na optických vláknech je ještě mnohem jednodušší: funkci TAPu je možné realizovat jednoduchým optickým děličem výkonu (splitter) viz obrázek 3.3, jehož cena se obvykle pohybuje kolem 10USD za kus, přičemž zpravidla potřebujeme 2 kusy.

Způsob provedení útoku:

Pokud se útočníkovi podaří instalovat optický splitter do přenosové trasy, např. v rozvaděči, kde se nachází konektorové spojení trasy, může odposlouchávat veškerý provoz, který touto trasou prochází. Schema zapojení tohoto typu odposlechu je na obrázku 3.4.

Vlastní zařízení, které shromažďuje a zpracovává zachycená data je možné připojit stejným typem optického vlákna a principielně může být od vlastního optického splitteru vzdálené stovky metrů nebo i několik kilometrů. K instalaci je samozřejmě



Obr. 3.2: Příklad monitorovacího příposlechu (TAPu) pro sítě typu gigabit ethernet na metalickém vedení výrobce Silicom Connectivity Solotions. Podobná zařízení je možné objednat např. z Amazonu v ceně kolem 250-300 USD.Obrázek převzat z webových stránek výrobce.

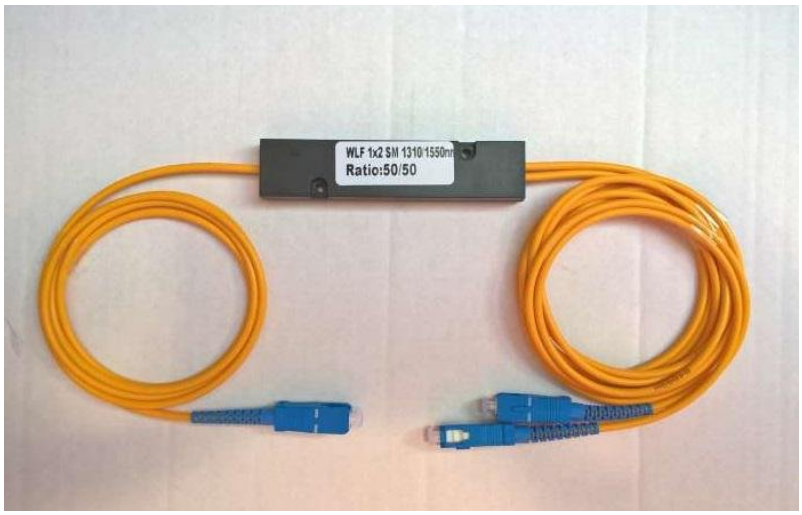
nutná jistá interní znalost: útočník musí znát detaily topologie optické kabelové trasy a mít fyzický přístup k rozvaděči, kde se nachází rozebiratelné spojení (konektory). Instalace optického splitteru se samozřejmě neobejde bez přerušení komunikace, je proto potřeba přizpůsobit čas instalace obvyklému provozu z důvodu minimalizace pravděpodobnosti odhalení.

Možnosti detekce:

Optický splitter je možné detekovat reflektometrickým měřením, které se používá např. při výstavbě a předávání optických tras. Pracujeme-li s citlivou aplikací, kde hrozí útok tohoto typu, je vhodné před uvedením spoje do provozu provést měření a u podezřelých bodů na trase provést inspekci na místě a zjistit, zda a proč jsou součástí trasy optické splittery. Instalace optického splitteru do trasy, která je v provozu, způsobí krátkodobý výpadek provozu. Jedná-li se o natolik kritickou aplikaci, doporučujeme po každém výpadku spojení, které je způsobeno na straně optokabelové trasy, opakovat reflktometrické měření trasy a porovnat náměry historickými záznamy. Toto měření samozřejmě znamená jisté náklady a to jak přímé v podobě strojového času OTDR a mzdových nákladů technických pracovníků, tak nepřímé v podobě delší doby nefunkčnosti spojení.

Vlastnosti optických splitterů z pohledu jejich obrazu v reflektogramu trasy byly podrobně zkoumány a výsledky jsou připraveny k publikování. Zde shrňme nejdůležitější výsledky.

OTDR pracuje se segnálem, který je v optovláknové trase odrážen buďto na materiálu optického vlákna (Rayleighův rozptyl), nebo na nehomogenitách trasy, jako např. svárech a konektorech. Vzdálenost zdroje odrazu se měří pomocí doby šíření signálu. Optický splitter zapojený ve směru měření se na reflektogramu jeví jako shluk poruch. Viz obrázy 3.5, kde je referenční náměr trasy a 3.6, kde je náměr



Obr. 3.3: Příklad pasivní optické odbočnice vhodné pro odposlech provozu na optických vláknech. Ilustrační obrázek převzat z www.i4wifi.cz

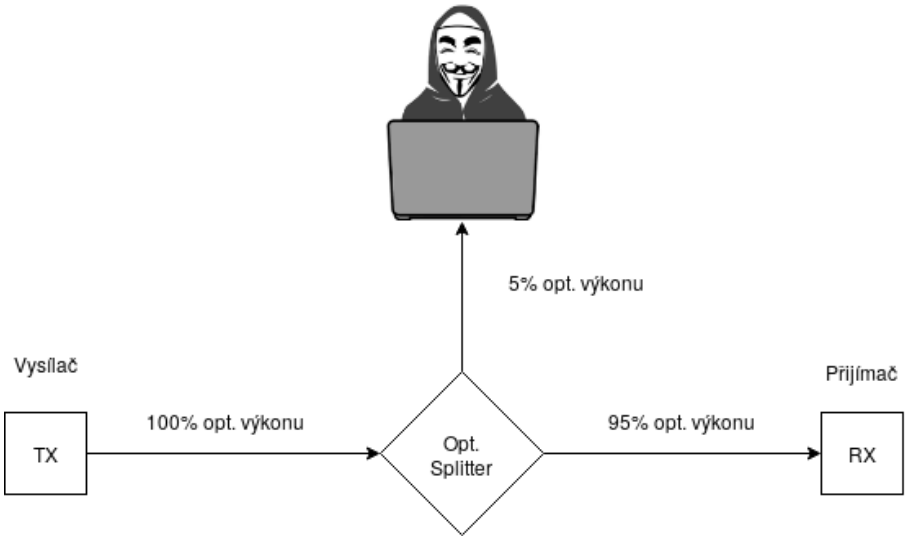
trasy se zapojeným optickým splitterrem s dělícím poměrem 5:95% přenášeného výkonu, který ještě poskytuje dostatečnou úroveň signálu ve větvi pro odposlech. Mměření pomocí ODTR nám tak umožnuje vytipovat podezřelá místa na trase, kde by mohlo docházet k odposlechu dat na fyzické vrstvě.

Z obrázků 3.5 a 3.6 je vidět, že rozdíl v náměrech trasy se zapojeným optickým splitterm a bez něj se neliší až tak výrazně a pokud se technik, který je analyzuje nesoustředí speciálně na vyhledávání míst s možným odposlechem, snadno tento rozdíl přehlédne. Viditelnost splitteru na reflektogramu je možné ze strany potenciálního útočníka ještě zhoršit zapojením jednoduchého izolátoru do odposlechové větve splitteru tak, je to jo znázorněno na obrázku 3.7. Reflektogram výše diskutované trasy se zapojeným optickým izolátorem je ne obrázku 3.8.

Samozřejmě se nabízí i možnost detekce vložení optického splitteru do trasy pomocí změny útlumu trasy. Tento útllum můžeme do určité míry měřit přímo na koncových zařízeních, pokud jsou vybaveny optickými moduly s DMI. Bohužel přesnot měření těchto modulů není příliš velká (obvykle +- 1-2 dBm) a to jak úrovně vysílaného, tak i úrovně přijímaného signálu (experimentálně ověřeno). Rovněž samotný útlum optické trasy během dne mírně kolísá vlivem mechanického namáhání kabelu díky změnám teploty, u kabelů závěsných i vlivem větru případně dalších činitelů. Experimentálně ověřené kolísání útlumu reálně trasy délky cca 40km sestavené na optických vláknech brněnské akademiské počítačové sítě je 0,7 dB, což je na úrovni vložného útlumu optického splitteru.

Možnosti prevence:

Jedinou možností prevence úniku dat na této úrovni je důsledná kontrola optických tras pomocí OTDR a to jak před uvedením do provozu, tak i po každém



Obr. 3.4: Schema zapojení optického splitteru pro odposlech datové komunikace.

výpadku trasy. Tato činnost je po časové odborné stránce relativně náročná a je proto vždy nutné zvážit, jaká je hodnota dat přenášených po síti vzhledem k nákladům. Navíc pravděpodobnost odhalení splitteru s dostatečně velkým poměrem průchozího a monitorovacího signálu je za předpokladu použití kvalitních konektorů nevelká.

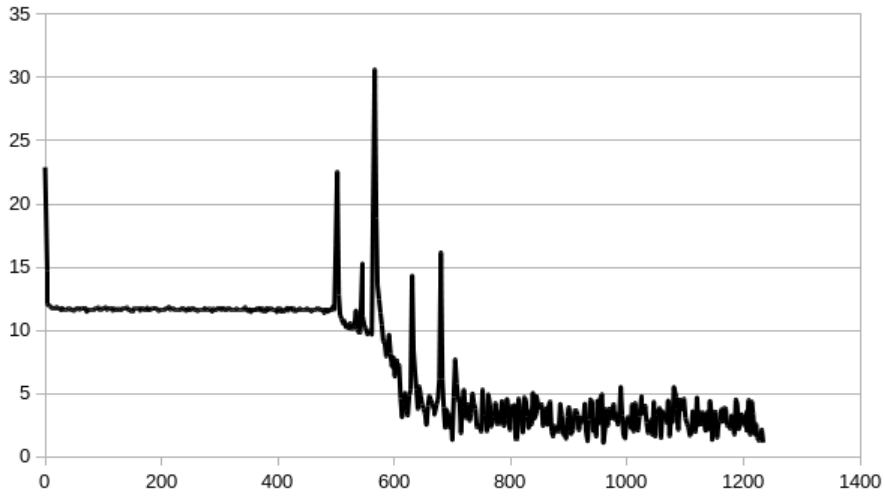
Možnost záměny:

Tato technologie se v praxi často používá i pro legitimní účely – zapojení sondy pro analýzu datového provozu. Jednou ze sond tohoto typu je i sonda flowmon jejíž původ je spojen s brněnskými akademickými institucemi. Typické schema zapojení této sondy je na obrázku 3.9.

Z tohoto obrázku je zřejmé, že zapojení legitimního nástroje na analýzu provozu a zařízení pro odposlech provozu je identické. Budeme-li tedy technickými prostředky detekovat na trase optické splittery, je nutné manuální zjišťování, k čemu je optický splitter využit.

Zde se samozřejmě nabízí otázka, jak se technicky liší sonda pro všasnou detekci útoků od zařízeni pro odposlech provozu. Sondy pro analýzu datových toků mají k dispozici kompletní datový provoz optické trasy, na které jsou nasazeny. Je otázka hardwarové a softwarové konstrukce sondy, jestli bude provoz pouze průběžně analyzovat a uživateli dodávat pouze agregované výstupy v podobě netflow nebo IPFIX statistik, nebo jestli uživateli zpřístupní i celý datový obsah zachyceného provozu. Tady samozřejmě vzniká otázka právních aspektů použití sondy pro monitorování provozu.

V případě prostého odposlechu je jasné, že ten kdo jej instaloval se pohybuje mimo rámec zákona. V případě HW sondy pro monitorování datových toků se uži-



Obr. 3.5: Referenční náměr trasy bez zapojení optického splitteru.

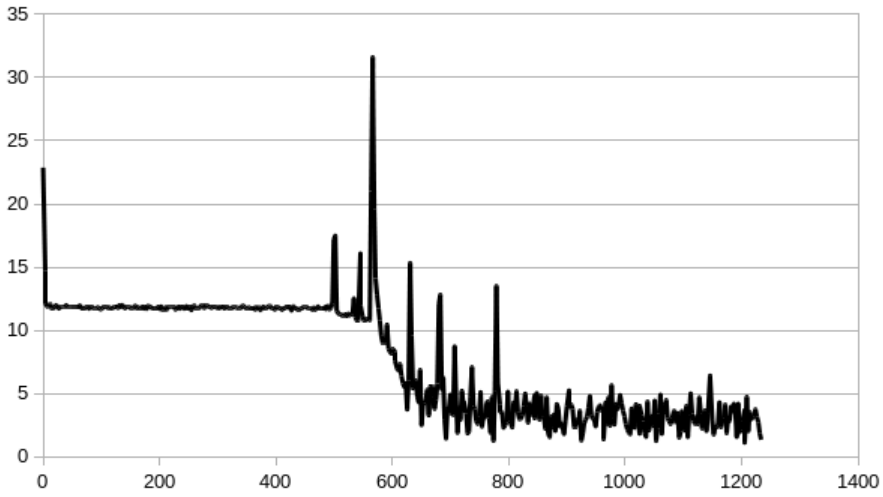
vatel nebo správce tohoto zařízení v principu může dostat ke stejným údajům, jako v případě použití odposlechu. V tomto případě však technika výrazně předběhla zákon, neboť použití sond pro monitorování datového provozu není (v době přípravy této práce) právně nijak ošetřeno. Je otázka, jak velký objem dat je sonda schopna dlouhodobě uchovávat (principielně je omezena pouze velikostí diskového prostoru), zda uchovává či může uchovávat (v závislosti na svém SW vybavení) i datový obsah paketů a zda jsou tyto data přístupná uživateli.

3.2 Útok na aktivní síťové prvky

Řada aktivních síťových prvků, zejména prakticky všechny moderní prvky typiu přepínač (switch) mají vestavěnou funkcionalitu, která umožní monitorovat přenášený provoz v zásadě obdobným způsobem, jako je instalace speciálního prvku popsana v předchozí kapitole. Tato funkcionalita se zpravidla nazývá port mirroring. Většina soudobých ethernetových přepínačů umožňuje zrcadlit datový provoz, který do přepínače vstupuje, případně z něj vystupuje daným rozhraním, nebo provoz dané VLAN sítě na definovaný výstup obvykle nazývaný SPAN (Switch Port Analyzer) port.

Tato fukcionalita byla původně vytvořena pro účely analýzy datového provozu, je však velmi snadno zneužitelná i pro odposlech provozu. Uživatel nemá žádnou možnost zjistit, že k tomuto typu odposlechu provozu dochází. Zapnutí či vypnutí zrcadlení vybraného provozu se na přenosových vlastnostech sítě nijak neprojeví a není zjistitelné jinak, než inspekcí konfiguračního nastavení aktivních prvků sítě.

Vedle nejběžněji používaného modelu zrcadlení provozu na fyzický port přepínače

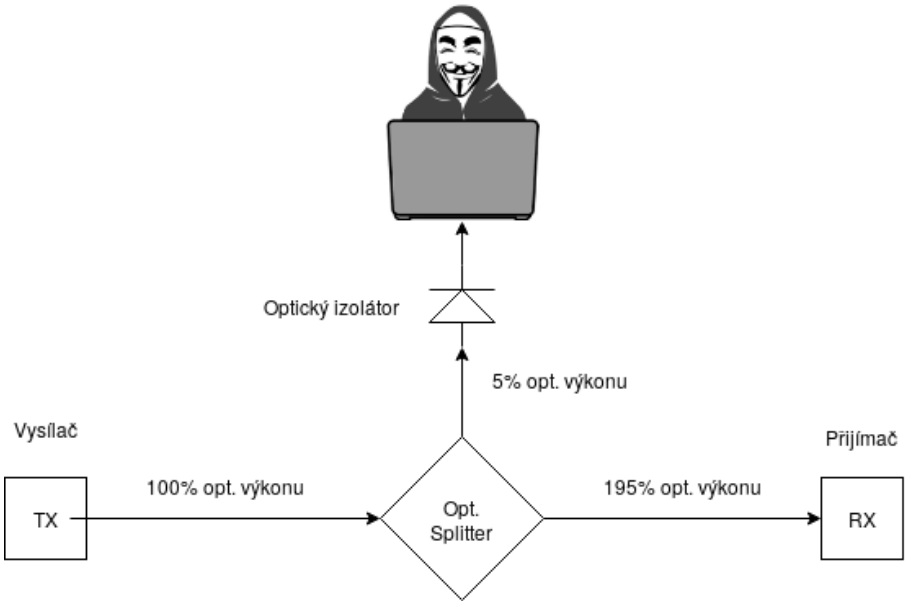


Obr. 3.6: Náměr stejné trasy se zapojeným optickým splitterem.

existuje i možnost zrcadlit provoz vzdáleně pomocí tzv. RSPAN (Remote SPAN) portu, kdy přepínač zrcadlí vybraný provoz do speciálně konfigurované VLAN sítě a posléze jej analyzovat ve vzdálené lokalitě. Nutnost speciálně konfigurované VLAN sítě je dána základními vlastnostmi přepínačů, kdy přepínač směruje ethernetové rámce pouze na rozhraní, za kterým se nachází zařízení s příslušnou cílovou MAC adresou (je-li toto rozhraní známé z analýzy provozu, který tímto přepínačem prochází). Pro přenos zrcadleného provozu je však tato jinak potřebná vlastnost přepínačů na škodu.

Tento typ útoku může provést buďto administrátor, který má oprávnění ke konfiguraci daného síťového prvku, nebo útočník, který získá neoprávněný přístup k tomuto prvku. Jedinou možností obrany proti tomuto typu útoku je dodržování disciplíny při konfiguraci aktivních síťových prvků a pravidelná analýza logu událostí a konfiguračních změn. Chceme-li síť v dostatečné míře zabezpečit proti útoku na konfiguraci aktivních prvků, je nutné (samozřejmě vedle dodržování základních pravidel konfigurace bezpečnostních parametrů těchto prvků) dobře definovat procesy změny konfigurace sítě, auditu a kontroly. Tyto formalizované postupy jsou známy pod pojmem procesní řízení a nejsou mezi správci sítě příliš populární pro netriviální objem administrativní části práce ve srovnání s vlastní konfigurací síťových prvků. Je to však jediná možnost, jak minimalizovat možnost získání neoprávněného přístupu k datům ze strany síťových administrátorů.

Správce aktivních síťových prvků je obecně velmi obtížně kontrolovatelný a z principu své práce je schopen datový provoz kopírovat. V zásadě jedinou možností je zajistit si loajalitu správců sítě netechnickými prostředky. Je zde jistá paralela k např. k leteckému provozu: pilot letadla má ve vzduchu stroj plně pod kontrolou



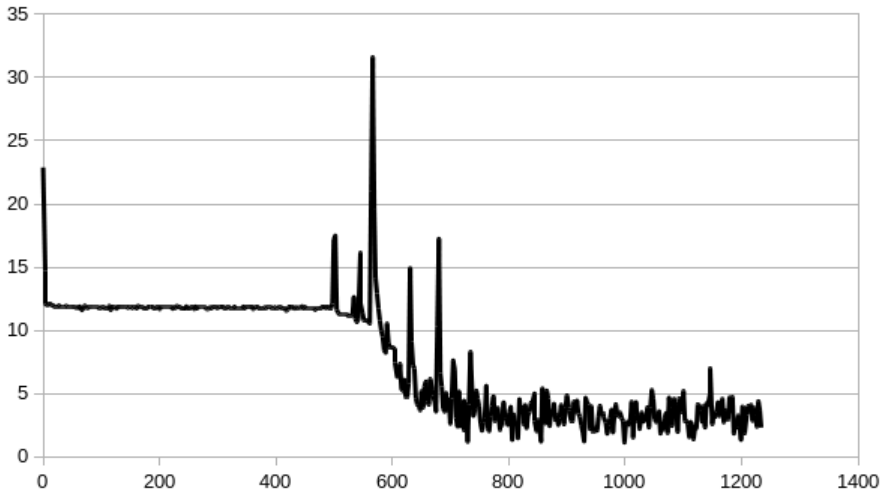
Obr. 3.7: Zapojení optického izolátoru pro snížení pravděpodobnosti odhalení optického splitteru.

a pokud se rozhodne s letadlem úmyslně havarovat, nezabrání mu v tom sebelepší přístrojové vybavení letadla a automatická kontrola čehokoli.

Na druhou stranu je třeba konstatovat, že odposlech dat, ať již na fyzické vrstvě, nebo na úrovni aktivních síťových prvků, poskytuje útočníkovi přístup pouze k těm údajům, které jsou právě přenášeny po síti. Útočník nemůže aktivně rozhodovat, jaká data získá. Naproti tomu útok na uživatelské úrovni poskytuje útočníkovi data dle jeho vlastního výběru. Proto přikládám tak velký význam monitorování chování uživatelů, které umožní včasné odhalení případného útoku, pokud by probíhal ve větším měřítku.

3.3 Softwarová a aplikační vrstva

Útoky na základní softwarové vybavení serverů patří k nejběžnějším bezpečnostním incidentům soudobých sítí. V literatuře byla zdokumentována řada různých typů útoků i způsobu obrany proti nim. Na toto téma existuje nepřeberné množství učebnich a nejrůznějšího studijního materiálu [88], [24], [43], [5]. Z pohledu zabezpečení celého systému je zajištění bezpečnosti operačního systému jak komunikačních serverů, tak i uživatelkých stanic jednou z hlavních priorit. Stejně tak je tomu i u základních služeb, jako je DNS, NTP, SNMP a webové servery. Na toto téma již byla a nepochybně ještě bude vypracována řada studií, návodů na vhodnou konfiguraci jednotlivých služeb. Proto nepovažuji za účelné tyto řady prací dále rozšiřovat, ale



Obr. 3.8: Náměr trasy se zapojeným optickým splitterem v případě použití izolátoru v odposlechové větvi splitteru.

soustředím se na uživatelskou rovinu, které je v případě nakládání s citlivými daty neméně důležitá.

3.4 Uživatelská rovina

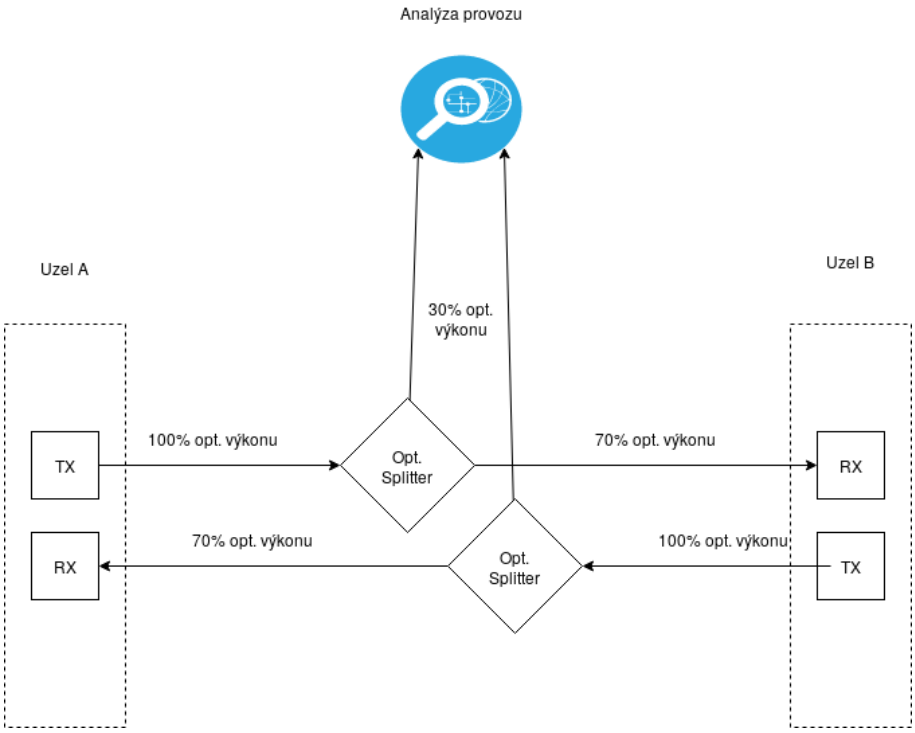
Nebezpečí útoku na uživatelské úrovni vzrůstá úměrně rostoucímu počtu uživatelů systému Redimed i nárůstu počtu přenášených studií. S nárůstem objemu přenášených dat roste přitažlivost systému pro případného útočníka. Jakýkoli pokus o prolomení zabezpečení systému stojí jisté úsilí a je spojeno i s určitou mírou rizika odhalení a případných trestněprávních následků. Proto je nepravděpodobné, že by se vyskytlo příliš mnoho útoků na systém, který neobsahuje dostatečné množství citlivých nebo jinak hodnotných dat.

Zároveň s nárůstem počtu uživatelů roste i riziko ztráty přihlašovacích údajů nebo úmyslného zneužití systému samotným uživatelem. V mnoha případech je prakticky nemožné rozlišit, zda daný datový přenos inicioval oprávněný uživatel systému, nebo zda došlo k úniku jeho přihlašovacích údajů a přenos inicioval neznámý útočník.

V zásadě existují jen dvě možnosti detekce úniku dat na této úrovni:

- Podrobný audit jednotlivých datových přenosů
- Matematická analýza datových toků

Audit jednotlivých přenosů není příliš představitelný vzhledem k množství přenášených obrazových studií. Navíc detailní kontrola přenosů by způsobovala administrativní zátěž pro uživatele a brzdila by další rozvoj využívání systému Redimed a spolupráce zdravotnických institucí.



Obr. 3.9: Typické zapojení sondy pro legitimní analýzu datového provozu, kterou často prování bezpečnostní oddělení.

Jedinou reálnou možností obrany proti úniku dat iniciovanému na uživatelské úrovni tedy zůstává matematická analýza datových toků a vyhledávání neobvyklých situací. Odchylka od obvyklého stavu může a nemusí znamenat nežádoucí únik dat. Počet přenášených studií za jednotku času přirozeným způsobem kolísá, čímž se snižuje spolehlivost určení toho, co je či nění obvyklý provoz.

Pro analýzu datových toků je zapotřebí dostatečný matematický aparát, který připomeneme v následující kapitole.

Matematickým zpracováním logů událostí dokážeme identifikovat situaci, kdy by došlo k významnému objemu nežádoucí komunikace, tj. situaci, kdy někdo kopíruje větší množství obrazových studií. Tímto postupem není možné zabránit úniku několika jednotlivých studií. Pokud máme za úkol ochránit jednotky velmi citlivých studií, není jiná možnost, než striktní evidence přístupu k nim.

Jiný problém, kterým jsme se v rámci projektu MeDiMed zabývali, je ochrana anonymizovaných studií užitých k výzkumným a výukovým účelům. V tomto případě se uživatelé obávají neautorizovaného užití jimi publikovaných výsledků. Pro takový případ je nutné zajistit publikované obrazové studie dodatečnou informací, např. vodoznakem [69], [68], [67].

4 Použité matematické nástroje

V této kapitole připomeneme matematické nástroje a postupy, které se dají použít pro analýzu logů událostí systému Redimed. Kvantitativní analýza logů událostí může pomici odhalit nežádoucí přenosy dat a přitom zachovat nezbytnou míru anonymity dat pacientů.

Jako překvapivě účinné se ukázaly základní nástroje popisné statistiky, jejichž přehled je uveden v následující podkapitole. Vedle těchto nejjednodušších nástrojů byly zkoumány i možnosti využití metod analýzy časových řad a spektrální analýzy pro vyhledávání periodických vzorů provozu.

4.1 Popisná a inferenční statistika

Slovo statistika má v kontextu kontextu zpracování dat minimálně dva různé významy. Jednak označuje vědní disciplínu, ale používá se též k označení některých vlastností sledované veličiny, např. aritmetický průměr je statistikou v tomto smyslu. Statistiku jakožto vědní disciplínu můžeme dále dělit na statistiku popisnou, která se zabývá numerickým popisem získaných dat, a statistiku induktivní, která se zabývá hledáním zákonitostí v získaných datech. V této kapitole přípomeneme nejdůležitější poznatky z teorie pravděpodobnosti a matematické statistiky, které budeme používat pro analýzu dat přenosového systému Redimed. Podrobnější informace je možno nalézt v klasických učebnicích [19], [50] [52], [62], vysokoškolských učebních textů [47], [89]. Velmi pěkný on-line přehled užité matematiky včetně elegantně zpracovaných kapitol o teorii pravděpodobnosti a matematické statistiky nabízí Ústav matematiky fakulty strojního inženýrství VUT [1].

Popisná statistika slouží pro kvantitativní popis vlastností tzv. statistického souboru. Statistickým souborem je množina měřených nebo pozorovaných dat číselného charakteru, v našem případě počet přenesených obrazových studií za jednotku času, objem přenesených dat za jednotku času nebo počet komunikujících partnerů v daném časovém úseku.

Popisná statistika se zabývá empiricky zjištěnými hodnotami a má svůj protějšek v teorii pravděpodobnosti, která pracuje s teoretickými matematickými modely. Základním pojmem teorie pravděpodobnosti je náhodný jev. Náhodný jev je výsledek nějakého pokusu nebo děje, který může či nemusí nastat. Může být popsán slovně, např. “při hodu kostkou padla šestka” (konec konců teorie pravděpodobnosti vznikala na popud hazardních her), nebo může mít číselný charakter, např. počet lidí ve frontě na zmrzlinu je vyšší než 10, nebo počet přenesených CT snímků za poslední hodinu je nižší než 5. Číselně kvantifikovatelný stav náhodného děje pak nazýváme náhodnou veličinou. (Např. počet přenesených medicínských studií za den).

Tato náhodná veličina má pak řady praktických realizací, tj. v našem případě zjištěných počtů přenesených snímků, které průběžně měříme každý den. Tím vznikne statistický soubor pozorovaných počtů přenesených snímků. V tomto místě se nám potkává rovina teoretická - náhodná veličina - s rovinou empirickou – statistický soubor praktických realizací této náhodné veličiny.

Pravděpodobnostní chování náhodné veličiny X popisujeme pomocí její distribuční funkce

$$F(x) = P(X < x). \tag{4.1}$$

Distribuční funkce $F(x)$ vyjadřuje pravděpodobnost, že náhodná veličina X nabývá hodnoty menší než x. Náhodná veličina X má diskrétní rozdělelní pravděpodobnosti, pokud nabývá nejvýše spočetně mnoha hodnot $x_1, x_2, ...$ V takovém případě rozdělelní pravděpodobnosti náhodné veličiny X popisujeme pomocí pravděpodobnostní funkce

$$p(x) = P(X = x), \tag{4.2}$$

přičemž musí platit, že

$$\sum_{i=1}^{\infty} p(x_i) = 1. \tag{4.3}$$

Druhým významným typem náhodných veličin jsou náhodné veličiny se spojitým rozdělelním pravděpodobnosti, tj. takové náhodné veličiny, jejichž distribuční funkce $F(x)$ je spojitá. Spojité náhodné veličiny charakterizujeme pomocí jejich hustoty $f(x)$, která je definována takto:

$$F(x) = \int_{-\infty}^{x} f(t)dt, \tag{4.4}$$

neboli

$$f(x) = F'(x). \tag{4.5}$$

Pro popis statistického souboru používáme dva základní typy charakteristik:

- míry polohy a
- míry variability

Jako míra polohy se nejčastěji používá aritmetický průměr, případně u některých náhodných veličin medián a u kategorických dat modus. Aritmetický průměr statistického souboru $x_1, x_2, ...x_n$ rozsahu n je definován jako

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{4.6}$$

Pro analýzu dat, které můžou mírně kolísat, např. počet přenesených obrazových studií za hodinu, může být výhodné číselnou řadu naměřených hodnot $x_1, x_2, ...x_n$

tzv. vyhladit, tj. nahradit posloupnost $x_1, x_2, ...x_n$ posloupností aritmetických průměrů několika /optimálně lichého počtu) sousedních hodnot, čímž získáme řadu $\hat{x}_2, \hat{x}_3, ...\hat{x}_{n-1}$, kde

$$\hat{x_j} = \frac{1}{3} \sum_{i=j-1}^{j+1} x_i. \tag{4.7}$$

V tomto případě mluvíme o tzv. klouzavém průměru. Klouzavý průměr jakkoli je ve své podstatě jednoduchý, je velmi vhodný nástroj pro detekci významných odchylek od obyklého stavu a kromě detekce potencionálně nežádoucích přenosů medicínských obrazových dat jsem jej s úspěchem využívali i pro analýzu logů událostí aktivních prvků datových sítí a podobné aplikace [80], [73], [74], [86].

U klouzavého průměru někdy využíváme i vážený aritmetický průměr

$$\overline{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}, \tag{4.8}$$

kde $w_i \geq 0$ jsou váhy, které vyjadřují relativní význam jednotlivých hodnot x_i.

Aritmetickému průměru, jakožto charakteristice empirických dat, odpovídá v teoretické rovině střední hodnota, která je mírou polohy odpovídající náhodné veličiny, jakožto teoretického modelu. Střední hodnota je definována jako

$$E(X) = \sum_{x} xp(x), \tag{4.9}$$

kde x probíhá celý definiční obor náhodné veličiny X v případě diskrétní náhodné veličiny a

$$E(X) = \int_{-\infty}^{\infty} f(x)dx \tag{4.10}$$

v případě náhodné veličiny se spojitým rozdělelním pravděpodobnosti.

Nevýhodou aritmetického průměru i střední hodnoty je jejich citlivost na byť i malý počet velmi odlehlých měření. Bude-li např. Dané zdravotnické zařízení přenášet každý den práve 10 obrazových studií po dobu devíti dnů a jeden den jich přenese 30, vyjde nám průměrná hodnota 12. Proto v některých případech používáme jinou charakteristiku polohy a sice medián. Medián je takový prvek x_k statistického souboru $x_1, x_2, ...x_n$, pro který platí, že počet prvků x_i takových, že $x_i < x_k$ je stejný, jako počet prvků x_j takových, že $x_j > x_k$. Tj. x_k leží přesně "uprostřed" hodnot $x_1, x_2, ...x_n$. V praxi je však výpočet mediánu výrazně pracnější, než výpočet střední hodnoty, resp. aritmetického průměru, proto jej používáme jen tam, kde aritmetický průměr není vhodný. Pro úplnost ještě dodejme poslední v praxi používanou charakteristiku polohy a tou je modus. Modus je nejčastěji se vyskytující hodnota. Tuto charakteristiku však využíváme spíše v případech, kde data kategorizujeme do skupin jako např. obrazové studie malého, středního a velkého rozsahu.

Pro popis variability měřených nebo pozorovaných dat používáme více charakteristik:

- Varianční rozpětí
- Kvantily
- Směrodatná odchylka
- Varienační koeficient

Variančním rozpětím R statistického souboru $x_1, x_2, ...x_n$ rozumíme rozdíl největší a nejmenší hodnoty tohoto souboru:

$$R = x_{max} - x_{min} \tag{4.11}$$

Kvantily jsou takové hodnoty, pro které platí, že příslušný počet prvků statistického souboru $x_1, x_2, ...x_n$ má hodnotu vyšší, resp. nižší, než daný kvantil. Hovoříme zpravidla o dolním či horním kvartilu jako o takové hodnotě, že 75% prvků statisitkého souboru $x_1, x_2, ...x_n$ má hodnotu vyšší, resp. nižší, než tento kvantil. Obdobně mluvíme o decilech v případě, že tuto vlastnost požadujeme pro 90% hodnot, případně o dalších percentilech. Obdobně jako varianční rozpětí můžeme definovat I percentilové rozpětí. V praxi se však příliš nepoužívá. Horní percentil, tj. číslo "nad kterým leží" jen 1% pozorovaných nebo očekávaných hodnot často používáme jako prahovou hodnotu pro určení toho, kdy již stav systému považujeme za neobvyklý a je vhodné vyvolat manuální inervenci.

Nejčastěji používanou charakteristikou míry variability statistického souboru se používá výběrová směrodatná odchylka. Přívlastek výběrová se používá pro odlišení empirické charateristiky a teoretické charakteristiky odpovídající náhodné veličiny. Výběrovou směrodatnou odchylku statistického souboru $x_1, x_2, ...x_n$ definujeme jako

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{4.12}$$

U teoretickým modelů se jako míra variability náhodné veličiny X používá zpravidla rozptyl

$$D(X) = E([X - E(X)]^2). \tag{4.13}$$

Protože rozptyl nemá stejnou jednotku, jako hodnoty náhodné veličiny X, resp. odpovídajícího statistického souboru $x_1, x_2, ...x_n$, definujeme i pro náhodnou veličinu X směrodatnou odchylku jako

$$\sigma(X) = \sqrt{D(X)}. \tag{4.14}$$

V praxi nás však často zajímá relativní velikost výběrové směrodatné odchylky vzhledem k hotnotě artimetického průměru. Proto zavádíme ještě další charakteristiku a tou je varianční koeficient

$$v_x = \frac{s_x}{\overline{x}}. \tag{4.15}$$

V oblasti teoretických charakteristik se ještě u náhodných veličin s nenulovým rozptylem používá koeficient šikmosti

$$A_3(X) = \frac{E([X - E(X)]^3)}{[\sigma(X)]^3} \tag{4.16}$$

a koeficient špičatosti

$$A_4(X) = \frac{E([X - E(X)]^4)}{[\sigma(X)]^4} - 3. \tag{4.17}$$

Tyto charakteristiky však uvádím pouze pro úplnost a v další práci je nebudeme potřebovat.

4.1.1 Statistická závislost dvou náhodných veličin

Pro analýzu adtových přenosů budeme využívat i statistickou závislost náhodných veličin. Statistická závislost nemuzí znamenat a v mnoha případech ani neznamená kauzalitu. Můžeme ji s výhodou použít v situacích, kdy potřebujeme zjisti, jestli se náš zkoumaný jev chová obdobným způsobem jako jiné jevy podobného charakteru. Např. rozdíl v objemu přenesených dat v den svtátního svátku a v den následující u nemocnice A bude korespondovat s rozdílem v objemu přenesených dat v těchto dnech u nemocnice B. Přitom mezi těmito jevy není příčinná souvislost, ale protože oba jevy souvisí se stejným kalendářem, je zde jistá statistická závislost.

Pro popis statistické závislosti dvou jenů používámě obvykle Pearsonův korelační keoficient. Existují i další možnosti, např. Spearmanův korelační koeficient, ale v našem případě je Pearsonův korelační koeficient (dále jen korelační koeficient) zcela vyhovující.

Korelační koeficient dvou statistických souborů $X = x_1, x_2, ...x_n$ a $Y = y_1, y_2, ...y_n$ je definován vztahem

$$r = \frac{\sum_{i=1}^{n}(x_i -)(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}. \tag{4.18}$$

Korelační koeficient vyjadřuje míru závislosti odchylky od průměrné hodnoty u dvou statistických souborů. Vztah 4.18 je možné upravit do tvaru jednoduššího pro výpočet

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{(n-1)s_x s_y}. \tag{4.19}$$

Korelační koeficient nabývá hodnot od -1 do 1, přičemž hodnoty blízké 1, resp. -1 znamenají silnou lineární závislost (přímou, resp. nepřímou) statistických souborů, zatímco hodnoty blízké 0 signalizují, že mezi sledovanými statistickými soubory lineární závislot není.

Koeficient korelace mezi dvěma statistickýmisouboru má samozřejmě i svůj protějšek v teorii pravděpodobnosti v podobě koeficientu korelace dvou náhodných veličin. Ten je pro náhodné jevy X a Y definován jako

$$\rho_{X,Y} = \frac{E((X - \mu_X)(Y - \mu_Y}{\sigma_X \sigma_Y}. \tag{4.20}$$

Využití korelace mezi objemem přenášených dat v jednotlivých nemocnicích nám umožní přčesněji identifikovat očekávaný průběh přenosu dat.

4.2 Analýza časových řad

Protože pozorované počty přijatých či odeslaných medicínských obrazových studií mají charakter časové řady, můžeme pro jejich analýzu použít i matematické nástroje pro analýzu časových řad.

Časovou řadou rozumíme posloupnost náhodných veličin Y_t, kde index $t \in \{0, 1, 2, ...\}$ má význam času.

Časovou řadu nazýváme stacionární, jestliže její rozdělení pravděpodobnosti je v čase neměnné, tj. společné rozdělení pravděpodobnosti náhodných veličin $Y_T, Y_{T+1}, Y_{T+2}, ...$ není závislé na T.

Pro analýzu časových řad existuje řada matematických nástrojů:

- Metoda dekompozice
- Boxova-Jenkinsonova metodologie
- Spektrální analýza

Metoda dekompozice

Dekompozicí časové řady rozumíme vyjádření časové řady jako součet systematických složek - trendové složky a periodických složek - a náhodné složky, která má charakter bílého šumu.

$$Y_t = T_t + S_t + \epsilon_t. \tag{4.21}$$

Trendem rozumíme dlouhodobý vývoj daného procesu. Periodické složky reflektují vývoj časové řady závislý na kalendáři.

Při analýze časové řady se snažíme vyjádřit nejdříve trendovou složku pomocí vhodné funkce závislé na co nejmenším počtu parametrů, zpravidla jako lineární, kvadratickou nebo exponenciální funkci. Odhady parametrů získáme metodou nejmenších čtverců. Další metodou je vyhlazení časové řady metodou klouzavých průměrů. Tato metoda je adaptivní a je možné j použít i v případě trendu, který v čase mění svůj charakter. Klouzavý průměr počítáme optimálně z lichého počtu po sobě jdoucích členů časové řady.

Po odečtení trendové složky zůstanou v časové řadě složky periodické a složka náhodná. Dále odstraňujeme jednotlivé periodické složky postupně od nejvyšších frekvencí. Periodických složek může časová řada obsahovat i více. Použitím klouzavých průměrů, které odpovídají délce nejkratší periody, odstraníme tuto periodickou složku.

Po odstranění všech systematických složek zůstává jen složka náhodná. Ta by měla mít charakter bílého šumu. Bílý šum má nulovou střední hodnotu a hodnoty ϵ_{t1} a ϵ_{t2} v libovolných ale různých časech $t1$, $t2$ jsou vzájemně nekorelované, tj. $Cov(\epsilon_{t1}, \epsilon_{t2}) = 0$.

Boxova-Jenkinsonova metodologie

Boxova-Jenkinsonova metodologie vychází z předpokladu, že všechny složky časové řady (včetně trendu a cyklické složky) mají náhodný charakter. Jejím těžištěm je korelační analýza. Výhodou této metody je flexibilita a rychlá adaptace na změnu charakteru modelovaného procesu. Nevýhodou je nemožnost jednoduché interpretace matematického modelu časové řady.

Boxova-Jenkinsonova metodologie předpokládá, že zkoumaná časová řada je stacionární, centrovaná (tj. střední hodnota jednotlivých náhodných veličin je rovna nule) a má konečný rozptyl.

Principem této metody je kombinace autoregresního modelu a modelu klouvavých součtů.

Autoregresní model $AR(p)$ časové řady Y_t pracuje se závislostí hodnoty y_t řady v čase t, na předchozích hodnotách časové řady

$$y_t = b_1 y_{t-1} + b_2 yt - 2 + ... + b_p yt - p + \epsilon_t, \tag{4.22}$$

kde ϵ_t je bílý šum.

Model klouzavých součtů $MA(q)$ modeluje časovou řadu jako

$$y_t = \epsilon + w_1 \epsilon_{t-1} + w_2 \epsilon t - 2 + ... + w_q \epsilon t - q, \tag{4.23}$$

kde ϵ_t je bílý šum.

Kombinací těchto procesů vzniká model $ARMA(p, q)$.

Pro analýzu logů událostí v systému MeDiMed se jako výrazně vhodnější ukázala metoda dekompozice časových řad popsaná v kapitole 4.2 jako výrazně vhodnější. Boxovu-Jenkinsonovu metodu uvádím jen pro úplnost a její vlastnosti a možnosti využití nebudeme dále podrobněji rozebírat.

Spektrální analýza

Spektrální analýza časových řad je založena na Fourierově analýze. Předpokládá, že časovou řadu je možné vyjádřit pomocí funkcí $sin(x)$ a $cos(x)$ o různých amplitudách a frekvencích. Spektrální analýzu je vhodné použít v případě časových řad, které obsahují periodické složky, jejichž frekvenci potřebujeme zjistit. V případě časových řad vázaných na běžný kalendář, což je případ jakýchkoli časových řad generovaných vědomou lidskou činností, lze předpokládat, že časová řada obsahuje složky s periodou denní, týdenní a roční, v některých případech i s periodou měsíční. V takovém případě není nutné určovat periodické složky časové řady pomocí spektrální analýzy.

4.3 Entropické modely

Existuje celá řada prací, které se pro detekci anomálního chování sítě, zejména pro detekci bezpečnostních útoků, snaží využít míru neuspořádanosti, nebo složitosti datových toků. Jednou z velmi zajímavých prací na toto téma je [92], kde autoři využili entropii k detekci tehdy aktuálních internetových útoků, tzv. červů (Nachi worm, Welchia worm, Blaster worm a další). Studiu chování tohoto typu škodlivého software se v té době věnovala řada prací, např. [54], a analýza síťového provozu a jeho anomálií byla přirozeným vyústěním [57], [87], [95], [36]. Následovala celá řada pokračovatelů, kteří se snažili optimalizovat vzorkování datových toků pro přesnější a zejména rychlejší detekci, jako např. [61], [38], [21], [45]. Současně se zkoumaly i další možnosti využití statistických metod pro analýzu datových toků [93], [53], včetně distribuce pravděpodobnosti [37], [58]. Velmi zajímavá je i práce, která zkoumá možnosti využití klasických partií matematické statistiky - testování hypotéz [51].

Nutno dodat, že v těchto pracech byla entropie použita jako odhad jiné míry služitosti či strukturovanosti zachycených vzorků dat, která by z teoretického pohledu více odpovídala situaci. Touto mírou je Kolmodorovská složitost (Kolmogorov Complexity) [42], [59].

Na rozdíl od entropie, která popisuje průměrný očekávaný informační obsah zprávy nebo symbolu, který je vybrán jistým nahodilým postupem z dané množiny zpráv, nebo symbolů, popisuje Kolmogorovská složitost informační obsah dané zprávy nebo symbolu. Kolmogorovskou složitost daného objektu můžeme formálně definovat jako minimální velikost popisu (slovního či algoritmického) tohoto objektu, tj. např. jako minimální velikost počítačového programu, kterým je možné daný objekt vygenerovat. Pro praxi je však přímé použití Kolmogorovké složitosti nepříliš vhodné, proto bývá k jejímu odhadu využívána právě entropie.

V mnoha praktických aplikacích, včetně [92] se pro odhad velikosti entropie datového vzorku využívají standardní kompresní algorigmy jako je např. Lempel–Ziv–Oberhumer, který používá všeobecně známý komprimační nástroj ZIP. Jako odhad entropie dat využijeme poměr velikosti původních a komprimovaných dat, který většina implementací tohoto algoritmu poskytuje.

Entropie byla původně využívána v oblasti termodynamiky pro popis rozložení energie v systému. Je-li energie rozložena rovnoměrně, je termodynamická entropie vysoká. Jsou-li naopak v systému místa s různým množstvím energie (např. místa s různou teplotou nebo tlakem), je termodynamická entropie nízká. Termodynamická entropie se tak využívá k vyjádření množství práce, kterou může systém vykonat. Je-li energie rozptýlena rovnoměrně, není způsob, jak ji uvnitř systému využít.

Tímto modelem se v průběhu dalších let inspirovala informatika a využila entropii k popisu množství přenášené informace. Používané entropické modely a možnosti jejich použití pro detekci nežádoucích aktivit v rámci systému přenosu medicínských obrazových dat budou podrobněji diskutovány v samostatných podkapitolách.

Entropie tak, jak ji chápeme v informatice se zpravidla odkazuje na původní práci Clauda Shannona [71] a tato disciplína byla dále rozvíjena v polovině minulého století [60]. Teorie entropie se samozřejmě rozvíjela i čistě matematickým směrem, velmi zajímavá je například práce [91], to už je ale mimo možnosti přímého využití pro detekci nežádoucích přenosů dat v medicínském prostředí. Zajímavý algebraický přístup k definici entropie je v článku [18]. Entropie exponenciálního typu, kam patří např. Tsallisova entropie, jsou diskutovány v [90]. To je jeden z důležitých modelů entropie, který je možné použít pro analýzu spektra příjemců snímků přenášených systémem Redimed. Entropie tohoto typu jsou parametrizovatelné, proto je možné je přizpůsobovat konkrétním aplikacím

entropy:tsallis-tuning.

Entropie není vhodná pro analýzu datových toků ve smyslu počtu přenesených studií a objemu přenesených dat, poskytuje ale velmi zajímavé výsledky v oblasti analýzy složení komunikujících partnerů. Nejdůležitější entropické modely, které jsem pro tuto analýzu využil, jsou popsány v následujících kapitolách.

V informatice je nejvíce využívaným modelem entropie je entropie nazvaná po Claude Elwoodovi Shannonovi. Pro systém S, který má konečnou množinu stavů $s_1, s_2, ...s_n$ s pravděpodobnostmi výskytu těchto stavů $P(s_i)$ definujeme entropii $H(S)$ systému vztahem

$$H(S) = -\sum_{i=1}^{n} P(s_i) log_2(P(s_i)). \tag{4.24}$$

Z fomálních důvodů zde definujeme

$$0.log_2 0 \equiv 0. \tag{4.25}$$

Entropie je maximální pro rovnoměrné rozložení pravděpodobnosti výskytu stavů s_i, tj. pro $P(s_i) = 1/n$ a minimální pro zcela deterministický systém.

Vedle Shannonovy entropie existuje ještě řada dalších modelů. Z těch známějších jmenujme alespoň Tsallisovu entropii a Rényiho entropii.

Rényiho entropii $H_\alpha(S)$ pro systém S s konečnou množinou stavů $s_1, s_2, ...s_n$ s pravděpodobnostmi výskytu těchto stavů $P(s_i)$ definujeme vztahem

$$H_\alpha(S) = \frac{1}{1-\alpha} log_2(\sum_{i=1}^{n} p_i^\alpha), \tag{4.26}$$

kde $\alpha \geq 0$ a $\alpha \neq 1$ Pro $\alpha -> 1$ Rényiho entropie konverguje k Shannonově entropii. Při $\alpha -> \infty$ tento model entropie zdůrazňuje stavy s nejvyšší pravděpodovností výskytu a pro $\alpha -> 0$ naopak význam častěji se vyskytujících stavů klesá.

Tsallisovu entropii $H_q(S)$ pro systém S s konečnou množinou stavů $s_1, s_2, ...s_n$ s pravděpodobnostmi výskytu těchto stavů $P(s_i)$ definujeme vztahem

$$H_q(S) = \frac{1}{1-q} (\sum_{i=1}^{n} p_i^q - 1), \tag{4.27}$$

kde $q \neq 1$. Zde podobně jako v případě Rényiho entropie parametrem q určujeme výslednou podobu modelu.

V další práci prozatím využívám Shannonovu entropii, možnosti a zejména výhody využití jiných modelů jsou předmětem dalšího výzkumu.

4.4 Práce s neúplnou a nepřesnou informací

Hranice toho, kdy už považovat počet odesílaných studií za podezřelý a kdy ještě za běžnou situaci je neostrá. Není možné jednoznačně rozhodnout, že 67 snímků za den je ještě v pořádku a při 68 už je podezření na nežádoucí únik dat. Při hledání vhodných matematických modelů, které by byly schopny pracovat i jistou mírou neurčitosti, nebo našeho přesvědčení, zda daná hodnota ještě je nebo už není v souladu s očekáváním, je přirozeným řešením využití fuzzy množin.

Zakladatelem teorie fuzzy množin je Lotfi Aliasker Zadeh, který první práci na toto téma publikoval již v roce 1965 v časopise Information and Control. V klasické teorii množin může ke každé množině přiřadit tzv. charakteristickou funkci, která pro libovolný prvek dává hodnotu 1, pokud tento prvek je prvkem dané množiny a hodnotu 0 v opačném případě. Zjednodušeně by se dalo říct, že fuzzy množiny jsou rozšířením klasické teorie množin v tom smyslu, že charakteristickou funkci zobecníme tak, aby mohla nabývat libovolné hodnoty z intervalu $< 0, 1 >$.

V této kapitole připomeneme základní pojmy a vlasnosti fuzzy množin. V této chvíli ještě nejsou pro analýzu datových toků radioligického komunikačního systému

Redimed využívány, ale možnosti jeich nasazení analyzujeme . Očekávaný přínos je zejména v možnosti poskytnout lidské obsluze dektčních systémů komplexnější informaci a podporu pro rozhodování.

Dále uvedený popis konceptu fuzzy množin je převzat z mé dizertační práce.

Nechť U je množina, μ_A je zobrazení U do $< 0, 1 >$. Fuzzy množinou A na univerzu U nazveme uspořádanou dvojici $A = (U, \mu_A)$. Zobrazení $\mu_A : U \longrightarrow \langle 0, 1\rangle$ nazýváme funkce příslušnosti fuzzy množiny A a její hodnotu $\mu_A(x)$ stupněm příslušnosti prvku $x \in U$ k fuzzy množině A.

Fuzzy množinu A můžeme formálně ztotožnit s její funkcí příslušnosti. Proto se v literatuře často funkce příslušnosti označuje stejným symbolem jako fuzzy množina.

Fuzzy množina je zobecněním pojmu množina. Vezmeme-li v předchozí definici takovou funkci μ_A, že $\mu_A(x) \in \{0, 1\}$, stane se funkce μ_A charakteristickou funkcí množiny. (Dokonce existují pokusy o axiomatickou výstavbu teorie fuzzy množin jako zobecnění Zermelovy-Fraenkelovy axiomatizace teorie množin.)

Nechť $A = (U, \mu_A)$ je fuzzy množina. Pak

- nosič fuzzy množiny A je klasická množina $supp\, A = \{x \in U | \mu_A(x) \neq \emptyset\}$,
- jádro fuzzy množiny A je klasická množina $ker A = \{x \in U | \mu_A(x) = 1\}$,
- fuzzy množina A se nazývá normální, jestliže $ker A \neq \emptyset$.

Speciálním případem fuzzy množin jsou tzv. intervalová čísla $\langle a, b\rangle$, kde $a, b \in \mathbf{R}$. Tj. fuzzy množiny na $\mathbf{R}$, takové, že $supp\, A = ker A\, \langle a, b\rangle$.

Základní vlastnosti fuzzy množin:

Nechť A,B jsou fuzzy množiny na univerzu U ; $\alpha, \beta \in \langle 0, 1\rangle$. Pak platí:

- $A_\alpha \supseteq A_\beta$ pro $\forall \alpha \leq \beta$
- $(A \cap B)_\alpha = A_\alpha \cap B_\alpha$
- $(A \cup B)_\alpha = A_\alpha \cup B_\alpha$
- $\overline{A}_\alpha = \overline{A_{(1-\alpha)+}}$

Nechť A_i jsou fuzzy množiny na univerzu U pro všechna $i \in I$, kde I je indexová množina. Pak platí:

- $\bigcup\limits_{i\in I}(A_i)_\alpha \subseteq (\bigcup\limits_{i\in I} A_i)_\alpha$
- $\bigcup\limits_{i\in I}(A_i)_{\alpha+} = (\bigcup\limits_{i\in I} A_i)_{\alpha+}$
- $\bigcap\limits_{i\in I}(A_i)_\alpha = (\bigcap\limits_{i\in I} A_i)_\alpha$
- $\bigcap\limits_{i\in I}(A_i)_{\alpha+} \subseteq (\bigcap\limits_{i\in I} A_i)_{\alpha+}$

Nechť A je fuzzy množina, $\alpha, \beta \in \langle 0, 1\rangle$. Pak platí:

- $A_\alpha = \bigcap\limits_{\beta<\alpha} A_\beta = \bigcap\limits_{\beta<\alpha} A_{\beta+}$
- $A_{\alpha+} = \bigcup\limits_{\beta>\alpha} A_\beta = \bigcup\limits_{\beta>\alpha} A_{\beta+}$

Na podkladě fuzzy množin postupně vyrostly úpravy mnoha matematických disciplín, jako je teorie pravděpodobnosti, nebo i entropické modely, které umožňují snadnější modelování práce s neúplnou nebo neurčitou informací. Neurčitost informace může být způsobena i subjektivním názorem člověka. Speciálně v případě matematické analýzy datových toků se zde skrývá subjektivní názor lidské obsluhy na to, kdy už považujeme datový tok za podezřelý a budeme ho podrobněji zkoumat po obsahové stránce. Existuje řaada prací, které se zabývají využitím fuzzy množin pro modelování míry přesvědčení člověka ve správnost údajů pro rozhodování. Vedle historických původních prací L.A.Zadeha Z nejdůležitějších pramenů citujme alespoň [97], [96], [55], [94], [70], nebo učebnici [98].

K nejzajímavějším využitím fuzzy množin z pohledu analýzy datových toků komunikačních systémů, jako je Redimed, patří fuzzy entropie: [46], [23], [39].

5 Netechnické aspekty detekce úniku dat

Pokud potřebujeme skutečně funkční systém pro detekci anomálních přenosů dat, musíme kromě sofistikovaných technických či matematických postupů vzít v úvahu i některé aspekty netechnické. Jedná se zejména o právní problematiku, neboť shromažďování a zpracování dat podléhá určitým právním úpravám a v neposlední řadě i o záležitosti psychologické, protože výstupy detekčního systému budou následně zpracovávány lidskou obsluhou a je proto třeba brát v úvahu i záležitosti např. reakci lidí na stereotypní výstupy a pod.

5.1 Právní aspekty

Úkol zabezpečit výpočetní a komunikační systém proti neoprávněné manipulaci s daty úzce souvisí i s právní problematikou. Hranice mezi tím, co reálně potřebujeme pro ajištění bezpečnosti provozu a tím, co vyžadují nejrůznější zákonné předpisy a normy je mnohdy velmi tenká a snadno se můžeme dosta za hranu zákona. Proto této kapitole přináším přehled základních právních norem, se kterými se při řešení úkolu zabezpečení počítačové sítě musíme vyrovnat.

Existuje celá řada právních norem, které upravují chování uživatelů v kyberprostoru a řeší počítačovou kriminalitu a to jak na národní tak i na evropské či mezinárodní úrovni. Právní předpisy definují pojem "Počítačová kriminalita"jako trestnou činnost, které se odehrává v kyberprostoru, tj. má souvislost s informačními a komunikačními technologiemi. Výraz kyberprostor (cyberspace) byl poprvé použit spisovatelem Williamem Gibsonem v románu Neuromancer [44] z roku 1984. Román pojem kyberprostor zpopularizoval natolik, že se začal používat jako odborný termín pro „prostor“ počítačových systémů a sítí v němž probíhá on-line komunikace. Z pohledu mezinárodního práva počítačovou kriminalitou rozumíme jednání, na která dopadá Úmluva o počítačové kriminalitě [10].

Právo rozlišuje dva základní druhy počítačové kriminality:

1. Počítač je předmětem útoku. Do této kategorie patří veškeré průniky do informačních a komunikačních systémů, prolamování jejich ochrany a na to navazující zásahy do dat v nich uložených. Rovněž sem patří jednání, kdy útočník sice nepronikne dovnitř ICT systému, nicméně dojde k vyřazení systému z provozu, případně je jeho funkčnost omezena. Tj. spadají sem i všechny DOS a DDOS útoky, neoprávněné využívání napadených počítačů k těžbě kryptoměny, nebo využívání napadených systémů k dalším útokům.
2. Počítač je jen nástrojem útoku. Do této kategorie patří zejména různé podvodné zprávy, které pod smyšlenou historkou nabízení vyplacení velké částky peněz za podmínky předchozího uhrazení výloh na finanční transakce, nebo

tzv. phishing, tj. snaha vylákat smyšleným sdělením zadání např. přihlašovacích údajů, které pak pachatel hodlá zneužít.

Samozřejmě mezi těmito kategoriemi je určitá prostupnost a vazba. Například pomocí phishingu může pachatel získat přístupové údaje, které následně využije k neoprávněnému vstupu do počítačového systému. Jednání v prvé kategorii je typicky možné považovat za počítačovou kriminalitu, zatímco u druhé kategore může být někdy otázkou, zda užití výpočetní techniky hraje takovou roli, aby se jednalo o počítačovou kriminalitu.

5.1.1 Právní předpisy

Úmluva předepisuje, aby státy zavedly trestnost následujícího jednání:

- Nezákonný přístup – úmyslný neoprávněný přístup k počítačovému systému.
- Nezákonný odposlech – úmyslný, neoprávněný, technickými prostředky provedený odposlech neveřejného přenosu počítačových dat z nebo do počítačového systému.
- Zasahování do dat - úmyslné neoprávněné poškození, vymazání, nebo pozměnění počítačových dat.
- Zasahování do systému - úmyslné neoprávněné závažné omezení funkčnosti počítačového systému. Dle vysvětlující zprávy k Úmluvě sem patří i DOS útoky [2].
- Zneužívání zařízení – Zde jde o postih úmyslné a neoprávněné tvorby, držení a zpřístupňování programů a technických zařízení uzpůsobených k trestných činů, shromažďování přístupových kódů, pomocí nichž lze získat neoprávněný přístup do počítačového systému.
- Počítačové padělání – úmyslné zasahování do počítačových dat, které povede k jejich nepravosti, a to s úmyslem, aby tato data byla považována za pravá.
- Počítačový podvod – úmyslné způsobení ztráty na majetku jinému, činěné zásahem do počítačových dat nebo do fungování počítačového systému, s úmyslem neoprávněně získat pro sebe nebo pro jiného majetkový prospěch.
- Trestné činy týkající se porušení autorského práva a práv souvisejících s právem autorským a porušení práv duševního vlastnictví chráněným mezinárodními smlouvami.

V rámci evropského práva se uplatňuje směrnice č. 2013/40/EU o útocích na informační systémy a nahrazení rámcového rozhodnutí Rady 2005/222/SVV [4], která po členských státech Evropské unie mimo jiné požaduje, aby uzákonily jako v zásadě všechny druhy útoků na počítačové systémy, u kterých to požaduje Úmluva o počítačové kriminalitě [10].

Evropské právo obsahuje i řadu dalších předpisů, které s problematikou počítačové kriminality souvisejí. Jejich obsah však není pro náš připad relevantní.

České právo se počalo zaměřovat na postih počítačové kriminality od počátku 90. let 20. století, kdy od 1. 1. 1992 se stal součástí tehdejšího trestního zákona §257a upravující trestný čin poškození a zneužití záznamu na nosiči informací.

Jádrem současné české právní úpravy je §230 trestního zákoníku [6] stanovující jako trestný čin neoprávněný přístup k počítačovému systému a nosiči informací.

Tento paragraf obsahuje dva základní odstavce. Podle odstavce 1 se trestného činu dopustí ten, kdo překoná bezpečnostní opatření, a tím neoprávněně získá přístup k počítačovému systému nebo k jeho části. V tomto bodě zákona je chráněna důvěrnost počítačových dat a počítačového systému. Teprve sekundárně jsou chráněny integrita a dostupnost počítačových dat a systémů.

Podle odstavce 2 se trestného činu dopustí ten, kdo získá přístup k počítačovému systému nebo k nosiči informací a

a) neoprávněně užije data v nich uložená,
b) taková data vymaže poškodí nebo změní,
c) data padělá nebo pozmění tak, aby byla považována za pravá,
d) neoprávněně vloží data do počítačového systému nebo na nosič informací.

Zákon operaju s pojmy "překonání bezpečnostních opatření"a "neoprávněnost--neoprávněné užití, neoprávněné vymazání-vložení atd. Přesná definice těchto pojmů je předmětem mnoha diskusí a právních komentářů publikovaných např. v [8], [20], [22], [7] a [9]. Další užitečné informace k problematice počítačové kriminality je možné nalézt v [56], [17] a [63].

Další paragraf, který se týká počítačové kriminality, je §182, kerý chrání tajemství dopravovaných zpráv. Relevantní jsou zejména odst. 1 písm. b), c), které, říkají, že trestný čin spách, kdo úmyslně poruší tajemství

b) datové, textové, hlasové, zvukové či obrazové zprávy posílané prostřednictvím sítě elektronických komunikací a přiřaditelné k identifikovanému účastníku nebo uživateli, nebo
c) neveřejného přenosu počítačových dat do počítačového systému, z něj nebo v jeho rámci.

Porušení tajemství je jakékoli neoprávněné narušení posílané zprávy nebo neveřejného přenosu počítačových dat se snahou zjistit jejich obsah. Podmínkou trestnosti není, aby tento obsah musel být někomu dalšímu sdělen. Složitější na posouzení je, kdy je porušení tajemství oprávněné: k určité míře sledování může být oprávněn například zaměstnavatel při kontrole činnosti zaměstnance.

Další otázkou je, jak široce se má chápat neveřejný přenos počítačových dat. Je chráněno pouze tajemství vlastního obsahu zprávy, kvůli kterému přenos probíhá, nebo i doprovodná technická data a údaje o probíhajícím datovém provozu? V tomto

bodě se právní otázky potkávají s největším právnickým fenoménem posledních let - GDPR (General Data Protection Regulation) [3].

GDPR představuje nový právní rámec ochrany osobních údajů v evropském prostoru. Cílem této právní úpravy je hájit práva občanů EU proti neoprávněnému zacházení s jejich daty, zejména osobními údaji. GDPR se týká nejen firem a institucí, ale i jednotlivců a online služeb, které zpracovávají data uživatelů. GDPR zavedlo astronomické pokuty za porušování pravidel a v mnoha případech vyvolává nejistotu ohledně toho, jaké údaje ještě můžeme zpracovávat.

Je např. IP adresa osobní údaj ve smyslu zákona? Podle způsobu přiřazení IP adresy konkrétnímu zařízení, způsobu využívání tohoto zařízení a způsobu a době uchovávání záznamů o přidělení IP adresy může ale také nemusí být z IP adresy zjistitelná (alepoň po nějakou dobu) identita jejího uživatele.

5.2 Psychologické aspekty

Jakýkoli systém detekce anomálií nemůže algoritmicky rozhodnout, zda detekovaná anomálie představuje např. bezpečnostní hrozbu, či poruchu technologie, nebo se jedná jen o odchylku od běžného stavu, která má racionální vysvětlení a nepředstavuje skutečný problém. Systém potřebuje lidskou obsluhu, která provede příslušné vyhodnocení a rozhodnutí a případně spustí odpovídající reakci. Pro správné fungování detekčních systémů je zapotřebí vzít v úvahu i určité známé vzorce lidského chování. Například, pokud se bude příliš často opakovat falešný poplach, lidská obsluha se v relativně krátkém čase na takovou situaci adaptuje tím, že poplachové zprávy bude apriori považovat za falešný poplach a bude je prostě ignorovat.

Mám dlouholetou zkušenost s provozem rozsáhlé datové sítě, s dohledovým centrem pracujícím v režimu nepřetržitého provozu. Není jednoduché přesvědčit obsluhu dohledového centra, aby každému hlášení dohledového systému věnovala dostatečnou pozornost. Hlavními problémy, se kterými jsem se v praxi setkal jsou:

- Často se opakující falešný poplach. Pokud se falešný poplach nebo hlášení o chybě častěji opakuje, obsluha na něj přestává reagovat. Typickým příkladem je plánovaná údržba rozložená do více etap, pokud není v předstihu nahlášena dohledovému centru. V praxi jsem se opakovaně setkal s následující situací: systém nahlásil výpadek spojení k zákazníkovi, obsluha po telefonickém rozhovoru zjistila, že zákazník provádí údržbu svého zařízení a jedná se tak o plánovaný, lež nenahlášený výpadek. Po pár dnech se situace opakovala. Při třetím, maximálně čtvrtém opakování situace již obsluha dohledového centra zákazníka nekontaktovala a prohlásila, že "zákazník určitě zase prování neohlášenou údržbu"aniž by zjišťovala skutečný stav věci.

- Často se opakující problémy, které se nakonec "vyřeší samy". Typickým příkladem jsou opakující se krátkodobé výpadky napájení. I v tomto případě obsluha dohledového centra velmi rychle dospěje do stavu, že v případě signalizace výpadku uzlu, který byl tímto problémem postižen, rostě prohlásí "To bude určitě zase výpadek napájení, počkáme, jestli se to za hodinu nespraví samo.".
- Příliš velká úroveň vnoření dohledovaných prvků. V situaci, kdy na přehledové mapě stavu sítě máme pod jednou ikonkou, která barevně signalizuje stav sítě, schovaný celý kampus, bude výpadek jednoho (třeba i nepodstatného) síťového prvku bude signalizován změnou barvy ikony pro celý kampus. Při změně barvy obsluha dohledá příčinu signalizace poruchy, dále však již stav zbytku kampusu nekontroluje s tím, že "červenou barvu této ikony způsobuje nefunkční dohledový modul UPS v rozvodné skříni na půdě".

Při konstrukci jakého koli systému pro detekci anomálních stavů přenosu dat musíme mít na zřeteli všechny tyto právní a psychologické apekty.

6 Analýza logů systému Redimed

Jak již bylo zmíněno v úvodu práce, ke dni odevzdání této práce měl medicínský komunikační systém Redimed něco přes 570 uživatelů. Ne všichni uživatelé používají Redimed stejným způsobem a stejnou měrou. Řada uživatelů je pouze pasivními příjemci dat. Jedná se zejména o privátní praxe radiologů, kteří se věnují vyhodnocování snímků zaslaných z jiných zdravotnických institucí. Dalšími typickými uživateli tohoto typu jsou praktičtí lékaři, kteří tak mají k dispozici obrazovou dokumentaci pacienta, kterého např. odeslali do jiného zdravotnického zařízení a mohou mu následně podrobněji vysvětlit způsob a průběh léčby v nemocnici apod.

Počet uživatelů, kteří aktivně odesílají data je mnohem menší, než celkový počet uživatelů systému.Vývoj počtu aktivních odesilatelů obrazových informací je zachycen v tabulce 6.1. Počet aktivních odesilatelů průběžně roste po celou dobu existence systému Redimed. Dá se říct, že v roce 2015 systém překonal počáteční fázi, kdy se uživatelé teprve seznamovali s jeho možnostmi a hledali vhodný způsob využití odpovídající právě jejích potřebám a režimu práce. Pro analýzu a predikci toho, jak by se měl systém chovat a co už je odchylka od očekávaného stavu, na kterou by bylo vhodné upozornit správce, proto použijeme data z let 2015 - 2018.

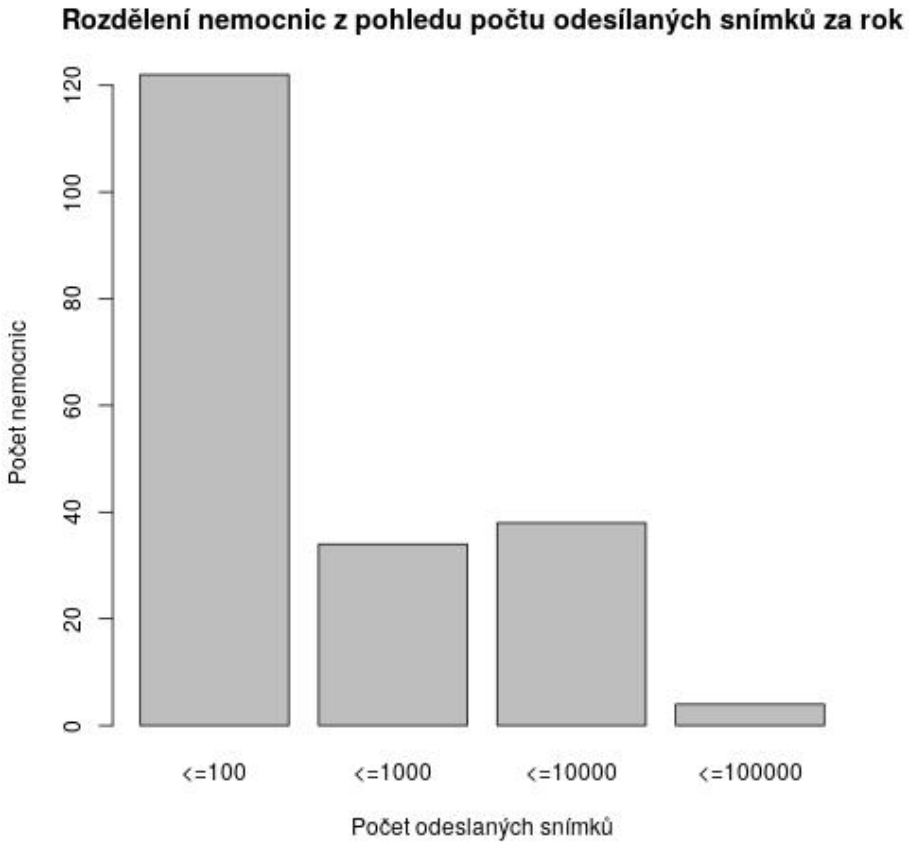
Tab. 6.1: Počet aktivních odesilatelů medicínských obrazových informací v systému Redimed.

Rok	2011	2012	2013	2014	2015	2016	2017	2018
Počet aktivních odesilatelů	83	105	105	127	145	172	183	198

I uživatelé, kteří aktivně odesílají snímky pomocí systému Redimed, využívají tento systém různým způsobem. Je zde řada uživatelů, kteří si systém Redimed chtěli jen vyzkoušet, nebo kteří mají jen velmi malé potřeby odesílat vlastní snímky spolupracujícím institucím. Uživatelé této kategorie odesílají nejvýše desítky až stovky snímků ročně. Pak jsou tady uživatelé střední velikosti, tj. uživatelé, kteří odesílají průměrně alespoň 5 snímků každý pracovní den, tj. zhruba 1000 a více snímků ročně. Systém Redimed má i určitý počet velkých uživatelů, kteří odesílají průměrně desítky snímků denně a některé speciální uživatele. Speciálními uživateli jsou například stanice určené k přeposílání nikoli celých snímků, ale pouze jejich hlaviček do systému pro výpočet radiační zátěže pacientů.

Rozložení uživatelů z pohldu počtu odesílaných snímků je nejlépe vidět z grafu na obrázku 6.1. Kompletní přehled ročního počtu odeslaných studií a objemu pře-

nesených dat za jednotlivé uživatele systému Redimed je uveden v příloze B. Jedná se o přehled za rok 2018. Spektru provozu dominují malí uživatelé, ale je zde silně zastoupená i skupina velkých uživatelů, kteří odesílají více než 10 tisíc snímků ročně. Matematický popis profilu provozu jednotlivých skupin uživatelů nejsnadněji ukážeme na příkladech.



Obr. 6.1: Rozložení uživatelů z pohledu počtu odesílaných snímků.

6.1 Analýza provozu malých uživatelů

U malých uživatelů je prakticky nemožné provádět nějaké statistické vyhodnocení jejich provozu způsobem, který by umožňoval predikovat očekávaný profil provozu a upozornit na neobvyklý stav. Důvodem je relativně velký rozptyl vyhodnocovaných dat, který by vedl buďto k situaci, že bude relativně málo citlivý na změnu počtu přenášených studií a tím i relativně benevolentní k případnému útočníkovi, nebo naopak příliš striktní a v tom případě by generoval příliš mnoho falešných poplachů. Pro ilustraci se podívejme na profil provozu polikliniky z jednoho mešího města (pod 10 tis. obyvatel), která odesílá zhruba 200 snímků ročně. Profil datového provozu

této polikliniky v jednotlivých letech je uveden v tabulce 6.2, počty odeslaných snímků v jednotlivých měsících roku 2018 v tabulce 6.3.

Tab. 6.2: Počet odeslaných studií příkladové polikliniky v jednotlivých letech.

Rok	2011	2012	2013	2014	2015	2016	2017	2018
Počet odeslaných studií	181	56	217	211	221	247	267	208

Tab. 6.3: Počet odeslaných studií příkladové polikliniky v průběhu roku 2018.

Měsíc	1	2	3	4	5	6	7	8	9	10	11	12
Počet odeslaných studií	32	21	47	21	22	25	14	25	1	0	0	0

Protože se snažíme najít především situace, kdy uživatel (naše příkladová poliklinika) odesílá více dat, než je obvyklé, můžeme pro analýzu provozu použít upravený vstupní soubor, v němž využijeme pouze data těch měsíců, ve kterých bylo odesláno alespoň 10 studií. Základní popisné statistiky ročního profilu provozu této polikliniky ve smyslu počtu přenesených studií v původním i upraveném souboru jsou shrnuty v tabulce 6.4.

Tab. 6.4: Analýza měsíčního počtu odesílaných studií příkladové polikliniky.

Statistika	**Původní soubor**	**Upravený soubor**
Rozsah souboru	12	8
Varianční rozpětí	47	33
Průměr	17,33	25,88
Směrodatná odchylka	11,83	9,92
Varianční koeficient	0,68	0,38

Z tabulky 6.4 lze snadno vidět, že počty odesílaných studií v jednotlivých měsících jsou velmi proměnlivé a i v případě, kdy neuvažujeme měsíce, ve kterých není žádný nebo jen minimální provoz, kolísá počet odeslaných studií o téměř 40%. K

této skutečnosti je třeba ještě přičíst fakt, že uživatel může kdykoli odeslat jednotky studií navíc, např. z důvodu testování spojení. V případě této polikliniky např. 5 studií navíc znamená o 10% vyšší počet studií, než je aktuální dosažené měsíční maximum.

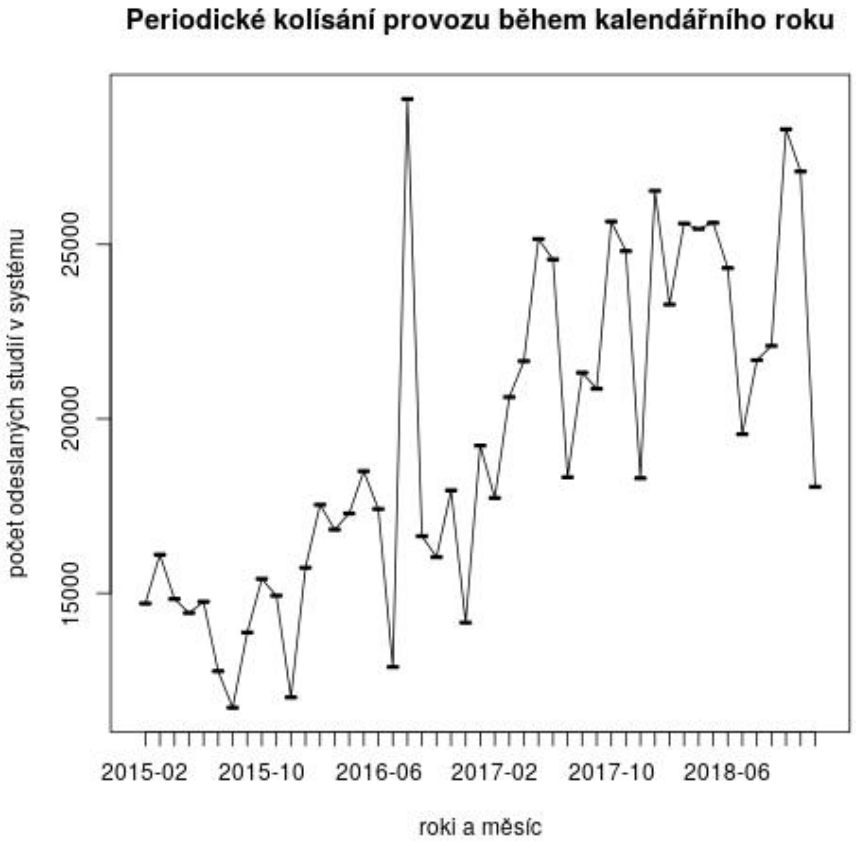
Z uvedeného přehledu je zřejmé, že v případě malých uživatelů radiologického komunikačního systému Redimed, sofistikovaná matematická analýza datových toků nedává příliš smysl. U uživatelů této velikosti je nejépe použitelnou metodou prosté sledování počtu odeslaných studií v každém měsíci. Pokud by počet počet odeslaných studií překročil vhodně stanovený násobek maximální hodnoty uplynulého roku je vhodné uživatele upozornit. Jako "bezpečná hodnota", tj. stav, kdy nebudeme generovat příliš mnoho planých poplachů a dokážeme zachytit nástup případného útoku, se jeví dvojnásobek maxima odeslaných studií za měsíc (měřeno v předchozím kalendářním roce). Takovéto nastavení kontroly dokáže vstřebat i průběžný mírný nárůst komunikace uživatele. Pro lepší představu o struktuře provozu této polikliniky je kompletní profil provozu za rok 2018 uveden v příloze C.

6.2 Analýza provozu velkých uživatelů

U velkých uživatelů, kteří odesílají dostatečné množství snímků, bývá provoz ustáleněší a relativně méně variabilní (ve smyslu velikosti variančního koeficientu), proto můžeme použít více matematických nástrojů. V případě velkých uživatelů množství odesílaných studií více koresponduje s běžným kalendářem a můžeme proto s výhodou použít některé metody analýzy časových řad a vzájemnou korelaci statistik z různých zdravotnických institucí pro zpřesnění odhadu očekávaného počtu odesílaných studií.

U dostatečně velkých uživatelů, tj. takových kteří odesílají v průměru alespoň 5 snímků denně (stačí v pracovní dny) už se projevují periodické znaky chování uživatelů svázané s kalendářem. Na průběhu grafu počtu odeslaných snímků je zřetelně vidět roční periodický průběh, kde se projevuje vliv letních prázdnin a vánočních svátků. Graf celkového počtu studií odesílaných systémem Redimed v průběhu posledních čtyř let je na obrázku 6.2. Na grafu je zřetelně vidět 8 lokálních minim provozu v době letních prázdnin a vánočních svátků.

Pro demonstraci možností matematické analýzy profilu datových toků jsem jsem vybral jednu z větších nemocnic v krajském městě. Nemocnice odesílá ročně zhruba 5000 snímků. Graf měsíčních úhrnů počtu snímků odeslaných z této nnemocnice je na obrázku 6.3. Graf zřetelně kopíruje charakteristické chování celého systému. Tuto vlastnost můžeme s výhodou využít pro modelování periodického kolísání počtu přenesených snímků během roku.

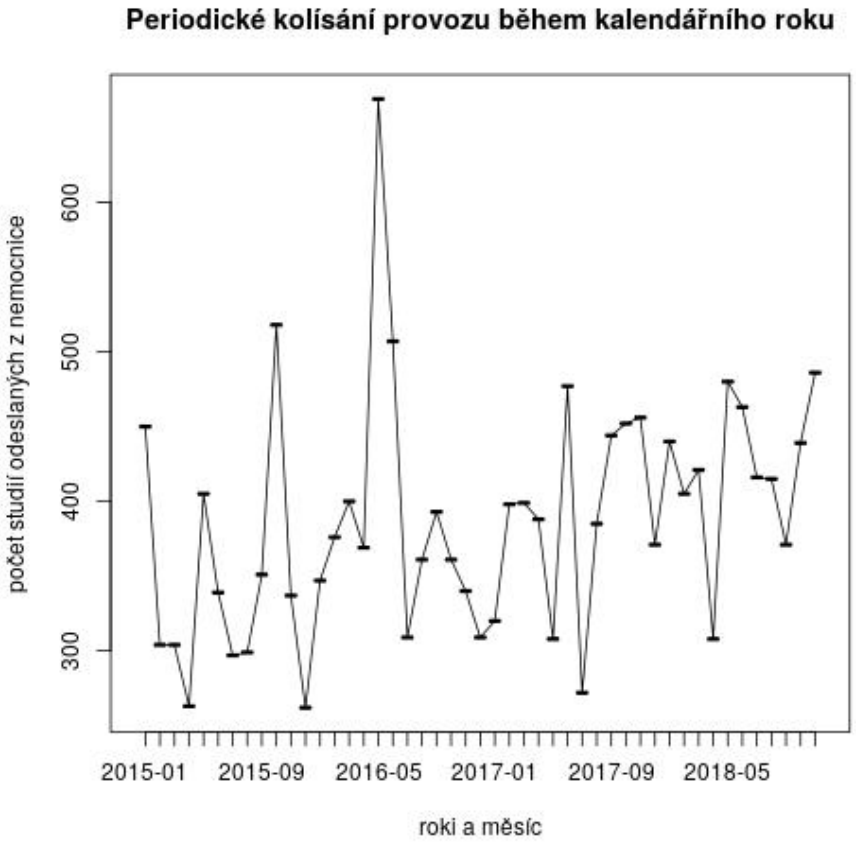


Obr. 6.2: Měsíční úhrny počtu přenesených zpráv v systému Redimed.

Pro srovnání s malou zdravotnickou organizací se podívejme, jak vypadá odesílaných studií během roku 2018. Počty odeslaných studií v jednotlivých měsících roku 2018 jsou uvedeny v tabulce 6.5 a statistická analýza těchto dat v tabulce 6.6. Varianční koeficient je v tomto případě výrazně menší, než u polikliniky z předchozího případu. Můžeme tedy zdánlivě přesněji předpovídat budoucí chování systému a citlivěji nastavit meze, při jejichž překročení bude systém kontaktovat správce dané nemocnice. Při podrobnějším pohledu ale zjistíme, že varianční rozpětí počtu snímků odesílaných z této nemocnice je stále relativně velké. Pokud nastavíme prahovou hodnotu pro generování upozornění správců systému dostatečně vysoko (tak abysme minimalizovali vznik falešných poplachů), zůstane mám ještě pčíliš široký prostor pro případného útočníka. Pokud by se počet odesílaných snímků za měsíc zvýšil jen o málo desítek, detekční systém by na takovou situaci nereagoval.

Proto potřebujeme ještě další podpůrné analýzy, aby bylo možné zpřesnit odhad toho, jestli je objem přenášených dat ještě v rámci požadovaného stavu.

Z tabulky 6.5 je jasně vidět, že v průběhu měsíce odesílá tato nemocnice více snímků, než poliklinika v menším městě za celý rok. To můžeme s výhodou využít pro



Obr. 6.3: Měsíční úhrny počtu zpráv odeslaných z vybrané nemocnice.

včasnější zachycení nežádoucích toků dat, neboť na analýzu provozu této nemocnice použijeme stejnou metodu, jako na analýzu provozu menší polikliniky za delší časové období. Profil datových toků během měsíce listopadu za tuto nemocnici k dispozici v tabulce 6.7 ve sloupci A.

Tab. 6.7: Profil provozu vělších nemocnic.

Datum	**Nemocnice A**	**Nemocnice B**	**Nemocnice C**
2018-11-01	35	38	26
2018-11-02	9	19	12
2018-11-03	4	1	1
2018-11-04	1	1	0
2018-11-05	10	27	33
2018-11-06	25	39	14
2018-11-07	33	41	4

2018-11-08	16	8	24
2018-11-09	30	26	15
2018-11-10	0	3	4
2018-11-11	1	2	2
2018-11-12	35	77	11
2018-11-13	15	19	16
2018-11-14	36	18	16
2018-11-15	13	18	3
2018-11-16	15	16	31
2018-11-17	0	3	0
2018-11-18	1	12	0
2018-11-19	24	36	39
2018-11-20	18	25	18
2018-11-21	21	40	17
2018-11-22	24	24	18
2018-11-23	22	19	16
2018-11-24	2	5	1
2018-11-25	2	15	0
2018-11-26	12	37	15
2018-11-27	16	24	13
2018-11-28	9	25	19
2018-11-29	21	21	20
2018-11-30	36	20	9

V tabulce 6.7 je dobře patrná týdenní perioda v počtu odeslaných studií, zároveň je zde ale vidět i to, že tato perioda neodpovídá kalendáři zcela přesně. Ještě lépe je to viditelné z grafu na obrázku **??**. Tyto nepravidelnosti jsou dány nepravidelnostmi v pracovním kalendáři. Prakticky neexistuje měsíc (snad s výjimkou srpna, který je ale ovlivněn prázdninovým provoze a čerpáním dovolené u řady lidí), ve kterém by se nevyskytoval alespoň jeden státní svátek, nebo krátkodobé školní prázdniny. Z toho důvodu je poměrně komplikované využít periodické chování uživatelů pro zpřesnění odhadu počtu snímků, které mají být v daný den přeneseny. Metody analýzy časových řad, kterým jsem analýzu počtu odesílaných studií taktéž podrobil, nám sice nabízení řešení pro vyrovnání odchylek v pracovním kalendáři, bohužel ale za cenu nižší spolehlivosti odhadu.

Pro řešení tohoto problému se m iosvědčilo využít korelační analýzu. Proto jsou v tabulce 6.7 také počty odeslaných studií ve stejném období od dvou dalších srovnatelně velkých nemocnic (sloupce B a C). Vlivy nepravidelnosti kalendáře se projevují

Tab. 6.5: Počet odeslaných studií příkladové velké nemocnice v průběhu roku 2018.

Měsíc	1	2	3	4	5	6	7	8	9	10	11	12
Počet odeslaných studií	331	486	439	371	415	416	463	480	308	421	405	440

Tab. 6.6: Analýza měsíčního počtu odesílaných studií příkladové větší nemocnice.

Rozsah souboru	12
Varianční rozpětí	178
Průměr	414,58
Směrodatná odchylka	55,06
Varianční koeficient	0,13

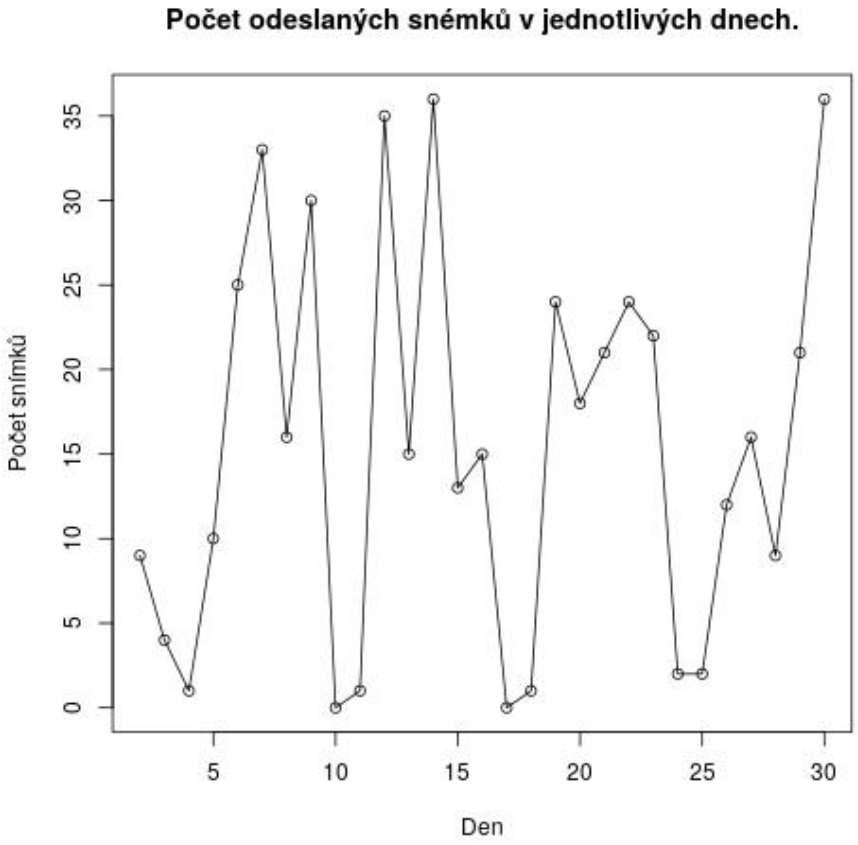
u všech nemocnic stejně. Pokud u některé nemocnice nastane výraznější pokles počtu odeslaných studií vlivem nepravidelnosti v pracovním kalendáři, dá se očekávat, že stejně budou reagovat i další nemocnice. Takový pokles nebo naopak nárůst počtu odeslaných snímků se neprojeví úplně stejně ve všech nemocnicích. Je to dáno jednak drobnými rozdíly v organizaci práce, samozřejmě rozdíly ve velikosti a dopravní dostupnosti, ale potřebu komunikovat ovlivňují i urgentní případy, které zpravidla v době státních svátků neošetřují všechny nemocnice.

Korelace nárůstu a poklesu počtu odeslaných snímků z vybraných nemocnic je z tabulky patrná na první pohled. Pro potřeby deteke anomálií v datových tocích ji však potřebujeme uchopit vhodným matematickým nástrojem. Korelační koeficienty pro ověření dostatečné statistické vazby mezi počtem odeslaných snímků tří příkladových nemocnic jsou zde:

```
r_{A,B} = 0,69

r_{A,C} = 0,69.
```

V obou případech vychází dostatečně vysoký korelační koeficient a proto můžeme datový provoz těchto nemocnic použít jako vzájemnou referenci.



Obr. 6.4: Graf počtu odeslaných studií příkladové nemocnice během měsíce listopadu roku 2018.

Pro odhad očekávaného počtu přenesených studií používáme relativní přírůstek nebo úbytek počtu přenesených studií mezi dvěma po sobě jdoucími dny vypočtený pomocí předpisu 6.1.

$$\delta(x_i) = \frac{x_i - x_{i-1}}{avg7(x_i)}, \tag{6.1}$$

kde $avg7(x_i)$ je klouzavý průměr počtu přenesených studií za poslední týden, tj.

$$avg7(x_i) = 1/7 \sum_{j=i-7}^{i-1} x_j. \tag{6.2}$$

Jako referenční hodnotu, ke které vztahujeme změny používáme týdenní průměr z několika důvodů: průměrná hodnota na přiměřeně dlouhé časové období do určité míry polačuje vliv odlehlých měření a současně by se mohlo stát, že daná nemocnice např. pře víkend žádné snímky neodesílá a bylo by nutné ošetřit stavy, které by při výpočtu mohly vést k pokusu o dělení nulou. Relativní přírůstky počtu přenesených studií pro naše tři příkladové nemocnice jsou k dispozi v tabulce

6.8. Použití relativních procentuálních přírůstků nám pomáhá vyrovnat sezónní kolísání počtu odesílaných snímků během roku a zjednodušuje srovnání podobných nemocnic, kterým řešíme nepravidelnosti v kalendáři způsobené státními svátky.

U velkých zdravotnických zařízení je zapotřebí sledovat i očekávaný pokles počtu odesílaných studií v době slabého provozu. Celkový objem přenášených dat je zde natokil významný, že případný útočník by mohl snadno využít provozního sedla a v době slabého provozu odeslat data, která potřebuje, aniž by si toho někdo všimnul.

V případě velkých nemocnic, které odesílají desítky snímků denně, má smysl se zavývat i rozložením provozu v průběhu dne. V tabulce je příklad rozložení odesílání snímků během dne u nemocnice, která byla diskotována v předchozím příkladu. V tomto případě jsem použil data ze srpna letošního roku, protože historické záznamy se s touto přesností neuchovávají.

U nemocnice této velikosti nemáme dostatečné množství dat k tomu, aby bylo možné konstruovat spolehlivé statistické modely rozložení datového provozu během dne. Přesto však máme k dispozici některé základní údaje, které nám mohou pomoci detekovat neobvyklé stavy. U analýzy provozu během dne se však přibližujeme hranicím toho, kde analýza dat začne narážet na právní překážky, zejména v podobě GDPR.

Z analýzy rozložení provozu během dne můžeme snadno zjisti odchylky od běžného stavu způsobené např. tím, že pracovník, který je za odesílání snímků zodpovědný, začal pracovat se zpožděním, měl delší obědovou pauzu a podobně. To je stav, který by mohl vést k řadě nepříjemností a zhoršování vztahů s uživateli systému. Účelem systému pro detekci neobvyklého provozu není zkoumat pracovní morálku zaměstnanců připojených institucí. Přesto má smysl se u velkých nemocnic zabývat i problémem rozložení provozu během dne. Při vyhodnocování analýzy je však třeba velké opatrnosti, protože zde lze očekávat určité množství anomálií, které mají přirozené vysvětlení a nepředstavují bezpečnostní hrozbu. Pokud např. laborant zodpovědný za odesílání snímků ráno zaspí, je pravděpodobné, že snímky, které během dne vznikly bude odesílat později, než obvykle a podle technických možností může být hustší provoz během dne, provoz i během obědové pauzy, nebo mohou být snímky odeslány později odpoledne nebo večer.

U velkých nemocnic je zajímavé sledovat i spektrum příjemců snímků, které daná nemocnice odesílá. Podrobný přehled spektra příjemců snímků je z důvodu velikosti tabulky uveden v příloze D. Pro analýzu spektra příjemců odesílaných snímků se jako nejvhodnější jeví použití entropie. Postupně jsem testoval několik matematických modelů postavených jak na tradiční Shonnonově entropii, tak i modely postavene na Rényiho nebo Tsallisově entropii s různými hodnotami parametru α, resp. q.

Jednotlivé nemocnice, které odesíllají obrazová data pomocí systému Redimed, používají různé způsoby práce: některé nemocnice mají stále spektrum partnerů,

kterým posílají snímky, jiné komunikují s širším spektrem partnerů, přičemž ale jednotlivým partnerům posílají jen malé množství snímků. V obou případech je možné použít pro analýzu spektra komunikujících partnerů Shannonovu entropii, případně i tradiční statistické metody.

Nejsložitější situace nastává u nemocnic, které mají širší spektrum komunikujících partnerů, přičemž jednomu či dvěma posílají obvukle větší množství dat, než ostatním. Občasné výkyvy v množství dat posílaných "větším"partnerů způsobují, že použití statistickým metod selhává z důvodu velkého rozptylu dat. Podobně je tomu i v případě Shannonovy entropie.

V tomto případě potřebujeme parametrizovatelné mdely entropie, které je možné přizpůsobit profilu provozu dané nemocnice. Po delsím testování Renyiho a Tsallisovy entropie jsem dospěl k závěru, vhodnější je použití Rényiho entropie. Parametr α je třeba přizpůsobit jak profilu provozu nemocnice, tak i počtu přenášených snímků.

Uvažujme množinu partnerů, kterým nemocnice odeslala snímky za jednotku času, v tomto případě jeden den. Pravděpodobnost jednotlivých stavů, tj. odeslání snímku danému příjemci můžeme popsat pomocí relativních četností. Vzorek provozu příkladové nemocnice za tři dny je v tabulce 6.11. Z analýzy záměrně vynechávám víkendový provoz, protože během víkendu je provo nimimální, někdy úplně nulový.

Při analýze struktury příjemců je třeba vzít v úvahu některé základní vlastnosti entropie:

- S rostoucím počtem komunikujících partnerů entropie obecne klesá.
- Entropie roste s rovnoměrností rozložení provozu mezi komunikující partnery.

Tyto vlastnosti je třeba v rámci konstrukce rozhodovacího kriteria vhodným způsobem vyrovnávat. Po mnoha experimentech, z nichž většina vedla k nepříliš uspokojivým výsledkům se podařilo optimalizovat kompenzace vlivu počtu málo frekventovaných příjemců následující modifikací Rényiho entropie:

$$H_{C\alpha}(S) = \frac{1}{1-\alpha} log_2(\sum_{i=1}^{n} p_i^{\alpha}) \frac{1}{log_2^{1,3}(n)}, \tag{6.3}$$

kde pro danou příkladovou nemocnici je experimentálně zjištěna optimální hodnota pramatru $\alpha = 0.8$. Korekčí činitel

$$\frac{1}{log_2^{1,3}(n)} \tag{6.4}$$

kompenzuje vliv počtu pozorování, tj. v tomto případě počtu komunikujících partnerů.

Vliv počtu komunikujících partnerů na hodnotu entropie můžeme snadno demonstrovat na jednoduchém příkladu:

Nemocnice posílá 8 snímků hlavnímu partnerovi a po jednom snímku jednomu až třem dalším partnerům. Hodnoty Rényiho entropie vypočtené pro hodnotu parametru $\alpha = 0.8$ v jednoduché a modifikované verzi s kompenzací popsanou vztahem 6.4 jsou v tabulce 6.10.

Hodnoty entropie pro vzorek provozu zachycený v tabulce 6.11 jsou shrnuty v tabulce 6.12. Příklad analýzy struktury komunikujících partnerů této nemocnice za celý měsíc je v příloze E.

Vypočtenou hodnotu entropie porovnáváme s horní a dolní prahovou hodnotou. Pro případ naší příkladové nemocnice je horní prahová hodnota 0,30 a dolní prahová hodnota hodnota 0,01. Překročení prahových hodnot ještě samo o sobě neznamená, že budem informovat uživatele o podezření na něco nekalého. Informaci o případném překročení prahových hodnot ještě kombinujeme s celkovým množstvím odeslaných snímků a množstvím komunikujících partnerů v daném časovém úseku. Mezní stavy, kdy např. v daný den odesíláme snímky jen jednomu z mála hlavních partnerů prostě jen z důvodu slabého provozu, není účelné se pokoušet zachytit matematickým výpočtem. Např. v případě entropických modelů bude v tomto případě entropie maximální možná. (Máme jen jednoprvkovou množinu stavů, tudíž pravděpodobnost výskytu daného jednoho stavu je 1.) Podobně je tomu v případě, že z nějakého důvodu v daný den neposíláme snímky většímu odběrateli. K tomu může být racionální důvod, např. externí radiolog, který pro danou nemocnici popisuje snímky, má dovolenou. V entropickém modelu se taková situace projeví maximalizací entropie.

Entropické modely nepoužíváme jen pro krásu této kapitoly matematiky. Účelem je odhalit neobvyklé rozložení struktury partnerů, kterým daná zdravotnická instituce odesílá snímky. Z neobvyklých stavů jsou významné jan takové, kdy by mohlo dojít k nežádoucímu úniku dat mimo zdravotnické zařízení. Tj. stav, kdy se nám objeví nový partner, kterému odesíláme větší množství snímků, nebo několik partnerů, kterým odesíláme středně velké množství snímků. Ostatní situace, které mohou vyvolat změnu entropie, nejsou z pohledu úniku dat relevantní a je třeba je detekovat jiným způsobem, abychom potlačili vznik falešných poplachů.

6.3 Největší uživatelé a speciální provoz

U největších uživatelů má smysl uvažovat o ještě podrobnější analýze dat v kratších časových intervalech, protože v tomto případě máme k dispozici dostatek dat k tomu, any příslušné metody produkovaly statisticky významné výsledky. V zásadě zde nepotřebujeme žádné další nebo nové metody oproti běžným velkým nemocnicím.

Speciálním provozem rozumíme například dedikované Redimed klienty, kterř slouží pro přeposílání dat do systému pro výpočet radiační zátěže pacientů. V tomto

případě se přeposílají hlavičky všech DICOM snímků, které v daném zdravotnickém zařízení vznikají. PRo výpočet radiační zátěže není nutné mít k dispozici celý snímek, všechny potřebné informace jsou obsaženy v DICOM hlavičce. Z pohledu analýzy datových toků odesílá příslušný Redimed klient data pouze jednomu příjemci a množství dat a jejich načasování závisí pouze na množství pacientů a na konfiguraci softwarového modulu, který data přeposílá.

V průběhu času se charakter provozu některých klientů může vyvíjet a je pravděpodobné, že budou vznika další aplikace, které komuniční systém Redimed využijí. Proto je třeba systém detekce neobvyklých stavů koncipovat jako otevřený a počítat s možností rozšíření o další speciální komponenty, které budou vhodné pro kontrolu právě těchto speciálních aplikací.

V případě speciálních služeb může být pro analýzu zajímavý i objem přenášených dat. Např. u měření radiační zátěže se posílají pouze DICOM hlavičky. Hlavičky sice mohou mít různou velikost, stáje jsou však výrazně menší, než celé snímky. U nových aplikací se dá očekávat podobně specifický profil objemu dat, nebo poměru obejmu dat a počtu studií.

Tab. 6.8: Relativní přírůstky počtu odesílaných snímků příkladových nemocnic.

Den	**Nemocince A**	**Nemocnice B**	**Nemocnice C**
2018-11-08	-101.71%	-139.16%	155.56%
2018-11-09	100.00%	92.65%	-71.59%
2018-11-10	-176.47%	-112.59%	-84.62%
2018-11-11	6.09%	-4.83%	-14.89%
2018-11-12	206.96%	359.59%	65.63%
2018-11-13	-100.00%	-207.14%	47.30%
2018-11-14	113.08%	-3.98%	0.00%
2018-11-15	-121.05%	0.00%	-103.41%
2018-11-16	10.77%	-8.59%	292.54%
2018-11-17	-91.30%	-59.48%	-261.45%
2018-11-18	6.09%	41.18%	0.00%
2018-11-19	140.00%	103.07%	354.55%
2018-11-20	-40.38%	-63.11%	-140.00%
2018-11-21	19.63%	82.03%	-6.54%
2018-11-22	22.83%	-74.67%	6.48%
2018-11-23	-13.59%	-22.44%	-11.38%
2018-11-24	-127.27%	-61.64%	-97.22%
2018-11-25	0.00%	43.48%	-6.42%
2018-11-26	61.95%	93.90%	96.33%
2018-11-27	27.72%	-55.15%	-16.47%
2018-11-28	-49.49%	4.27%	52.50%
2018-11-29	96.55%	-18.79%	8.54%
2018-11-30	125.00%	-4.79%	-91.67%

Tab. 6.9: Analýza denního rozložení provozu odesílaných snímků.

Hodina	**Počet odeslaných studií**
2019-08-19 07:00:00	9
2019-08-19 10:00:00	3
2019-08-19 11:00:00	5
2019-08-20 05:00:00	1
2019-08-20 06:00:00	8
2019-08-20 07:00:00	8
2019-08-20 08:00:00	1
2019-08-20 09:00:00	2
2019-08-20 11:00:00	15
2019-08-21 06:00:00	5
2019-08-21 07:00:00	4
2019-08-21 09:00:00	10
2019-08-21 10:00:00	5
2019-08-22 06:00:00	5
2019-08-22 11:00:00	4
2019-08-22 15:00:00	1

Tab. 6.10: Příklad vlivu počtu komunikujících partnerů na hodnotu entropie.

Počet partnerů	H_α	$H_{C\alpha}$
8+1	0,572	0,572
8+1+1	1,030	0,566
8+1+1+1	1,409	0,572

Tab. 6.11: Vzorek struktury odesílaných snímků a komunikujících partnerů příkladové nemocnice.

Datum	**Příjemce**	**Objem dat**	**Počet odeslaných snímků**
2018-11-01	M3756	5962784920	18
2018-11-01	M3738	2260845823	10
2018-11-01	M3722	567571326	3
2018-11-01	M3228	369526528	1
2018-11-01	M3712	218431813	1
2018-11-01	M3460	201452313	2
2018-11-02	M3738	697737676	8
2018-11-02	M3700	31231301	1
2018-11-03	M3738	1309182904	4
2018-11-04	M3738	396473569	1
2018-11-05	M3756	935890724	3
2018-11-05	M3738	304781668	5
2018-11-05	M3687	19407822	1
2018-11-05	M3405	11868972	1
2018-11-06	M3738	2399480094	8
2018-11-06	M3756	536652348	6
2018-11-06	M3094	235534067	2
2018-11-06	M3617	233844546	4
2018-11-06	M3014	14094502	5

Tab. 6.12: Entropie vzorku provozu zachyceného v tabulce 6.11

Datum	**Počet odeslaných snímků**	**Relativní četnost**	**Entropie**
2018-11-01	18	0.514	
	10	0.286	
	3	0.086	
	1	0.029	
	1	0.029	
	2	0.057	
			0.044
2018-11-02	8	0.889	
	1	0.111	
			0.036
2018-11-05	3	0.300	
	5	0.500	
	1	0.100	
	1	0.100	
			0.159
2018-11-06	8	0.211	
	6	0.158	
	2	0.053	
	4	0.105	
	5	0.132	
	13	0.342	
			0.065

7 Závěr

S růstem popularity radiologického komunikačního systému Redimed a počtem jeho uživatelů roste i nebezpečí, že se v řadách uživatelů (resp. v případě větších nemocnic jejich zaměstnanců) najde někdo, kdo bude chtít tento systém zneužít pro neoprávněné kopírování lékařské dokumentace pacientů. Proto je zapotřebí systém Redimed vybavit automatizovanými nástroji pro odhalování nežádoucí komunikace dříve, než nastane reálný pokus o zneužití tohoto systému.

Existuje široká škála matematických nástrojů, které jsou vhodné k analýze množství přenášených obrazových studií a odhalování neobvyklých datových toků. Žádný z těchto nástrojů však nedokáže pokrýt celou šíři způsobů, jakými uživatelé systém Redimed využívají. Teprve kombinací několika metod a postupů je možné vytvořit systém, který by přiměřeně citlivě reagoval na neobvyklé situace a přitom nevyvolával více než malé množství falešných poplachů. Falešné poplachy otupují pozornost lidské obsluhy, která jediná dokáže signály generované automatickým systémem analýzy dat posoudit a rozlišit situace, které jsou skutečně problematické od neobvyklých, ale přitom legitimních stavů.

Jakýkoli automatický detekční systém (pokud nemá vyvolávat enormní množství falešných poplachů) má určitou minimální hladinu citlivosti a není možné jej použít pro detekci úniku dat v množství menším, než je tato hranice. Automatický detekční systém není možné použít pro odhalení neoprávněného odesílání jednotek medicínských obrazových studií, přesto však je účelné takové systémy vyvíjet a v praxi používat, protože mohou odhalit stavy, kdy by docházelo k masivnímu úniku dat a to jsou právě ty situace, které by měl být schopen odhalit provozovatel komunikačního systému.

Matematická analýza datových toků je jen jednou ze složek zabezpečení komunikačního systému. Řeší jen jednu třídu možných útoků a neměl by odvést pozornost od zabezpečení dalších prvků celého komunikačního řetězce. Na druhou stranu matematická analýza datových toků byť nezajišťuje přímo filtrování provozu, ale slouží jako signalizace neobvyklých stavů, patří k těm nejzajímavějším metodám detekce útoků na komunikační systém. Umožňuje reagovat nejen na dosud známé typu útoků, ale upozorní i na nové typy útoků, pokud při nich dochází k úniku většího objemu dat. To jsou právě ty útoky, které jsou v případě práce s medicínskými informacemi ty nejkritičtější.

Matematické metody detekce neovyklých datových toků je třeba neustále vyvíjet a přizpůsobovat přirozeným změnám chování uživatelů komunikačního systému, růstu objemu přenášených dat, zapojování dalších uživatelů do systému a vzniku a nasazování nových aplikací. Na druhou stranu je zde i dostatečný prostor pro další vývoj. Stále je možné zpřesňovat hranici toho, kdy je už vhodné datový tok

považovat ze neobvyklý a informovat lidskou obsluhu.

Vyhledávání neobvyklých datových toků, které by mohly znamenat napadení systému vnějším či vnitřním nepřítelem je nikdy nekončící proces. V okamžiku, kdy útočník s dostatečným telekomunikačním vzděláním, zjistí, jak přesně detekční systém funguje, může upravit model útoku tak, aby zabránil nebo alespoň výrazně ztížil detekci útoku. Na druhou stranu zdokonalování metod detekce vede ke zmenšování objemu dat, které dokáže útočník získat aniž by byl odhalen.

Literatura

[1] URL <https://mathonline.fme.vutbr.cz/default.aspx>

[2] URL <http://www.unodc.org/documents/organized-crime/UNODC_CCPCJ_EG.4_2013/CYBERCRIME_STUDY_210213.pdf?source=post_page---------------------------->

[3] The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years.
URL <https://eugdpr.org/>

[4] Lex Access to European Union law.
URL <https://eur-lex.europa.eu/legal-content/CS/ALL/?uri=CELEX:32005F0222>

[5] Rady pro administrátory (RSS).
URL <https://csirt.cz/news/advice_for_admins/>

[6] směrnice č. 2013/40/EU o útocích na informační systémy a nahrazení rámcového rozhodnutí Rady 2005/222/SVV.
URL <https://eur-lex.europa.eu/>

[7] usnesení Nejvyššího soudu sp. zn. 3 Tdo 162/2014.
URL <http://www.nsoud.cz/>

[8] usnesení Nejvyššího soudu sp. zn. 5 Tdo 781/2017.
URL <http://www.nsoud.cz/>

[9] usnesení Nejvyššího soudu sp. zn. 7 Tdo 731/2015.
URL <http://www.nsoud.cz/>

[10] Úmluva č. 104/2013 Sb.m.s. o počítačové kriminalitě.
URL <https://aplikace.mvcr.cz/sbirka-zakonu/>

[11] 2016.
URL <http://www.hl7.org/implement/standards/index.cfm?ref=nav>

[12] 2019.
URL <http://dicom.nema.org/medical/dicom/current/output/html/part01.html>

[13] 2019.
URL <https://www.dicomstandard.org/current/>

[14] Historie národní sítě pro vědu, výzkum a vzdělávání. 2019.
URL <https://www.cesnet.cz/sdruzeni/dokumenty/historie-narodni-site-pro-vedu-vyzkum-a-vzdelavani/>

[15] Historie národní sítě pro vědu, výzkum a vzdělávání. 2019.
URL <https://www.cesnet.cz/sdruzeni/dokumenty/historie-narodni-site-pro-vedu-vyzkum-a-vzdelavani/>

[16] Topologie sítě CESNET2. 2019.
URL <https://www.cesnet.cz/sluzby/pripojeni/topologie/>

[17] Ales, Z.; David, B.: *Kyberkriminalita*. Wolters Kluwer, 2017.

[18] Amblard, P.-O.; Vignat, C.: A note on bounded entropies. *Physica A: Statistical Mechanics and its Applications*, roèník 365, è. 1, 2006: str. 50–56, doi:10.1016/j.physa.2006.01.002.

[19] Andel, J.: Matematická statistika. *SNTL/Alfa, Praha*, roèník 346, 1978.

[20] Bellia; L., P.: A Code-Based Approach to Unauthorized Access Under the Computer Fraud and Abuse Act. Apr 2017.
URL <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2955742>

[21] Choi, B.-Y.; Park, J.; Zhang, Z.-L.: Adaptive random sampling for traffic load measurement. In *IEEE International Conference on Communications, 2003. ICC'03.*, roèník 3, IEEE, 2003, s. 1552–1556.

[22] Craig, A. N.; Shackelford, S. J.; Hiller, J. S.: Proactive Cybersecurity: A Comparative Industry and Regulatory Analysis. *American Business Law Journal*, roèník 52, è. 4, 2015: str. 721–787, doi:10.1111/ablj.12055.

[23] Criado, F.; Gachechiladze, T.: Entropy of fuzzy events. *Fuzzy Sets and Systems*, roèník 88, è. 1, 1997: str. 99–106, doi:10.1016/s0165-0114(96)00073-5.

[24] Death, D.: *Information security handbook: develop a threat model and incident response strategy to build a strong information security framework*. Packt Publishing, 2017.

[25] Dostál, O.: Metropolitní archív medicínských obrazových informací. *Zpravodaj ÚVT MU*, roèník XII, è. 5, 2002.

[26] Dostál, O.; Filka, M.; Slavíček, K.: Brněnská akademická ATM síť. In *Sborník mez. konference COFAX*, Bratislava: DT Bratislava, 1998, ISBN 80-233-0405-4, s. 79–82.

[27] Dostál, O.; Filka, M.; Šárek, M.: Optická síť VŠ. In *Sborník referátů konference Optické komunikace - OK 94*, Praha, 1994.

[28] Dostál, O.; Javorník, M.; Slavíček, K.: Management of Interhospital Processing of Medical Multimedia Data. In *Contemporary Trends in Top Management Education: How to Accomodate Demand and Suplply*, Brno: B.I.B.S.,a.s., 2004, ISBN 80-86575-74-8, s. 70–70.

[29] Dostál, O.; Javorník, M.; Slavíček, K.: Opportunity of Current ICT in the Processing of Medical Image Information. In *In Proceedings of the International Conference International Association of Science and Technology for Development. Mexico: EASTED*, Mexico: The International Association of Science and Technology for Development, 2006, ISBN 0-88986-545-0, s. 193–195.

[30] Dostál, O.; Javorník, M.; Slavíček, K.; aj.: MEDIMED-Regional Centre for Archiving and Interhospital Exchange of Medicine Multimedia Data. In *Proceedings of the Second IASTED International Conference on Communications, Internet, and Information Technology*, Scottsdale, Arizona, USA: International Association of Science and Technology for Development- IASTED, 2003, ISBN 0-88986-398-9, s. 609–614.

[31] Dostál, O.; Javorník, M.; Slavíček, K.; aj.: Developement of Regional Centre for Medical Multimedia Data Processing. In *Communications, Internet, and Information Technology*, St. Thomas (USA): ACTA Press, 2004, ISBN 0-88986-445-4, s. 632–636.

[32] Dostál, O.; Javorník, M.; Slavíček, K.; aj.: Integration of Telemedicine Activities in the Czech Republic. In *4th International Conference on Innovations in Information Technology, Innovations '07. IEEE*, Dubai, United Arab Emirates: UAE University, 2007, ISBN 978-1-4244-1840-4, s. 532–536.

[33] Dostál, O.; Slavíček, K.: Wireless Technology in Medicine Applications. In *Personal Wireless Communications*, Praha: Springer Verlag, 2007, ISBN 978-0-387-74158-1, s. 316–324.

[34] Dostál, O.; Slavíček, K.; Javorník, M.: System for Effective Collaboration in the Area of Medical Imaging. In *International Conference on Advanced Computer Science and Information Systems*, Bali: Faculty of Computer Science, Universitas Indonesia, Depok, 16424, 2010, s. 207 – 212.

[35] Dostál, O.; Slavíček, K.; Javorník, M.; aj.: *ICT Systems Monitoring*. Saarbrucken: LAMBERT Academic Publishing, 2012, ISBN 978-3-8473-7231-8.

[36] Duffield, N.; Lund, C.; Thorup, M.: Properties and prediction of flow statistics from sampled packet streams. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, ACM, 2002, s. 159–171.

[37] Duffield, N.; Lund, C.; Thorup, M.; aj.: Estimating flow distributions from sampled flow statistics. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, ACM, 2003, s. 325–336.

[38] Duffield, N.; aj.: Sampling for passive internet measurement: A review. *Statistical Science*, roèník 19, è. 3, 2004: s. 472–498.

[39] Fan, J.-L.; Ma, Y.-L.: Some new fuzzy entropy formulas. *Fuzzy Sets and Systems*, roèník 128, è. 2, 2002: str. 277–284, doi:10.1016/s0165-0114(01)00127-0.

[40] Filka, M.: *Optoelektronika pro telekomunikace a informatiku*. Prof. Ing. Miloslav Filka, Csc. a kol., 2017.

[41] Filka, M.; Dostál, O.; Slavíček, K.: ATM síť Brněnských vysokých škol. In *Sborník přednášek celostátní konference s mezinárodní účastí TELEKOMUNIKACE 98*, Brno: VUT Brno, 1998, ISBN 80-214-1140-6, s. 35–37.

[42] Fortnow, L.: Kolmogorov complexity. *Aspects of Complexity*, 2001, doi:10.1515/9783110889178.73.

[43] Gold, S.; Cornwall, H.: *Hugo Cornwalls new hackers Handbook*. Century, 1990.

[44] Haven, T. D.; Gibson, W.; Gibson, W.; aj.: *Neuromancer*. Alpha-Comic Verlag, 1990.

[45] Hohn, N.; Veitch, D.: Inverting sampled traffic. *IEEE/ACM Transactions on Networking*, roèník 14, è. 1, 2006: s. 68–80.

[46] Huang, G.-S.: A New Fuzzy Entropy for Intuitionistic Fuzzy Sets. *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*, 2007, doi:10.1109/fskd.2007.76.

[47] Jaroslav, M.: *Matematicka statistika pro informatiky: urceno pro posl. fak. prirodoved.* SPN, 1987.

[48] Javorník, M.; Dostál, O.; Slavíček, K.: Regional Medical Imaging System. *World Academy of Science, Engineering and Technology*, roèník 7, 2011, ISSN 2010-376X.

[49] Javorník, M.; Dostál, O.; Slavíček, K.; aj.: Knowledge Management and Cost - Effectiveness in the Area of Medical Image Data. In *The 38th International Conference on Computers Industrial Engineering*, Beijing, China: University, Beijing,China, 2008, ISBN 978-7-121-07437-0, s. 883–887.

[50] Jiri, A.: *Statisticke metody*. Matfyzpress, 2007.

[51] Jung, J.; Paxson, V.; Berger, A. W.; aj.: Fast portscan detection using sequential hypothesis testing. In *IEEE Symposium on Security and Privacy, 2004. Proceedings. 2004*, IEEE, 2004, s. 211–225.

[52] Karel, Z.; Josef, S.: *Pravdepodobnost a matematicka statistika*. Matfyzpress, 2012.

[53] Keiner, L. E.; Yan, X.-H.: A neural network model for estimating sea surface chlorophyll and sediments from thematic mapper imagery. *Remote sensing of environment*, roèník 66, è. 2, 1998: s. 153–165.

[54] Kim, J.; Radhakrishnan, S.; Dhall, S. K.: Measurement and analysis of worm propagation on Internet network topology. In *Proceedings. 13th International Conference on Computer Communications and Networks (IEEE Cat. No. 04EX969)*, IEEE, 2004, s. 495–500.

[55] Klir, G. J.; Yuan, B. (editoøi): *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh*. Singapore: World Scientific, první vydání, 1996, ISBN 981-02-2422-2.

[56] Kolouch, J.: *CyberCrime*. CZ.NIC, z.s.p.o., 2016.

[57] Lakhina, A.; Crovella, M.; Diot, C.: Diagnosing network-wide traffic anomalies. In *ACM SIGCOMM computer communication review*, roèník 34, ACM, 2004, s. 219–230.

[58] Lakhina, A.; Crovella, M.; Diot, C.: Mining anomalies using traffic feature distributions. In *ACM SIGCOMM computer communication review*, roèník 35, ACM, 2005, s. 217–228.

[59] Li, M.; Vitányi, P.; aj.: *An introduction to Kolmogorov complexity and its applications*, roèník 3. Springer, 2008.

[60] Lindley, D. V.: Information Theory and Statistics. Solomon Kullback. New York: John Wiley and Sons, Inc.; London: Chapman and Hall, Ltd.; 1959. Pp. xvii, 395. $12.50. *Journal of the American Statistical Association*, roèník 54, è. 288, 1959: s. 825–827, doi:10.1080/01621459.1959.11691207, <https://doi.

org/10.1080/01621459.1959.11691207>.
URL <https://doi.org/10.1080/01621459.1959.11691207>

[61] Mai, J.; Sridharan, A.; Chuah, C.-N.; aj.: Impact of Packet Sampling on Portscan Detection. *IEEE J.Sel. A. Commun.*, roèník 24, è. 12, Prosinec 2006: s. 2285–2298, ISSN 0733-8716, doi:10.1109/JSAC.2006.884027.
URL <https://doi.org/10.1109/JSAC.2006.884027>

[62] Marie, S.; Zdenek, K.: *Pravdepodobnost a matematicka statistika*. PC-DIR, 1997.

[63] Martin, V.: *Pocitace a kriminalita: trestnepravni a kriminologicke aspekty*. Academia, 1989.

[64] Masarykova: VÝUKA A VÝZKUM V OBLASTI MEDICÍNSKÝCH OBRAZOVÝCH INFORMACÍ. 2019.
URL <https://www.medimed.cz/>

[65] Novák, V.; Pužmanová, R.; Slavíček, K.: *Czech DWDM National Research Network-Case Study*. USA: International Engineering Consortium, 2007, ISBN 1-931695-53-9, s. 351–353.

[66] Novák, V.; Slavíček, K.; Cihlář, J.; aj.: Designand Deployment of CESNET2 DWDM Core Network. In *CESNET Conference 2006*, Praha: CESNET,z.s.p.o., 2006, ISBN 80-239-6533-6, s. 43–53.

[67] Roček, A.; Javorník, M.; Slavíček, K.; aj.: Reversible Watermarking in Medical Imaging with Zero Distortion in ROI. In *Proceedings of 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2017)*, New York: IEEE, 2017, ISBN 978-1-5386-1911-7, s. 356–359, doi:http://dx.doi.org/10.1109/ICECS.2017.8292071.
URL <https://ieeexplore.ieee.org/document/8292071/>

[68] Roček, A.; Slavíček, K.; Dostál, O.; aj.: A new approach to fully-reversible watermarking in medical imaging with breakthrough visibility parameters. *Biomedical Signal Processing and Control*, roèník 29, 2016, ISSN 1746-8094, doi:http://dx.doi.org/10.1016/j.bspc.2016.05.005.
URL <https://doi.org/10.1016/j.bspc.2016.05.005>

[69] Roček, A.; Slavíček, K.; Javorník, M.: RONI Size and another Attributes of Representative Sample of Medical Images in Common Hospital Operation, Related to Securing by Watermarking Methods. In *International Conference on Image Processing, Production and Computer Science (ICIPCS'16)*, editace

P. O. M. M. Ahamed, London: URENG, 2016, ISBN 978-93-84422-62-2, s. 44–51.

[70] Schulte, U.: *Einfuehrung in Fuzzy Logik*. Muenchen: Franzis, 1993.

[71] Shannon, C. E.: A Mathematical Theory of Communication. *Bell System Technical Journal*, roèník 27, è. 3, 1948: s. 379–423, doi: 10.1002/j.1538-7305.1948.tb01338.x, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x>.
URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1948.tb01338.x>

[72] Slavíček, K.: Dark Fiber in Cesnet Backbone. *WSEAS Transactionson Communications*, roèník 5, 2006, ISSN 1109-2742.

[73] Slavíček, K.; Dostál, O.; Javorník, M.: Mathematical Processing of Event Logs. In *New Information and Multimedia Technologies NIMT - 2010*, Brno: VUT Brno, 2010, ISBN 978-80-214-4126-2, s. 58–61.

[74] Slavíček, K.; Dostál, O.; Javorník, M.: Mathematical processing of event logs from nerwork devices. In *2010 International Conference on Intelligent Network and Computing*, Chengdu: IEEE, 2010, ISBN 978-1-4244-8271-9.

[75] Slavíček, K.; Dostál, O.; Javorník, M.; aj.: MEDIMED - Regional Centre for Medicine Image Data Processing. In *Knowledge Discovery and Data Mining*, USA: IEEE Computer Society, 2010, ISBN 978-0-7695-3923-2, s. 310 – 313.

[76] Slavíček, K.; Javorník, M.; Dostál, O.: Technology backround of international collaboration on medicine multimedia knowledge base establishment. In *Proceedings of the 2nd WSEAS International Conference on COMPUTER ENGINEERING and APPLICATIONS(CEA'08)*, Acapulco, Mexico, January 25-27, 2008: Published by WSEAS Press, 2008, ISBN 978-960-6766-33-6, s. 137–142.

[77] Slavíček, K.; Javorník, M.; Dostál, O.: Redundancy in Processing of Medical Image Data. In *Fourth International Conference on Computer Sciences and Convergence Information Technology*, Seoul, Korea: IEEE Computer Society Conference Publishing Services, 2009, ISBN 978-1-4244-5244-6, s. 519–523.

[78] Slavíček, K.; Javorník, M.; Dostál, O.: Extension of the Shared Regional PACS CenterMeDiMed to Smaller Healthcare Institutions. In *The Eleventh International Conference on Networks*, editace P. L. T. G. I. Pozniak-Koszalka, Saint Gilles, Reunion Island: IARIA, 2012, ISBN 978-1-61208-183-0, s. 83–87.

[79] Slavíček, K.; Javorník, M.; Dostál, O.: *MEDIMED Shared Regional PACS Center*. Croatia: InTech, první vydání, 2013, ISBN 978-953-51-1102-3, s. 43–62. URL <http://dx.doi.org/10.5772/55904>

[80] Slavíček, K.; Ledvinka, J.; Javorník, M.; aj.: Mathematical Processing of Syslog Messages from Routers an Switches. In *Information and Automation for Sustainability*, Colombo: IEEE Catalog Number CFP0809B, 2008, ISBN 978-1-4244-2900-4, s. 463–468.

[81] Slavíček, K.; Novák, V.: Single Fiber Lines in CESNET Backbone. In *Proceedings of the WSEAS International Conferences ISCOCO'05*, Tenerife: WSEAS, 2005, ISBN 960-8457-39-4, s. 400–404.

[82] Slavíček, K.; Novák, V.: Fiber Optics Transport Infrastructure of Cesnet Backbone. In *Proceedings of 6th WSEAS International Conferenceon Applied Computer Science*, Tenerife,Spain: WSEAS, 2006, ISBN 960-8457-57-2, s. 323–328.

[83] Slavíček, K.; Novák, V.: Cesnet Backbone Transport Network. *WSEAS Transactionson Communications*, roèník 6, 2007, ISSN 1109-2742.

[84] Slavíček, K.; Novák, V.: Introduction of Alien Wavelength into Cesnet DWDM Backbone. In *Sixth International Conference on Information, Communications and Signal Processing*, Singapore: IEEE, 2007, ISBN 978-1-4244-0982-2, s. 977–981.

[85] Slavíček, K.; Novák, V.; Ledvinka, J.: CESNET Fiber Optics Transport Network. In *The Eight International Conference on Networks*, Gosier, Guadeloupe/France: IEEE, 2009, ISBN 978-0-7695-3552-4, s. 403–408.

[86] Slavíček, K.; Schindler, V.; Dostál, O.; aj.: Kalman filter improvement for gyroscopic mouse movement smoothing. In *IIE Int'l Proceedings of International Conference on Research in Science, Engineering and Technology*, editace P. S. Z. Thaweesak, Kuala Lumpur: International Institute of Engineers, 2013, ISBN 978-93-82242-47-5, s. 43–48.

[87] Theriault, K.; Vukelich, D.; Farell, W.; aj.: Network traffic analysis unsing behaviour-based clustering. Whitepaper, BBN Technologies.

[88] Vacca, J. R.: *Computer and information security handbook*. Morgan Kaufmann Publishers, an imprint of Elsevier, 2017.

[89] Vaclav, D.; Marie, H.: *Pravdepodobnost a matematicka statistika*. Karolinum, 2013.

[90] Vignat, C.; Plastino, A.; Plastino, A.: Correlated Gaussian systems exhibiting additive power-law entropies. *Physics Letters A*, roèník 354, è. 1-2, 2006: str. 27–30, doi:10.1016/j.physleta.2006.01.041.

[91] Voigt, J.: Stochastic operators, information, and entropy. *Comm. Math. Phys.*, roèník 81, è. 1, 1981: s. 31–38.
URL `<https://projecteuclid.org:443/euclid.cmp/1103920158>`

[92] Wagner, A.; Plattner, B.: Entropy based worm and anomaly detection in fast IP networks. In *14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise (WETICE'05)*, June 2005, s. 172–177, doi:10.1109/WETICE.2005.35.

[93] Xu, K.; Zhang, Z.-L.; Bhattacharyya, S.: Profiling internet backbone traffic: behavior models and applications. In *ACM SIGCOMM Computer Communication Review*, roèník 35, ACM, 2005, s. 169–180.

[94] Yager, R. (editor): *Fuzzy sets and possibility theory*. Pergamon Press, 1982.

[95] Yegneswaran, V.; Barford, P.; Ullrich, J.: Internet intrusions: Global characteristics and prevalence. *ACM SIGMETRICS Performance Evaluation Review*, roèník 31, è. 1, 2003: s. 138–147.

[96] Zadeh, L.: Fuzzy Sets as a Basis for Possibility. *Fuzzy Sets and Systems*, roèník 1, 1978: s. 3–28.

[97] Zadeh, L. A.: Fuzzy Logic. *Computer*, roèník 21, è. 4, Duben 1988.

[98] Zimmermann, H.-J. (editor): *Fuzzy Sets and Systems*, Amsterdam, 1993.

[99] ŠÁREK, M.: Brněnská akademická síť. *Zpravodaj ÚVT MU*, roèník IV, è. 2, 1993.

Přílohy

Seznam příloh

A Protokol DICOM

Protokol DICOM je sice obecně používaným protokolem v soudobé radiologii, nicméně mimo medicínské prostředí se neuplatňuje. V této příloze proto připomínám jak obecné principy, tak i nejdůležitější body definice protokolu DICOM. Na úvod si připomeňme základní pojmy:

- NIS/RIS Nemocniční informační systém/ Radiodiagnostický informační systém Počítačový systém určený pro ukládání, zpracování a využívání informací týkající se administrativy a klinických aspektů medicínských služeb v nemocnici resp. speciálně na radiologickém oddělení.
- PACS - Picture Archiving and Communication System. PACS řeší rozhraní pro jednotlivé typy akvizičních modalit, komunikaci s radiologickým informačním systémem, dlouhodobou archivaci obrazové informace, zobrazení informace na diagnostických prohlížecích stanicích a další aspekty zpracování medicínských obrazových dat.
- DICOM Digital Imaging Communication in Medicine. Světový komunikační standard pro výměnu medicínských obrazových dat.
- DICOM aplikační entita Zdroje dat pro PACS (akviziční modality), jako například počítačový tomograf, magnetická rezonance, ultrazvuk a pod. a zobrazovací stanice sloužící pro diagnostické účely, klinické účely případně další přístroje vybavené digitálním obrazovým rozhraním DICOM .

První koncepce digitální obrazové komunikace byla představena profesorem Heinzem U. Lemkem v roce 1979. První konference o technologii PACS se uskutečnila v roce 1982. V roce 1990 se konalo ve Francii setkání věnované PACS technologiím a sponzorované NATO. Výsledkem setkání byl vznik prvního projektu a realizace PACS pro armádu USA. Jednotlivé systémy se postupně nezávisle vyvíjely v Evropě, Asii i USA než došlo ke vzniku společného řešení založeného na standardu DICOM. Vznik standardu DICOM byl zásadním milníkem a podnětem pro rozvoj digitálního zpracování medicínských obrazových dat.

DICOM (Digital Imaging and Communication in Medicine) je mezinárodní průmyslový standard popisující formát dat a komunikaci v oblasti zpracování medicínských obrazových informací. Tento standard je vyvíjený společným úsilím ACR (American College of Radiology) a NEMA (National Electrical Manufactures Association) ve spolupráci s dalšími standardizačními organizacemi jako jsou IEEE, HL7 či ANSI. První verze 1.0 tohoto standardu byla publikována v roce 1985 (ACR-NEMA Standard Publication No. 300-1985), aktuálně používaná verze 3.0 byla vydána v roce 1993. Standardnizační skupina neplánuje vznik dalších verzi standardu, nicméně tento standard kontinuálně doplňuje o definice nových zažízení a služeb [13].

Standard DICOM definuje:

- Sadu protokolů pro síťovou komunikaci
- Syntaxi a sémantiku příkazů a přídavných informací, které můžou být vyměňovány s využitím těchto protokolů.
- Sadu služeb pro komunikaci s paměťovými médii
- Způsob ověření shody zařízení se standardem (Conformance tests) a prohlášení o shodě zařízení se standardem.

Standard je koncipován jako strukturovaný dokument. Zavádí explicitně tzv. informační objekty, pro obrázky, grafiku, křivky, reporty, tisky, atd. A specifikuje zavedenou techniku pro jednoznačnou identifikaci informačních objektů v rámci sítě. Samotný standard DICOM se skládá z mnoha druhů služeb, většinou vyžadujících přenos dat po síti. Základní částí Standardu DICOM je formalizace a zajištění síťové komunikace mezi entitami. Stamdard rovněž definuje typy a datové formáty pro ukládání obrazové a pacientské dokumentace na výměnných paměťových médiích. Byl vyvinut a optimalizován pro potřeby práce s diagnostickou obrazovou informací v radiologii, kardiologii a příbuzných disciplínách.

Na obrázku A.1 (převzato z [12]) je komunikační model standardu DICOM jak pro síťovou komunikaci (on-line), tak i komunikaci s paměťovými médii (off-line).

Standard DICOM sestává z šestnácti částí [13], značených PS 3.1 - PS 3.18 (části PS 3.9 a PS 3.13 byly vypuštěny):

PS 3.1 Úvod. Tato část popisuje historii a všeobecnou strukturu standardu.

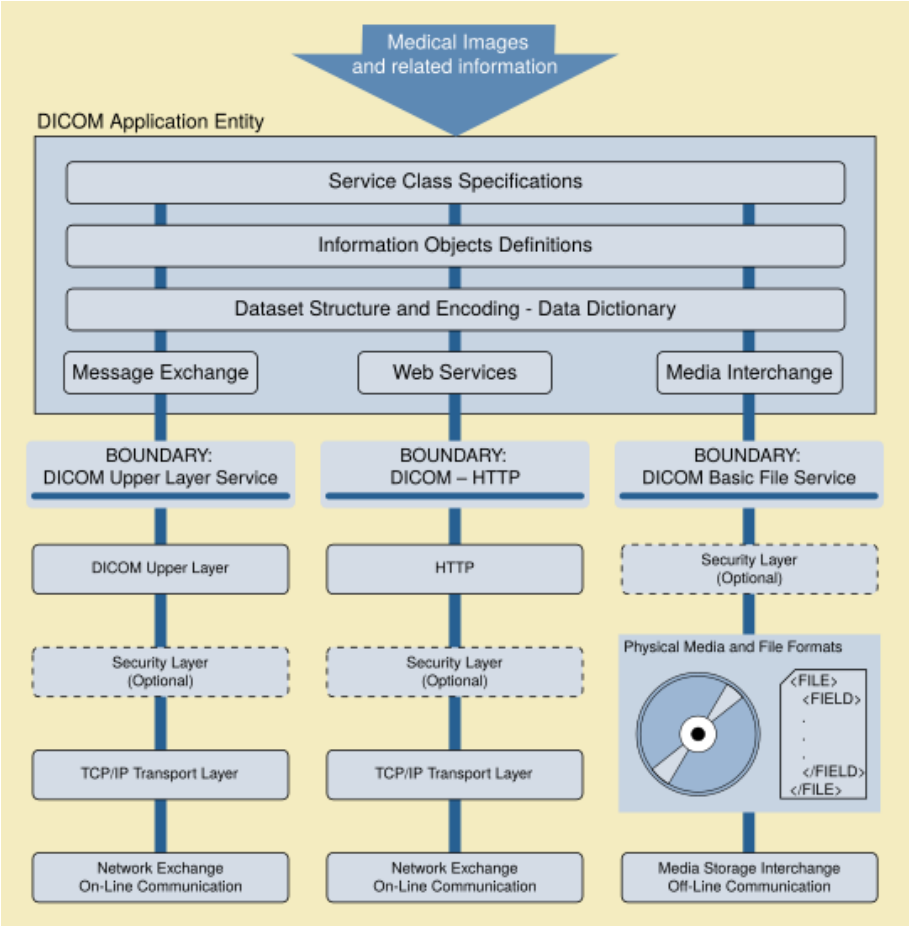
PS 3.2 Shoda. Zde jsou specifikovány všeobecné požadavky na implementaci vedoucí ke shodě se standardem, zahrnující funkce, příkazy, data. Tato část rovněž specifikuje strukturu dokumentu „Prohlášení o shodě“ (DICOM Conformance Statement). Specifikována je požadovaná struktura tohoto dokumentu a informace, kterou musí dokument obsahovat. Realizátor v tomto dokumentu specifikuje úroveň shody implementace se standardem.

PS 3.3 Definice informačních objektů. Zde je specifikována struktura definice informačních objektů (Information Object Definition) a jejich atributy. Každá IOD je abstraktní definicí nezbytných informací pro popis IOD, vyjadřuje vztah k reálným objektům relevantním k IOD a atributy, které popisují charakteristiky IOD.

PS 3.4 Specifikace servisních tříd. Servisní třída specifikuje vztah jednotlivých informačních objektů a jednoho nebo více příkazů, které můžou být nad těmito informačními objekty vykonávány.

PS 3.5 Datové struktury a kódování. Tato část dokumentu specifikuje kódování obsahu datových zpráv, specifikaci kompresních technik, způsob jednoznačné identifikace informace apod.

PS 3.6 Slovník dat. Tato část je registrem tzv. DICOM datových elementů. Datový



Obr. A.1: Komunikační model standardu DICOM. Převzato z [12].

element představuje elementární informační jednotku, která je specifikována návěštím, které se skládá z čísla skupiny a čísla elementu v této skupině, jménem, typem hodnoty, hodnotovou multiplicitou.

PS 3.7 Výměna zpráv. Kapitola definuje služby a protokoly pro výměnu zpráv mezi aplikacemi, způsobu používání příkazů, případně posloupnosti příkazů v rámci DICOM komunikace.

PS 3.8 Podpora síťové komunikace pro výměnu zpráv. Tato kapitola definuje služby a protokoly využívané pro výměnu zpráv přímo přes sítové prostředí mezi DICOM aplikacemi. Specifikuje komunikační služby a protokoly nejvyšší komunikační vrstvy pro komunikaci mezi DICOM aplikacemi.

PS 3.10 Paměťová média a formáty souborů. Tato část specifikuje všeobecný model pro ukládání obrazových dat na různé typy výměnných médií

PS 3.11 Aplikační profily paměťových médií. V rámci aplikačního profilu je definována specifická podmnožina aplikací DICOM Standardu pro výměnu obrazových informací s využitím paměťových médií.

PS 3.12 Formáty médií a fyzická média pro výměnu dat. Tato kapitola se týká podpory

výměny informací mezi medicínskými aplikacemi. Popisuje vzájemné vztahy mezi obecným modelem paměťového média a formátem specifických fyzických médií.

PS 3.14 Zobrazovací funkce standardní stupnice šedi. Zde jsou specifikovány standardizované zobrazovací funkce potřebné pro zobrazování obrazové informace založené na stupnici šedi. Cílem je zajistit konzistenci v prezentaci obrazů na různých médiích (displeje, tiskárny apod.) Popisuje exaktní funkce pro kalibraci zobrazovacích systémů pro prezentaci obrazů.

PS 3.15 Bezpečnost a profily management systému. Obsahem kapitoly je specifikace základních bezpečnostních pravidel, která musí obsahovat implementace pro dosažení shody aplikace se standardem. Zahrnuje problematiku kryptovacích schémat, veřejných klíčů a smart karet.

PS 3.16 Mapování obsahových zdrojů. V této části dokumentu jsou specifikovány vzory strukturovaných dokumentů jako DICOM informačních objektů, množinu kódovaných termínů používaných informačními objekty a překlady kódovaných termínů specifických pro jednotlivé země.

PS 3.17 Vysvětlující informace. Jsou zde vysvětleny jednotlivé pojmy používané ve standardu.

PS 3.18 DICOM standard specifikuje prostředky pro Webovský přístup k DICOM objektům (WADO). Popisuje způsob přístupu k DICOM objektům pomocí protokolu HTTP.

Standard DICOM definuje protokoly (protocols), objekty (objects), služby (service) a požadavky na soulad s normou (conformance requirements). Informační objekty (Information Object Definitions IOD) vyjadřují abstraktní definici nezbytných informací pro popis IOD, včetně vztahu k reálným objektům a obsahují atributy, které charakterizují IOD. Základními typy objektů jsou Image - Snímek, Series - Série (např. skupina snímků generovaná jednou modalitou), Study - Obrazová studie. Obrazová studie obsahuje jednu nebo více sérií vyšetření a to případně i na několika modalitách.

Dalšími nezbytnými atributy informačních objektů (IOD) jsou jednoznačná identifikace instance DICOM objektu (Instance UID), obrazové studie (study UID) a série obrazové studie (series UID Unique Identifier). V rámci standardu jsou definovány třídy služeb (Service Classes), specifikující typ operací nad objekty. Rozlišujeme operace typu COMPOSITE (Verification, Storage, Query/Retrieve, Study Content Notification) a NORMALIZED (Patient Management, Study Management, Results Management, Basic Print Management). Komplexní služby jsou vytvářeny pomocí elementů služeb, nazývaných DIMSEs (DICOM message service elements).

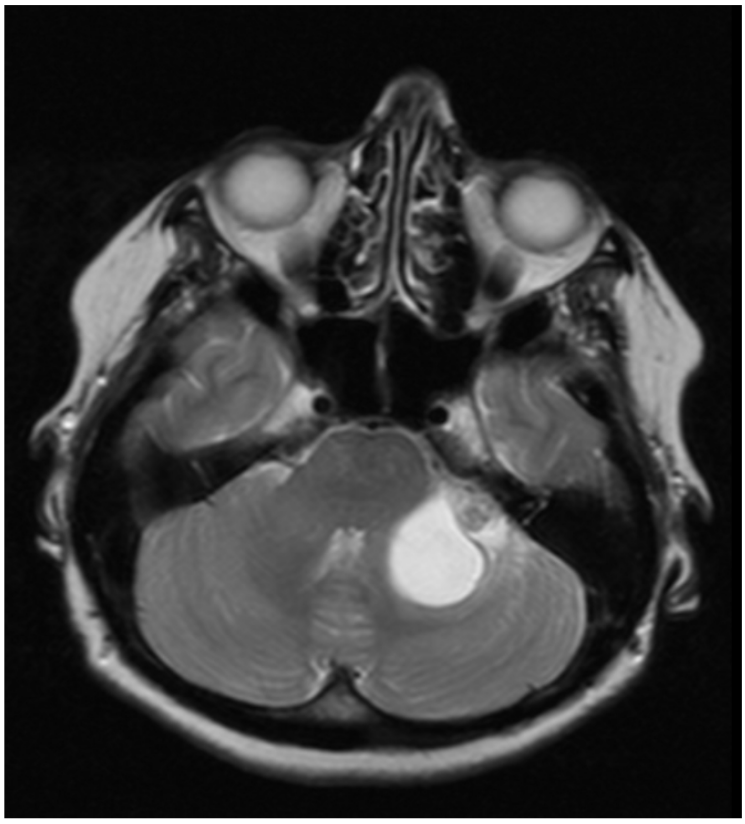
Servisní třídy a informační objekty vytvářejí spolu tzv. service object pair classes (SOP). SOP pak vyjadřují třídu operací nad informačním objektem. SOP třída

je elementární jednotkou DICOM. Na obrázku A.2 je příklad DICOM hlavičky (Header), za ní následují vlastní obrazová data. DICOM obrazová data mohou být nekomprimovaná a nebo komprimovaná. Společné uložení obrazových a identifikačních údajů podstatně snižuje pravděpodobnost jejich vzájemné záměny nebo ztráty.

První část hlavičky obsahuje formátovací informace, kde jsou popsány rozměry obrazu, pomocné textové informace ke snímku, identifikace modality, na které snímek vznikl a její nastavení. Obrazová data jsou uložena ve stejném souboru za hlavičkou.

DICOM vyžaduje 128 bytovou preambuli. Za ní následuje písmena ´D´, ´I´, ´C´, ´M´. Následují informace organizované ve skupinách obsahující údaje o případné komprimaci dat, informace o fotometrické interpretaci (monochromatický snímek, barevný snímek, stupně šedi, korekce sytosti / jasu snímku, barevná paleta, atd.), hodnoty jasu a kontrastu a celá řada dalších údajů.. Velikost hlavičky kolísá podle množství informací, které hlavička obsahuje.

Tag ID	VR	VM	Length	Description	Value
(0002,0000)	UL	1	4	File Meta Information Group Length	194
(0002,0001)	OB	1	2	File Meta Information Version	00 01
(0002,0002)	UI	1	26	Media Storage SOP Class UID	1.2.840.10008.5.1.4.1.1.4
(0002,0003)	UI	1	56	Media Storage SOP Instance UID	1.3.12.2.1107.5.2.6.14061.30000008111006092878100000883
(0002,0010)	UI	1	20	Transfer Syntax UID	Explicit VR Little Endian [1.2.840.10008.1.2.1]
(0002,0012)	UI	1	30	Implementation Class UID	1.3.6.1.4.1.20468.0.1.1.6.0.1
(0002,0013)	SH	1	8	Implementation Version Name	TMDTK160
(0008,0005)	CS	1	10	Specific Character Set	ISO_IR 100
(0008,0008)	CS	4	22	Image Type	ORIGINAL\PRIMARY\M\ND
(0008,0016)	UI	1	26	SOP Class UID	1.2.840.10008.5.1.4.1.1.4
(0008,0018)	UI	1	56	SOP Instance UID	1.3.12.2.1107.5.2.6.14061.30000008111006092878100000883
(0008,0020)	DA	1	8	Study Date	20081110
(0008,0021)	DA	1	8	Series Date	20081110
(0008,0022)	DA	1	8	Acquisition Date	20081110
(0008,0023)	DA	1	8	Content Date	20081110
(0008,0030)	TM	1	14	Study Time	094828.921000
(0008,0031)	TM	1	14	Series Time	095302.421000
(0008,0032)	TM	1	14	Acquisition Time	094910.890009
(0008,0033)	TM	1	14	Content Time	095302.812000
(0008,0050)	SH	1	10	Accession Number	2105/08-6
(0008,0060)	CS	1	2	Modality	MR
(0008,0070)	LO	1	8	Manufacturer	SIEMENS
(0008,0080)	LO	1	20	Institution Name	[redacted]
(0008,0081)	ST	1	36	Institution Address	
(0008,0090)	PN	0	0	Referring Physician's Name	
(0008,1010)	SH	1	6	Station Name	USAMR1
(0008,1030)	LO	1	12	Study Description	head^general
(0008,103E)	LO	1	14	Series Description	pd+t2_tse_tra
(0008,1050)	PN	1	10	Performing Physician's Name	[redacted]
(0008,1070)	PN	1	14	Operators' Name	
(0008,1090)	LO	1	8	Manufacturer's Model Name	Symphony
(0008,1140)	SQ	0	0	Referenced Image Sequence	
(FFFE,E000)		1	Undefined	Item	
(0008,1150)	UI	1	26	Referenced SOP Class UID	1.2.840.10008.5.1.4.1.1.4
(0008,1155)	UI	1	56	Referenced SOP Instance UID	1.3.12.2.1107.5.2.6.14061.30000008111006092878100000855
(FFFE,E00D)		0	0	Item Delimitation Item	
(FFFE,E000)		1	Undefined	Item	
(0008,1150)	UI	1	26	Referenced SOP Class UID	1.2.840.10008.5.1.4.1.1.4



Obr. A.2: Příklad struktury DICOM hlavičky a přenášené obrazové informace.

B Počty studií odeslaných jednotlivými nemocnicemi za rok 2018

V této příloze uvádím pro možnost zpětné kontroly či použití jiných metod vyhodnocení počty odeslaných studií a objem přenesených dat jednotlivých uživatelů systému Redimed v roce 2018. Jedná se o provoz za celý rok 2018, systémové identifikátory příjemců studií byly nahrazeny anonymníi identifikátory M1 - M198.

Tab. B.1: Počty studií a objem dat přenesených radiologickým komunikačním systémem Redimed v roce 2018.

Zdravotnické zařízení	**Počet odeslaných snímků**	**Objem přenesených dat**
M1	2234298194	68486
M2	317052754381	27729
M3	718845729366	27636
M4	142552910676	11227
M5	62727863807	10209
M6	160029600653	8350
M7	444382288854	8152
M8	521036864436	7113
M9	529589304601	6894
M10	760314299040	6157
M11	456570326050	5335
M12	939781487988	4975
M13	406625368245	4632
M14	696508842979	4321
M15	23613153370	4241
M16	427399212584	4115
M17	22757738680	4057
M18	401366936548	3906
M19	32412081812	3800
M20	148535474179	3595
M21	167023179748	3159
M22	401232354002	3155
M23	25393711578	2981
M24	287057991405	2897
M25	205730760812	2857

M26	102830595161	2643
M27	283536466021	2526
M28	434626670012	2325
M29	58610406334	2240
M30	356654983848	2083
M31	33660862460	2058
M32	135985444161	1906
M33	453277013552	1855
M34	31266715641	1848
M35	285333273826	1623
M36	10336331129	1603
M37	12536521852	1462
M38	42635635805	1458
M39	101428001870	1356
M40	221324173175	1319
M41	251557261090	1239
M42	295696534067	1183
M43	12601004498	952
M44	49848465981	944
M45	228692060606	891
M46	11164641214	795
M47	161939554654	764
M48	109383777663	760
M49	77667223652	721
M50	17685886862	710
M51	77232648651	703
M52	40454370066	651
M53	3943838750	496
M54	6411657885	489
M55	14127321276	408
M56	5575191038	398
M57	69479939946	374
M58	12466128033	346
M59	27593513845	318
M60	57083107504	310
M61	51885910257	310
M62	6667798776	303

M63	28356749609	271
M64	30094527599	260
M65	8519059287	251
M66	31265261603	228
M67	6472749357	216
M68	3894664421	208
M69	33273891256	204
M70	11433959812	191
M71	847656138	170
M72	4082979193	133
M73	821915750	131
M74	13854006481	126
M75	1143101583	124
M76	19643758057	115
M77	14834539702	100
M78	8084677721	99
M79	15847940749	97
M80	12022887927	89
M81	712677291	84
M82	1490816720	81
M83	34746568897	73
M84	11961677210	67
M85	6961615170	64
M86	10970789552	62
M87	13883306078	59
M88	5691642577	55
M89	4102135034	54
M90	10352331378	52
M91	309085155	52
M92	641618115	49
M93	618160596	49
M94	990096584	48
M95	3043735958	47
M96	20139932669	46
M97	11041299097	46
M98	6289098349	45
M99	4911911189	45

M100	4456837318	45
M101	1972138426	43
M102	181191371	41
M103	7489150712	35
M104	7196885385	33
M105	2604621207	33
M106	1930064946	33
M107	2075080620	32
M108	7268882747	31
M109	1895377270	30
M110	5248670260	28
M111	2181964340	28
M112	1549411106	27
M113	4458424596	26
M114	2793176739	25
M115	17214615501	21
M116	3065620718	20
M117	1429537533	20
M118	224750549	20
M119	8097213640	19
M120	2087361160	19
M121	1900031191	19
M122	1667374771	19
M123	5207857455	18
M124	3038017140	18
M125	2849700098	18
M126	175907171	18
M127	2454594033	16
M128	1556237176	15
M129	1288213858	15
M130	758974021	15
M131	3244135320	14
M132	954922338	14
M133	5491037395	13
M134	947087168	13
M135	3660564255	12
M136	629981424	12

M137	7737734934	11
M138	1826928398	11
M139	1264698544	11
M140	634158244	11
M141	2723130739	10
M142	1079417933	10
M143	907132571	10
M144	781848265	10
M145	60176536	10
M146	583752493	9
M147	1927249080	8
M148	1879963825	8
M149	56694207	8
M150	1428154632	7
M151	796746666	7
M152	725538609	7
M153	311744393	7
M154	3110324998	6
M155	777837409	6
M156	569173091	6
M157	498668937	6
M158	1879256	6
M159	1705126295	5
M160	366723124	5
M161	182600435	5
M162	734170925	4
M163	594239653	4
M164	514611048	4
M165	472400734	4
M166	138878965	4
M167	96895653	4
M168	5265758254	3
M169	1055868765	3
M170	869496226	3
M171	790323763	3
M172	726428264	3
M173	597167677	3

M174	27069407	3
M175	2165451446	2
M176	1097101409	2
M177	748641131	2
M178	679452741	2
M179	184001547	2
M180	191372	2
M181	962071784	1
M182	358576349	1
M183	356493393	1
M184	240481656	1
M185	99200938	1
M186	96228229	1
M187	63320438	1
M188	59280163	1
M189	37604738	1
M190	35610593	1
M191	34783156	1
M192	19851478	1
M193	15680126	1
M194	13548798	1
M195	10313285	1
M196	610323	1
M197	2638	1
M198	839	1

C Profil provozu polikliniky menšího města

V této příloze uvádím pro možnost zpětné kontroly či použití jiných metod vyhodnocení profil datového provozu polikliniky analyzované v kapitole 6.1. Jedná se o provoz za celý rok 2018, systémové identifikátory příjemců studií byly nahrazeny písmeny.

Poliklinika za sledované období odeslala celkem 208 studií o celkovém objemu necelé 4 GB dohromady sedmi partnerům.

Tab. C.1: Profil provozu polikliniky menšího města

Datum	**Příjemce**	**Objem dat**	**Počet snímků**
2018-09-05	B	89047313	1
2018-08-31	A	50636015	4
2018-08-31	B	17168899	2
2018-08-30	A	26475587	2
2018-08-28	A	32048744	3
2018-08-27	A	21478860	2
2018-08-09	A	38170412	2
2018-08-08	A	8400417	3
2018-08-08	B	4082357	1
2018-08-07	A	129266983	6
2018-07-31	A	25701989	1
2018-07-30	A	27889733	1
2018-07-25	A	21795813	1
2018-07-23	A	9974604	2
2018-07-18	A	14358878	1
2018-07-16	A	17651086	1
2018-07-12	A	4315598	1
2018-07-10	A	37433432	2
2018-07-09	C	16856878	1
2018-07-09	A	16856878	1
2018-07-04	A	16956206	1
2018-07-03	A	47711427	1
2018-06-28	A	19416204	2
2018-09-05	B	89047313	1
2018-08-31	A	50636015	4
2018-08-31	B	17168899	2

2018-08-30	A	26475587	2
2018-08-28	A	32048744	3
2018-08-27	A	21478860	2
2018-08-09	A	38170412	2
2018-08-08	A	8400417	3
2018-08-08	B	4082357	1
2018-08-07	A	129266983	6
2018-07-31	A	25701989	1
2018-07-30	A	27889733	1
2018-07-25	A	21795813	1
2018-07-23	A	9974604	2
2018-07-18	A	14358878	1
2018-07-16	A	17651086	1
2018-07-12	A	4315598	1
2018-07-10	A	37433432	2
2018-07-09	C	16856878	1
2018-07-09	A	16856878	1
2018-07-04	A	16956206	1
2018-07-03	A	47711427	1
2018-06-28	A	19416204	2
2018-09-05	B	89047313	1
2018-08-31	A	50636015	4
2018-08-31	B	17168899	2
2018-08-30	A	26475587	2
2018-08-28	A	32048744	3
2018-08-27	A	21478860	2
2018-08-09	A	38170412	2
2018-08-08	A	8400417	3
2018-08-08	B	4082357	1
2018-08-07	A	129266983	6
2018-07-31	A	25701989	1
2018-07-30	A	27889733	1
2018-07-25	A	21795813	1
2018-07-23	A	9974604	2
2018-07-18	A	14358878	1
2018-07-16	A	17651086	1
2018-07-12	A	4315598	1

2018-07-10	A	37433432	2
2018-07-09	C	16856878	1
2018-07-09	A	16856878	1
2018-07-04	A	16956206	1
2018-07-03	A	47711427	1
2018-06-28	A	19416204	2
2018-09-05	B	89047313	1
2018-08-31	A	50636015	4
2018-08-31	B	17168899	2
2018-08-30	A	26475587	2
2018-08-28	A	32048744	3
2018-08-27	A	21478860	2
2018-08-09	A	38170412	2
2018-08-08	A	8400417	3
2018-08-08	B	4082357	1
2018-08-07	A	129266983	6
2018-07-31	A	25701989	1
2018-07-30	A	27889733	1
2018-07-25	A	21795813	1
2018-07-23	A	9974604	2
2018-07-18	A	14358878	1
2018-07-16	A	17651086	1
2018-07-12	A	4315598	1
2018-07-10	A	37433432	2
2018-07-09	C	16856878	1
2018-07-09	A	16856878	1
2018-07-04	A	16956206	1
2018-07-03	A	47711427	1
2018-06-28	A	19416204	2
2018-09-05	B	89047313	1
2018-08-31	A	50636015	4
2018-08-31	B	17168899	2
2018-08-30	A	26475587	2
2018-08-28	A	32048744	3
2018-08-27	A	21478860	2
2018-08-09	A	38170412	2
2018-08-08	A	8400417	3

2018-08-08	B	4082357	1
2018-08-07	A	129266983	6
2018-07-31	A	25701989	1
2018-07-30	A	27889733	1
2018-07-25	A	21795813	1
2018-07-23	A	9974604	2
2018-07-18	A	14358878	1
2018-07-16	A	17651086	1
2018-07-12	A	4315598	1
2018-07-10	A	37433432	2
2018-07-09	C	16856878	1
2018-07-09	A	16856878	1
2018-07-04	A	16956206	1
2018-07-03	A	47711427	1
2018-06-28	A	19416204	2

D Měsíční přehled struktury příjemců snímků od příkladové nemocnice

V této příloze je kompletní přehled struktury příjemců snímků od příkladové nemocnice za jeden měsíc. Interní kódy označující příjemce byly nahrazeny anonymními.

Tab. D.1: Analýza struktury příjemců snímků.

Datum	**Příjemce**	**Objem přenášených dat**	**Počet odeslaných snímků**
2018-11-30	M3756	3738369591	9
2018-11-30	M3094	1014895682	7
2018-11-30	M3739	742537688	5
2018-11-30	M3738	724356476	10
2018-11-30	M3730	89412355	1
2018-11-30	M3623	24150731	1
2018-11-30	M2920	17271235	1
2018-11-30	M3014	12754051	2
2018-11-29	M3536	1209865505	4
2018-11-29	M3714	1205468257	4
2018-11-29	M3756	715564983	2
2018-11-29	M3738	538603733	4
2018-11-29	M3700	91174351	2
2018-11-29	M3730	89936787	1
2018-11-29	M3739	80186136	1
2018-11-29	M3641	42411094	1
2018-11-29	M3405	40414326	1
2018-11-29	M2879	8113073	1
2018-11-28	M3648	563561188	1
2018-11-28	M3739	363916406	1
2018-11-28	M3738	274256160	5
2018-11-28	M3700	47510498	2
2018-11-27	M3756	1435088705	7
2018-11-27	M3700	1302144580	6
2018-11-27	M3730	712356446	2
2018-11-27	M3014	9034126	1
2018-11-26	M3738	2944780139	9
2018-11-26	M3756	655730733	1

2018-11-26	M3405	178627498	1
2018-11-26	M3460	53975053	1
2018-11-25	M3687	47231613	2
2018-11-24	M3700	1690796637	2
2018-11-23	M3738	862607593	8
2018-11-23	M3756	635205563	1
2018-11-23	M3739	416410939	3
2018-11-23	M2879	264496428	3
2018-11-23	M3617	211557511	2
2018-11-23	M3604	52428704	3
2018-11-23	M3700	525335	1
2018-11-23	M3625	481893	1
2018-11-22	M3756	2784270583	8
2018-11-22	M3738	666353125	7
2018-11-22	M3724	372310950	1
2018-11-22	M3746	65472682	1
2018-11-22	M3734	38214358	1
2018-11-22	M3014	17656003	2
2018-11-22	M3460	17299635	1
2018-11-22	M3739	11123179	1
2018-11-22	M3089	7786254	1
2018-11-22	M3536	4062766	1
2018-11-21	M3738	2556715448	9
2018-11-21	M3756	1354583982	6
2018-11-21	M3739	1003780188	2
2018-11-21	M3700	123256468	2
2018-11-21	M3641	46768131	1
2018-11-21	M3014	525607	1
2018-11-20	M3756	519324413	6
2018-11-20	M3405	352930864	3
2018-11-20	M3228	352783984	3
2018-11-20	M3460	351786857	3
2018-11-20	M3724	218889209	2
2018-11-20	M3123	525383	1
2018-11-19	M3738	3133758705	13
2018-11-19	M3623	1410132314	5
2018-11-19	M3722	183834945	1

2018-11-19	M3641	62965567	1
2018-11-19	M3700	15596232	4
2018-11-18	M3730	133999930	1
2018-11-16	M3738	1864418650	7
2018-11-16	M3756	1768288459	4
2018-11-16	M3722	1167426214	1
2018-11-16	M2866	341571870	1
2018-11-16	M3730	27386943	1
2018-11-16	M3700	3504990	1
2018-11-15	M3617	1375271599	3
2018-11-15	M3756	1028270021	3
2018-11-15	M3094	845319723	1
2018-11-15	M3738	486726571	5
2018-11-15	M3700	55699936	1
2018-11-14	M3738	7091100034	24
2018-11-14	M3756	2610247298	4
2018-11-14	M3712	934177151	2
2018-11-14	M3739	278368225	1
2018-11-14	M3727	96473009	2
2018-11-14	M3123	7043598	1
2018-11-14	M3724	6933918	1
2018-11-14	M3641	2360830	1
2018-11-13	M3756	2873602128	11
2018-11-13	M3722	389838273	1
2018-11-13	M3687	99881689	1
2018-11-13	M3014	14399214	1
2018-11-13	M3739	525351	1
2018-11-12	M3738	3930658230	14
2018-11-12	M3756	3019572678	9
2018-11-12	M3572	620717305	2
2018-11-12	M3700	295821313	1
2018-11-12	M3641	236348777	5
2018-11-12	M3094	232051194	2
2018-11-12	M3712	135556013	1
2018-11-12	M3460	47010877	1
2018-11-11	M3405	193670842	1
2018-11-09	M3756	6060396136	14

2018-11-09	M3722	970494831	1
2018-11-09	M3623	865250112	4
2018-11-09	M3738	763483736	6
2018-11-09	M3712	117824505	3
2018-11-09	M3094	48966175	1
2018-11-09	M3700	26019132	1
2018-11-08	M3405	3063508290	2
2018-11-08	M3234	1036906053	1
2018-11-08	M3616	1036614892	1
2018-11-08	M3453	1035041580	1
2018-11-08	M3756	1009303960	5
2018-11-08	M3738	997325686	6
2018-11-07	M3738	4214242725	13
2018-11-07	M2871	3284953744	3
2018-11-07	M3405	1636560838	3
2018-11-07	M3756	1341034566	9
2018-11-07	M3739	93375867	1
2018-11-07	M3572	68496813	1
2018-11-07	M3746	40969537	1
2018-11-07	M3014	19964165	1
2018-11-07	M3700	18526508	1
2018-11-06	M3738	2399480094	8
2018-11-06	M3756	536652348	6
2018-11-06	M3094	235534067	2
2018-11-06	M3617	233844546	4
2018-11-06	M3014	14094502	5
2018-11-05	M3756	935890724	3
2018-11-05	M3738	304781668	5
2018-11-05	M3687	19407822	1
2018-11-05	M3405	11868972	1
2018-11-04	M3738	396473569	1
2018-11-03	M3738	1309182904	4
2018-11-02	M3738	697737676	8
2018-11-02	M3700	31231301	1
2018-11-01	M3756	5962784920	18
2018-11-01	M3738	2260845823	10
2018-11-01	M3722	567571326	3

2018-11-01	M3228	369526528	1
2018-11-01	M3712	218431813	1
2018-11-01	M3460	201452313	2

E Analýza struktury příjemců snímků od příkladové nemocnice

V této příloze je analýza struktury příjemců snímků od příkladové nemocnice za jeden měsíc.

Tabulka E.1 obsahuje počty odeslaných snímků a počty komunikujících partnerů v jednotlivých dnech, tabulka E.2 rozpad na jednotlivé partnery.

Výočet entropie spektra příjemců snímků v jednotlivých pracovních dnesh je uveden v tabulce E.3.

Tab. E.1: Počet odesílaných snímků a jejich příjemců u příkladové nemocnice.

Den	**Počet odeslaných snímků**	**Počet příjemců**
2018-11-01	35	6
2018-11-02	9	2
2018-11-03	4	1
2018-11-04	1	1
2018-11-05	10	4
2018-11-06	25	5
2018-11-07	33	9
2018-11-08	16	6
2018-11-09	30	7
2018-11-10	0	0
2018-11-11	1	1
2018-11-12	35	8
2018-11-13	15	5
2018-11-14	36	8
2018-11-15	13	5
2018-11-16	15	6
2018-11-17	0	0
2018-11-18	1	1
2018-11-19	24	5
2018-11-20	18	6
2018-11-21	21	6
2018-11-22	24	10
2018-11-23	22	8
2018-11-24	2	1

2018-11-25	2	1
2018-11-26	12	4
2018-11-27	16	4
2018-11-28	9	4
2018-11-29	21	10
2018-11-30	36	8

Tab. E.2: Rozpad počtu snímků na jednotlivé příjemce.

Den	**Příjemce**	**Počet příjemců**
2018-11-01	M3756	18
2018-11-01	M3738	10
2018-11-01	M3722	3
2018-11-01	M3228	1
2018-11-01	M3712	1
2018-11-01	M3460	2
2018-11-02	M3738	8
2018-11-02	M3700	1
2018-11-03	M3738	4
2018-11-04	M3738	1
2018-11-05	M3756	3
2018-11-05	M3738	5
2018-11-05	M3687	1
2018-11-05	M3405	1
2018-11-06	M3738	8
2018-11-06	M3756	6
2018-11-06	M3094	2
2018-11-06	M3617	4
2018-11-06	M3014	5
2018-11-07	M3738	13
2018-11-07	M2871	3
2018-11-07	M3405	3
2018-11-07	M3756	9
2018-11-07	M3739	1
2018-11-07	M3572	1
2018-11-07	M3746	1
2018-11-07	M3014	1

2018-11-07	M3700	1
2018-11-08	M3405	2
2018-11-08	M3234	1
2018-11-08	M3616	1
2018-11-08	M3453	1
2018-11-08	M3756	5
2018-11-08	M3738	6
2018-11-09	M3756	14
2018-11-09	M3722	1
2018-11-09	M3623	4
2018-11-09	M3738	6
2018-11-09	M3712	3
2018-11-09	M3094	1
2018-11-09	M3700	1
2018-11-11	M3405	1
2018-11-12	M3738	14
2018-11-12	M3756	9
2018-11-12	M3572	2
2018-11-12	M3700	1
2018-11-12	M3641	5
2018-11-12	M3094	2
2018-11-12	M3712	1
2018-11-12	M3460	1
2018-11-13	M3756	11
2018-11-13	M3722	1
2018-11-13	M3687	1
2018-11-13	M3014	1
2018-11-13	M3739	1
2018-11-14	M3738	24
2018-11-14	M3756	4
2018-11-14	M3712	2
2018-11-14	M3739	1
2018-11-14	M3727	2
2018-11-14	M3123	1
2018-11-14	M3724	1
2018-11-14	M3641	1
2018-11-15	M3617	3

2018-11-15	M3756	3
2018-11-15	M3094	1
2018-11-15	M3738	5
2018-11-15	M3700	1
2018-11-16	M3738	7
2018-11-16	M3756	4
2018-11-16	M3722	1
2018-11-16	M2866	1
2018-11-16	M3730	1
2018-11-16	M3700	1
2018-11-18	M3730	1
2018-11-19	M3738	13
2018-11-19	M3623	5
2018-11-19	M3722	1
2018-11-19	M3641	1
2018-11-19	M3700	4
2018-11-20	M3756	6
2018-11-20	M3405	3
2018-11-20	M3228	3
2018-11-20	M3460	3
2018-11-20	M3724	2
2018-11-20	M3123	1
2018-11-21	M3738	9
2018-11-21	M3756	6
2018-11-21	M3739	2
2018-11-21	M3700	2
2018-11-21	M3641	1
2018-11-21	M3014	1
2018-11-22	M3756	8
2018-11-22	M3738	7
2018-11-22	M3724	1
2018-11-22	M3746	1
2018-11-22	M3734	1
2018-11-22	M3014	2
2018-11-22	M3460	1
2018-11-22	M3739	1
2018-11-22	M3089	1

2018-11-22	M3536	1
2018-11-23	M3738	8
2018-11-23	M3756	1
2018-11-23	M3739	3
2018-11-23	M2879	3
2018-11-23	M3617	2
2018-11-23	M3604	3
2018-11-23	M3700	1
2018-11-23	M3625	1
2018-11-24	M3700	2
2018-11-25	M3687	2
2018-11-26	M3738	9
2018-11-26	M3756	1
2018-11-26	M3405	1
2018-11-26	M3460	1
2018-11-27	M3756	7
2018-11-27	M3700	6
2018-11-27	M3730	2
2018-11-27	M3014	1
2018-11-28	M3648	1
2018-11-28	M3739	1
2018-11-28	M3738	5
2018-11-28	M3700	2
2018-11-29	M3536	4
2018-11-29	M3714	4
2018-11-29	M3756	2
2018-11-29	M3738	4
2018-11-29	M3700	2
2018-11-29	M3730	1
2018-11-29	M3739	1
2018-11-29	M3641	1
2018-11-29	M3405	1
2018-11-29	M2879	1
2018-11-30	M3756	9
2018-11-30	M3094	7
2018-11-30	M3739	5
2018-11-30	M3738	10

2018-11-30	M3730	1
2018-11-30	M3623	1
2018-11-30	M2920	1
2018-11-30	M3014	2

Tab. E.3: Entropie spektra příjemců. Vypočítáno jen pro pracovní dny.

Datum	**Počet odeslaných snímků**	**Relativní četnost**	**Entropie**
11-1	18	0.514	
	10	0.286	
	3	0.086	
	1	0.029	
	1	0.029	
	2	0.057	
			0.044
11-2	8	0.889	
	1	0.111	
			0.036
11-5	3	0.300	
	5	0.500	
	1	0.100	
	1	0.100	
			0.159
11-6	8	0.211	
	6	0.158	
	2	0.053	
	4	0.105	
	5	0.132	
	13	0.342	
			0.065
11-7	13	0.394	
	3	0.091	
	3	0.091	
	9	0.273	
	1	0.030	

	1	0.030	
	1	0.030	
	1	0.030	
	1	0.030	
			0.075
11-8	2	0.125	
	1	0.063	
	1	0.063	
	1	0.063	
	5	0.313	
	6	0.375	
			0.141
11-9	14	0.467	
	1	0.033	
	4	0.133	
	6	0.200	
	3	0.100	
	1	0.033	
	1	0.033	
			0.063
11-12	14	0.400	
	9	0.257	
	2	0.057	
	1	0.029	
	5	0.143	
	2	0.057	
	1	0.029	
	1	0.029	
			0.067
11-13	11	0.733	
	1	0.067	
	1	0.067	
	1	0.067	
	1	0.067	
			0.050
11-14	24	0.667	

	4	0.111	
	2	0.056	
	1	0.028	
	2	0.056	
	1	0.028	
	1	0.028	
	1	0.028	
			0.029
11-15	3	0.231	
	3	0.231	
	1	0.077	
	5	0.385	
	1	0.077	
			0.169
11-16	7	0.467	
	4	0.267	
	1	0.067	
	1	0.067	
	1	0.067	
	1	0.067	
			0.125
11-19	13	0.542	
	5	0.208	
	1	0.042	
	1	0.042	
	4	0.167	
			0.059
11-20	6	0.333	
	3	0.167	
	3	0.167	
	3	0.167	
	2	0.111	
	1	0.056	
			0.147
11-21	9	0.429	
	6	0.286	

	2	0.095	
	2	0.095	
	1	0.048	
	1	0.048	
			0.095
11-22	8	0.333	
	7	0.292	
	1	0.042	
	1	0.042	
	1	0.042	
	2	0.083	
	1	0.042	
	1	0.042	
	1	0.042	
	1	0.042	
			0.125
11-23	8	0.364	
	1	0.045	
	3	0.136	
	3	0.136	
	2	0.091	
	3	0.136	
	1	0.045	
	1	0.045	
			0.121
11-26	9	0.750	
	1	0.083	
	1	0.083	
	1	0.083	
			0.058
11-27	7	0.438	
	6	0.375	
	2	0.125	
	1	0.063	
			0.100
11-28	1	0.111	

	1	0.111	
	5	0.556	
	2	0.222	
			0.160
11-29	4	0.190	
	4	0.190	
	2	0.095	
	4	0.190	
	2	0.095	
	1	0.048	
	1	0.048	
	1	0.048	
	1	0.048	
	1	0.048	
			0.207
11-30	9	0.250	
	7	0.194	
	5	0.139	
	10	0.278	
	1	0.028	
	1	0.028	
	1	0.028	
	2	0.056	
			0.086