

LATVIJAS UNIVERSITĀTE

BAKALAURA DARBS

RĪGA 2009

LATVIJAS UNIVERSITĀTE
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE
DATORIKAS NODAĻA

ĪSZIŅU SŪTĪŠANAS SISTĒMA “SMS.ID.LV”

BAKALaura DARBS

Autors: **Kirils Solovjovs**

Stud. apl. ks05020

Darba vadītājs: doc. Dr. mat. Paulis Ķikusts

RĪGA 2009

ANOTĀCIJA

Bakalaura darbā “Īsziņu sūtīšanas sistēma “sms.id.lv”” autors apraksta darbu ar šo sistēmu, sniedzot vēsturisku ieskatu risinājumu izvēlē, sistēmas izstrādē un ieviešanā. Darbā arī aprakstīts padarītais un plānotais, lai atjaunotu un uzlabotu sistēmas funkcionalitāti.

Sistēmas “sms.id.lv” būtiska īpašība ir iespēja sūtīt īsziņas, lietotājam neievadot grafiskos šifrus (CAPTCHA-testus), kas to ļauj lietot arī cilvēkiem ar īpašām vajadzībām, līdz ar to daļa uzmanības ir veltīta teksta simbolu atpazīšanai (OCR).

Autors piedāvā risinājumu, kā cilvēki ar pasliktinātu redzi spēj lietot bezmaksas īsziņu sūtīšana no Interneta. Darba mērķis ir atjaunot īsziņu sūtīšanas sistēmas “sms.id.lv” funkcionalitāti, modernizēt un attīstīt šo sistēmu, kā arī dokumentēt sistēmas izstrādi, no jauna veiktos darbus un iegūtos rezultātus.

Atslēgvārdi: īsziņas, Internets, CAPTCHA, sms.id.lv, pieejamība, OCR.

ABSTRACT

In bachelor thesis “SMS system “sms.id.lv”” author describes the work on the SMS system “sms.id.lv” providing a historical insight into the choice of solutions, development and implementation of the system. The thesis also describes completed and planned work to restore and improve the functionality of the system.

An essential feature of system “sms.id.lv” is the possibility to send text messages without the need for the user to solve a graphical puzzle (CAPTCHA). That permits people with special needs to use the system.

Author offers a solution for people with reduced vision to be able to send short mobile messages free of charge using Internet. The goal of the thesis is to restore the functionality of the system, upgrade and develop the system and document the development and the results obtained.

Keywords: SMS, Internet, CAPTCHA, sms.id.lv, availability, OCR.

SATURS

APZĪMĒJUMU SARAKSTS.....	1
IEVADS.....	3
1.TEORĒTISKĀ BĀZE UN SISTĒMAS IZSTRĀDE.....	5
1.1.Teorētiskā bāze.....	5
1.1.1.Pakalpojumi sniedzēju izvēle par labu grafiskajiem CAPTCHA-testiem.....	5
1.1.2.Iespāids uz lietotājiem.....	8
1.1.3.Piedāvātie risinājumi.....	9
1.2.Sistēmas izstrāde.....	10
1.2.1.Alternatīvu izpēte.....	10
1.2.2.Iemesli sistēmas izveidei.....	11
2.VĒSTURISKAIS IESKATS UN SISTĒMAS IEVIEŠANA.....	13
2.1.Datu bāze.....	16
2.1.1.CAPTCHA-testu tabulas.....	17
2.1.2.Vārtejas atļauju tabula “senders”.....	18
2.1.3.Īsziņu rindas tabulas.....	19
2.1.4.Telefona numuru un identifikatoru tabulas.....	20
2.1.5.Bloķēto konfigurāciju tabulas.....	21
2.1.6.Žurnāla tabulas.....	23
2.1.7.Starpniekserveru tabulas.....	25
2.2.Kodols.....	26
2.2.1.Kļūdu kodu tekstuāla atšifrēšana.....	27
2.2.2.Statiskā operatora piederības noteikšana.....	29
2.2.3.Īsziņu ierindošana.....	30
2.2.4.Bloķēto konfigurāciju testēšana.....	31
2.2.5.Automātiskā konfigurāciju bloķēšana.....	31
2.2.6.Reģistrēšana žurnālā.....	32
2.2.7.Citi uzdevumi.....	33
2.3.Maršrutēšana.....	33
2.4.Operatoru bloki.....	33
2.4.1.Operatora noteikšanas bloks.....	34
2.4.2.Paplašinātais bloks.....	35
2.5.Tīmekļa saskarne “sms.id.lv”.....	35
2.6.Vārteja.....	37
2.7.WAP serviss.....	37
2.8.SMTP serviss.....	38
2.9.IRC serviss.....	39
2.10.Automātiskie servisi.....	40
2.10.1.CAPTCHA-testu iegūšana un atpazīšana.....	40
2.10.2.Īsziņu rindas pārvaldnieks.....	41
2.10.3.Starpniekserveru datubāzes atjaunināšana.....	42
2.10.4.Rezerves kopēšana.....	43
2.11.Administrēšanas rīki.....	43
2.11.1.Bloķēto konfigurāciju administrēšanas panelis.....	44

2.11.2.Īsziņu plūsmas pārraudzības rīks.....	44
2.11.3.OCR pārraudzības rīks.....	45
2.11.4.Statistikas rīks.....	46
3.SISTĒMAS UZTURĒŠANA UN ATTĪSTĪŠANA.....	47
3.1.Mūsdienās pieejamās alternatīvas.....	47
3.1.1.“sms.lv – Bezmaksas SMS sūtīšana”.....	47
3.1.2.“esms.lv – tavs personīgais īsziņu centrs”.....	49
3.1.3.“smsbox.lv – SMS par brīvu”.....	50
3.1.4.“bezmaksassms.com – nosūtīt bezmaksas īsziņas uz Latviju”.....	52
3.1.5.Citas alternatīvas.....	53
3.2.Problēmu identificēšana.....	53
3.2.1.Operatoru bloki ir novecojuši.....	53
3.2.2.CAPTCHA-testu atpazīšana nefunkcionē.....	53
3.2.3.SMTP serviss nedarbojas.....	54
3.2.4.IRC serviss tērē pārāk daudz resursu.....	54
3.2.5.Citas nepilnības.....	54
3.3.Iespējamie attīstības scenāriji.....	54
3.3.1.Neatjaunot sistēmas darbību.....	55
3.3.2.Atjaunot sistēmas darbību caur maksas vārtejām.....	55
3.3.3.Atjaunot sistēmas darbību, bāzējot to uz operatoru e-pasta vārtejām.....	55
3.3.4.Atjaunot sistēmas darbību pēc vecā modeļa.....	55
3.4.Izstrādātie risinājumi.....	56
3.4.1.Operatoru bloki.....	56
3.4.2.CAPTCHA-testu atpazīšana.....	59
3.4.3.SMTP serviss.....	64
3.4.4.Citi darbi.....	66
3.5.Sistēmas testēšana.....	67
3.6.Iesāktie darbi.....	68
3.6.1.Operatoru bloki.....	68
3.6.2.CAPTCHA-testu apstrādes moduļi.....	69
3.8.Plānotie uzlabojumi.....	71
3.8.1.Iespēja personalizēt sistēmu.....	71
3.8.2.Midleti.....	71
3.8.3.Citi iespējamie uzlabojumi.....	72
4.REZULTĀTI.....	73
5.SECINĀJUMI.....	75
PATEICĪBAS.....	77
IZMANTOTĀ LITERATŪRA UN AVOTI.....	78
PIELIKUMU SARAKSTS.....	81
PIELIKUMS NR. 1: “VĀRTEJAS PIESLĒGUMA SKRIPTS”.....	82
PIELIKUMS NR. 2: “STATISTIKAS RĪKA PIRMKODS”.....	83
PIELIKUMS NR. 3: “SISTĒMAS TESTĒŠANAS REZULTĀTI”.....	84
PIELIKUMS NR. 4: “LIETOŠANAS STATISTIKA”.....	88

APZĪMĒJUMU SARAKSTS

sistēma – *šajā gadījumā* īsziņu sūtīšanas sistēma “sms.id.lv”

pārraudzība (monitoring) – process, kurā īpaši pilnvarots lietotājs pārbauda un/vai novēro sistēmas darbību

vārteja (gateway) – aparatūra vai programmatūra, kas nodrošina sasaisti starp atšķirīgiem tīkliem

OCR (Optical Character Recognition) – metode, ar kuras palīdzību iespējams optiski atpazīt simbolus attēlos, pārvēršot tos par tekstu

Tjūringa tests – tāds uzdevums, kuru spēj veikt cilvēks, bet datorprogramma nespēj; ar tā palīdzību var noskaidrot, vai uzdevuma veicējs ir cilvēks vai datorprogramma

CAPTCHA-tests – programma, kas spēj uzģenerēt Tjūringa testus un novērtēt atbildi uz tiem

HTML (HyperText Markup Language) – valoda, kas nosaka dokumenta saturu un struktūru (globālajā tīmeklī)

CSS (Cascading Style Sheets) – valoda, kas apraksta dokumenta izskatu un formātu (globālajā tīmeklī)

JavaScript – valoda, kas ļauj radīt dinamiskas globālā tīmekļa lapas, kontrolējot lapas uzvedību

PHP (PHP: Hypertext Preprocessor) – plaši izplatīta valoda, kas parasti tiek izmantota, lai programmētu globālā tīmekļa aplikācijas

GD – PHP bibliotēka, kas ļauj manipulēt ar dažāda formāta attēlu failiem

midlets (MIDlet) – valodā Java rakstīta programma, kas paredzēta lietošanai iegultajās ierīcēs (piemēram, mobilajos telefonos)

HTTP (HyperText Transfer Protocol) – protokols, kas nodrošina informācijas apmaiņu starp klientu un serveri globālajā tīmeklī

IRC (Internet Relay Chat) – reālā laika tekstuālās saziņas metode (protokols) Internetā, kas paredzēts grupveida saziņai

SMTP (Simple Mail Transfer Protocol) – protokols, kas paredzēts elektroniskā pasta pārsūtīšanai starp datoriem Internetā

SPF (Sender Policy Framework) – metode, kas ļauj norādīt datorus, kuriem atļauts izsūtīt elektroniskā pasta vēstules no noteikta e-pasta domēna, kā arī ļauj pārbaudīt, vai saņemtā elektroniskā pasta vēstule ir sūtīta no datora, kam atļauts to darīt

iesākumi (headers) – metadati, kas novietoti pirms attiecīgās informācijas un apraksta to

W3C (World Wide Web Consortium) – galvenā starptautiskā organizācija, kas atbildi par globālā tīmekļa standartizēšanu

starpniekserveris (proxy server) – serveris, kas loģiskajā savienojumā atrodas starp klienta datoru un serveri

maršrutētājs (router) – tīkla iekārta, kuras uzdevums ir nodrošināt datu pārraides maršruta izvēli

MiB(1) – mebibaits – 1048576 baiti

LMT – Latvijas Mobilais Telefons SIA, mobilo telefonu operators

TELE2 – Tele2 SIA, mobilo telefonu operators

Amigo – ZetCOM SIA, mobilo telefonu operators

BITE – Bite Latvija SIA, mobilo telefonu operators

Triatel – Telekom Baltija A/s, mobilo telefonu operators

OKarte – LMT veidota priekšapmaksas sarunu karte

Programmas koda fragments, teksta faila vai skripta saturs darbā tiks apzīmēts šādi.

Fails *dati.txt*:

Hello, world!

Loģiskajās shēmās tiks lietoti šādi apzīmējumi:

A —————> B "A ir B sastāvdaļa"

A ————— B "A ir saistīts ar B"

A - - - - -> B "A ir atkarīgs no B"

A <————> B "Starp A un B norisinās apbūsēja ziņojumu nodošana"
A —————> B "A nodod ziņojumus B"

A.1. att. Loģiskajās shēmās izmantoto simbolu atšifrējums

IEVADS

Mūsdienās globālā tīmekļa lapu izstrādē arvien populārāks kļūst modelis, kas atdala saturu no dizaina un darbībām (HTML + CSS + JavaScript), līdz ar to arī nedaudz palielinās šādu lapu pieejamība cilvēkiem ar īpašām vajadzībām (piemēram, traucējumiem uztvert krāsas vai rakstītu tekstu), taču to atsvēr arvien biežāks grafisko CAPTCHA-testu lietojums, kas neļauj šiem cilvēkiem piekļūt dažādiem pakalpojumiem. Viens no šādiem pakalpojumiem ir bezmaksas īsziņu sūtīšana no globālā tīmekļa lapām un Interneta vispār.

Bakalaura darbā “Īsziņu sūtīšanas sistēma “sms.id.lv”” autors apraksta savu vairāku gadu ilgo darbu pie īsziņu sūtīšanas sistēmas “sms.id.lv”, sniedzot vēsturisku ieskatu risinājumu izvēlē, sistēmas izstrādē un ieviešanā. Ņemot vērā to, ka bakalaura darba izstrādes **sākumposmā** sistēma diemžēl nav funkcionāla, autors veltīs atsevišķu nodaļu bakalaura darba izstrādes gaitā paveiktajam, lai atjaunotu sistēmas funkcionalitāti.

Sistēma “sms.id.lv” ir veidojusies vairāku gadu gaitā un ir piedzīvojusi vairākas pamatīgas restrukturizācijas. Šobrīd tā sastāv no sistēmas kodola, datu bāzes, mobilo sakaru operatoru vārteju blokiem, lietotāja saskarnes, pārraudzības programmatūras u.c. daļām.

Ņemot vērā sistēmas izmēru un ilgstošo darbību ir pašsaprotami, ka atsevišķas sistēmas daļas autors ir veidojis nevis individuāli, bet gan sadarbībā ar citām personām. Šāda sadarbība ir atbilstoši norādīta vietās, kur ir stāstīts par attiecīgajām sistēmas daļām. Ja sistēmas daļai nav īpaši minēti autori, tad saprotams, ka šo sistēmas daļu izstrādājis bakalaura darba autors viens pats.

Šī darba **pētāmā problēma** ir aprakstīt veidu, kā cilvēki, kas personisko īpašību vai tehnisku ierobežojumu dēļ nevar atrisināt grafiskos CAPTCHA-testus, spēj lietot bezmaksas īsziņu sūtīšana no Interneta.

Izvirzītā **hipotēze** – grafiskā CAPTCHA-testa vietā iespējams lietot arī lietotāju uzvedības analīzi.

Darba mērķis ir atjaunot īsziņu sūtīšanas sistēmas “sms.id.lv” funkcionalitāti, modernizēt un attīstīt šo sistēmu, kā arī dokumentēt sistēmas izstrādi, no jauna veiktos darbus un iegūtos rezultātus. Šajā darbā autors izmanto sekojošas **metodes mērķa sasniegšanai**:

- sistēmas līdzšinējās attīstības dokumentēšana;
- nefunkcionālo sistēmas daļu identificēšana ar testēšanas paņēmieni;
- alternatīvu izpēti teksta režīma pārlūkprogrammā;

- sistēmas nefunkcionālo daļu pārprogrammēšana un konfigurēšana;
- veikto uzlabojumu dokumentēšana.

Darbā izmantoti šādi **faktoloģiskā materiāla avoti**:

- mobilo operatoru lapās esošo grafisko CAPTCHA-testu paraugi;
- Internetā atrodamās starpniekserveru datubāzes;
- pārraudzības procesā iegūtie dati (statistika);
- sistēmas testi.

Ir jāņem vērā, ka sistēma nav komerciāls projekts – tam nav piešķirts finansējums, kā arī netiek gūti ienākumi no reklāmām, līdz ar to finansiālā ziņā projekts ir stipri ierobežots savā attīstībā. Tātad nav iespējams iegādāties nekādu aparatūru, kas speciāli piemeklēta sistēmas vajadzībām – jāaprobežojas vien ar programmatūras risinājumiem uz esošās aparatūras.

Bakalaura darbs ir sadalīts nodaļās. Darba sākumā tiek apskatīta teorētiskā bāze un sistēmas izstrādes principi, tajā skaitā uz to brīdi pieejamās alternatīvas, tad seko sistēmas ieviešanas vēsture, atsevišķās apakšnodaļās atdalot dažādās sistēmas daļas. Tam seko iespējamie attīstības scenāriji un uzlabojumu ieviešana, apskatot arī šobrīd pieejamās alternatīvas. Darbs tiek noslēgts ar rezultātiem un secinājumiem.

1. TEORĒTISKĀ BĀZE UN SISTĒMAS IZSTRĀDE

Šajā nodaļā autors apraksta darba teorētisko bāzi un izpēta situāciju, kas bija sistēmas ieviešanas brīdī – pirms vairākiem gadiem, kā arī analizē, kas autoru *toreiz* pamudināja uzsākt sistēmas izstrādi.

1.1. Teorētiskā bāze

Šajā apakšnodaļā autors apraksta darba teorētisko bāzi un atbild uz jautājumiem:

1. Kādēļ ir ieviesti CAPTCHA-testi?
2. Ko tie var sniegt pakalpojumu sniedzējam?
3. Kā tie var traucēt lietotājam?
4. Kādas alternatīvas iespējams atrast?

CAPTCHA-tests ir programma, kas spēj uzģenerēt testus un novērtēt atbildi uz tiem. Būtiska CAPTCHA-testa īpašība ir tā, ka lielākā daļa cilvēku uz to spēj sniegt pareizu atbildi, taču datorsistēmas nespēj.(9)

1.1.1. Pakalpojumi sniedzēju izvēle par labu grafiskajiem CAPTCHA-testiem

Kamēr globālais tīmeklis bija tikai vienvirziena komunikācijas kanāls (informācijas izdevēji publicēja, bet lasītāji lasīja tekstu, skatījās attēlus utt.), būtībā bija vienalga, kā informācijas satura ziņā uzvedas lasītājs, jo informācijas izdevēju vai citus lasītājus tas nekādā veidā nespēja ietekmēt.

Ātri vien globālais tīmeklis kļuva interaktīvs. Parādījās iespēja veidot atgriezenisko saiti, ko daudzi arī izmantoja – tika veidotas aptaujas, Interneta komentāru sistēmas un pat sarežģītāka tīmekļa programmatūra. Pagājušā gadsimta beigās lasītāji (jau *toreiz* “lietotāji”) sāka to izmantot ļaunprātīgi. Viņi rakstīja programmatūru, kas viņu vietā veica atkārtotu darbību, piemēram, balsoja aptaujās vai minēja lietotāju paroles globālajā tīmeklī. Saprotams, ka tas bija jāaptur un *toreiz* šķita, ka viens no lietotājus mazāk ierobežojošiem risinājumiem ir CAPTCHA-tests. Lai gan mūsdienā acīm (kad ir pieejami arī citādāki risinājumi) raugoties abas šīs problēmas varēja atrisināt arī citādi:

- paroļu minēšanu varēja atrisināt *arī* ar
 - a) sistēmas bloķēšanos uz dažām sekundēm pēc katra nepareizā paroles ievada;
 - b) IP adreses vai lietotāja bloķēšanu uz dienu vai ilgāk;

- aptauju rezultātu viltošanu varēja atrisināt *arī* ar
 - a) sīkdatņu palīdzību, atzīmējot kurš lietotājs vēl **nav** balsojis;

! Komentārs: nav jēgas ar sīkdatņu palīdzību atzīmēt, kurš **ir** balsojis, jo lietotājs sīkdatnes var izdzēst, turklāt daži pārlūki tās automātiski pēc kāda laika vai arī beidzot darbu izdzēš. Taču lietotāja marķēšana ar unikālu kodu **vēl pirms** aptaujas sākuma un šī koda atzīmēšana kā nederīga nobalsošanas brīdi varētu daļēji atrisināt problēmu.

- b) lietotāja uzvedības analīzi (vai lietotājs ir iegājis lapā tikai nobalsot, vai lietotājs veltīja laiku, lai iepazītos ar aptaujas jautājumu un izdomātu atbildi u.tml.)

Pakalpojumu sniedzējiem (pirms atgriezeniskās saites ieviešanas saukti arī par informācijas izdevējiem) patika CAPTCHA-testa ideja un vienkāršība. Tas universāli derēja jebkurai sistēmai. Pakalpojumu sniedzēji varēja tik vien kā pievienot lietotāja ievada formai papildus pārbaudes parametru un pirms tam uzģenerēt papildus testu. Vienalga vai tā bija balsošana, lietotāja pieteikšanās sistēmā vai kāds cits pakalpojums, ko tā sniedzējs vēlētos aizsargāt no tiem, kas nav cilvēki.(6)

Šādi testi tobrīd nodrošināja ļoti augstu izšķiršanas pakāpi starp cilvēku un datoru. *Moni Naor* 1996. gadā uzskaitīja vairākus Tjūringa testus, lai noteiktu, vai lietotājs ir cilvēks.(7) Vairākumu no viņa piedāvātajiem testiem autors nesauktu par CAPTCHA-testiem, jo nav iespējams automatizēti izveidot pāri – pareizo atbildi un uzdevumu:

- lietotājam tiek parādīts cilvēka sejas attēls, un viņam jāatbild, vai attēlotais ir vīrietis vai sieviete (šim testam ir divas iespējamās atbildes, no kurām viena ir pareiza);
- lietotājam tiek parādīts cilvēka sejas attēls, un viņam jāatbild, vai cilvēks smaida vai nē (šim testam ir divas iespējamās atbildes, no kurām viena ir pareiza);
- lietotājam tiek parādīts zīdītāja attēls un nosaukts ārējs orgāns, cilvēkam jāuzklikšķina uz attiecīgā orgāna attēlā (šim testam ir ļoti daudz iespējamo atbilžu, no kurām pareizas ir atbildes, kad lietotājs uzklikšķina uz vai netālu no apgabala, kurā redzams attiecīgs orgāns);
- lietotājam tiek parādītas vairākas bildes ar cilvēkiem, no kuriem vienā bildē attēlotais cilvēks ir kails, lietotāja uzdevums ir noteikt, kurš cilvēks ir kails (šim testam ir vairākas iespējamās atbildes, no kurām viena ir pareiza);

- lietotājam tiek parādīts stāstiņa teksts, un viņam jānosauc, ko šajā stāstiņā apzīmē vārds “tas”; piemēram, “Suns apgāza vāzi un uz galda izlija ūdens. Tagad tas ir slapjš.” (šim testam ir konkrētajā piemērā ir astoņas iespējamās atbildes, no kurām viena ir pareiza – konkrētajā piemērā tas ir “galds”).

Dažiem *Moni Naor* piedāvātajiem testiem gan ir iespējams pieņēmamā kvalitāte automatizēti izveidot uzdevuma un pareizās atbildes pāri, kas nozīmē, ka tie ir CAPTCHA-testi:

- lietotājam tiek parādīts no vienkāršiem poligoniem sastāvošs divu dimensiju zīmējums, un viņam jāizvēlas vārds, kas vislabāk nosauc zīmējumā attēloto, no dota piecu vārdu saraksta; lai ģenerētu šādus testus, iespējams radīt datu bāzi, kas apraksta, kādas īpašības piemīt konkrētu objektu zīmējumiem, un katrai no īpašībām piekārto zīmēšanas darbības ar noteiktu nejaušības pakāpi (šim testam ir piecas iespējamās atbildes, no kurām viena ir pareiza);
- lietotājam tiek parādīts ar roku rakstīts vārds, ko lietotājam ir jāspēj uzdrukāt; papildus iespējams attēlu papildināt ar tādiem trokšņiem, kas netraucē cilvēkam saskatīt vārdu; lai ģenerētu šādus testus nepieciešams ar datora palīdzību imitēt rokrakstam līdzīga teksta attēlu izveidi; šis princips ir pamatā lielai daļai mūsdienu grafisko CAPTCHA-testu (šim testam ir bezgalīgi daudz iespējamo atbilžu, no kurām viena ir pareiza);
- lietotājam tiek atskaņots skaņas ieraksts, kas satur vairākus vārdus un lietotājam tie ir jāpieraksta; šāda ieraksta radīšanai var lietot runas sintēzes tehniku, turpretī šī testa veiksmīgai iziešanai ar datoru var lietot teksta atpazīšanas tehnisku; skaņas atpazīšanas princips ir pamatā daudziem populāriem dzirdes CAPTCHA-testiem (šim testam ir bezgalīgi daudz iespējamo atbilžu, no kurām viena ir pareiza; citā variācijā var tikt piedāvāti atbilžu varianti, tādā veidā atvieglojot testa izpildi gan lietotājam, gan datoram);

Protams, visus trīs testus iespējams pielāgot tā, lai tiem būtu bezgalīgi daudz atbilžu varianti vai arī noteikts limitēts skaits atbilžu variantu. Tomēr, neiedziļinoties šajā faktorā, bet, balstoties pamatā uz vienkāršību ieviešanā un vidējā lietotāja ērtībām, sākotnēji izplatījās tāds CAPTCHA-tests, kurš funkcionēja pēc sekojošas sistēmas:

1. Pakalpojuma sniedzēja serveris no vārdnīcas (vai alfabēta) izvēlas kādu vārdu vai skaitli, iespējams, nomainot tajā kādu simbolu. Rezultējošā simbolu virkne tiek saglabāta vēlākām vajadzībām.

2. Šis vārds tiek sagatavots kā attēls, lietojot dažādus fontus.
3. Attēls tiek kropļots (piemēram, pielietojot nobīdes kartes, papildinot ar līnijām u.tml.)
4. Attēlam tiek pievienots dažāda veida troksnis un papildus traucējumi, ko cilvēkam atšifrēt nesagādā problēmas, taču datoram gan.
5. Attēls tiek izvadīts lietotājam uz ekrāna, un viņš tiek lūgts precīzi ievadīt attēloto tekstu.
6. Lietotāja ievadītais tiek salīdzināts ar saglabāto simbolu virkni. Ja šie dati sakrīt, tad tiek secināts, ka lietotājs ir cilvēks, pretējā gadījumā – dators.

1.1.2. Iespāids uz lietotājiem

Toreiz, kad CAPTCHA-testi tika ieviesti un pakāpeniski padarīti grūtāki, nebija aktuāli domāt par pieejamību, un, lai gan daži pacēla šo jautājumu gaismā, pieejamība tomēr palika novārtā.

Cilvēks var grafisko CAPTCHA-testu, kas aprakstīts iepriekšējās apakšnodaļas beigās, nespēt atrisināt vai spēt atrisināt ar grūtībām vairāku iemeslu dēļ, kas sadalāmi trijās grupās:

- Cilvēks ar īpašām vajadzībām savu fizisko trūkumu dēļ nespēj redzēt vai arī nespēj pilnvērtīgi redzēt. Šādā gadījumā viņš spēj lietot Internetu un globālo tīmekli, izmantojot īpašas programmas, kas visu saturu, kas nolasāms ar datoru, attēlo Braila rakstā vai skaņā, vai arī kā citādāk. Jāņem vērā, ka grafisko CAPTCHA-testu nav iespējams šādi attēlot, jo tā mērķis ir izveidot grafiska veida barjeru, lai to nespētu nolasīt datorprogrammas.
- Datortehnikas vai programmatūras ierobežojumu dēļ cilvēks neredz attēlus un citus vizuālos elementus, bet tikai tekstu (un papildus tam, bet ne obligāti – viņa pārlūkprogramma neapstrādā JavaScript kodu). Izrietošās sekas šeit būtībā ir tādas pašas kā iepriekšējai grupai, jo tieši tos pašus elementus “nerāda” arī pārlūkprogrammas, ko lieto neredzīgie vai vājredzīgie lietotāji.
- Cilvēkam nešķiet ērti vai lietderīgi katru reizi risināti CAPTCHA-testu. Šāda nostāja ir saprotama, jo tas ir monotons darbs, turklāt, gadiem ejot un datorprogrammām kļūstot arvien “gudrākām”, arī CAPTCHA-testi ir padarīti arvien sarežģītāki, un dažus nespēj atrisināt pat cilvēks. Vidēji viena CAPTCHA-testa izpildei cilvēkam nepieciešamas 10 sekundes.

Vienā dienā cilvēki visā pasaulē atrisina apmēra 0,2 miljardus CAPTCHA-testu, kopumā dienā tiek iztērēti 150 tūkstoši stundu, kuras varēja izlietot lietderīgāk.(10) Tā kā 150 tūkstošos stundu ir 18750 darba dienas, bet gadā ir apmēram 240 darba dienas, tas nozīmē, ka kopumā visā pasaulē katru gadu šai nodarbei tiek veltīti 78 gadi, ko varētu pavadīt strādājot pilnas slodzes darbu. Tas būtībā ir divu cilvēku ieguldījums savā darbavietā mūža garumā.

Šai problēmai tiek piedāvāts risinājums – *reCAPTCHA*, kas šo darbu izmanto lietderīgi, t.i. atpazīstot teksta fragmentus, ko pagaidām nespēj atpazīst standarta OCR metodes. *reCAPTCHA* darbojas, uzturot neatpazīto vārdu attēlu datu bāzi un lietotājam parādot divus vārdus – vienu no tiem (kontroles vārds) sistēma jau ir atpazinusi, bet otru ne. Lietotājam jāievada abi, un, ja kontroles vārds ir pareizs, tad sistēma pieņem, ka arī otrs vārds ir pareizs. Tādā veidā *reCAPTCHA* arvien papildina savu datu bāzi un atpazīst tekstu fragmentus, kas citādi nav atpazīstami.(10)

Tomēr autors uzskata, ka ir pielietojumi, kur var vispār iztikt bez CAPTCHA-testiem vai vismaz grafiskajiem CAPTCHA-testiem, un piedāvā tam alternatīvus risinājumus nākamajā apakšnodaļā. Līdzīgi spriež arī W3C darba grupa, kas iezīmē to, ka ir bieži vien grafiskos CAPTCHA-testus spēj atrisināt specializēta datorprogramma, taču cilvēki nevar.(11) Viņi arī piedāvā vairākus iespējamus risinājumus, no kuriem autora piedāvātajam risinājumam vistuvāk atbilst neinteraktīvās pārbaudes.

1.1.3. Piedāvātie risinājumi

Darba autora hipotēze ir, ka lielākajā daļā gadījumu ir iespējams panākt labākus rezultātus, analizējot lietotāja uzvedību un neprasot no viņa nekādu papildus mijiedarbību ar pakalpojuma sniedzēja sistēmu, un netērējot lietotāja laiku ar CAPTCHA-testiem. Tas iespējams, ja saņemto datu klāsts ir gana plašs.

Acīmredzams, ka, piemēram, veicot aptauju globālajā tīmeklī, var būt grūti savākt pietiekoši lielu datu kopumu, jo vienam lietotājam aptauju ir paredzēts aizpildīt vienu reizi. Vairums pieprasījumu aizpildīt aptauju izskatīsies līdzīgi vai pat vienādi. Tātad joprojām eksistē pielietojumi, kur CAPTCHA-testi var būt noderīgi, taču varētu diskutēt par iespējām iztikt ar tādiem CAPTCHA-testiem, kas nav grafiski.

Savukārt, ja ir kāda darbība, ko lietotājs var un drīkst veikt atkārtoti, piemēram, īsziņu sūtīšana, tad ir iespējams savākt pietiekoši lielu informācijas daudzumu par konkrētā lietotāja uzvedību vai pat visu sistēmas lietotāju uzvedību noteiktā laika momentā, lai šos datus varētu izmantot ļaunprātīgo lietotāju vai datorprogrammu bloķēšanai un garantētu sistēmas drošību.

Autors uzskata, ka nav pareizi un lietderīgi censties nodalīt datorprogrammas no cilvēkiem. Svarīgi ir tikai darbību, ko šis cilvēks vai datorprogramma veic, klasifikācija – vai tās ir “labas” vai “sliktas” darbības. Ja vien šāds dalījums ir iespējams, tad tas būtu viss, kas nepieciešams pakalpojumu sniedzējiem. Piemēram, ir svarīgi noskaidrot:

- vai lietotājs (datorprogramma) aptaujā nobalso tikai vienu reizi;
- vai lietotāja kontā kāds mēģina ielauzties (nav svarīgi vai tas ir cilvēks vai datorprogramma);
- vai kāds mēģina izsūtīt masveida reklāmu vai cita veida ziņojumus;
- utt.

Nākamajā apakšnodaļā autors apskatīs vajadzību veidot sistēmu, bet bakalaura darba 2. nodaļā pievērsīsies konkrētai realizācijai, ar kuras palīdzību ļaunprātīgas vai kaitnieciskas darbības tiek pamanītas un bloķētas sistēmā “sms.id.lv”.

1.2. Sistēmas izstrāde

Šajā apakšnodaļā autors pēta vajadzību veidot sistēmu, kā arī atbild uz jautājumiem:

1. Kādi līdzīgi pakalpojumi jau bija pieejami?
2. Kas autoru pamudināja veidot savu sistēmu?
3. Kādas bija sākotnējās sistēmas iespējas?

Nopietns darbs pie sistēmas norit kopš 2004. gada. Sistēmas iedīgļi radās jau 2000. gada rudenī.

1.2.1. Alternatīvu izpēte

Sistēmas izstrādes brīdi bija pieejami vairāki īsziņu sūtīšanas servisi, kas visi funkcionēja, saņemot datus no lietotāja un nosūtot tos tālāk operatoriem caur e-pasta vārtejām. Tobrīd Latvijā darbojās divi mobilo telefonu operatori – LMT, kā arī tikko ienākušais Baltkom GSM (tagad TELE2) – un pāris peidžeru sistēmas.

Pionieris īsziņu sūtīšanas no globālā tīmekļa jomā un populārākais no šāda veida servisiem bija uz servera “gobbo.caves.lv” izvietotais pakalpojums.⁽¹²⁾ Šī tīmekļa lapa ļāva nosūtīt īsziņas tikai uz peidžeru sistēmām un mobilo telefonu operatoru LMT. Tajā bija dažas reklāmas un, protams, nebija nekādu CAPTCHA-testu.

Ziņa uz LMT mobilo telefonu, Info un BaltkomPlus peidperiem







(Send Message(SMS) to LMT mobile phones, Info and BaltkomPlus pagers)
Pamēini saglabāt un atjaunot parametrus!!!

Ziņojums: (garums: 0)	Servisa piegādātājs: LMT GSM
	Numurs : 9XXXXXX
	Tavs vārds:
	Tavs e-mail
	(kur sūtīt apstiprinājumu)
	☐ Atsūtīt apstiprinājumu ☒ Atcerēties parametrus
Obligāti jāaizpilda: ziņojums, Numurs	
Atjaunot	ooo Sūtīt zinu ooo Izdzest

1.1. att. Tīmekļa lapas “<http://gobbo.caves.lv/~cikis/phone/all/>” fragments (12)

Ir grūti šobrīd spriest par sistēmas darbības kvalitāti, taču noderīgs papildinājums bija laika ziņu un vārda dienu pieejamība lietotājiem, kas vēlējās tās saņemt.

1.2.2. Iemesli sistēmas izveidei

Šajā apakšnodaļā autors pamato iemeslus, kāpēc viņš uzskatīja par vajadzīgu radīt savu īsziņu sūtīšanas sistēmu, un apraksta sistēmas sākotnējās iespējas.

Būtiski stimuli jaunas sistēmas radīšanai bija trīs:

- autora dzīvē pamazām ienāca mobilie sakari; lai gan autoram un viņa ģimenes locekļiem sākotnēji nebija personīgā mobilā telefona numura un mobilā telefona aparāta, tāds bija dažiem viņa draugiem, kā rezultātā pavērās iespēja izpētīt šo salīdzinoši jauno tehnoloģiju;
- laika ziņu un vārda dienu serviss no “gobbo.caves.lv” pārstāja funkcionēt, tika pamesta novārtā arī īsziņu sūtīšana, kas ātri vien izgāja no ierindas, līdz ar to bija nepieciešama kāda kvalitatīva alternatīva;
- autoru interesēja, kā šādas sistēmas var funkcionēt, kā arī tobrīd vienkāršotas īsziņu sūtīšanas (no globālā tīmekļa) sistēmas radīšana nebija sarežģīts process.

Jaunizveidotajai sistēmai tika uzstādītas dažas ļoti vienkāršas prasības attiecībā uz tās pieejamību un kvalitāti:

- sistēmai jāpiedāvā lietotājiem katru dienu saņemt laika ziņas un vārda dienas;

- sistēmai jānodrošinās pret svešu mobilo tālruņa numuru pierakstīšanu to numuru sarakstā, kas izteikuši vēlmi saņemt ziņas ikdienā;
- sistēmai jābūt automatizētai – gan reģistrācijas procesam, gan ikdienas ziņu izsūtīšanas procesam jānotiek bez cilvēka iejaukšanās;
- lietotājam jāspēj attiekties no sistēmas lietošanas, ja tas vairs nevēlas lietot pakalpojumus;
- sistēmai jābūt pieejamai mobilo telefonu lietotājiem LMT un TELE2 tīklā;
- sistēmai jāveic darbību (t.sk. IP adresu) reģistrēšanu žurnālā, lai vajadzības gadījumā būtu iespējams identificēt pārkāpējus.

2. VĒSTURISKAIS IESKATS UN SISTĒMAS IEVIEŠANA

Šajā nodaļā autors apraksta kā soli pa solim sistēma tika ieviesta darbā, kā arī apraksta un izskaidro katru būtisko sistēmas sastāvdaļu, kas eksistēja līdz bakalaura darbā noteikto mērķu izpildei.

Sākotnēji sistēmas mērķis bija lietotājiem, kas uz to parakstījušies, izsūtīt laika ziņas un vārda dienas vienu reizi dienā – no rīta. Tobrīd īsziņu sūtīšana no LMT un TELE2 publiski pieejamajiem Interneta servisiem notika bez papildu grafiskajiem vai citādiem CAPTCHA-testiem. Turklāt tie bija vienīgie mobilo sakaru operatori, kas pieejami Latvijas teritorijā, un par numuru saglabāšanu, mainot operatoru, nebija pat runas. Līdz ar to sistēma sākotnēji tiešām bija vienkārša, lai gan paralēli laika ziņu un vārda dienu pakalpojuma izmantošanai sistēmas lietotājiem bija iespējams arī nosūtīt parastu īsziņu. Visu savu līdzšinējo dzīves ciklu, visas sistēmas daļas darbojās, izmantojot operētājsistēmu Linux Slackware vai kādu citu Linux operētājsistēmu.

Sākotnējās sistēmas iespējas iekļāva:

- īsziņu sūtīšanu LMT un TELE2 lietotājiem e-pasta režīmā (2.1. attēls);
- iespēju pieteikties un attiekties būt par sistēmas abonentu (2.2. attēls);
- automatizētu laika ziņu savākšanu no globālā tīmekļa;
- vārda dienu atrašanu datu bāzē;
- vārda dienu un laika ziņu nosūtīšanu sistēmas abonentiem;
- nosūtīto īsziņu reģistrēšanu žurnālā;
- vienkāršotu statistikas pārlūku.



The screenshot shows a web interface for sending SMS messages. It features a 'Phone:' label followed by a text input field. Below this is a 'Text' label followed by a larger text area. To the left of the text area is a 'Characters left' label with a value of '142'. At the bottom of the form are two buttons: 'Send' and 'Clear'. Below the buttons, there are two lines of small text: '*SMSes are being logged for security reasons.' and '*Because of users' impatience send is now controlled by JavaScript'. At the very bottom, it says 'This page has been visited 650 times.'

2.1. att. Tīmekļa lapas “<http://www.ps.id.lv/mobile/sms.php>” fragments (13)

When you have **subscribed and validated** your phone in *SMSlist* system, you'll receive the **weather forecast** (for Riga) and **namedays** in Latvian every day at about 07.00. If the server won't be able to send them, it will try to resend them later.

If you want to subscribe, please make sure your mobile phone is with you and type in your phone number, you'll be sent a validation code. Save this code, you'll need it if you wish to unsubscribe later.

NB! This may fail, if you have not enabled the service "SMS from e-mail".

Phone (local format, like 9000000):

This page has been visited **2394** times.

362 users are currently registered.

8 of them haven't validated the account yet.

2.2. att. Tīmekļa lapa “<http://www.ps.id.lv/mobile/news.php>” (14)

Vairākus mēnešus vēlāk, redzot, ka sūtīšana e-pasta režīmā (caur mobilo sakaru operatoru e-pasta vārtejām) nav stabila un uzticama, sistēma tika modificēta sūtīšanai caur mobilo sakaru mājaslapās pieejamajām vārtejām.

Neilgi pēc tam īsziņu sūtīšanai uz LMT parādījās pirmie CAPTCHA-testi, kas sastāvēja no diviem simboliem fiksētās koordinātēs. Šai tendencei drīz sekoja arī TELE2.

Mainoties ārējiem apstākļiem, 2003. gadā parādījās Amigo sarunu karte nācās pievienot atbalstu arī šim virtuālajam operatoram, kas darbojas LMT tīklā(15). Diemžēl tobrīd Amigo nedz bija populārs, nedz arī piedāvāja īsziņu nosūtīšanu no Interneta, tamdēļ autors samierinājās ar faktu, ka Amigo pagaidām netiks atbalstīts sistēmā.

2004. gadā LMT ieviesa iespēju sūtīt īsziņas uz klientu izvēlētām iesaukām. Tam tika ieviests atbalsts sistēmā. Šajā gadā arī tiek ieviesta **automātiskās bloķēšanas sistēma, kas pēc lietotāja uzvedības noskaidro, vai īsziņas būtu bloķējamas**. Tā ir jauna alternatīva, kas varētu aizvietot CAPTCHA-testus.

2005. gada beigās tika slēgts laika ziņu serviss, taču darbs pie īsziņu sūtītāja turpinājās.

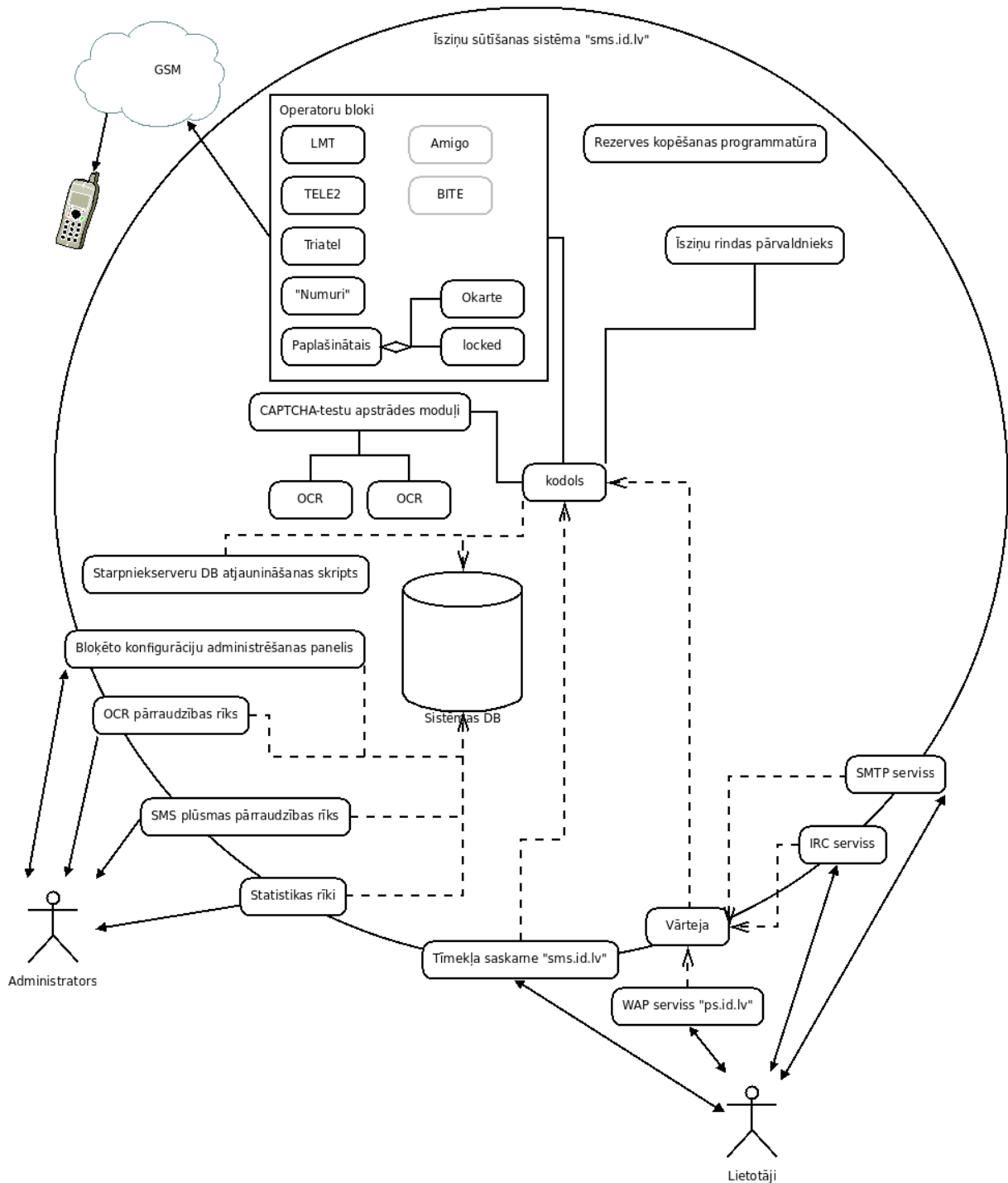
Nākamajos gados Latvijā parādījās jaunie operatori – Triatel, BITE un arī citi. Latvija pārslēdzās uz 8-ciparu numerāciju.

2007. gadā ieviesta arī iespēja saglabāt numuru, mainot operatoru, kam arī tika veltītas gana daudz koda rindiņas. Sistēmā ieviesta arī publisko starpniekserveru datubāze, lai pamanītu, kad lietotājs izmanto neaizsargātu starpniekserveri, un bloķētu šādus mēģinājumus.

Vēl vairākas reizes tika mainīti grafiskie CAPTCHA-testi operatoru mājaslapās, un 2008. gada vidū sistēma pārtrauca darbu.

Šis bakalaura darbs ir apliecinājums autora vēlmei atjaunot sistēmas darbību, lai cilvēki, kas nespēj saredzēt grafiskos CAPTCHA-testus, un tie, kam CAPTCHA-testu risināšana šķiet neērta, varētu sūtīt īsziņas no Interneta.

Uzsākot bakalaura darbu, sistēma sastāv no vairākām daļām, kas savā starpā sadarbojas kā attēlots shēmā (2.3. attēls). Tās tiek aprakstītas sekojošās apakšnodaļās.



2.3. att. Sistēmas loģiskā shēma

2.1. Datu bāze

Sistēmas datu bāze darbojas uz servera programmatūras “MySQL 5.0.67”. Datu bāze ir veidota no 17 tabulām, ko var sagrupēt vairākās funkcionālās grupās:

1. CAPTCHA-testu tabulas
 - *ocrcodes*
 - *ocrfailure*
 - *stats*
2. Vārtejas atļauju tabula
 - *senders*
3. Īsziņu rindas tabulas
 - *queue*
 - *dropped*
4. Telefona numuru un identifikatoru tabulas
 - *aliases*
 - *savemoved*
 - *virtuals*
5. Bloķēto konfigurāciju tabulas
 - *blockids*
 - *blocked*
 - *banedit*
6. Žurnāla tabulas
 - *processed*
 - *archived*
7. Starpniekserveru tabulas
 - *proxy*
 - *proxy_socks*
 - *proxy_tor*

Katru no šīm tabulām autors apraksta turpmākajās apakšnodaļās.

2.1.1. CAPTCHA-testu tabulas

Šo tabulu uzdevums ir nodrošināt struktūru atpazīto CAPTCHA-testu glabāšanai, kā arī datiem par kodiem, kas nav atpazīti.

2.1.1.1. Tabula “ocrcodes”

Tabula nepieciešama, lai glabātu atrisinātos CAPTCHA-testus.

2.1. tabula

CAPTCHA-testu tabula “ocrcodes”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
id	bigint(20)	PRIMARY	auto_increment		Identifikators
keyid	varchar(64)	UNIQUE			Dati, kas identificē, kādai sesijai ir derīga attiecīgā atslēga
OCR	varchar(6)				CAPTCHA-testa atbilde
in	datetime			0000-00-00 00:00:00	Atrisināšanas laiks
out	datetime			0000-00-00 00:00:00	Risinājuma izmantošanas laiks
who	enum('LMT', 'TELE2')			LMT	Operators, kam CAPTCHA-tests derīgs

2.1.1.2. Tabula “ocrfailure”

Tabula satur vēsturisku informāciju par neveiksmīgi atrisinātajiem CAPTCHA-testiem.

2.2. tabula

CAPTCHA-testu tabula “ocrfailure”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
whensent	datetime			0000-00-00 00:00:00	CAPTCHA-testa atbildes iesniegšanas laiks
sender	varchar(128)				Lietotāja saskarnes nosaukums, caur kuru sūtīta īsziņa
imcode	varchar(32)				Faila, kurā glabājas CAPTCHA-tests, nosaukums
keyid	varchar(64)			NULL	Dati, kas identificē, kādai sesijai bija derīga attiecīgā atslēga
smscode	varchar(6)			NULL	CAPTCHA-testa (nederīgā) atbilde
oper	enum('LMT', 'TELE2')			LMT	Operators, kas ģenerējis CAPTCHA-testu

2.1.1.3. Tabula “stats”

Tabula šobrīd glabā statistiku par veiksmīgo un neveiksmīgo CAPTCHA-testu iziešanas skaitu katram operatoram, taču var tikt pielāgota, lai glabātu arī citu statistikas informāciju.

2.3. tabula

CAPTCHA-testu tabula “stats”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
name	varchar(255)				Datu vienības nosaukums
value	varchar(255)				Datu vienības vērtība

Šīs tabulas šībrīža saturs ir redzams 2.4. tabulā. Datu vienības apzīmē attiecīgi LMT (*lmt*) un TELE2 (*_tele2*) veiksmīgo (*codeguessed_*) un neveiksmīgo (*codefailed_*) CAPTCHA-testu atrisināšanas skaitu. Atrisinātie CAPTCHA-testi, kuru rezultāts nav zināms, šajā tabulā netiek uzskaitīti. Datu vienība *poweroff* apzīmē sistēmas stāvokli – CAPTCHA-testu izmantošana ir ieslēgta (“0”) vai izslēgta (jebkura cita pozitīva skaitliska vērtība).

2.4. tabula

Tabulas “stats” saturs

Lauks “name”	Lauks “value”
codeguessed_lmt	2310512
codefailed_lmt	72897
poweroff	0
codeguessed_tele2	3841042
codefailed_tele2	8362

2.1.2. Vārtejas atļauju tabula “senders”

Tabulā atrodas saraksts ar ārējām saskarnēm, kas drīkst izmantot sistēmu, lai sūtītu īsziņas. Šeit glabājas, piemēram, SMTP saskarnes, WAP saskarnes, kā arī citu globālā tīmekļa lapu atļaujas pieslēgties sistēmai caur vārteju.

2.5. tabula

Vārtejas atļauju tabula “senders”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
id	bigint(20)	PRIMARY	auto_increment		Identifikators
sender	varchar(128)	UNIQUE			Saskarnes nosaukums
pass	varchar(32)				Saskarnes pilnvarojuma kods, šifrēts ar MD5 algoritmu

ip	varchar(15)				Saskarnes IP adrese
owner	varchar(255)				Atbildīgais par saskarni
c_mail	varchar(255)				Atbildīgā e-pasts
c_phone	char(8)				Atbildīgā mobilais telefona nr.
active	tinyint(1)			1	Saskarne ir aktīva(1) vai nē(0)

2.1.3. Īsziņu rindas tabulas

Šīs tabulas ir nepieciešamas, lai nodrošinātu, ka īsziņas, kuras nav izdevies nosūtīt ar pirmo reizi, tiktu nosūtītas vēlāk.

2.1.3.1. Tabula “queue”

Tabula satur īsziņas, kas gaida rindā uz izsūtīšanu.

2.6. tabula

Īsziņu rindas tabula “queue”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
uniqid	bigint(20)	PRIMARY	auto_increment		Identifikators
adreee	varchar(20)				Saņēmēja numurs
text	varchar(160)				Sūtāmās īsziņas teksts
rr	tinyint(4)			0	Pēdējais kļūdas kods, kuras dēļ īsziņa nonāca/palika rindā
mail	tinyint(1)			0	Vai īsziņu jāsūta e-pasta režīmā
qin	datetime			0000-00-00 00:00:00	Kad īsziņa nonākusi rindā

2.1.3.2. Tabula “dropped”

Tabula satur informāciju par īsziņām, kuras ir bijušas rindā, taču tās nav izdevies nosūtīt nelabojamas kļūdas dēļ.

2.7. tabula

Īsziņu rindas tabula “dropped”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
uniqid	bigint(20)	PRIMARY	auto_increment		Identifikators
adreee	varchar(20)				Saņēmēja numurs
text	varchar(160)				Nenosūtītās īsziņas teksts
rr	tinyint(4)			0	Kļūdas kods, kuras dēļ īsziņa tika izmesta no rindas

mail	tinyint(1)			0	Vai īsziņa tika sūtīta e-pasta režīmā
qin	datetime			0000-00-00 00:00:00	Kad īsziņa nonākusi rindā
uniqid	datetime			0000-00-00 00:00:00	Kad īsziņa izmesta no rindas

2.1.4. Telefona numuru un identifikatoru tabulas

Šīs tabulas nepieciešamas, lai sistēmai būtu informācija par adresātu īpašajiem gadījumiem (iesaukām, virtuālajiem numuriem), kā arī, lai atcerētos informāciju par telefona numuriem, kuriem mainīts operators.

2.1.4.1. Tabula “aliases”

Tabulā glabājas informācija par adresātu iesaukām. Piemēram, pateicoties šai tabulai, sūtot īsziņu adresātam “~kirils” īsziņa tiks nosūtīta uz darba autora telefona numuru.

2.8. tabula

Telefona numuru un identifikatoru tabula “aliases”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
alias	bigint(19)	UNIQUE			Aizstājvārds
number	varchar(20)				Telefona numurs
info	mediumtext				Informācija par personu, vārds, uzvārds un citi dati
used	tinyint(4)			1	Aizstājvārds ir spēkā(1) vai nav(0)

2.1.4.2. Tabula “savemoved”

Tabula satur informāciju par mobilā telefona numuriem, kuri pārcelti pie cita operatora. Šādā veidā sistēmai bez vajadzības nav atkārtoti jāpieprasa informācija ārējā numuru datubāzē.

2.9. tabula

Telefona numuru un identifikatoru tabula “savemoved”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
number	char(8)	PRIMARY			Telefon numurs
oper	varchar(16)				Jaunais operators
regmove	datetime				Datums, kad sistēmā reģistrēta operatora maiņa

2.1.4.3. Tabula “virtuals”

Tabula satur informāciju par numuriem, kas izmanto paplašināto operatoru bloku.

2.10. tabula

Telefona numuru un identifikatoru tabula “virtuals”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
number	varchar(20)	UNIQUE			Mobilā telefona numurs
virtual	varchar(128)				Vērtība, kas aizvieto šo mobilā telefona numuru paplašinātajā operatoru blokā
extension	varchar(16)				Paplašinājuma nosaukums
try	smallint(1)			0	Pirms paplašinātā operatoru bloka izmantošanas, mēģināt(1) vai nemēģināt(0) sūtīt īsziņu, neizmantojot paplašināto operatoru bloku

2.1.5. Bloķēto konfigurāciju tabulas

Tabulas satur informāciju par bloķētajām konfigurācijām, kā arī par administrēšanas paneļa lietotājiem. Šīs trīs tabulas autors izstrādājis sadarbībā ar Jāni Timmu.

2.1.5.1. Tabula “blockids”

Tabula satur vispārīgu aprakstu par katru bloķēto konfigurāciju. Šī tabula nesatur konkrētu konfigurāciju (to veido ieraksti citā tabulā).

2.11. tabula

Bloķēto konfigurāciju tabula “blockids”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
blockid	bigint(20)	PRIMARY	auto_increment		Bloka identifikators
author	varchar(255)			admin	Persona vai saskarne, kas uzlikusi bloku
setdate	datetime			0000-00-00 00:00:00	Bloka uzlikšanas datums
reason	varchar(255)				Bloka iemesls (tiek parādīts lietotājam)
notes	mediumtext				Piezīmes par bloka uzlikšanu (redzamas tikai administratoriem)

2.1.5.2. Tabula “blocked”

Tabula satur konkrētas parametru konfigurācijas, kas, grupējot pēc bloka identifikatora, kopā veido bloķētās konfigurācijas.

2.12. tabula

Bloķēto konfigurāciju tabula “blocked”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
id	bigint(20)	PRIMARY	auto_increment		Identifikators
rec	varchar(128)				Īsziņas saņēmējs
sms	varchar(32)				Īsziņas teksts
whensent	varchar(15)				Īsziņas nosūtīšanas laiks
operator	varchar(255)				Saņēmēja mobilā telefona operators
viamail	tinyint(1)				Vai īsziņa tiek sūtīta e-pasta režīmā
sender	varchar(128)				Saskarne, caur kuru iesūtīta īsziņa
realip	tinyint(1)			1	Īsziņas nosūtītāja identifikators (IP adrese)
agent	varchar(128)				Programmatūra, kas tiek lietota īsziņas sūtīšanai
neg	set('rec','sms','whensent','operator','sender','realip')				Lauki, kuru vērtības jāuzskata par apgrieztām (piemēram, šis bloks aktivizējas nevis, kad attiecīgais lauks atbilst maskai, bet kad neatbilst)
deny	tinyint(1)			1	Sakrītots šai parametru konfigurācijai bloķēt īsziņu(1), nekavējoties atļaut īsziņu(-1) vai pāriet pie nākamās parametru konfigurācijas izvērtēšanas(0).
blockid	bigint(20)			0	Bloka identifikators

2.1.5.3. Tabula “banedit”

Tabula satur informāciju par administratoriem, kas drīkst izmantot bloķēto konfigurāciju administrēšanas paneli. Tabulu pilnībā veidojis Jānis Timma.

2.13. tabula

Bloķēto konfigurāciju tabula “banedit”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
ID	int(11)	UNIQUE	auto_increment		Identifikators

Login	char(32)				Lietotāja vārds
Password	char(32)				Lietotāja parole, šifrēta ar MD5 algoritmu
Random	int(11)			0	Lietotājs šobrīd nav pieteicies sistēmā (0) vai ir (jebkura cita vērtība)
EditTime	int(11)			0	Laiks, kad lietotājs sācis rediģēt bloku
EditID	int(11)			0	Bloks, ko lietotājs šobrīd rediģē

2.1.6. Žurnāla tabulas

Tabulas paredzētas likumiskajām vajadzībām, lai reģistrētu datus žurnālā.

2.1.6.1. Tabula “processed”

Šī tabula satur visas īsziņas, kas tikušas sūtītas caur sistēmu pēdējā mēneša laikā. Tabula saturs tiek izmantots statistikai, lai novērstu sistēmas nepilnības un uzlabotu drošību, kā arī, lai izmeklētu incidentus un sadarbotos ar tiesībsargājošām struktūrām.

Svarīgs tabulas uzdevums ir arī datu nodrošināšana uzbrukumu analīzei reālā laika režīmā – salīdzinot pēdējās stundas laikā nosūtītās īsziņas ar īsziņu, kas šobrīd tiek sūtīta, iespējams noteikt, vai lietotāja uzvedība ir pieņemama.

2.14. tabula

Žurnāla tabula “processed”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
rec	varchar(20)	INDEX			Īsziņas saņēmējs
sms	varchar(160)	INDEX			Īsziņas teksts
whensent	datetime	INDEX		0000-00-00 00:00:00	Īsziņas nosūtīšanas laiks
operator	varchar(16)			NULL	Saņēmēja mobilā telefona operators
status	smallint(6)			0	Nosūtīšanas kļūdas kods
viamail	tinyint(1)			0	E-pasta režīma lietojums
sender	varchar(128)				Saskarne, caur kuru nosūtīta īsziņa
realip	varchar(255)	INDEX		NULL	Īsziņas nosūtītāja identifikators (IP adrese)
agent	varchar(255)	INDEX		NULL	Programmatūra, kas tika lietota īsziņas sūtīšanai
keyid	varchar(64)			NULL	Satur datus, kas identificē, kādai sesijai ir derīga attiecīgā atslēga

smscode	varchar(6)			NULL	CAPTCHA-testa atbilde
whenqueued	datetime			0000-00-00 00:00:00	Datums, kad īsziņa tika ielikta rindā (ja vispār)
pduration	decimal(6,5)			0.00000	Cik ilgi lietotājam bija jāgaida, lai uzzinātu sūtīšanas rezultātu (sekundēs)

2.1.6.2. Tabula “archived”

Šī tabula satur informāciju par visām caur sistēmu nosūtītajām īsziņām. Tabula sistēmā nepieciešama statistiskām un vēsturiskām vajadzībām.

2.15. tabula

Žurnāla tabula “archived”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
id	bigint(20)	PRIMARY	auto_increment		Identifikators
rec	varchar(20)	INDEX			Īsziņas saņēmējs
sms	char(32)				Nosūtītās īsziņas saturs, kas apstrādāts ar vienvirziena algoritmu
whensent	datetime	INDEX		0000-00-00 00:00:00	Īsziņas nosūtīšanas laiks
operator	varchar(16)			NULL	Saņēmēja mobilā telefona operators
status	smallint(6)			0	Nosūtīšanas kļūdas kods
viamail	tinyint(1)			0	E-pasta režīma lietojums
sender	varchar(128)				Saskarne, caur kuru nosūtīta īsziņa
realip	varchar(255)	INDEX		NULL	Īsziņas nosūtītāja identifikators (IP adrese)
agent	varchar(255)	INDEX		NULL	Programmatūra, kas tika lietota īsziņas sūtīšanai
keyid	varchar(64)			NULL	Satur datus, kas identificē, kādai sesijai ir derīga attiecīgā atslēga
smscode	varchar(6)			NULL	CAPTCHA-testa atbilde
whenqueued	datetime			0000-00-00 00:00:00	Datums, kad īsziņa tika ielikta rindā (ja vispār)
pduration	decimal(6,5)			0.00000	Cik ilgi lietotājam bija jāgaida, lai uzzinātu sūtīšanas rezultātu (sekundēs)

2.1.7. Starpniekserveru tabulas

Šīs tabulas sistēma izmanto, lai identificētu, vai lietotājs īsziņu sūtīšanai pielieto kādu neaizsargātu publisku starpniekserveri.

2.1.7.1. Tabula “proxy”

Tabula satur informāciju par HTTP starpniekserveriem. Vēsturiski šī tabula kādreiz tika (un, iespējams, vēl tiks) izmantota arī sistēmas darbības nodrošināšanai, tāpēc šobrīd lielākā daļa lauku tajā ir it kā lieki.

2.16. tabula

Starpniekserveru tabula “proxy”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
ID	int(10)	PRIMARY	auto_increment		Identifikators
IP	varchar(15)	UNIQUE			IP adrese
Port	int(11)			0	Ports
Type	varchar(100)				Klase
Country	varchar(50)				Valsts
Date	int(11)			0	Ievades datums
Status	tinyint(3)			2	Statuss: nedarbojas(2), darbojas(1), nav pārbaudīts(0)
Speed	float			0	Saziņas ātrums (baiti sekundē)
HTTP_X_FORWARDED_FOR	varchar(255)			-X-	Ko starpniekserveris norāda savienojuma vērtībā HTTP_X_FORWARDED_FOR
HTTP_VIA	varchar(255)			-X-	Ko starpniekserveris norāda savienojuma vērtībā HTTP_VIA
HTTP_PROXY_CONNECTION	varchar(255)			-X-	Ko starpniekserveris norāda savienojuma vērtībā HTTP_PROXY_CONNECTION
HTTP_USER_AGENT	varchar(255)			-X-	Ko starpniekserveris norāda savienojuma vērtībā HTTP_USER_AGENT
REMOTE_ADDR	varchar(255)			-X-	Ko starpniekserveris norāda savienojuma vērtībā REMOTE_ADDR

2.1.7.2. Tabula “proxy_socks”

Tabula satur informāciju par SOCKS starpniekserveriem. SOCKS ir protokols, kas darbojas tīkla saziņas 5. līmenī (sesijas līmenī) un nodrošina klientam iespēju caur starpniekserveri, ne tikai pārlūkot globālo tīmekli, bet lietot Internetu kā tādu.

2.17. tabula

Starpniekserveru tabula “proxy_socks”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
ip	char(15)	PRIMARY			IP adrese
port	mediumint(9)			0	Ports

2.1.7.3. Tabula “proxy_tor”

Tabula satur informāciju par *tor* gala mezgliem. Tas ir samērā jauns protokols, kas nodrošina lielu anonimitātes pakāpi. Tabulā iespējams glabāt informāciju arī par citām IP adresēm, kas ir pieejamas tādā mērā, ka tās iespējams publiski izmantot, atrodoties dažādās pasaules malās.

2.18. tabula

Starpniekserveru tabula “proxy_tor”

Nosauk.	Tips	Indeksi	Īpašības	Noklusētā vērtība	Apraksts
ip	char(15)	PRIMARY			IP adrese
type	enum('tor','other')				IP adreses tips

2.2. Kodols

Kodols tika izveidots 2004. gadā, un tā pirmsākumi meklējami sākotnējā īsziņu sūtīšanas formā, kas kādreiz atradās uz servera “ps.id.lv”. Tā mērķis ir apvienot CAPTCHA-testu apstrādes moduļus, operatoru bloku un īsziņu rindas pārvaldnieku, lai kopīgi ar datu bāzi veidotu labi funkcionējošu sistēmu. Datu bāze sistēmai tiek piesaistīta tieši caur kodolu. Darbības nodrošināšanai kodolu izmanto vārteja un tīmekļa saskarne “sms.id.lv”.

Kodols ir atbildīgs par daudziem būtiskiem sistēmas aspektiem, piemēram statisko operatora piederības noteikšanu, automātisko konfigurāciju bloķēšanu u.c. Kodols satur arī vispārīgo īsziņas sūtīšanas funkciju, kas sadarbojas ar citām sistēmas daļām. Tabulu ar kļūdu kodiem un to tekstuālo skaidrojumu arī ir kodolā. Tas ir rakstīts valodā PHP.

2.2.1. Kļūdu kodu tekstuāla atšifrēšana

Kodola funkcija $r2str(\$id, \$l='en')$ ($\$id$ – kļūdas kods, $\$l$ – valoda) spēj pārvērst kļūdas kodu par paskaidrojumu latviešu vai angļu valodā. Paskaidrojumi latviešu valodā iekļauti 2.19. tabulā. Tie var noderēt šī darba lasītājam.

Šī funkcija ir būtiska, lai lietotāji un administratori spētu saprast kļūdas cēloni. Pašai sistēmai ir vieglāk strādāt ar skaitļiem, tāpēc iekšēji kļūdas vienmēr tiek glabātas kā skaitļi.

2.19. tabula

Sistēmas kļūdu kodu atšifrējumi latviešu valodā

Kļūdas kods	Lietotājam redzamais paskaidrojums	Kļūdas iemesls
0	Serveris: Tukšs dokuments. Īsziņas statuss nezina.	Operatora bloks nav saņēmis nekādu atbildi no operatora, līdz ar to nav iespējams noteikt, vai īsziņa nosūtīta vai nē.
-1	Vispārēja kļūda. Vairāk informācijas nav pieejams.	Operatora bloks nespēja noskaidrot konkrētu iemeslu, kāpēc īsziņa nav nosūtīta, taču spēja secināt, ka tā nav nosūtīta.
-2	Serveris: Īsziņa pārāk gara.	Operators paziņoja, ka īsziņas garums pārsniedz atļauto.
-3	Serveris: Savienojums taimautoja.	Operatora bloks nespēja sagaidīt atbildi no operatora.
-4	Nevaru uzminēt autorizācijas kodu. Mēģini pēc pāris stundām.	Operators atkārtoti nepieņēma sistēmas sagatavoto atbildi uz CAPTCHA-testu.
-5	Īsziņa pārāk gara.	Kodols ir saņēmis īsziņu, kuras garums pārsniedz atļauto.
-6	Neizdevās atrast atslēgas koda tagu. Mēģini pēc pāris stundām.	OCR skripts nav spējīgs atrisināt pietiekoši daudz CAPTCHA-testu un šobrīd tie ir pieejami ļoti mazā skaitā.
-7	Serveris: Nespēju atvērt savienojumu.	Operatora bloks nespēja izveidot savienojumu ar operatoru.
-8	Serveris: Nespēju nosūtīt datus.	Operatora bloks nespēja nosūtīt datus operatoram.
-9	Šis numurs nepieder saņēmēja tīklam.	Sistēma īsziņas nosūtīšanai noteica konkrētu operatoru, taču šis operators atteicās pieņemt īsziņu, aizbildinoties, ka tas nav viņa numurs.
-10	Lietotājs nav aktivizējis īsziņu saņemšanu no interneta.	Īsziņas saņēmējs nav atļāvis saņemt īsziņas no Interneta.
-11	Pietrūkst datu.	Lietotājs nav ievadījis īsziņas tekstu vai saņēmēja numuru.
-12	Kodu kombinācija vairs nav derīga. Mēģini pēc pāris stundām.	CAPTCHA-testa atbilde, ko sistēma izvēlējās izmantot šīs īsziņas nosūtīšanai, vairs nav derīga izmantošanai.
-13	Šis nav vietējs publiskā mobilā telefonu tīkla numurs.	Numurs, ko lietotājs ievadījis nepieder nevienam operatoram un nav arī aizstājvārds.
-14	Īsziņu sūtīšana uz šo operatoru šobrīd nav pieejama.	Operatoram, kam pieder numurs, pagaidām neeksistē labi strādājošs operatora bloks.
-15	Pārsniegts īsziņu limits no vārtejas. Lūdzu gaidīt.	Sistēma ir pārsniegusi īsziņu skaitu, ko operators spēj pieņemt noteiktā laika posmā.
-16	Nepareizs telefona numurs.	Operatora bloks secinājis, ka numura formāts neatbilst operatora prasībām.
-17	Lietotājs pārsniedzis saņemamo īsziņu limitu.	Īsziņas adresāts ir sasniedzis operatora uzstādītu diennakts īsziņu skaita ierobežojumu.

-18	Serveris: Iekšēja kļūda.	Operatora serveris ir pārstājis korekti funkcionēt.
-19	Autorizācijas kods nav pieejams. Mēģini pēc pāris stundām.	Vairs netiek izmantots.
-20	Sūtītāja limits sasniegts. Lūdzu gaidīt.	Šobrīd netiek izmantots.
-21	Filtrs: plūdi - šāda īsziņa jau tikko ir nosūtīta.	Adresātam nesen jau nosūtīta tāda vai līdzīga īsziņa.
-22	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Lietotājs vairumā izsūta līdzīgas vai nesakarīgas īsziņas.
-23	Filtrs: dalīšana - daudzu īsu ziņojumu izsūtīšana.	Adresāts no lietotāja īsā laikā saņem daudz īsas īsziņas, lai gan tās būtu iespējams apvienot vienā garākā.
-24	Filtrs: dauzīšanās - nepārtraukta sūtīšanas pogas spaidīšana.	Lietotājs ignorē kļūdas paziņojumus un vairākkārtēji cenšas nosūtīt īsziņu.
-25	Filtrs: Šāda īsziņas konfigurācija ir bloķēta. Īsziņa netiks nosūtīta. Lai vienotos par bloka noņemšanu, sazinieties ar prohibit [@at] ctrl.sms.id.lv pa e-pastu.	Konfigurācija, kādā lietotājs sūta īsziņu, ir bloķēta.
-26	Aizstātvārds nav lietošanā.	Aizstātvārds, uz kuru lietotājs cenšas nosūtīt īsziņu neeksistē vai nav lietošanā.
-27	Filtrs: Maksimālais īsziņu skaits - <i>\$maxsmsipday</i> dienā - sasniegts. Nāc atpakaļ rīt.	Lietotājs ir nosūtījis tik īsziņas, cik sistēma pieļauj vienā dienā.
-28	Nederīga virtuālā numura deklarācija.	Vērtība, kas aizvieto numuru paplašinātajā operatora blokā nav derīga.
-29	Admins: Virtuālā paplašinājuma funkcija nav definēta.	Neeksistē tāds paplašinātais bloks, kāds nepieciešams sūtot uz šo virtuālo numuru.
-30	Serveris: Nevar autorizēties ārējā kontā.	Operatora bloks nespēj identificēties ārējā kontā.
-31	SIM karte, kurai piesaistīts šis numurs, ir tikusi nozagta, pazaudēta vai vairs nav lietošanā.	SIM karte, kurai piesaistīts šis numurs, ir tikusi nozagta, pazaudēta vai vairs nav lietošanā.
2	Īsziņa pievienota rindai un tiks nosūtīta vēlāk.	Īsziņu nebija iespējams nosūtīt, taču lietotājs norādīja, ka šādā gadījumā, tā jāmēģina nosūtīt vēlāk.
1	Īsziņa nosūtīta veiksmīgi.	Operatora bloks saņēma atbildi, ka īsziņa ir pieņemta nosūtīšanai uz operatora tīklu.

2.2.2. Statiskā operatora piederības noteikšana

Kodolā ir masīvs ar numerācijas prefiksiem katram no populārajiem operatoriem. Šis masīvs ir balstīts uz Sabiedrisko pakalpojumu regulēšanas komisijas datiem(16). Par šo funkcionalitāti ir atbildīga funkcija *GSMprovider(\$line,\$dbsi=false)* (*\$line* – adresāts, *\$dbsi* norāda, vai šis ir atkārtots mēģinājums sūtīt uz virtuālo numuru). Pirms pārbaudīt, kuram operatoram pieder adresāta numurs, sistēmas kodols pārbauda telefona numuru un identifikatoru tabulas, lai pārliecinātos, vai adresāts īstenībā nav aizstājvārds, vai tas nav definēts kā virtuāls numurs, sūtīšanai uz kuru jālieto paplašinātais operatoru bloks, kā arī vai sistēma gadījumā jau nav reģistrējusi faktu, ka šim numuram ir mainīts operators. Tiek arī pārbaudīts, vai ievadītais adresāts atbilst LMT adreses identifikatora prasībām.(17) Ja neviens no šiem priekšnosacījumiem neizpildās, tad pēdējā pārbaude ir, vai numurs atbilst Latvijas mobilo tālrunu numuru formātam. Ja numurs neatbilst šim formātam, tad operatoru, protams, noskaidrot nav izdevies, bet pretējā gadījumā tālāk norisinās konkrēta operatora noskaidrošana.

Sistēmas kodola funkcijas *GSMprovider(\$line, \$dbsi=false)* fragments:

```
if($line[0]=="~")
    return "alias";

list($iS,$iSt)=mysql_fetch_array(mysql_query("select `extension`,`try` from virtuals
where (number='".addslashes($line)."'",$dbhaGSM));
if($iS&&!(($dbsi&&$iSt)))
    return "virtual";

list($opmove)=mysql_fetch_array(mysql_query("select `oper` from savemoved where
(number='".addslashes($line)."'",$dbhaGSM));
if($opmove)
    return $opmove;

if(eregi("[a-z][0-9a-z._-]{4,19}$",$line))
    return "LMTnick";

if(!ereg("^2[0-9]{7}$",$line))
    return false;
```

Vajadzības gadījumā kodols ir atbildīgs arī par iesauku atšifrēšanu par numuriem:

```
if(GSMprovider($to)=="alias"){
    $ro=mysql_query("select number from aliases where used=1 and alias =
'".addslashes(substr($to,1)).'",$dbhaGSM);
    list($ro)=mysql_fetch_row($ro);
    if(!$ro)
        $reason=-26;
    else
        $to=$ro;
}
```

2.2.3. Īsziņu ierindošana

Lai nodrošinātu lietotājam pēc iespējas vienkāršāku un produktīvāku darbu ar sistēmu, viņam tiek piedāvāts ierindot (ievietot rindā) īsziņu gadījumā, ja to neizdosies nosūtīt uzreiz. Šādā gadījumā lietotājam tīmekļa saskarnē “sms.id.lv” ir jāatzīmē iespēja “Nosūtīt vēlāk, ja rodas īslaicīgas problēmas tūlītējā nosūtīšanā.” Saskarnēs un servisos, kur šāda iespēja nav paredzēta, tā darbojas automātiski. Par īsziņu ierindošanu atbildīgs tieši kodols.

Saraksts ar kļūdām, kas var būt par iemeslu, lai ierindotu īsziņu ir definēts kodolā. Tie visi ir pagaidu iemesli: -3, -7, -8, -18, -19, -30 (skatīt 2.19. tabulu). Līdzīgi kodolā eksistē arī saraksts ar nelabojamām kļūdām, kad īsziņa no rindas būtu dzēšama: -1, -2, -5, -9, -10, -11, -13, -14, -16, -26, -28. Vispārējā īsziņu sūtīšanas funkcija pēc rezultāta noskaidrošanas pārbauda, vai kļūda ir to starpā, kas ļauj ierindot īsziņu un vai lietotājs to ir vēlējies:

```
if( (!$noqueue) && in_array($reason,$queuer) )  
    if(queueup($to,$sms,$reason,$viamail))  
        $reason=2;
```

Ierindošanas funkcija *queueup(\$to,\$sms,\$lr,\$viamail=0)* (*\$to* – adresāts, *\$sms* – īsziņas teksts, *\$lr* – kļūdas kods, *\$viamail* – izmantot e-pasta režīmu) savā būtībā ir ļoti vienkārša:

```
function queueup($to,$sms,$lr,$viamail=0){  
    global $dbhaGSM;  
    if($viamail)  
        $viamail=1;  
    else  
        $viamail=0;  
  
    $sms=addslashes($sms);  
    $to=addslashes($to);  
    $lr=$lr+0;  
  
    return mysql_query("insert into `queue` (adrese, text, mail, `qin`, rr)  
values('$to', '$sms', '$viamail', NOW(), '$lr')", $dbhaGSM);  
}
```

Par to, kas ar šo īsziņu notiks tālāk kodols vairs “neuztraucas”, kamēr vien īsziņu rindas pārvaldnieks to nemēģina nosūtīt atkārtoti (taču to kodols uztvers jau kā citu īsziņu un citu uzdevumu).

2.2.4. Bloķēto konfigurāciju testēšana

Pirmā lieta, ko izdara sistēma pēc tam, kad lietotājs cenšas nosūtīt īsziņu, ir bloķēto konfigurāciju testēšana. Šeit kodols paņem visus sūtīšanas pieprasījumu raksturojošos parametrus (pašu īsziņu, adresātu, sūtītāja IP adresi, sūtīšanas laiku utt.), kas kopā tiek saukti par konfigurāciju, un pārbauda pret katru bloķēto konfigurāciju. Ja izdodas atrast tādu bloķēto konfigurāciju, kas atbilst konkrēta pieprasījuma konfigurācijai, tad sistēma atgriež kļūdu -25 un īsziņu nesūta.

2.2.5. Automātiskā konfigurāciju bloķēšana

Tā kā šī sistēma saviem lietotājiem neprasa atrisināt CAPTCHA-testus, lai nodrošinātu kārtību un novērstu kaitnieciskas darbības ir jālieto citas metodes. Autors pirmajā nodaļā piedāvāja unikālu risinājumu kā cīnīties ar ļaundabīgām darbībām – lielākajā gadījumā svarīgi ir, ko dara lietotājs, nevis kas viņš ir (dators vai cilvēks). Lai gan tīmekļa lietotāja uzvedības analīze kā aizsardzības metode pret, piemēram, nevēlamiem ziņojumiem šobrīd jau tiek pētīta, pagaidām pētījumi neiekļauj lietotāja sūtīto saturu. *“In order to analyze the behavior pattern of Web users, we collected Web access log from July 1st, 2007 to August 26th, 2007 with the help of a commercial search engine company. No private information is included in these access logs but user sessions can be identified by different session IDs.”*(8) Tātad Tsinghua Universitātes Datorzinātnes un tehnoloģijas laboratorijas pētījumā, kas šajā jomā ir viens no jaunākajiem, tika izmantota tika sekojoša informācija:

- lietotāja veiktā pieprasījuma laiks;
- lietotāja pieprasītā tīmekļa adrese;
- lietotāja sesija;
- iespējams, lietotāja IP adrese.

Tas acīmredzami liecina, ka šajā pētījumā netika analizēts saturs, taču autora piedāvājums ir analizēt tieši lietotāja veidoto saturu (kopā ar notikuma laiku), bet daudzos gadījumos ignorēt IP adresi vai pat sesiju.

Sistēmā “sms.id.lv” tas darbojas tā, ka pēc tam, kad saņemts pieprasījums nosūtīt īsziņu, uzreiz tiek pārbaudītas esošās bloķētās konfigurācijas un tad notiek automātiskā konfigurāciju bloķēšana (ja lietotājs uzvedas neatbilstoši). Jāpiebilst, ka pirms automātiskas bloķēšanas lietotājs tiek vairākas reizes brīdināts par draudošajām nepatīkšanām, lai dotu viņam iespēju labprātīgi pārtraukt nekārtības.

Uzvedības analīze ir iespējama pateicoties pilnīgai informācijas uzkrāšanai tabulā “processed”. Lai parādītu, kā kodols noskaidro, ka lietotājs neuzvedas kā vajadzīgs, autors sāks ar vienkāršu piemēru. Ja pēdējās stundas laikā uz konkrēto tālruņa numuru jau tikusi veiksmīgi nosūtīta tieši tāda īsziņa, tad numura īpašnieks droši vien negribētu vēlreiz ko tādu saņemt. Kā viegli iedomāties, šo nosacījumu realizēt ir triviāli:

```
if(sqlr1v("select count(*) from processed where status>0 and whenqueued=0 and
rec='$Xrec' and sms='$Xsms' and (UNIX_TIMESTAMP(NOW())-
UNIX_TIMESTAMP(whensent))<3600")){
    $reason=-21.1;
    $comnts="to number $0to with text \"'$Xsms'\"";
}
```

Jau sarežģītāks piemērs būs noteikt, ka adresāts īsā laikā saņem daudz īsas īsziņas, lai gan tās būtu iespējams apvienot vienā garākā. Tomēr arī uz to matemātika ir spējīga, ja vien ir iespējams nodrošināt vajadzīgos datus:

```
$zzp=sqlr1v("select AVG(LENGTH(sms)),count(sms) from processed where status>0 and
whenqueued=0 and rec='$Xrec' and (UNIX_TIMESTAMP(NOW())-UNIX_TIMESTAMP(whensent))<600
order by whensent desc limit 10",1);
if((strlen($sms)<$maxlen/2)&&(($maxlen/2-$zzp[0])*(($zzp[1]+1)>=200)){
    $reason=-23;
    $comnts="{ $zzp[1]} sms (last of them with text \"'$Xsms'") with average length
{$zzp[0]} to number $0to";
}
```

Protams, ka šīs nav vienīgās pazīmes, ko kodols meklē, analizējot sistēmas lietotāju uzvedību. Kā redzams, tad nevienā no piemēriem netiek ņemta vērā sūtītāja IP adrese un kopumā netiek vērtēts, vai sūtītājs ir cilvēks vai programma. Arī ar lietotāja uzvedības analīzi ir iespējams nodrošināt efektīvu sistēmas aizsardzību.

2.2.6. Reģistrēšana žurnālā

Pēc tam, kad kodols ir pabeidzis īsziņas sūtīšanu (vai nu veiksmīgi, vai neveiksmīgi), pirms atbildes sniegšanas sistēmai un lietotājam kodolam atliek vienīgi pierakstīt paveikto žurnālā. Tabulā “processed” tiek saglabāta pilna pieprasījuma konfigurācija, sūtīšanas rezultāts (klūdas kods), informācija par noteikto operatoru un izmantotajiem CAPTCHA-testiem un laiks, cik sekundes kodols (kopā ar pārējām sistēmas sastāvdaļām) ir patērējis šīs īsziņas apstrādei.

Jāpiebilst, ka automātiskie servisi parūpējas, lai īsziņas teksts žurnālā neglabātos mūžīgi. Ātrdarbības nolūkā vismaz mēnesi vecās īsziņas tiek pārvietotas uz tabulu “archived”, kur īsziņas teksts tiek neatgriezeniski šifrēts.

2.2.7. Citi uzdevumi

Ir arī citas lietas, par ko ir atbildīgs tieši sistēmas kodols. Piemēram, rupjo īsziņu pamanīšana. Kodols nodrošina funkciju, kas, balstoties uz nelielu teksta datubāzi, spēj pateikt, vai īsziņa satur lamuvārdus. Šo funkciju izmanto īsziņu plūsmas pārraudzības rīks, lai šādas īsziņas reālā laika režīmā parādītu administratoram.

Arī pamanīt, vai lietotājs gadījumā neizmanto publisku anonīmu starpniekserveri, lai slēptu savas pēdas, ir uzticēts kodolam. Tas izmanto starpniekserveru tabulas un, pamanot, ka lietotājs izmanto šāda veida starpniekserveri, liek sistēmai vienkārši atslēgties un neatbildēt, simulējot, ka lietotājam zudis savienojums ar starpniekserveri vai Internetu.

Kodols sadarbībā ar tīmekļa saskarni “sms.id.lv” dažreiz spēj pamanīt, ja lietotājs (vai dators) īsziņas tekstu nevis ieraksta, bet iekopē teksta laukā. Tas tiek darīts, salīdzinot laiku, kad saskarnes forma tika pieprasīta, un laiku, kad tā tika nosūtīta atpakaļ jau aizpildīta.

2.3. Maršrutēšana

Tā kā mobilie operatori papildus CAPTCHA-testiem arī analizē lietotāju uzvedību (autors uzskata, ka tas gan notiek ļoti primitīvā veidā – tikai pēc pieprasījumu skaita no IP adreses laika vienībā), tad, lai izvairītos no tā, ka īsziņas netiek nosūtītas dēļ šī ierobežojuma, starp sistēmu un Internetu atrodas maršrutētājs, kas reizi 15 sekundēs nomaina IP adresi no kuras tiek sūtīts pieprasījums uz operatoru serveriem. Maršrutēšana realizēta ar programmas “iptables v1.4.2” palīdzību, balstoties uz *bash* skriptiem.

2.4. Operatoru bloki

Kad kodols ir noskaidrojis, kāds operators ir numuram, uz ko sūtāma īsziņa, kodols izsauc kādu no operatora blokiem. Tie parasti ir atbildīgi par īsziņas nodošanu operatoram, lai gan divi no tiem ir īpaši – operatora noteikšanas bloks un paplašinātais bloks par kuriem pastāstīts apakšnodaļās.

Standarta operatora blokā ietilpst divas funkcijas. Viena – īsziņas sūtīšanai e-pasta režīmā un otra – īsziņas sūtīšanai caur HTTP protokolu. Visi operatoru bloki šobrīd ir realizēti valodā PHP.

Īsziņas nosūtīšana e-pasta režīmā ir triviāla. Piemēram, šeit ir viss kods, kas vajadzīgs, lai nosūtītu īsziņu uz TELE2, izmantojot e-pasta režīmu. Funkcija *sendSMSTELE2viamail(\$to,\$sms)* (*\$to* – adresāts, *\$sms* – īsziņas teksts):

```
function sendSMSTELE2viamail($to,$sms){
    $rander=rand(1,999);
    while (strlen($rander)<5)
        $rander="0".$rander;
    $malle="371".$to."@sms.tele2.lv";
    if(!@mail($malle,'',html_entity_decode($sms),"From: ".$rander."@c.sms.id.lv"))
        return(-1);
    else
        return(1);
}
```

Kā redzams, šajā gadījumā tiek vienkārši nosūtīts e-pasts uz konkrētu TELE2 e-pasta adresi, taču sūtīšana caur HTTP protokolu jau ir interesantāka. Standarta gadījumā, algoritmu var raksturot šādos soļos:

1. No tabulā “ocrcodes” esošajiem atrisinātajiem CAPTCHA-testiem tiek iegūts risinājums un sesijas dati, kas identificē, kurai sesijai pieder šis risinājums.
2. Tiek uzbūvēts HTTP pieprasījums, iekļaujot visu to pašu, ko saturētu pieprasījums no attiecīgā operatora mājaslapas.
3. Pieprasījumam tiek pievienoti dati par sesiju, kā arī CAPTCHA-testa risinājums.
4. Pieprasījums tiek nosūtīts uz serveri un sagaidāta atbilde.
5. Ja atbildi izdevies identificēt, tā tiek nokodēta par kādu no sistēmas kļūdas kodiem un nosūtīta kodolam; ja atbildi nav izdevies atpazīt, tādā gadījumā kodolam tiek nosūtīts kļūdas kods 0 un visa atbildes tiek saglabāta uz servera, lai to vēlāk būtu iespējams izpētīt un programmēt operatora bloku šīs atbildes atpazīšanai.

2.4.1. Operatora noteikšanas bloks

Šis ir bloks, kas tika ieviests 2007. gadā, kad Latvijā kļuva pieejama iespēja saglabāt telefona numuru, bet mainīt operatoru. Bloka uzdevums ir noteikt, pie kura operatora tad patiesībā šobrīd atrodas attiecīgais numurs, uz kuru jāsūta īsziņa. Operatoru noteikšanas bloku izsauc kodols, kad tam neizdodas nosūtīt īsziņu un tas saņem paziņojumu, ka numurs nepieder saņēmēja tīklam, vai arī gadījumos, kad sistēma atgriež kļūdu -14.

Šādos gadījumos tiek izsaukta operatora noteikšanas bloka funkcija *findprov(\$num)* (*\$num* – mobilā telefona numurs), kas būtība funkcionē kā jau jebkurš operatora bloks, taču grafiskā CAPTCHA-testa vietā tajā ir vājš HTML un CSS CAPTCHA-tests. Ja tiek noskaidrots, ka numura operators ir mainījies, šī informācija tiek saglabāta tabulā “savemoved”, lai turpmāk nebūtu jāveic atkārtota operatora noteikšana (tik ilgi, kamēr numura operators nemainīsies atkal).

2.4.2. Paplašinātais bloks

Paplašinātais bloks satur vairākas sūtīšanas funkcijas, kas neatbilst reāliem operatoriem. Šis bloks tiek iedarbināts, ja tabulā “virtuals” pretī adresāta numuram ir ieraksts par paplašinātā bloka lietošanu. Šāds ieraksts saturēs paplašinājuma (funkcijas) nosaukumu un vērtību, kas jāliek adresāta vietā. Paplašinātos blokus iespējams izmantot arī kā rezerves variantu, kā atkāpi (*fallback*).

Lūk vienkāršs paplašinātais bloks “locked”:

```
function sendSMSvirtual_locked($to,$sms,$ignored=0,$params=NULL){  
    return ($to+0);  
}
```

Šī bloka uzdevums ir neko nesūtīt, bet kā kļūdas kodu atgriezt noteiktu vērtību (piemēram, -31). Atgriežamā vērtība tiek novietota adresāta laukā, līdz ar to konkrētu vērtību ir iespējams uzdot tabulā “virtuals”.

Otrs paplašinātais operatora bloks, kas ir pieminēšanas vērts, ir “Okarte”. Šis bloks izmanto LMT sistēmu “OKteilis”, kas gan šobrīd nefunkcionē(18), taču kādreiz caur to varēja neierobežotā daudzumā nosūtīt ziņas OKartes lietotājiem bez grafiskā CAPTCHA-testa izpildes.

2.5. Tīmekļa saskarne “sms.id.lv”

Tīmekļa saskarne realizēta valodā PHP un nodrošina lietotājam ērtu veidu, kā nosūtīt īsziņas un ir pa tiešo (nevis caur vārteju) pieslēgta sistēmas kodolam. Pēc būtības tā ir vienkārša forma, kas savāc visu tās saturu (īsziņu, adresātu un lietotāja lūgumu nosūtīt vai nesūtīt īsziņu vēlāk, ja tas neizdodas uzreiz) un ar šiem parametriem izsauc kodola funkciju, kas atbildīga par īsziņas nosūtīšanu. Taču šai tīmekļa saskarnei ir vēl divas svarīgas funkcijas.

Pirmkārt, tai ir JavaScript papildkods, kas nav obligāti nepieciešams pamatfunkciju veikšanai, taču, ar kura palīdzību lietotājs redz atlikušo brīvo simbolu skaitu īsziņas tekstā un ar kura palīdzību līdz rezultāta saņemšanai par iepriekšējo sūtījumu tiek bloķēta sūtīšanas poga un parādīts nomierinošs uzraksts “Sūtīšana...”, kas katru sekundi papildnās ar vēl vienu punktu, tādā veidā simbolizējot notiekošu procesu. Tas palīdz nervoziem lietotājiem gaidot nesūtīt vienu īsziņu vairākas reizes.

Otra svarīgā funkcija ir tā, ka lietotājam pieprasot sūtīšanas formu, tiek uzģenerēta diezgan sarežģīta forma, taču lietotājam šajā procesā nav jāiejaucas. Automātiski tiek pierēģistrēta sesija, un pārlūks automātiski apstrādā samērā sarežģīto formu. Savukārt, ja notiek mēģinājums viltot formu, uzdodot to par savu, tad atrisināt to nav tik vienkārši. Turklāt, ja tīmekļa saskarne pamana, ka lietotājs īstenībā neatrodas tīmekļa lapā “sms.id.lv”, tā lietotāju uz turieni pārsūta, kā arī paziņo, ka “Sūtīšanu nodrošina sms.id.lv, kas nav nekādā veidā saistīts ar lapu, no kuras Jūs šeit nokļuvāt.”

Šādi var izskatīties tīmekļa saskarnes “sms.id.lv” formas kods:

```
<form method=post action=/send/index.php onsubmit="go(this);">
<table><tr><td>Adresāts:<td><input type=text name=to value="" maxlength=20>
<tr><td valign=top>Teksts:<td><textarea name=sms rows=5 cols=16 wrap=soft
onKeyUp='magicCount(this);'></textarea>

<tr><td align=right><span id="lenleft">143</span><input type=checkbox name=queue
class=checkbox checked><td><small>Nosūtīt vēlāk, ja rodas īslaicīgas problēmas tūlītējā
nosūtīšanā.</small>
<tr><td><input type=submit name=s value="Sūtīt"><td><input type=submit name=reset
value="Dzēst" onclick="eras=1;">
</table>
<INPUT NAME=dyna      VALUE='3c981f12e956433fdd4891c98d10f2cc'    type=hidden>
<input type="hidden"
value="vultus" name="sec4"><input type="hidden" value="synagoga" name="sec1"><input
type="hidden" value="edificium" name="sec2"><input type="hidden" value="abeo"
name="sec0"><!--
<input type="hidden" value="rememdiun" name="sec3"> --><input type="hidden"
value="tonsor" name="sec8"><input type="hidden" value="quilibet" name="sec5"><!--
<input type="hidden"
value="vigor" name="sec9"> --><input type="hidden" value="silva" name="sec7"><input
type="hidden" value="thymbra" name="sec6"><input value='insec' name="single"
type="hidden">
<input type="hidden" value="code537" name="static">
</form>
```

Lietotāja pārlūkprogrammai nododot tādus datus, tīmekļa saskarne “sms.id.lv” sagaidīs ļoti precīzi atbildi no tās un par visām viltojumu aizdomām ziņos administratoram apkopotā veidā.

2.6. Vārteja

Šī sistēmas daļa ir nepieciešama, lai sistēmai varētu pieslēgt papildus saskarnes (arī attālinātas). Tā rakstīta valodā PHP un no vienas puses ir pieslēgta sistēmas kodolam, bet no otras puses tā ir publiski pieejama visā Internetā caur HTTP protokolu. Līdz ar to iespējams sūtīt īsziņas no jebkura servera, ar kuru noslēgta vienošanās par pakalpojumu izmantošanu. Programmas koda, kas nepieciešams, lai pieslēgtos sistēmas vārtejai, paraugs, ir pieejams pielikumā nr. 1.

2.7. WAP serviss

Serviss atrodas uz cita servera (starp citu, "ps.id.lv", no kura kādreiz cēlās "sms.id.lv"), ir rakstīts valodā PHP un sistēmai pieslēdzas caur tās vārteju.

WAP serviss noder cilvēkiem, kas vēlas sūtīt īsziņas uz telefonu no telefona, taču caur Internetu. Lai gan autoram šī ideja pati par sevi šķita mazliet dīvaina, autors tomēr nolēma uzturēt šādu servisu, tā kā tas neprasīja būtiskus resursus un pēc tā bija pieprasījums.

Šis ir pilnīgs sistēmas kods, kur funkcija *sendSMS* ir tā, kas visus datus padod uz vārteju un atpakaļ saņem sūtīšanas rezultātu:

```
<?
require "wap.php";
require "gsm.php";
?>
<wml>
<card id="main_card" title="SMS">
<do type="prev" label="Back">
<prev/>
</do>
<p>
<?
if($p||$r){
//here goes the magic

$d=stripslashes($d);
$r=sendSMS($p?$p:$r,htmlentities($d),true,true);

if (($r->id)>0)
    echo("Your SMS was sent successfully: {$r->text}<br/>");
else
    echo("Your SMS was NOT sent successfully: {$r->text}<br/>");
}
?>
</p>
<p>
<fieldset title="test">
Phone: <input name="phone" size="8" format="*N" maxlength="8"/>
or alias (LMT users only): <input name="phones" size="7" format="*m"
maxlength="20"/><br/>
Text: <input name="text" size="30" format="*m" maxlength="<?=$smsmaxlen; ?>"/>
</fieldset>
<anchor>Send
    <go href="http://wap.ps.id.lv/sms.php" method="post">
```

```

        <postfield name="p" value="$(phone)"/>
        <postfield name="r" value="$(phones)"/>
        <postfield name="d" value="$(text)"/>
    </go>
</anchor>
</p>
<do type="accept" label="Send" name="summ">
    <go href="http://wap.ps.id.lv/sms.php" method="post">
        <postfield name="p" value="$(phone)"/>
        <postfield name="r" value="$(phones)"/>
        <postfield name="d" value="$(text)"/>
    </go>
</do>
</card>
</wml>

```

Īpaša koda iezīme ir tā, ka tam it kā ir divas dažādas sūtīšanas pogas “Send”. Daudzi vecāki telefoni spēj attēlot tikai vienu no tām, tāpēc šādā veidā tiek nodrošināta maksimāla saderība ar dažādu lietotāju aparāturu.

2.8. SMTP serviss

SMTP serviss nodrošina to, ka lietotājiem ir iespēja nosūtīt īsziņu, izmantojot e-pasta klienta programmatūru. To iespējams izmantot arī, lai paziņotu sev par tikko pienākušiem e-pastiem, pāradresējot ienākušos e-pastus uz mobilo tālruni. Lai izmantotu šo pakalpojumu, e-pasti jāšūta uz adresi *numurs@sms.id.lv*, kur *numurs* ir attiecīgā mobilā telefona numurs, kam īsziņa piegādājama.

Šo sistēmas daļu, izmantojot valodu Free Pascal, programmējis Jānis Timma kopā ar bakalaura darba autoru. Tā darbojas saskaņā ar SMTP protokola standarta prasībām.(2) Pirmkārt, sistēmai ir DNS ieraksts “sms.id.lv. 3600 IN MX 5 gate.sms.id.lv.”, kas burtiski nozīmē, ka “@sms.id.lv” e-pastus apkalpo serveris gate.sms.id.lv. Uz šī servera 25 porta atrodas serviss, kas ar klientu “sarunājas” SMTP valodā. Lai varētu nosūtīt īsziņu, nosūtītāja serverim ir jābūt pareizi nokonfigurētai DNS sistēmai, ieskaitot MX ierakstus, vai arī jānorāda sava e-pasta adrese nevis domēnā (piemēram, “kirils@yahoo.com”), bet IP adresē (“kirils@[159.148.105.100]”). Kad formalitātes būs izpildītas, SMTP serviss pieņems adresātu un īsziņas tekstu. Pēc visu datu saņemšanas, tie tiks padoti vārtejai, kam īsā laikā būtu jāsniedz atbildi par īsziņas likteni.

2.9. IRC serviss

Serviss izstrādāts valodā PHP un balstīts uz IRC protokola standartu.(3) Šis serviss paredzēts, lai lietotāji, kas atrodas *IRCnet* tīklā un IRC serveros “chat.118.lv”, “chat.delfi.lv”, varētu nosūtīt īsziņu no tā paša tīkla. Serviss uztur vienu procesa instanci katram tīklam un lietotājam izskatās kā cits IRC lietotājs ar vārdu “irc2sms” *IRCnet* tīklā un vārdu “SMSbot” citos tīklos. Servisam ir vairāki piekļuves līmeņi, kas tiek noteikti nevis ar piekļuves koda palīdzību, bet gan pēc lietotāja datiem (IP adreses, vārda) attiecīgajā IRC serveri. Ja ar šo servisu cenšas sarunāties kāds, kam neuzrādās īsta IP adrese, serviss jau laicīgi paziņo: “Ar slēptu IP adresi / hostu nestrādāšu! Kā lai Tevi atpazīstu? (Raksti '/mode \$w -x')”. Šis serviss, klausās, ko cilvēki saka IRC kanālos, kuros viņš atrodas, kā arī, ko cilvēki saka viņam privāti.

IRC kanālos tas reaģē uz savu vārdu, atbildot ar tekstu “Ko? Vienkārši raksti: !sms <numurs> <teksts> .” Tur iespējams veikt arī aptuveni tās pašas darbības, ko iespējams veikt privāti sarunājoties ar IRC servisa datorprogrammu.

Serviss saprot un spēj izpildīt vairākas komandas:

- !restart
 - × Šī komanda izslēdz un uzreiz atkal ieslēdz servisu attiecīgajā serverī.
- !act <raw>
 - × Šī komanda servisa vārdā nosūta komandu <raw> uz IRC serveri, it kā to būtu izpildījis pats serviss.
- !opme
 - × Piešķir izsaucējam operatora privilēģijas attiecīgajā IRC kanālā.
- !kickme
 - × Izsviež izsaucēju no attiecīgā IRC kanāla.
- !mm
 - × Lūdz servisu sazināties ar lietotāju privāti. Serviss uz šo parasti noreagēs, uzsākot privāto saraksti ar tekstu “yes, user?”
- !explain <code>
 - × Serviss atbildēs ar sistēmas kļūdas koda <code> skaidrojumu latviešu valodā.
- !stats < all | [today|yesterday] [robots|humans] [[!]irc[!]mail] >
 - × Serviss atbildēs ar statistiku par īsziņu skaitu, kas sūtīta caur sistēmu.

- !uptime
 - × Serviss atbildē norādīs laiku, cik ilgi tas jau darbojas (kopš pēdējā pārtraukuma).
- !sms <numurs> <teksts>
 - × Serviss centīsies nosūtīt īsziņu ar saturu <teksts> adresātam <numurs> un paziņos rezultātu.

2.10. Automātiskie servisi

Šī apakšnodaļa apraksta automātiskos sistēmas servissus, kas darbojas ar noteiktu intervālu bez lietotāja vai administratora iejaukšanās. Automātiskie servisi ir būtiska sistēmas sastāvdaļa, un katru no tiem vajadzīgajā brīdī iedarbina programma “cron 2.3.3”, kas paredzēta noteiktu komandu izpildei iepriekš norādītā regulārā laikā (piemēram, reizi gadā, katru otro dienu, 20. un 50. stundas minūtē u.tml.)

2.10.1. CAPTCHA-testu iegūšana un atpazīšana

Katram operatora blokam, kam nepieciešama operatora lapas iepriekšēja pieprasīšana (neatkarīgi no CAPTCHA-testu esamības lapā), sistēmā eksistē valodā PHP rakstīts modulis, kas šo lapu ielādē un saglabā sesijas datus (un OCR rezultātu no CAPTCHA-testa) tabulā “ocrcodes”. Varētu teikt, ka tas ir operatoru bloka “brālis”, jo abu darbības ir vērstas uz vienu mērķi.

Ja operatora lapā ir grafiskais CAPTCHA-tests, jau ielādes laikā tiek noskaidrota šī CAPTCHA-testa attēla adrese, tas lejupielādēts un atpazīts. Atpazīšanai tiek izmantoti OCR skripti, kas arī programmēti valodā PHP. LMT OCR skripts darbībai izmanto vienu Jāņa Timmas rakstītu funkciju.

Pēc OCR veikšanas tiek pārbaudīts, vai CAPTCHA-testa atbilde satur vajadzīgo simbolu skaitu un satur tikai atļautos simbolus. Ja ne – CAPTCHA-tests tiek izmests, bet, ja viss ir kārtībā, tas tiek saglabāts tabulā “ocrcodes”.

Tas, ka kodi tiek pieprasīti un atšifrēti iepriekš, ļauj sistēmai strādāt zibenīgi, kad lietotājs ir ieradies un vēlas nosūtīt īsziņu. Blakus efekts ir tāds, ka moduļiem ir jāspēj aprēķināt, kāds būs maksimālais pieprasītais īsziņu daudzums katram operatoram nākamajā minūtē. Piemēram, LMT nepieciešamo kodu daudzums tiek rēķināts šādi – pēdējā minūtē iztērēto kodu daudzums reizināts ar 4, plus 10:

```
list($weneed)=mysql_fetch_row(mysql_query("select count(*)*4+10 from `processed` where ((UNIX_TIMESTAMP(`whensent`)>UNIX_TIMESTAMP(NOW())-$secsx) and ( (keyid>'') or (status=-6) )) and (`operator`='LMT')"));
$weneed=max($weneed,10);
```

Tātad, ja pēdējā minūtē ir iztērēti 3 LMT kodi un palikuši vēl 5 neizmantoti un derīgi, kopā ir nepieciešams savākt $3*4+10=22$ kodus, tas ir vēl $22-5=17$ kodus.

Šie moduļi tiek iedarbināti katru minūti.

2.10.2. Īsziņu rindas pārvaldnieks

Šis modulis nodrošina to, ka īsziņas, kuras ir ierindojis kodols, kādreiz tiek arī apstrādātas. Vienā reizē tiek paņemtas ne vairāk kā 100 rindā esošas īsziņas to ievietošanas secībā (*FIFO – First In First Out*), noskaidrots saņēmēja operators un, ja attiecīgajam operatoram ir vismaz 5 brīvi kodi (vai arī ir mēģināts izsūtīt mazāk par 2 īsziņām), mēģināts izsūtīt īsziņu vēlreiz.

Neatkarīgi no tā, vai tas izdevies vai nē, tālāk algoritms ņem nākamo īsziņu, taču gadījumā, ja īsziņas sūtīšanas kļūdas kods ietilpa nelabojamo kļūdu sarakstā, īsziņa no rindas tiek dzēsta, un par šo faktu izveidots ieraksts tabulā “dropped”.

Īsziņu rindas pārvaldnieka *pushsmsqueue.php* fragments, kas nodrošina pamata funkcionalitāti:

```
<?
$cdc=0;

$o=mysql_query("select * from `queue` order by `qin` asc limit 100");
while($c=mysql_fetch_array($o)){

    // nodrošināties, lai queue neapēd visus brīvos kodus - kaada tad jeega?
    $pre=addslashes(GSMprovider($c[adrese]));

    list($cho)=mysql_fetch_row(mysql_query("select count(*) from `ocrcodes` where ('out`=0) and (who='$pre') and (UNIX_TIMESTAMP(`in`)>1800)"));
    if(($cho<5)&&($cdc>=2))
        continue;

    unset($ext);
    $ext->mail=$c['mail'];
    $ext->queue=0;
    $ext->sender='RIEXC-queue';
    $ext->inqueued=$c['qin'];

    $op=sendSMS($c[adrese],$c[text],$ext);
```

```

$cdc++;
if($op->id==1)
    mysql_query("delete from `queue` where `uniqid`='{${c[uniqid]}}'");
else {
    if ( in_array($op->id,$failer) ) {
        if(mysql_query("insert into `dropped`
values('','".addslashes($c[1])."', '".addslashes($c[2])."', '{${op-
>id}}', '{${c[4]}}', '{${c[5]}}', NOW())"))
            mysql_query("delete from `queue` where `uniqid`='{${c[uniqid]}}'");
        else
            echo('ERR while dropping from queue'.mysql_error()."\n");
    } else
        mysql_query("update `queue` set rr='{${op->id}}' where `uniqid`='{${c[uniqid]}}'");
    }
}
?>

```

Kā jau tas redzams, īsziņu rindas pārvaldnieks ir rakstīts valodā PHP. Šis automātiskais serviss papildus saviem tiešajiem pienākumiem ir atbildīgs arī par vairs nevajadzīgo CAPTCHA-testa attēlu izdzēšanu no sistēmas, kā arī paziņojuma nosūtīšanu uz administratora mobilo telefonu gadījumā, ja pārtraukusi funkcionēt kāda no CAPTCHA-testu atpazīšanas sistēmām.

Modulis tiek iedarbināts reizi 5 minūtēs.

2.10.3. Starptīkserveru datubāzes atjaunināšana

Šis modulis ir rakstīts valodā PHP un ir atbildīgs par starptīkserveru saraksta iegūšanu no “samair.ru” datubāzes(19) un starptīkserveru kvalitātes pārbaudi. Modulis šobrīd ir atbildīgs tikai par HTTP starptīkserveru datubāzes uzturēšanu (tabula “proxy”).

Autors uzskata, ka programmēšanas ziņā starptīkserveru saraksta iegūšana nav nekas īpaši aizraujošs, taču kvalitātes pārbaude gan ir pieminēšanas vērta. Tā norisinās, caur katru starptīkserveri pieprasot divas uz sistēmas servera esošas datnes, no kurām viena ir attēls, bet otra ir PHP fails. Tad tiek pārbaudīts, vai abi faili ir saņemti bez izmaiņām (ja ne – starptīkserveris tiek atzīmēts kā nestrādājošs), un izmērīts laiks, kas patērēts šo failu saņemšanai. No tā tiek aprēķināts ātrums, un visi iegūtie dati tiek saglabāti tabulā “proxy”.

Savukārt PHP skripts, kas tiek pieprasīts caur starptīkserveri, tabulā reģistrē starptīkservera nosūtītos HTTP iesākus HTTP_X_FORWARDED_FOR, HTTP_VIA, HTTP_PROXY_CONNECTION, HTTP_USER_AGENT un REMOTE_ADDR un atgriež starptīkserverim HTML failu, ko pārbaudīs sistēma.

PHP skripts, kas tiek pieprasīts caur starpniekserveri:

```
<?
require 'mysql.php';

$dati=array(
'HTTP_X_FORWARDED_FOR',
'HTTP_VIA',
'HTTP_PROXY_CONNECTION',
'HTTP_USER_AGENT',
'REMOTE_ADDR',
);

$test=true;
$id=addslashes($id);

list($cnt)=mysql_fetch_row(mysql_query("select count(*) from proxy where `id`='".$id'"));
if($cnt!=1)die('bad_id');

for($i=0;$i<sizeof($dati);$i++){
    $t=addslashes($dati[$i]);
    $test&=mysql_query("update proxy set `".$dati[$i].`='".$t.'" where `id`='".$id'");
}

readfile('paraugs.htm');
?>
```

Modulis tiek darbināts vienu reizi dienā. Tabulas “proxy_socks” un “proxy_tor” šobrīd automātiski atjauninātas netiek.

2.10.4. Rezerves kopēšana

Šis modulis parūpējas par to, lai visi sistēmas faili un datubāze (izņemot tabulu “archived”) tiktu saglabāti drošā vietā. Rezerves kopēšanas skripts rakstīts *bash* valodā un faili tiek saspiesti ar *gzip* un *tar* programmām.

Jāpiebilst, ka rezerves kopēšanas serviss atbild arī par to, lai īsziņu teksti žurnālā neglabātos mūžīgi. Ātrdarbības nolūkā vismaz mēnesi vecās īsziņas tiek pārvietotas uz tabulu “archived”, kur īsziņas teksts tiek neatgriezeniski šifrētas, izmantojot vienvirziena šifrēšanas algoritmu. Primitīvākais pierādījums, ka šis algoritms ir vienvirziena, ir tas, ka no vismaz 160 baitiem informācijas tiek iegūti un saglabāti tikai 32 baiti.

Rezerves kopēšanas modulis darbojas reizi dienā.

2.11. Administrēšanas rīki

Administrēšanas rīki tiek izmantoti, lai atvieglotu sistēmas pārvaldību – novērotu sistēmu un dzēstu/pievienotu bloķētās konfigurācijas. Protams, visi šie rīki prasa cilvēka līdzdalību.

2.11.1. Bloķēto konfigurāciju administrēšanas panelis

ID	Piev. dat.	Autors	Iemesls	Piezīmes	Vārti	IP*	Pārlūks
Numurs	SMS	Datums	Operators	Mail			
						213.175.88	

1 ieraksts no 60 (1.67%)

D	Numurs	SMS	Datums	Operators	Mail	Vārti	IP	Pārlūks
15279	2009-05-24 15:45:50	GATE sms.id.lv	autoban: complex flood	2 sms with difference factor compared to last sms with text "loxgd" less than 5 to number 288261		[e] [p] [x]		
+	-	-	-	-	-	-	213.175.88	-

2.4. att. Bloķēto konfigurāciju administrēšanas panelis

Bloķēto konfigurāciju administrēšanas paneļa uzdevums ir ļaut sistēmas administratoram viegli pārvaldīt (dzēst, modificēt, pievienot) bloķētās konfigurācijas. Panelis ir rakstīts valodā PHP. To pamatā izveidojis Jānis Timma un nedaudz savu roku pielicis arī bakalaura darba autors. Šī iemesla dēļ darbā netiks detalizēti apskatīts šis rīks.

2.11.2. Īsziņu plūsmas pārraudzības rīks

```

SMS 34b86585cd14359a53f3fe9755d16967 (22/143)
FROM 85.15.2 (Mozilla/5.0 (Windows; U; Windows NT 6.0; en-US; rv:1.9.0.10) Gecko/2009042316 Firefox/3.0.10) via sms.id.lv
TO 282 24 (TELE2) via forx with code 3 5 (a5) at 2009-05-29 11:1
RESULT is 1: "Sent successfully." in 0.41252s

```

2.5. att. Īsziņu plūsmas pārraudzības rīks

Šis rīks reālajā laikā attēlo īsziņu konfigurācijas, kas iet cauri sistēmai. Jāpievērš uzmanība, ka pats īsziņas teksts tiek attēlots tikai tajā gadījumā, ja sistēma ir pamanījusi tajā kaut ko rupju, vai arī teksts ir nesakarīgs. Rīks rakstīts valodā PHP.

2.11.3. OCR pārraudzības rīks

SMS queue has 0 SMSes.

0 were dropped today.

ar Sat May 23 03:19:46 EEST 2009 tiek raadi jaunie stati

Cached: 58 (40 free), newest: 6 k

LMT code detection suckage [75 _=-_ 3050] percentage: 2.4%

TELE2 code detection suckage [47 _=-_ 4429] percentage: 1.05%

5 biggest failures of the recent time ;p

MORE RECENT vvvvvv

t/s	sid	img	f()	keyid	o
2009-05-29 08:5	TEST-S1		3	C n	TELE2
2009-05-29 08:5	TEST-S1		3	C n	TELE2
2009-05-29 07:0	sms.id.lv		K	L	LMT
2009-05-28 16:0	sms.id.lv		7	L	LMT
2009-05-28 15:4	sms.id.lv		6	9 n	TELE2
2009-05-28 14:0	sms.id.lv		7	G n	TELE2
2009-05-28 13:2	sms.id.lv		B	4F n	TELE2
2009-05-28 12:3	sms.id.lv		L	7 n	TELE2
2009-05-28 12:2	sms.id.lv		8	9 n	TELE2
2009-05-28 11:4	sms.id.lv		P	L	LMT

LESS RECENT ^^^^^^

Last 5 userspace failures on code

2009-05-28 15: (-4) TELE2

2009-05-28 14: (-4) TELE2

2009-05-28 13: (-4) TELE2

2009-05-28 12: (-4) TELE2

2009-05-28 12: (-4) TELE2

2.6. att. OCR pārraudzības rīks

Šis rīks ir veidots PHP un attēlo dažādus statistikas datus. Tā funkcijas ir sekojošas:

- pieprasa informāciju no tabulas “queue”, lai noskaidrotu, cik īsziņu šobrīd ir rindā;
- pieprasa informāciju no tabulas “dropped”, lai noskaidrotu, cik īsziņu šodien ir izmestas no rindas;
- attēlo tabulā “stats” esošo informāciju par veiksmīgiem un neveiksmīgiem OCR mēģinājumiem;
- pieprasa informāciju no tabulas “ocrcodes”, lai pāradītu, cik CAPTCHA-testi ir atrisināti un gatavi lietošanai;
- parāda pēdējos 10 tabulas “ocrfailures” ierakstus;
- parāda pēdējās 5 reizes, kad OCR vai CAPTCHA-testa kļūda ir reāli ietekmējusi lietotāju.

2.11.4. Statistikas rīks

2009-05-21	laikā 0	dažādi cilvēki sūtīja 15	īszīpas, no kurām 0	tika apsolīts nosūtīt vēlāk, 15	netika nosūtītas; 0	reizes sūtīšana atteikta dēļ bloka. Uzlikti 0	automātiskie bani.
2009-05-22	laikā 0	dažādi cilvēki sūtīja 9	īszīpas, no kurām 0	tika apsolīts nosūtīt vēlāk, 9	netika nosūtītas; 0	reizes sūtīšana atteikta dēļ bloka. Uzlikti 0	automātiskie bani.
2009-05-23	laikā 735	dažādi cilvēki sūtīja 1706	īszīpas, no kurām 4	tika apsolīts nosūtīt vēlāk, 893	netika nosūtītas; 124	reizes sūtīšana atteikta dēļ bloka. Uzlikti 11	automātiskie bani.
2009-05-24	laikā 1535	dažādi cilvēki sūtīja 3918	īszīpas, no kurām 4	tika apsolīts nosūtīt vēlāk, 1725	netika nosūtītas; 376	reizes sūtīšana atteikta dēļ bloka. Uzlikti 26	automātiskie bani.
2009-05-25	laikā 341	dažādi cilvēki sūtīja 1244	īszīpas, no kurām 3	tika apsolīts nosūtīt vēlāk, 407	netika nosūtītas; 30	reizes sūtīšana atteikta dēļ bloka. Uzlikti 7	automātiskie bani.
2009-05-26	laikā 302	dažādi cilvēki sūtīja 1202	īszīpas, no kurām 2	tika apsolīts nosūtīt vēlāk, 335	netika nosūtītas; 27	reizes sūtīšana atteikta dēļ bloka. Uzlikti 5	automātiskie bani.
2009-05-27	laikā 283	dažādi cilvēki sūtīja 1105	īszīpas, no kurām 14	tika apsolīts nosūtīt vēlāk, 306	netika nosūtītas; 32	reizes sūtīšana atteikta dēļ bloka. Uzlikti 6	automātiskie bani.
2009-05-28	laikā 306	dažādi cilvēki sūtīja 1103	īszīpas, no kurām 3	tika apsolīts nosūtīt vēlāk, 323	netika nosūtītas; 25	reizes sūtīšana atteikta dēļ bloka. Uzlikti 4	automātiskie bani.

2.7. att. Statistikas rīks

Šis rīks ir rakstīts valodā PHP un parāda statistikas pārskatu pa dienām. Tiek izmantotas tabulas “processed” un “blockids”, un par katru dienu tiek parādīts:

- datums;
- lietotāju skaits;
- īszīņu skaits;
- nenosūtīto īszīņu skaits;
- ierindoto īszīņu skaits;
- dēļ bloka nenosūtīto īszīņu skaits;
- konfigurāciju skaits, ko automātiski nobloķējis sistēmas kodols.

Rīka pirmkods pieejams pielikumā nr. 2.

3. SISTĒMAS UZTURĒŠANA UN ATTĪSTĪŠANA

Tā kā sistēma nedarbojas, bet šī darba mērķis ir atkal atjaunot tās funkcionalitāti, tad šajā nodaļā uzreiz pēc esošo alternatīvu izpēti tiek veikta iespējamo sistēmas attīstības scenāriju izpēte un labākā scenārija izvēle, kā arī ieviešana. Nodaļā tiek aprakstīti padarītie un iesāktie darbi, kā arī ieskicētas nākotnes iespējas un plāni.

3.1. Mūsdienās pieejamās alternatīvas

Kad 2008. gadā pārtrauca funkcionēt sistēma “sms.id.lv”, tukšo nišu ātri vien masveidā aizpildīja dažādi citi pakalpojumu sniedzēji. Meklējot šādas sūtīšanas iespējas Internetā, izmantojot Interneta meklētājus, var redzēt, ka pieejamas diezgan daudz alternatīvas, tāpēc autors uzskatīja par būtisku tās izpētīt.

Vairākas no šīm alternatīvām nav pieejamas lietotājiem, kas nespēj vai tehnisku ierobežojumu dēļ nevar redzēt attēlus. Šādu lietotāju īpatsvars ir vērā ņemams, lai arī viņi nav pārsvarā. Lai pārliecinātos par šādu lietotāju iespēju nosūtīt īsziņas no dažādām mājaslapām, izmantoju pārlūkprogrammu “lynx 2.8.6rel.4”. Šo programmu iespējams pielāgot darbam neredzīgajiem lietotājiem, piemēram, ieviešot simbolu attēlošanu Braila rakstā vai vārdu pārvēršanu skaņā. Līdz ar to programma labi atspoguļo to, kādā mērā lapu var uztvert lietotājs ar redzes traucējumiem.

Zemāk apkopoti fakti par mūsdienās pieejamajiem populārākiem bezmaksas īsziņu sūtīšanas servisiem no Interneta.

3.1.1. “sms.lv – Bezmaksas SMS sūtīšana”

Šis pakalpojumu sniedzējs piedāvā nogādāt īsziņas operatoriem LMT, TELE2 un BITE, taču, lai to lietotu noteikti jāatrisina CAPTCHA-tests, kas atkarīgs no izvēlēta operatora. Turklāt, operatora noteikšana ir programmēta JavaScript, kas autora testā ar pārlūkprogrammu *lynx* nedarbojās.

Tīmekļa lapā, no kuras iespējama sūtīšana, atrodas viena neuzbāzīga grafiskā reklāma. Sūtītājs pieejams latviešu un krievu valodā.(20)

Īsziņa uz mobilo telefonu pienāk dažu sekunžu laikā, un pirms lietotāja ievadītā teksta ir pievienots teksts “sms.lv”. Īsziņa uzrādās kā sūtīta no Interneta.

Nosūtīt īsziņu

Drošības kods nesakrīt

Telefona numurs
 → [Uzzināt operatoru](#)

Operators
☒ LMT ☐ Tele2 ☐ BITE

Īsziņas teksts (paliek simbolu: 123)

Jūsu vārds

Drošības kods 

Nosūtīt

* sadarbībā ar  LMT



3.1. att. Tīmekļa lapa “sms.lv” (20)

```
kirils@laptop: ~
File Edit View Terminal Tabs Help
<<< SMS.lv - Bezmaksas SMS sutisana (p1 of 2)

[head_logo.jpg] RU | LV

Nosūtīt īsziņu

Izveleties operatoru

Telefona numurs

+371 26036916 [hand.gif] Uzzinat operatoru

IFRAME: ifrm

Operators
(*) LMT ( ) Tele2 ( ) BITE
Iszinas teksts (paliek simbolu: 123)

(NORMAL LINK) Use right-arrow or <return> to activate.
Arrow keys: Up and Down to move. Right to follow a link; Left to go back.
H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

3.2. att. Tīmekļa lapa “sms.lv” caur pārlūkprogrammu *lynx* (20)

3.1.2. “esms.lv – tavs personīgais īsziņu centrs”

Šo pakalpojumu piedāvā SIA SalviNET, un patīkami, ka šeit lietotājam nav jārisina CAPTCHA-testi, lai sūtītu īsziņas. Viss veiksmīgi darbojas arī, izmantojot *lynx*, turklāt īsziņas iespējams sūtīt uz jebkuru mobilo telefonu operatoru Latvijā. Tiek piedāvāts arī atlikt sūtīšanu līdz konkrētam laikam, taču testēšanas brīdī šī funkcionalitāte nedarbojas. Tīmekļa lapā nav reklāmu, un tai ir patīkama un viegli lietojama saskarne, kas pieejama latviešu valodā.(21)

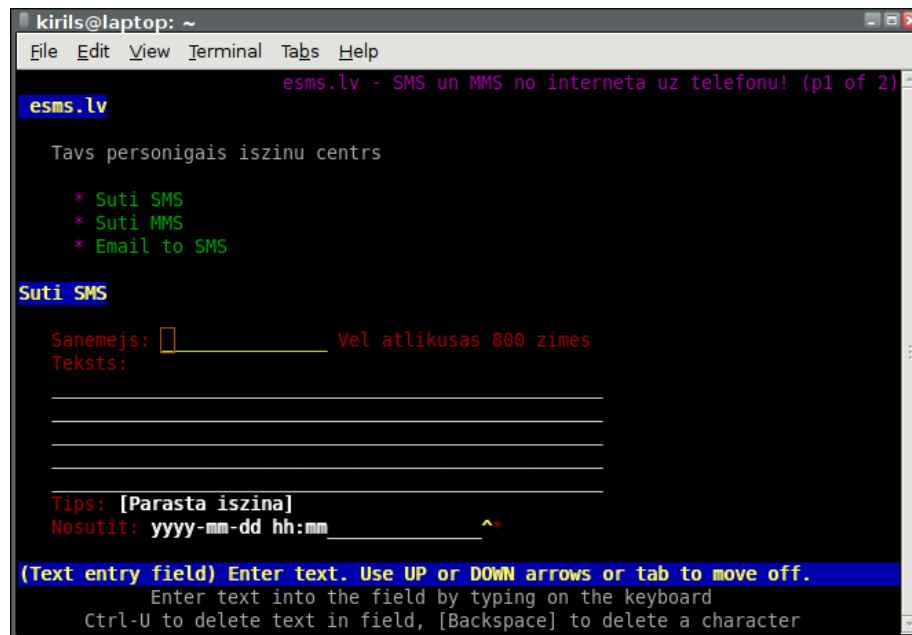
Īsziņa uz mobilo telefonu pienāk dažu sekunžu laikā un pēc lietotāja ievadītā teksta ir pievienots teksts “* *Si sms ir nosutita no www.esms.lv. SIA SalviNET neatbild par sis sms saturu.* *”. Kā īsziņas sūtītājs uzrādās numurs +37122071473 vai +37122071472 (iespējams, arī citi).

Neskatoties uz visai pozitīvo sākotnējo novērtējumu, šim pakalpojumam ir būtiski trūkumi – ar tā palīdzību no vienas IP adreses ir iespējams nosūtīt tikai vienu īsziņu piecās minūtēs, kas var būt stipri par maz, ja operatīvi jāsazinās ar vairākiem cilvēkiem. Otru draudu autors saskata tajā, ka “Saņēmējam ir iespēja bloķēt bezmaksas īsziņas nosūtot komandu “STOP” (No apmaksātam SMS ar reklāmu nevarēs atteikties)”(22), kas visdrīzāk norāda uz to, ka nākotnē ir plānota reklāmas izsūtīšana šī pakalpojuma lietotājiem.



The screenshot shows the 'esms.lv' website interface. At the top left is a logo with a cartoon character and the text 'esms.lv tavs personīgais īsziņu centrs'. Below the logo are three tabs: 'sūti sms' (selected), 'sūti mms', and 'email to sms'. The main content area is divided into two sections. The left section, titled 'Šis tas', contains a list of links: 'Lietošanas noteikumi', 'Izmaiņu saraksts', and 'Darāmo Darbu Saraksts'. The right section, titled 'Sūti SMS', contains a form for sending an SMS. It includes a red warning message: 'SMS nav nosūtīta, jo viena sms reizi piecas minutes ...'. Below this is a field for the phone number 'Saņēmējs: 26210987' and a status 'Vēl atlikušas 800 zīmes'. There is a large text area for 'Teksts:'. Below the text area is a dropdown menu for 'Tips: Parasta īsziņa'. At the bottom of the form is a field for 'Nosūtīt: yyyy-mm-dd hh:mm' with an asterisk, and a blue button labeled 'nosūtīt'. At the very bottom of the page, there is a footer with copyright information: 'esms.lv ©2009 Visas Tiesības Aizsargātas. • Pakalpojumu nodrošina SIA SalviNET • Hostingu nodrošina No.Lv.'

3.3. att. Tīmekļa lapa “esms.lv” (21)



3.4. att. Tīmekļa lapa “esms.lv” caur *lynx* (21)

3.1.3. “smsbox.lv – SMS par brīvu”

Neskatoties uz to, ka šis pakalpojums prasa, ka lietotājs atrisina CAPTCHA-testu(23), lai sūtītu īsziņas, šī testa rezultāts tiek šķietami ignorēts – ja vien atbildes laukā kaut kas vispār ir ievadīts, lietotāja īsziņa tiek nosūtīta neatkarīgi no tā, vai pēc CAPTCHA-testa secināms, ka sūtītājs ir cilvēks vai datorprogramma. Diemžēl tests ar *lynx* nenotika tik veiksmīgi kā cerēts. Īsziņu nosūtīšanai nepieciešams JavaScript, jo nosūtīšanas forma bez tā nestrādā. Īsziņu nosūtīšanas lapā ir viena teksta reklāma, un pats pakalpojums pieejams latviešu, krievu un angļu valodā. Šis pakalpojumu sniedzējs piedāvā nosūtīt īsziņas uz LMT, TELE2 un BITE tīklu, taču satur arī slēptu iespēju sūtīt īsziņas uz jebkuru mobilo telefonu operatoru Latvijā.

Īsziņa uz mobilo telefonu pienāk apmēram desmit sekunžu laikā un pirms teksta, ko lietotājs vēlējies nosūtīt, ir ietverts teksts “*smsBox.lv#*”. Testu rezultātā noskaidrots, ka īsziņa var uzrādīties gan kā sūtīta no Interneta, gan arī kā sūtīta no numura +37120079476.

Šīs sistēmas trūkumi ir jau pieminētā nespēja nosūtīt īsziņu, izmantojot *lynx*, un nesaskaņota reklāma – jau dažas dienas pēc tam, kad autors pirmo reizi bija izmēģinājis *smsbox.lv*, uz autora tālruņa numuru pienāca reklāma, kuras sūtītājs sevi identificēja kā “*smsbox.lv*”. Pēc autora uzskatiem tas ir ļoti liels mīnuss, kam būtu jāattur lietotājus no šī pakalpojuma izmantošanas.

smsBox

LV RU EN

Send SMS in Latvia for free

Phone number:

26030303

☐ LMT
☐ Tele2
☒ BITE
☐ None

SMS Message:

Hello GSM world.

SMS sender name:

Character count: 114

Security code

COMPANIES rusting

Send SMS

Free Forex Demo Account

Learn To Trade Forex Risk-Free With A Free Practice Account. Start Now!

<>

Ads by Google

Get your start in the Forex Market - soon will added daily Forex Snapshot, Forex live currency rates and more trading tools.

3.5. att. Tīmekļa lapa “smsbox.lv” (23)

kirils@laptop: ~

File Edit View Terminal Tabs Help

SMS par brīvu (p1 of 2)

LV RU EN

Send SMS in Latvia for free

Phone number:

(*) LMT () Tele2 () BITE () None

SMS Message:

SMS sender name:

Character count: 144

Security code

(Text entry field) Enter text. Use UP or DOWN arrows or tab to move off.

Enter text into the field by typing on the keyboard

Ctrl-U to delete text in field, [Backspace] to delete a character

3.6. att. Tīmekļa lapa “smsbox.lv” caur lynx (23)

3.1.4. “bezmaksassms.com – nosūtīt bezmaksas īsziņas uz Latviju”

Lai gan, lietojot šo pakalpojumu, lietotājam nav jārisina CAPTCHA-testi, tas ir tālu no pilnības. Pirmkārt, tests ar *lynx* gāja ļoti lēni, jo lapā ir grūti pārvietoties, un tas apstājās pie adresāta norādīšanas īsziņai – to nebija iespējams izdarīt, jo šī funkcionalitāte realizēta izmantojot JavaScript. Testējot ar ikdienišķāku pārlūkprogrammu, īsziņu izdevās “nosūtīt”, lai gan tikai pēc tam, kad autors bija adresātu pievienojis savai adrešu grāmatai (kas tiek dzēsta katru reizi, aizverot pārlūkprogrammu) un ievadījis tā vārdu, kā arī norādījis diezgan daudz datus par sevi.

Diemžēl īsziņas vietā saņēmējs uz telefona aparātu saņēma reklāmu un aicinājumu samaksāt par īsziņas piegādi. Līdz ar to, lai gan tehniski šis ir risinājums, ar kura palīdzību iespējams **nosūtīt** bezmaksas īsziņas, to saņemšana gan maksā naudu.

Tīmekļa lapā ir diezgan plašs reklāmas laukums un tā pieejama latviešu valodā ar diezgan daudz gramatikas kļūdām, kas jau tā grūti lietojamo lapu padara vēl grūtāk lietojamu.

(24)

BEZMAKSAS.SMS

SMS Latvija Filiālēm vairum SMS

Adrešu grāmatu
Pievienot

Uz : <Izvēlieties kontaktu no jūsu adrešu grāmatu>
No : <Ievadiet savu tālruna numuru>
Ziņa :
Sūtīt savu īsziņu!

Reklāma

Ātrs kredīts
Aizpildiet pieteikumu Internetā
un saņemiet naudu pēc pāris
minūtēm!
www.ferratum.lv

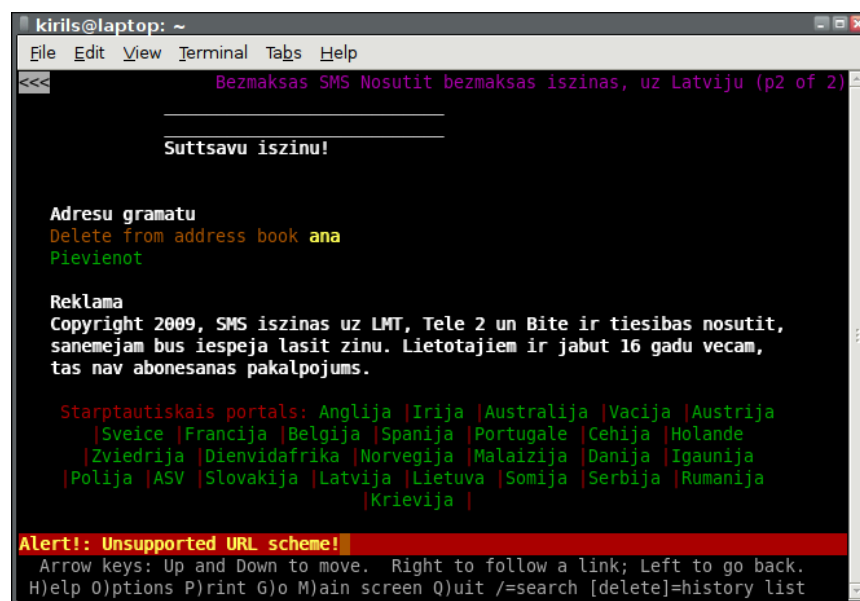
Ātrākais Naudas Kredīts
Sūti SMS un saņem naudu
Ātrākais kredīts internetā
www.smscredit.lv

Sūti SMS par Velti
11(1,9815/n). Piesakies un sūti
15 Bezmaksas SMS Ikdienā.
www.freeTonesforYourMobile

Mobilais Mārketingas
Plašākā segmentētā datu bāze
Saņem savu mērķauditoriju
smsbuzz.lv/reklama

Google reklāmas

3.7. att. Tīmekļa lapa “bezmaksassms.com” (24)



3.8. att. Tīmekļa lapa “bezmaksassms.com” caur *lynx* (24)

3.1.5. Citas alternatīvas

Autoram izdevās novērot arī citu alternatīvu tīmekļa lapu, ar kuru palīdzību paredzēts nosūtīt bezmaksas īsziņas no Interneta, eksistenci, taču neviena no tām nebija pietiekami funkcionāla, lai būtu pieminēšanas vērta.

3.2. Problēmu identificēšana

Veicot vienību testēšanu, tika identificētas problēmas daudzās sistēmas sastāvdaļās. Vēlāk veicot sistēmas akcepttestēšanu tika konstatētas arī mazākas funkcionālas nepilnības.

3.2.1. Operatoru bloki ir novecojuši

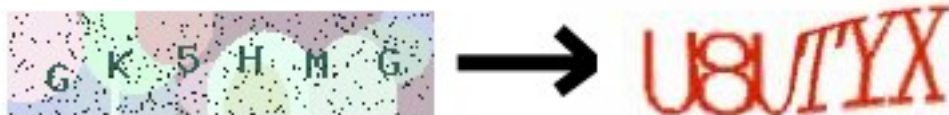
Operatoru bloki, kas atbildīgi par sasaisti ar katra operatora vārteju vairs nedarbojas, jo operatoru mājaslapas un vārteju parametri ik pa laikam mēdz mainīties, un tāpēc nepieciešams nepārtraukts darbs, lai operatoru bloki būtu funkcionāli un atbilstu mūsdienu situācijai.

3.2.2. CAPTCHA-testu atpazīšana nefunkcionē

Teksta atpazīšana (OCR) no grafiskajiem CAPTCHA-testiem nedarbojas LMT un TELE2 atpazīšanas moduļos, jo LMT un TELE2 ir nomainījuši (attiecīgi 3.9. un 3.10. attēli) CAPTCHA-testus.



3.9. att. LMT lietoto CAPTCHA-testu maiņa



3.10. att. TELE2 lietoto CAPTCHA-testu maiņa

3.2.3. SMTP serviss nedarbojas

Tā kā ir mainījušies sistēmas “sms.id.lv” vārtejas parametri, arī SMTP serviss vairs nedarbojas korekti. Turklāt, veicot testēšanu, autors atklāj, ka SMTP serveris, kas nodrošina sasaisti starp SMTP protokolu un – caur vārteju – sistēmu “sms.id.lv”, nav drošs.

3.2.4. IRC serviss tērē pārāk daudz resursu

Šobrīd, IRC servisam atrodoties vairākos IRC serveros, tiek nelietderīgi tērēti sistēmas resursi, jo serviss, kas ļauj lietotājiem nosūtīt īsziņu no IRC, tiek izmantots ļoti reti. Katra IRC servera lietotāji tiek apkalpoti atsevišķā procesā, kas ir nepārtraukti aktīvs.

3.2.5. Citas nepilnības

Statiskā operatora piederības datubāze ir novecojusi, jo Sabiedrisko pakalpojumu regulēšanas komisija ir piešķīrusi jaunus numerācijas diapazonus mobilo sakaru operatoriem.

Īsziņu rindā ir vairāk kā 10 tūkstoši īsziņu, kas tur sakrājušās pēc tam, kad pārstāja funkcionēt sistēmas operatoru bloki.

Joprojām nav iespējas nosūtīt īsziņas uz Amigo un BITE. Nedarbojas paplašinātais operatoru bloks “Okarte”.

Starptiekserveru datubāzes atjaunināšana nefunkcionē, taču testēšana joprojām strādā.

3.3. Iespējamie attīstības scenāriji

Balsoties uz sistēmas “sms.id.lv” esošo stāvokli, kā arī pieejamajām alternatīvām, autors redz vairākus iespējamus pētāmās problēmas risinājumus.

3.3.1. Neatjaunot sistēmas darbību

Viena iespēja būtu likvidēt sistēmu, taču vienoties ar kādu alternatīvajiem pakalpojumu sniedzējiem par viņu pakalpojuma izmantošanu. Piemēram, *esms.lv* būtu derīgs variants, ja tiktu palielināts īsziņu skaits, ko drīkst nosūtīt piecās minūtēs.

Taču šāds risinājums būtu pretrunā ar uzstādīto bakalaura darba mērķi un turklāt lietotājiem nāktos pieciest reklāmas, lietojot pakalpojumu un lasot īsziņas savā telefona aparātā.

3.3.2. Atjaunot sistēmas darbību caur maksas vārtejām

Otra iespēja ir aizvietot sistēmas operatora blokus ar vienu vienīgu vārtejas bloku, kas savienots ar kāda īsziņu pakalpojumu sniedzēja vārteju. Tādā veidā varētu iegūt ļoti augstu sistēmas uzticamību un pielāgojamību, taču tas būtu maksas pakalpojums un pie nosacījuma, ka sistēma dienā apkalpo 5-10 tūkstošus īsziņu, varētu izmaksāt vismaz vienu tūkstoti latu mēnesī.

Tā kā sistēma “*sms.id.lv*” ir nekomerciāls projekts, šādi finanšu resursi nav pieejami, lai gan, vērtējot pēc pārējiem aspektiem, tas būtu labākais variants.

3.3.3. Atjaunot sistēmas darbību, bāzējot to uz operatoru e-pasta vārtejām

Trešā iespēja ir izmainīt operatora blokus tā, lai izejošo īsziņu sistēma ģenerētu tikai e-pasta formā. Tas ļautu ātri atjaunot vairāku operatoru atbalstu, taču

- vairākiem mobilo sakaru operatoriem nav pieejams pakalpojums “īsziņa no e-pasta (norādot saņēmēja numuru) uz tālruni”, līdz ar to nebūtu iespējams sūtīt īsziņas uz šo operatoru tīkliem;
- e-pasta nosūtīšana neparedz veidu kā laicīgi uzzināt, vai īsziņu ir veiksmīgi izdevies nosūtīt vai nav, līdz ar to sistēmas uzticamība samazinātos;
- jau šobrīd ir pieejamas daudz alternatīvas, kas izmanto tieši šo metodi, taču to zemās kvalitātes dēļ tās netiek lietotas.

3.3.4. Atjaunot sistēmas darbību pēc vecā modeļa

Vēl vienas scenārijs ir pēc iespējas novērst visas problēmas un atjaunot sistēmas darbību analogiski vecajam modelim. Tas prasītu stipri lielāku darba ieguldījumu nekā citi scenāriji, taču iegūtā sistēmas kvalitāte stipri pārsniegtu izmaksas.

Vērtējot problēmu būtiskumu, ir viennozīmīgi skaidrs, ka jāatjauno operatoru bloki un CAPTCHA-testu izpilde. Izpētot statistiku (pielikums nr. 4.) redzams, ka SMTP ir plaši izmantots, bet IRC – ne pārāk. Tā kā šī brīža IRC realizācija prasa atsevišķu klausītāj-procesu uz katru IRC serveri, bet SMTP – tikai vienu pamata procesu, tad pareizais virziens būtu salabot SMTP saskarni, bet slēgt IRC saskarnes darbību.

3.4. Izstrādātie risinājumi

Lai sasniegtu bakalaura darba mērķi, izmantojot ievadā aprakstītās metodes, tika veikti šajā apakšnodaļā uzskaitītie uzlabojumi pēc pēdējā scenārija. Pēdējais scenārijs izvēlēts tāpēc, ka tas vislabāk atbilst darba mērķim, tam ir zemas izmaksas un salīdzinoša augsta kvalitāte, bet katrs no pārējiem scenārijiem neatbilda vismaz vienam no šiem svarīgajiem parametriem.

3.4.1. Operatoru bloki

Identificējot problēmas, tika noskaidrots, ka Triatel operatora bloks joprojām darbojas, taču ne TELE2, ne LMT operatoru bloks nedarbojas. Tam par iemeslu kalpojis tas, ka LMT un TELE2 ir mainījuši savu mājaslapu informācijas dizainu un izvietojumu. Operatoru bloki atjaunoti, izpētot jauno mājaslapu uzbūvi. Šai problēmai nebija pieejami nekādi alternatīvi risinājumi, lai gan autors tos meklēja. Pētot, vai panākt nepieciešamo (piegādāt īsziņu attiecīgā operatora mobilā telefona numuram) nav iespējams citādi, autors secināja, ka diemžēl šobrīd tās ir vienīgās iespējas, ja neskaita 3.1. nodaļā apkopotās alternatīvās sistēmas.

Pēc labojuma, TELE2 operatora bloka algoritms ir aprakstāms šādi:

1. No datubāzes tabulas “ocrcodes” tiek pieprasīts jaunākais neizmantotais CAPTCHA-testa atrisinājums, kas der operatoram TELE2.
 - a) Ja tāda nav, tiek atgriezts sistēmas kļūdas kods -6 un operatora bloks beidz darbu.
 - b) Ja tāds ir atrasts, operatora bloks tabulā “ocrcodes” atzīmē, ka šis kods ir izmantots, norādot tā izmantošanas laiku, un turpina darbu.
2. Oriģinālais CAPTCHA-tests tiek nokopēts uz kļūdu direktoriju (gadījumam, ja sistēma apstāsies pusceļā).
3. Tiek pārbaudīts vai CAPTCHA-testa atrisinājums atbilst pareizajam garumam.
 - a) Ja nē, tiek atgriezts sistēmas kļūdas kods -4 un darbs tiek beigts.
 - b) Ja jā, operatora bloks sagatavo īsziņas nosūtīšanu.

4. Sistēma izveido vajadzīgos HTTP iesākus un sagatavo nosūtāmos datus. Tiek noteikts īsziņas garums, pievienota informācija par sesiju no oriģinālās sesijas, kas saglabāta tabulā “ocrcodes”, kad tika iegūts CAPTCHA-tests, kā arī pievienoti standarta nosūtāmie dati.

5. Kļūdu direktorijā tiek izveidots fails, kurā uzreiz pēc saņemšanas tiek ierakstīts viss, ko operatora bloks saņem no operatora servera.

6. Sistēma pieslēdzas serverim “utils.trigger.ee” un nosūta HTTP pieprasījumu ar sagatavotajiem iesākumiem un datiem.

7. Sistēma lasa atbildi no servera un raksta saņemto failā līdz

- a) serveris aizver savienojumu;
- b) sistēma atpazīst atbildi, kas liecina par sūtījuma statusu.

8. Sistēma var atpazīst vairākas atbildes pēc to fragmentiem:

- a) “Jūsu īsziņa tika pievienota īsziņu rindai.”, sistēma atgriež kļūdas kodu 1.
- b) “The message you entered is too long.” – kļūdas kods -2.
- c) “Lūdzu, ierakstiet ” – kļūdas kods -11.
- d) “Kods nav pareizs” – kļūdas kods -4.
- e) “numurs nav pareizs” – kļūdas kods -16.
- f) “kļūda, pievienojot SMS īsziņu rindai” – kļūdas kods -1.
- g) “Īsziņas tiek sūtītas tikai Tele2 tīklā” – kļūdas kods -9.
- h) “numurs atrodas neatpazītā tīklā” – kļūdas kods -16.

9. Ja no sistēmas ir saņemta kaut viena no augstākminētajām atbildēm, tad tiek izdzēsts fails ar saņemtajiem datiem no servera (pretējā gadījumā tas paliek, lai to vēlā var analizēt sistēmas administrators).

10. Ja tika atgriezts kļūdas kods -4, tad oriģinālais CAPTCHA-tests tiek saglabāts, bet pretējā gadījumā – dzēsts.

11. Operatora bloks atgriež kļūdas kodu atpakaļ kodolam.

Redzams, ka operatora bloks galvenās funkcijas pilda, izmantojot ļoti vienkāršu, pat primitīvu piegājienu, kas nodrošina augstu sistēmas stabilitāti, taču blokā iekļauta arī sarežģītāka aizsardzība pret kļūdām, kā arī īpaši atvieglota iespēja administratoram novērst to atkārtosanos.

LMT operatora bloka algoritms pēc uzlabojumiem ir samērā līdzīgs un darbojas sekojoši:

1. No datubāzes tabulas “ocrcodes” tiek pieprasīts jaunākais neizmantotais CAPTCHA-testa atrisinājums, kas der operatoram LMT.
 - a) Ja tāda nav, tiek atgriezts sistēmas kļūdas kods -6 un operatora bloks beidz darbu.
 - b) Ja tāds ir atrasts, operatora bloks tabulā “ocrcodes” atzīmē, ka šis kods ir izmantots, norādot tā izmantošanas laiku, un turpina darbu.
2. Oriģinālais CAPTCHA-tests tiek nokopēts uz kļūdu direktoriju (gadījumam, ja sistēma apstāsies pusceļā).
3. Tiek pārbaudīts vai CAPTCHA-testa atrisinājums atbilst pareizajam garumam.
 - a) Ja nē, tiek atgriezts sistēmas kļūdas kods -4 un darbs tiek beigts.
 - b) Ja jā, operatora bloks sagatavo īsziņas nosūtīšanu.
4. Sistēma izveido vajadzīgos HTTP iesākus un sagatavo nosūtāmos datus, kas šajā gadījumā ir tikai trīs – īsziņas teksts, adresāts un CAPTCHA-testa atrisinājums. Tiek arī pievienota informācija, kas identificē oriģinālo sesiju, kā arī pievienoti standarta nosūtāmie dati.
5. Kļūdu direktorijā tiek izveidots fails, kurā uzreiz pēc saņemšanas tiek ierakstīts tas, ko operatora bloks saņem no operatora servera.
6. Sistēma pieslēdzas serverim “www.okarte.lv” un nosūta HTTP pieprasījumu ar sagatavotajiem iesākumiem un datiem.
7. Sistēma nolasa atbildi no servera un ieraksta saņemto failā (pirmos 4096 baitus).
8. Sistēma atpazīst atbildes pēc LMT sistēmas atgrieztā četr ciparu koda:
 - a) Saņemot kodu “1000”, sistēma atgrieztā kļūdas kodu 1.
 - b) “2000” vai “2001” – -11.
 - c) “2002” vai “2004” – -4.
 - d) “2003” – -2.
 - e) “2005” – -10.
 - f) “2006” – -17.
 - g) “2007” – -16.
 - h) “2008” – -15.

i) “2009” vai “2010” – -1.

9. Ja pazūd savienojums ar LMT serveri, tad tiek sistēma atgriezīs kļūdas kodu -3.

10. Ja savienojums ar serveri nav pazudis un no sistēmas ir saņemta kaut viena no augstākminētajām atbildēm, tad tiek izdzēsts fails ar saņemtajiem datiem no servera (pretējā gadījumā tas paliek, lai to vēlā var analizēt sistēmas administrators).

11. Ja tika atgriezts kļūdas kods -4, tad oriģinālais CAPTCHA-tests tiek saglabāts, bet pretējā gadījumā – dzēsts.

12. Operatora bloks atgriež kļūdas kodu atpakaļ kodolam.

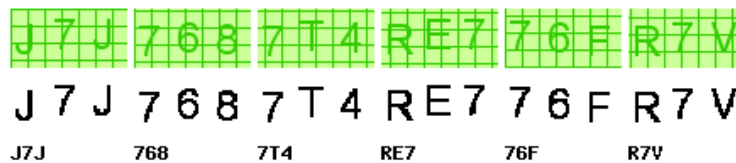
Lai informētu lietotājus par pieejamajiem operatoru blokiem un mazinātu lietotāju apjukumu, kad, tiem sūtot īsziņu uz Amigo vai BITE mobilā telefonu numuru, tā netiek nosūtīta, tīmekļa saskarnē “sms.id.lv” tika izvietots paziņojums “Sūtītājs darbojas testa režīmā. Šobrīd var sūtīt SMS uz LMT, TELE2 un Triatel tīkliem. Atnāc vēlāk un varbūt būs iespēja sūtīt arī kur citur. :)”

3.4.2. CAPTCHA-testu atpazīšana

Kā redzams 3.9. attēlā, LMT joprojām piedāvā izvēlēties vienu no diviem dažādajiem CAPTCHA-testiem, kas ļauj vieglāk uzrakstīt algoritmu, kas nepieciešams tā atrisināšanai ar datoru. Autors šajā gadījumā arī apskatīja divas alternatīvas – dažādos CAPTCHA-testus, ko piedāvā risināt LMT. No tiem autors izvēlējās atrisināt 3.11. attēlā redzamo, jo tas ir stipri līdzīgāks iepriekšējam CAPTCHA-testam, kas nozīmē lielu varbūtību, ka izdosies saglabāt iepriekšējo OCR kvalitāti.

Izvēloties otru alternatīvu, nebūtu redzami ieguvumi, jo abi CAPTCHA-testi ir vienādi funkcionāli un atļauj veikt vienu un to pašu darbību vienādā mērā, taču būtu acīmredzams trūkums, kas parasti ir ieviešot jaunu OCR – zems pareizi atpazīto attēlu līmenis.

LMT CAPTCHA-testu atpazīšanas kods tika uzrakstīts no jauna, lietojot vairākas jau iepriekš eksistējošas funkcijas (vienu no tām bija rakstījis Jānis Timma). Kopumā kodi tiek atpazīti ļoti labi (skatīt 3.11. att.) un arī ātrums ir apmierinošs. Pārbaudot sistēmu lielā noslodzē, nekādi būtiski trūkumi LMT OCR skriptā netika konstatēti, lai gan OCR skripts nesasniedz 100% pareizību (skatīt pielikumu nr. 4.)



3.11. att. Jauno LMT CAPTCHA-testu OCR pārbaude

OCR, ar ko tiek risināti LMT CAPTCHA-testi darbojas pēc algoritma ar šādu shēmu:

1. CAPTCHA-testa attēls tiek saglabāts un ar GD bibliotēkas palīdzību ielādēts PHP.
2. Attēls tiek pārvērsts par vienkāršainu melnbaltu attēlu.
3. Attēlam tiek noņemts režģis.
4. Tiek sameklēts pirmais simbols attēlā un tad salīdzināts ar izmantotā fonta ("Arial") simbolu datu bāzi. Tā kā ir zināms, ka šis CAPTCHA-tests izmanto tikai simbolus no virknes "ACEFJKLMNPRTVX23456789", tad salīdzināšana notiek tikai pret simboliem no šīs virknes.
5. No datu bāzes tiek izvēlēts simbols, kas rastra ziņā visvairāk atbilst attēla pirmajam simbolam.
6. Vēl divas reizes tiek atkārtots 4. un 5. solis.
7. Savienojot iepriekšējos soļos iegūtos simbolus, tiek iegūts triju simbolu garš rezultāts.

Lai demonstrētu sasniegto, tiek uzrādīts pseidokods LMT CAPTCHA-testa atrisināšanā izmantotai funkcijai, kuras mērķis ir noņemt koordināšu asīm paralēlo režģi no attēla:

```
OCRdegrid(input)
im ← input
spacer ← 10
im ← pārvērst im par melnbaltu attēlu
lx ← attēla im platums
ly ← attēla im garums
px ← tuvākā krāsa attēla, kas līdzīga melnai (0x00,0x00,0x00)
wh ← tuvākā krāsa attēla, kas līdzīga baltai (0xFF,0xFF,0xFF)

FOR y ← 0..ly-1
  HF ← TRUE
  FOR x ← 0..lx-1
    Ja attēla im krāsa koordinātēs (x,y) nav px, tad hf ← hf & FALSE
  IF hf
    fh ← y
    BREAK
fh ← fh MOD spacer-1
FOR x ← 0..lx-1
  HF ← TRUE
  FOR y ← 0..ly-1
    Ja attēla im krāsa koordinātēs (x,y) nav px, tad hf ← hf & FALSE
  IF hf
    fv ← x
```

```

BREAK

fv ← fv MOD spacer-1
x ← fv
WHILE x < lx
  FOR y ← 0..ly-1
    Ja [attēla im krāsa koordinātēs (x-1,y) nav px vai x = 0] un
    [attēla im krāsa koordinātēs (x+1,y) nav px vai x = lx-1]
      Iekrāsot attēla im koordinātes (x,y) krāsā wh
      x ← x + spacer
  y ← fh
  WHILE y < ly
    FOR x ← 0..lx
      Ja [attēla im krāsa koordinātēs (x,y-1) nav px vai y = 0] un
      [attēla im krāsa koordinātēs (x,y+1) nav px vai y = ly-1]
        Iekrāsot attēla im koordinātes (x,y) krāsā wh
        y ← y + spacer
OCRdegrid ← im

```

Lai gan šo kodu varētu viegli modificēt, lai tas automātiski pielāgotos, piemēram, jebkāda izmēra režģim, šis kods tāpat kā pārējās sistēmas daļas ir paredzēts konkrētam mērķim – LMT izvēlēta veida CAPTCHA-testu atpazīšanai. Tam par iemeslu ir ātrums un kvalitāte. Universāls kods darbotos lēnāk un nekvalitatīvāk, jo tam nebūtu papildus “priekšzināšanu” par CAPTCHA-testa uzbūvi.

Izpētot TELE2 gadījumu, šķiet, ka to ir viegli atrisināt. Ņemot vērā, ka piemērā (3.12. attēls) visu simbolu **normālais** augstums būtu vienāds, var secināt, ka uzraksts tiek deformēts izmantojot divas taisnes, no kurām viena atrodas tieši virs simboliem, bet otra – tieši zem. Katras taisnes katrs punkts tiek pārvietots attēla plaknē, lai iegūtu līkni. Līdz ar taisnes punktiem tiek pārvietotas arī attēla malas.



3.12. att. TELE2 lietotā CAPTCHA-testa deformācijas analīze

Tā kā deformācija notiek šādā veidā, autora piegājiens situācijas atrisināšanai ir – deformēt attēlu, kamēr līknes gar tā malām atkal kļūst taisnas. Tālākā rīcība gan prasītu nelielu uzmanību, jo attēla simboli saplūst viens ar otru, kas ir plaši izplatīts paņēmiens, lai padarītu teksta atpazīšanu grūtāku datorprogrammām. Taču šajā gadījumā acīmredzami tiek izmantots fonts, kura visi simboli ir vienāda platuma, kā arī CAPTCHA-testā vienmēr ir tieši 6 simboli, kas būtībā nozīmē, ka nākamais un pēdējais solis būtu horizontāli sadalīt iegūto attēlu 6 vienādās daļās un tad salīdzināt, kuram burtam, kura daļa atbilst.

Līdz ar to OCR darbotos pēc sekojošas shēmas:

1. CAPTCHA-testa attēls tiek saglabāts un ar GD bibliotēkas palīdzību ielādēts PHP.
2. Tā kā novērojumi liecina, ka attēli mēdz būt trijās krāsās – sarkanā, zilā un zaļā, attēls tiek pārvērsts par melnbaltu daudzkrāsainu (pelēku) attēlu (lai nepazaudētu iespēju viegli noteikt simbolu robežas – kā redzams tās ir nedaudz blāvākas).
3. Zem un virs uzraksta tiek uzzīmētas divas citas krāsas taisnes. Šīs taisnes lēnām tiek tuvinātas uzrakstam, nešķērsojot uzrakstu, bet apliecoties ap to (3.12. attēls), tādā veidā, lai liekumi nebūtu pārāk asi.
4. Ja tas nav iespējams, tad attiecīgais ekstrēma punkts tiek ignorēts (acīmredzot, tas ir simbols “Q”).
5. Tad šīs taisnes tiek iztaisnotas, līdz deformējot visu attēla plakni, kā rezultātā tiek iegūts taisnstūris. Taisnes tiek dzēstas, izmantojot krāsu filtrāciju.
6. Attēlam tiek noteikta kreisā un labā mala, līdzīgi kā tas aprakstīts 3. punktā, taču neveicot apliekšanu, bet apstājoties, tikko kāds taisnes punkts sasniedz kādu simbola punktu.
7. Attēls tiek apgriezts no kreisās un no labās malas.
8. Tiek izmērīts attēla platums un tas sadalīts 6 vienāda platuma joslās, katra no kurām tiek salīdzināta ar izmantotā fonta datu bāzi. Tā kā ir zināms, ka šis CAPTCHA-tests izmanto simbolus, kas atrodas simbolu virknē “ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789”, tad salīdzināšana notiek tikai pret simboliem no šīs virknes (t.i. visi cipari un lielie burti).
9. No datu bāzes tiek izvēlēti simboli, kas visvairāk atbilst katrai no 6 joslām.
10. Šo simbolu virkne attiecīgajā secībā ir rezultāts.

Paralēli CAPTCHA-testa datorizētas risināšanas procesam, meklējot alternatīvas, autoram izdevās atklāt ko negaidītu – nomainot adresi, no kuras sistēma ielādē CAPTCHA-testu, izdevās iegūt iepriekšējo attēla versiju, kas darbojās arī jaunajās sesijās. Vecajai attēla versija bija viens būtisks trūkums, kas ļāva datoram to vienkārši atrisināt – tā lietoja indeksēto krāsu paleti un 1. krāsas indekss vienmēr bija vienīgais, ko vajadzēja ņemt vērā. Pārējos varēja vienkārši atnest. Līdz ar to, salīdzinot alternatīvas, autors nolēma pagaidām iztikt ar iepriekšējo paša rakstīto grafiskā CAPTCHA-testa atrisināšanas kodu, taču noteikti turpināt risināt jauno CAPTCHA-testu, jo nav zināms, cik ilgi vēl saglabāsies iespēja izmantot veco attēlu formātu.

Lai demonstrētu TELE2 CAPTCHA-testa atrisināšanas tehniku, tiek uzrādīts pseidokods tā atrisināšanā izmantotai pamatfunkcijai, kas ievadā saņem attēla resursu, bet izvadā atgriež tā atrisinājumu (masīvs “burtti” satur katra alfabēta simbola reprezentāciju, kā 8 reiz 10 bitu masīvu):

```
decodeLameOCR(imgr)
burtti ← ()
burtti[0] ← ( 24, 60,102,195,195,195,195,102, 60, 24)
burtti[1] ← ( 24, 56,120, 24, 24, 24, 24, 24, 24,126)
burtti[2] ← ( 60,102,195, 3, 6, 12, 24, 48, 96,255)
burtti[3] ← (124,198, 3, 6, 28, 6, 3, 3,198,124)
burtti[4] ← ( 6, 14, 30, 54,102,198,255, 6, 6, 6)
burtti[5] ← (254,192,192,220,230, 3, 3,195,102, 60)
burtti[6] ← ( 60,102,194,192,220,230,195,195,102, 60)
burtti[7] ← (255, 3, 3, 6, 12, 24, 48, 96,192,192)
burtti[8] ← ( 60,102,195,102, 60,102,195,195,102, 60)
burtti[9] ← ( 60,102,195,195,103, 59, 3, 67,102, 60)
burtti[A] ← ( 24, 60,102,195,195,195,255,195,195,195)
burtti[B] ← (252,198,195,198,252,198,195,195,198,252)
burtti[C] ← ( 62, 99,193,192,192,192,192,193, 99, 62)
burtti[D] ← (252,198,195,195,195,195,195,195,198,252)
burtti[E] ← (254,192,192,192,252,192,192,192,192,254)
burtti[F] ← (255,192,192,192,252,192,192,192,192,192)
burtti[G] ← ( 62, 99,192,192,192,199,195,195, 99, 62)
burtti[H] ← (195,195,195,195,255,195,195,195,195,195)
burtti[I] ← (126, 24, 24, 24, 24, 24, 24, 24, 24,126)
burtti[J] ← ( 30, 6, 6, 6, 6, 6, 6, 6, 70,108, 56)
burtti[K] ← (195,198,204,216,240,240,216,204,198,195)
burtti[L] ← (192,192,192,192,192,192,192,192,192,254)
burtti[M] ← (195,231,255,219,219,219,195,195,195,195)
burtti[N] ← (195,227,243,243,219,219,207,199,199,195)
burtti[O] ← ( 60,102,195,195,195,195,195,195,102, 60)
burtti[P] ← (254,195,195,195,254,192,192,192,192,192)
burtti[Q] ← ( 60,102,195,195,195,195,219,207,102, 61)
burtti[R] ← (254,195,195,195,254,248,204,198,195,195)
burtti[S] ← (126,195,192,192,126, 3, 3, 3,195,126)
burtti[T] ← (255, 24, 24, 24, 24, 24, 24, 24, 24, 24)
burtti[U] ← (195,195,195,195,195,195,195,195,102, 60)
burtti[V] ← (195,195,195,102,102,102, 60, 60, 24, 24)
burtti[W] ← (195,195,195,195,219,219,219,255,231,195)
burtti[X] ← (195,195,102, 60, 24, 24, 60,102,195,195)
burtti[Y] ← (195,195,102, 60, 24, 24, 24, 24, 24, 24)
burtti[Z] ← (254, 6, 6, 12, 24, 48, 96,192,192,254)

laime ← ''
```

```

started ← 0
ib ← 0
lc ← 1
yCONST ← (1000,1000,1000,1000,1000,1000,1000,1000,1000,1000)
xCONST ← ()
FOR x ← 0..attēla imgr platums
  topm ← -1
  FOR y ← 0..attēla imgr augstums
    Ja attēla imgr krāsas indekss koordinātēs (x,y) ir lc
      topm ← y
      BREAK
  IF started > 0
    started ← started + 1
  IF topm ≠ -1
    IF started = 0
      xCONST[ib] ← x
      started=1
    yCONST[ib] ← MIN(topm,yCONST[ib])
  ELSE
    IF started > 10
      started ← 0
      ib ← ib +1

ib ← 0
Katram xCONST elementam _ ar saturu x
  pozb ← ''
  maxwin ← -1;
  mcolor ← -1
  Katram burti elementam burts ar saturu kods
    sader ← 0
    totl ← 0
    lettercolor ← lc
    FOR y ← 0..9
      yIM ← y + yCONST[ib]
      FOR b ← 0..7
        Ja b-tais bits no kods[y] ir 1
          totl ← totl + 1
        Ja [attēla imgr krāsas indekss koordinātēs (x+b,yIM) ir/nav lettercolor] un
          attiecīgi [b-tais bits no kods[y] ir/nav 1]
          sader ← sader + 1
    sader ← sader + totl/100;
    IF sader > maxwin
      pozb ← burts
      maxwin ← sader
      mcolor ← lettercolor

IF maxwin >= 70
  laime ← laime + pozb
  ib ← ib + 1
decodeLameOCR ← laime

```

3.4.3. SMTP serviss

Tā kā šis serviss vairs nedarbojas, jo mainījušies sistēmas vārtejas parametri, autors lēma, ka šis samērā plaši izmantotais serviss ir jāatjauno darba kārtībā. Ķeroties pie darba, autors secināja, ka SMTP servisa kodā (kā jau iepriekš minēts, tas rakstīts valodā Free Pascal) tiek lietoti Free Pascal moduļi (piemēram, “inet”), kas ir novecojuši un vairs netiek atbalstīti vai iekļauti šajā programmēšanas valodā, un līdz ar to programmu nav iespējams nokompilēt.

Šādā situācijā bija trīs izvēles iespējas:

- pārrakstīt SMTP servisu citā programmēšanas valodā;
- sameklēt Free Pascal modulus, kas funkcionāli varētu aizvietot vai gandrīz aizvietot vecos modulus, un pārrakstīt programmas tīkla daļu;
- kombinēt Free Pascal ar kādu citu programmēšanas valodu, lai trūkstošās (vecās) daļas izsauktu no tās.

Pēdējā alternatīva ir nevajadzīgi sarežģīta, un autors uzskata, ka pie tās būtu jāķeras tikai tad, kad nav nevienas citas iespējas.

Varētu pārrakstīt visu SMTP servisu citā valodā, taču autora zināšanas tīkla programmēšanas jomā robežojas ar tādu valodu kā PHP, Pascal utml. izmantošanu. Lai gan autors bieži programmē arī citās programmēšanas valodās, autoram nav ievērojamas pieredzes **tīkla programmu rakstīšanā** citās nopietnās programmēšanas valodās, kā piemēram C++.

Tamdēļ sākotnēji tiek izmēģināta alternatīva, kas sevī ietver diezgan lielu darbu – sameklēt Free Pascal modulus un pārrakstīt programmas tīkla daļu. Tas arī tiek izdarīts. Izmantojot jaunāku kompilatora versiju, kurā ir iebūvētas tīkla funkcijas un pieslēdzot moduli “math”, SMTP servisu nodrošinošās programmatūras “smsmaild” tīkla daļa ir pārrakstīta praktiski no nulles. Tā kā “smsmaild” uzlabotā versija balstās uz iepriekšējās, kuras izstrādē aktīvu dalību ņēmis Jānis Timma, autors atļausies šīs sistēmas daļas kodu nepublicēt.

Šobrīd cīņai ar nevēlama satura vēstulēm (īszinām), kā arī, lai novērstu sūtītāja identitātes viltojumus, SMTP serviss balstās uz bakalaura darba autora izveidoto algoritmu, kas pārbauda, vai IP adrese, no kuras tiek sūtīts e-pasts, ir piederīga domēnam, kas tiek uzrādīts kā e-pasta autora domēns. Šī sistēma ir līdzīga SPF(4) ar ierakstu “a mx -all”. SPF kļuva plaši izplatīts dažus gadus pēc SMTP servisa ieviešanas sistēmā “sms.id.lv”. (Ar šo autors nekādā gadījumā necenšas paziņot, ka SMTP servisa ieviešana kaut kādā mērā ietekmēja SPF izveidi, taču arī pretējs process nenotika.)

Šāda pieeja darbojas mazām e-pasta sistēmām, taču lielām sistēmām ar serveru fermām domēnam piederīgie serveri, kas ir “tiesīgi” izsūtīt e-pastus, var atrasties pat pilnīgi dažādos IP adrešu apgabalos.

Tāpēc, lai novērstu to, ka “gmail.com” lietotāji šobrīd nespēj izmantot sistēmas “sms.id.lv” pakalpojumus, fails *trusthosts.smd* papildināts tā, lai tas pieņemtu daudzus “gmail.com” e-pasta serverus:

216.239.32.0/19:gmail.com
64.233.160.0/19:gmail.com

```
66.249.80.0/20:gmail.com  
72.14.192.0/18:gmail.com  
209.85.128.0/17:gmail.com  
66.102.0.0/20:gmail.com  
74.125.0.0/16:gmail.com  
64.18.0.0/20:gmail.com  
207.126.144.0/20:gmail.com
```

Šis fails satur IP apgabalu un domēnu (pa pāriem) sarakstu, kam pienākas atviegloti nosacījumi attiecībā uz prasību pret sūtītāja adresi izpildi.

Izlabotas vairākas nopietnas drošības kļūdas, kas būtu ļāvušas uzbrucējam iegūt kontroli pār sistēmu, kā arī novērstas vairākas sīkas kļūdas. Tagad pēc īsziņas sūtīšanas serviss spēj pienācīgi uztvert sistēmas kļūdas kodu un nosūtīt to atpakaļ sūtītājam.

3.4.4. Citi darbi

Nemot vērā zemo pieprasījumu pēc IRC servisa, tas tika slēgts. Tādēļ servera noslodze (*load average*) samazinājās no vidēji 5 procesiem uz 1,5 procesiem, taču lietotājus IRC servisa slēgšana būtiski neietekmēja.

Īsziņu rindā bija sakrājušās vairāk kā 10 000 īsziņas no 2008. gada maija un arī jūnija, un septembra. Tās visas pārceltas uz tabulu “dropped”, kas pēc būtības nozīmē to, ka tās tika izmestas, nemēģinot nosūtīt. Kā alternatīvs variants tika apskatīts visas šīs īsziņas apstrādāt ar SMS rindas pārvaldnieku, tas ir – nosūtīt. Šādā gadījumā normas robežās tiktu uzturēts kritērijs, kas uzrādās pārraudzības rīkos un droši vien norāda uz kvalitāti – zems “nomesto” īsziņu skaits. Tomēr, vērtējot situāciju reāli – ne sūtītāji, ne saņēmēji nebūtu priecīgi par gadu vecu īsziņu saņemšanu. Tas izraisītu apjukumu un nepatiku vai labākajā gadījumā smieklus.

Darba gaitā tika pamanīts, ka rezerves kopiju izmērs ir krasi pieaudzis, tāpēc tika veikti pasākumi, lai minimizētu rezerves kopiju izmēru. Šobrīd autors ir apstājies pie gandrīz 60 MiB liela rezerves kopijas izmēra samazinājuma, taču, rodoties vajadzībai, tiks meklētas papildu iespējas atkārtoti samazināt rezerves kopiju izmēru.

Lai nodrošinātu starpniekserveru datubāzes atjaunināšanu, modificēts starpniekserveru saraksta iegūšanas skripts. To nācās darīt, jo bija būtiski mainījusies “samair.ru” lapas struktūra.

Ir atjaunota statistiskā operatora piederības datu bāze, iekļaujot vairākus jaunus numerācijas diapazonus, kā arī mainot numerācijas diapazonu ar prefiksu 262 piederību no Amigo uz LMT. Ja agrāk numerācijas datubāzi uzturēja Sabiedrisko pakalpojumu regulēšanas komisija, tad sākot ar 2008. gada beigām numerācijas datubāzi apkopo Elektronisko sakaru direkcija(5), tāpēc turpmāk sistēmā esošā datu bāze tiks atjaunota no Elektronisko sakaru direkcijas(25).

Rakstot dokumentāciju un pārskatot sistēmas kodolu un pārējos sistēmas modulus, tika atklātas vairākas programmēšanas nepilnības un sīkas kļūdas, kas tagad ir izlabotas.

3.5. Sistēmas testēšana

Pēc aprakstīto darbu pabeigšanas sistēma tika testēta, un autors ievēroja un novērsa vairākas kļūdas.

Adresāts: 26123456

Teksts:

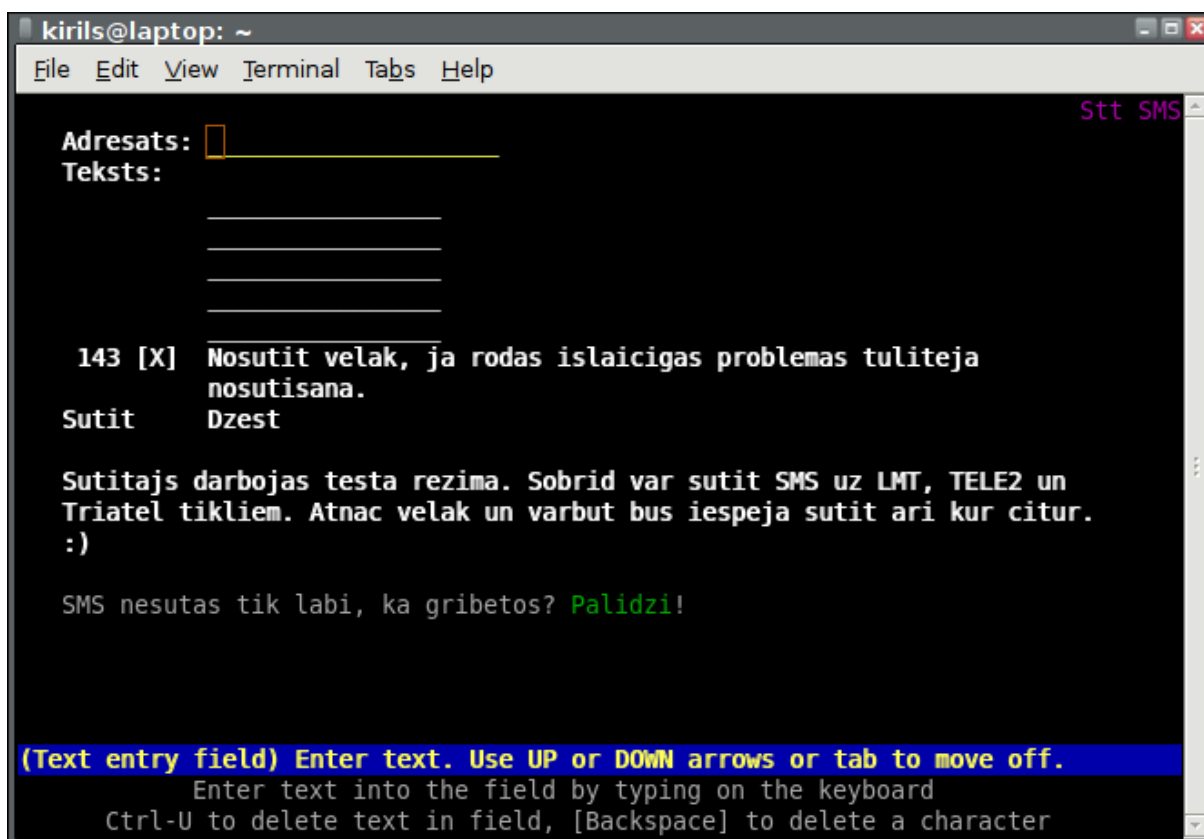
1 43 ☒ Nosūtīt vēlāk, ja rodas īslaicīgas problēmas tūlītējā nosūtīšanā.

Sūtīt Dzēst

Īsziņa nosūtīta veiksmīgi.

Sūtītājs darbojas testa režīmā. Šobrīd var sūtīt SMS uz LMT, TELE2 un Triatel tīkliem.

3.13. att. Tīmekļa lapa "sms.id.lv"



3.14. att. Tīmekļa lapa “sms.id.lv” caur lynx

Noslēdzotās testēšanas (pēc kļūdu novēršanas) kopsavilkums redzams pielikumā nr. 3.

3.6. Iesāktie darbi

Šobrīd sistēmas uzlabošana nav apstājusies, un ir iesākts darbs pie Amigo un BITE operatoru blokiem un CAPTCHA-testu apstrādes moduļiem.

3.6.1. Operatoru bloki

Ir izveidoti darbspējīgi un Amigo un BITE operatoru bloki. Darbs pie šī nav bijis grūts, taču tā kā abās operatoru lapās ir CAPTCHA-testi, tiem vēl nepieciešams uzrakstīt kvalitatīvu OCR.

BITE ir pieejama arī īsziņu sūtīšana no e-pasta(26), līdz ar to šim operatora blokam vajadzētu izveidot arī e-pasta režīmu, lai gan lietotāji atsauksmes liecina, ka tas nav visai ērts veids kā saņemt īsziņas. Amigo nav pieejama īsziņu sūtīšana no e-pasta.

3.6.2. CAPTCHA-testu apstrādes moduļi

Gan Amigo, gan BITE ir ieviesuši grafiskos CAPTCHA-testus. Lai varētu sūtīt īsziņas uz šiem operatoriem, ir jāizstrādā OCR algoritms, kas spēs atrisināt šos CAPTCHA-testus. Autors jau ir sācis darbu pie tā.

3.6.2.1. Amigo

Šie grafiskie CAPTCHA-testi ir salīdzinoši vienkārši. Kā redzams 3.15. attēlā, visi burti ir vienādā leņķī un vienā fontā un krāsā, lai gan dažādos izmēros.

Lielākās grūtības šobrīd sagādā burtu pārklāšanās un nogrieztās malas.

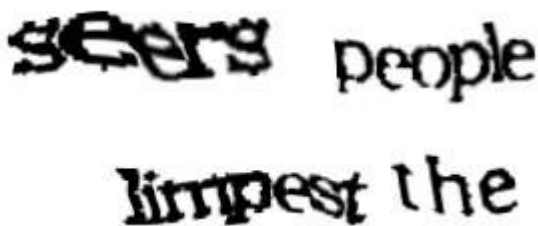


3.15. att. Amigo grafiskā CAPTCHA-testa paraugi

Īpaši traucējoši ir tas, ka pārklāšanās gadījumā, pārklājuma vietā nemainās krāsojuma intensitāte (3.15. attēls pa vidu). Tādēļ ar tiešām metodēm nav iespējams identificēt laukumus, kas pieder vairāk kā vienam simbolam.

3.7. BITE

BITE kā grafisko CAPTCHA-testu acīmredzami izmanto slaveno un jau iepriekš pieminēto *reCAPTCHA*. Šis ir plaši izmantots CAPTCHA-tests(10), kas skaitās viens no labākajiem, jo līdz šim vēl nav radīta datorprogramma, kas spētu to ticami atrisināt(27). Taču neskatoties uz to, Pauls Lamērs (*Paul Lamere*) ir izpētījis *reCAPTCHA* darbības shēmu un piedāvājis algoritmu kā darbu iespējams samazināt uz pusi (vismaz, ja šo CAPTCHA-testu risina cilvēks).



3.16. att. BITE grafiskā CAPTCHA-testa paraugi

Pauls Lamērs apraksta, ka viens no diviem vārdiem (3.16. att.), ko piedāvās *reCAPTCHA* vienmēr būs lieks. Viens būs īsts – ģenerēts ar datoru, uz ko *reCAPTCHA* zinās atbildi, bet otrs neīsts – ieskenēts no grāmatas, uz ko sistēma vēlēsies uzzināt atbildi. Pietiek pareizi (vai pat vispār) ievadīt tikai īsto vārdu.(28)

Viņš arī piedāvā vairākus veidus, kā atšķirt īsto vārdu no neīstā:

- īstais vārds nesatur ciparus, simbolus vai lielos burtus;
- visi īsie un apcirstie vārdi ir neīsti;
- ja vārdam apkārt ir punktiņi, tad tas ir neīsts;
- utt.

Autors par šo izteikumu patiesumu pārliecinājās, tos pārbaudot praksē.

Lai izpētītu turpmākās iespējas, autors izmēģināja pārbaudīt, kā ar šāda veida CAPTCHA-testiem tiks galā serviss, kas nodarbojas ar CAPTCHA-testu atrisināšanu. Pārskatot dažādu CAPTCHA-testu un ar to saistīto pakalpojumu sarakstu(29), autors noskaidroja, ka “CAPTCHA killer”(30) piedāvā bez maksas izmēģināt viņu slēgtā koda sistēmu. Veicot pārbaudi, tika secināts, ka serviss strādā vidēji ilgā laikā (patērējot vidēji divas minūtes uz vienu *reCAPTCHA* ģenerētu attēlu) un kvalitāte ir pārsteidzoši laba, taču ne pietiekama (skatīt 3.17. attēlu). Ja neņem vērā atstarpes simbola iztrūkumu atbildē, CAPTCHA-testu pilnīgi pareizi izdevās atrisināt vienā no četriem gadījumiem, bet pieņemamā līmenī (tādā, ka *reCAPTCHA* sistēma uzskatīja testu par izietu) tas izdevās divos no četriem gadījumiem, ņemot vērā to, ka pietiek pareizi atrisināt īsto vārdu.

	300	300x57	1 min 21 sec 4 min 7 sec ago
RESULT: Ditcayman http://bite.lv/			
	300	300x57	1 min 5 sec 4 min 13 sec ago
RESULT: parisianDctroit http://bite.lv/			
	300	300x57	2 min 3 sec 4 min 21 sec ago
RESULT: theturtoring http://bite.lv/			
	300	300x57	3 min 22 sec 6 days 8 hours ago
RESULT: bestnd http://recaptcha.com/			

3.17. att. “CAPTCHA killer” rezultātu logs (30)

Nelielas šaubas par “CAPTCHA killer” lietderību *reCAPTCHA* atrisināšanas progresa veicināšanā radīja arī fakts, ka šis algoritms uz vienādiem ievaddatiem mēdz sniegt dažādus izvaddatus. Tas tika apliecināts gan autora veiktajos testos, gan arī noskaidrojot citu lietotāju pieredzi.

Pamatojoties uz šajā nodaļā aprakstīto pieredzi un noskaidrotajām CAPTCHA-testa īpatnībām autors turpina izpēti, taču vēl nav pietiekami pietuvojies šīs problēmas risinājumam.

3.8. Plānotie uzlabojumi

Sasniedzot darba mērķi un atjaunojot sistēmas darbību, autors pamanīja, ka sistēmu iespējams vēl vairāk uzlabot, taču tajā būtu jāiegulda stipri vairāk, nekā šī darba izstrādēs laikā var atļauties.

Šie uzlabojumi pagaidām ir tikai idejas, taču palīdzēs lietotājiem justies ērtāk un vieglāk izmantot sistēmu dažādos dzīves brīžos.

3.8.1. Iespēja personalizēt sistēmu

Lietotājam vajadzētu piedāvāt iespēju reģistrēties, izmantojot savu mobilo telefonu. Tādā gadījumā viņam būtu personīga identifikācija, uz ko mazākā mērā attiektos bloķētās konfigurācijas. Lietotājam būtu iespēja saņemtās (no sistēmas sūtītās) īsziņas apskatīt arī tīmekļa pārlūkā, būtu iespēja veidot vēsturi un izmantot adresu grāmatu.

Lai nodrošinātu lietotāju lojalitāti, kārtību un atbildību, reģistrēšanās sistēma varētu darboties tikai uz ielūgumu pamata.

3.8.2. Midleti

WAP vietā būtu ieviešami midleti, ko lietotāji varētu lietot no saviem mobilajiem telefoniem. Tas atvieglotu gan adresu grāmatas lietošanu, gan arī padarītu par iespējamu pieteikties sistēmā, ja tiek realizēts iepriekšējā apakšnodaļā aprakstītais par sistēmas persnolizēšanu.

3.8.3. Citi iespējamie uzlabojumi

SMTP servisam būtu jāievieš SPF atbalsts, jo tas kļūst arvien izplatītāks un tas ir vienkāršs veids, kā noskaidrot, kādus tad patiesība izejošos e-pasta serverus lieto attiecīgais domēns. Ne vienmēr tie sakrīt ar ienākošajiem pasta serveriem, uz ko norāda MX ieraksts pie domēna.

Pašreizējā metode kā serviss cīnās ar nevēlama satura vēstulēm un novērš sūtītāja identitātes viltojumus, nedarbojas lielām sistēmām, kā tas bija redzams gadījumā ar “gmail.com”.

Šībrīža SMTP protokolā sūtītāja adresei ir tieši tāda pati funkcija kā atpakaļadresē reālā pasta sūtījumā – uzrakstot vēstuli un aiznesot to uz pastu nosūtīšanai, atpakaļadrese netiek pārbaudīta. Tā tiek rakstīta tikai tāpēc, lai varētu saņemt atbildi, bet būtībā nav nekādu problēmu norādīt citu adresi, no kuras ir sūtīta vēstule.

Tas pats attiecas arī uz e-pastu. Taču, ņemot vērā to, cik daudz mūsdienās balstās uz pieņēmumu, ka e-pasta atpakaļadrese ir korekta, atpakaļadreses viltošana ir reāla problēma.

SPF ir eksperimentāls, taču plaši izplatīts protokols, kas atrisina e-pasta atpakaļadresu viltošanu. Tas darbojas, domēnu īpašniekiem modificējot sava domēna ierakstu – pievienojot ierakstam tā saucamo SPF ierakstu, kurā domēna īpašnieks pasaka, kādi datori (IP adreses, citi domēna ieraksti) drīkst sūtīt e-pastu no viņa domēna. Ka saņēmējs saņem e-pastu, viņam atliek automātiski pārbaudīt, vai attiecīgais domēns satur SPF ierakstu, un, ja satur, pārliecināties, ka dators, no kuras ir nācis šis e-pasts, ir iekļauts atļauto datoru sarakstā.(4)

SPF ieviešana SMTP servisam ir vienkārši realizējama, jo ir pieejami gatavi SPF moduļi dažādās valodās (valodai Free Pascal, kurā rakstīts SMTP serviss, šobrīd gan nav pieejams tāds modulis).

4. REZULTĀTI

Izstrādājot bakalaura darbu, izdevās atjaunot īsziņu sūtīšanas sistēmas “sms.id.lv” darbību. Tagad lietotāji var atkal sūtīt īsziņas uz TELE2, LMT un Triatel mobilo sakaru tīkliem. Ir dokumentēti īsziņu sūtīšanas sistēmas “sms.id.lv” moduļi un datubāze, kas ļaus turpmākā nākotnē vajadzības gadījumā vieglāk iesaistīt projektā jaunus programmētājus, jo ar sistēmas uzbūvi, kā arī moduļu nozīmi tie varēs iepazīties šajā dokumentā. Sistēmas dokumentēšana autoram arī palīdzēja saprast, kuras sistēmas daļas ir liekas un nevajadzīgas, kā arī atrast vairākas slēptas kļūdas.

Izstrādājot darbu, tika noskaidrots, ka IRC serviss ir lieks, jo lietotāji to gandrīz neizmanto, tāpēc tas tika atslēgts. Šī servisa atslēgšana ļāva ievērojami ietaupīt servera resursus, kas paātrināja sistēmas darbību.

Ir atkal atjaunota SMTP servera darbība, pielāgojot to vārtejas lietotajam protokolam, un jau tagad lietotāji to ir atsākuši aktīvi izmantot. Tas viņiem nozīmē iespēju viegli un ātri nosūtīt īsziņu, nepārslēdzoties no e-pasta programmatūras uz sistēmas lapu globālajā tīmeklī.

Lielākajai daļai mūsdienu globālajā tīmeklī esošo lapu pēc būtības nav svarīgi, vai lietotājs ir dators vai cilvēks; svarīgi ir vien zināt, vai darbības, ko veic šis lietotājs nerada kaitējumu pakalpojuma sniedzējam vai citiem lietotājiem un nepārkāpj noteiktās normas. Tā kā sistēma “sms.id.lv” jau vairākus gadus ir balstījies uz lietotāju uzvedības analīzi, tā vietā, lai lūgtu tiem iziet CAPTCHA-testus, un gan darba autors, gan lietotāji, gan īsziņu saņēmēji ir bijuši apmierināti ar pakalpojuma kvalitāti un sistēmas slodzi, kā arī zemo traucējošo īsziņu skaitu, tad autors ir empīriski **pierādījis, ka grafiskā CAPTCHA-testa vietā iespējams lietot arī lietotāju uzvedības analīzi**, kas arī ir svarīgākais šī darba rezultāts.

Darba rezultāts tika ieviests dzīvē un tā rezultātā tika strauji atgūta lietotāju uzticība, kas bija zaudēta laikā, kamēr sistēma nefunkcionēja. Pēc darba rezultāta ieviešanas sistēmu vienā nedēļā lietoja aptuveni 3 tūkstoši lietotāju, nosūtot vairāk kā 6 tūkstošus īsziņu.

Pēc darba pabeigšanas **kopējais** sistēmas pirmkods aizņem aptuveni 9 tūkstošus rindas kopā ar pārbaudes, izstrādes skriptiem, iekļaujot komentārus, un sastāv no fragmentiem, kas lielākoties rakstīti valodās PHP, Free Pascal un *bash*.

Nākotnē sistēmā vajadzētu ieviest iespēju sūtīt īsziņas uz Amigo un BITE. Iespējams, ieviest personalizāciju, kas ļautu īsziņas, ko izsūtījuši citi sistēmas lietotāji, apskatīt arī tīmekļa pārlūkā, ne tikai mobilajā telefonā, būtu iespēja aplūkot savu darbību vēsturi un izmantot adresu grāmatu. SMTP servisa darbības uzlabošanai vajadzētu mainīt mehānismu, kas aizsargā pret izsūtītāja adreses viltojumiem – jālieto SPF, nevis jāanalizē MX ieraksti.

Šis bakalaura darbs ir noderīgs lasītājiem, kas strādā ar globālā tīmekļa interaktīvo lapu izstrādi, kurus uztrauc neērtības, ko lietotājiem sagādā lieku darbību veikšana, un kuri ir atvērti jaunām idejām par alternatīvām, kas spētu aizvietot CAPTCHA-testus.

5. SECINĀJUMI

Ir izdevies paveikt vairākas svarīgas lietas. Salīdzinot īsziņu sūtīšanas sistēmu “sms.id.lv” pirms darba uzsākšanas un tagad, ir notikuši acīmredzami uzlabojumi. Pirmkārt, sistēma funkcionē, bet iepriekš tā nedarbojās. Taču ir jāņem vērā tas, ka tagad ir parādījušies vairāki konkurējošie projekti, no kuriem praktiski visi savā būtībā ir komerciāli, līdz ar to sistēmas “sms.id.lv” lietotāju loks, iespējams, būs mazāks nekā pirms vairākiem gadiem, kad sistēmu lietoja vairāki tūkstoši cilvēku dienā, taču to rādīs laiks. Turklāt autors uzskata, ka nav viennozīmīgi vērtējams, vai lietotāju bāzes neliela samazināšanās nekomerciālam projektam ir pozitīvs vai negatīvs rezultāts.

Otrkārt, sastādot sistēmas dokumentāciju autors spēja novērtēt sistēmas apjomus un secināt, kur vēl ir vieta attīstībai un uzlabojumiem.

Vērtējot straujo alternatīvo pakalpojumu piedāvājuma celšanos, sistēmai neesot pieejamai tikai aptuveni pus gadu, jājautā, kāpēc tas notika? Autora saredz divus iespējamus motīvus, kāpēc kādam būtu izdevīgi piedāvāt šādus pakalpojumus komerciāli. Tā ir iespēja iegūt Latvijas iedzīvotāju mobilo telefonu numuru datu bāzi, ko tālāk var izmantot mārketingam. Divos gadījumos par to liecina tas, ka firma, kas piedāvā pakalpojumu, piedāvā arī izsūtīt reklāmu īsziņu veidā. Diemžēl jāatzīst, ka vienā no gadījumiem, tas pierādījās praksē. Otrs motīvs varētu būt reklāma pakalpojuma lapā, kas viennozīmīgi ir mazāk nekaitīgs. Izpētot šos motīvus, autors secina, ka, iespējams, arī īsziņu sūtīšanas sistēmā “sms.id.lv” varētu ieviest neuzbāzīgas reklāmas.

Par nopietnāko draudu un neveiksmi darbā autors uzskata to, ka pagaidām nav izdevies ar datorprogrammas palīdzību atrisināt CAPTCHA-testu, ko savā tīmekļa lapā izmanto BITE. Iemesls tam ir tāds, ka šis CAPTCHA-tests izmanto *reCAPTCHA* pakalpojumu, kas tiek izmantots visā pasaulē, taču, neskatoties uz to, vēl nav veiksmīgi atrisināts ar datoru. Turklāt, vismaz viens no diviem vārdiem, kas ietilpst katrā *reCAPTCHA* ģenerētajā CAPTCHA-testā, ir fotogrāfija no drukāta dokumenta, ko nav spējušas atpazīt viņu lietotās OCR. Tas nozīmē, ka šajā jomā būs ļoti grūti panākt kādu reālu progresu, taču cerības nav atmestas.

Autora personiskais ieguvums no darba izpildes ir dziļāka izpratne par CAPTCHA-testiem un to attīstību, kā arī ērta īsziņu sūtīšanas sistēma, kuru autors pats labprāt izmanto ikdienā.

Nobeigumā autors secina, ka **eksistē lietojumi, kur grafiskā CAPTCHA-testa vietā ir iespējams lietot arī neinteraktīvās metodes**, piemēram, lietotāju uzvedības analīzi. Turklāt nebeidzamā cīņa starp CAPTCHA-testu izgudrotājiem un cilvēkiem, kas apmāca datorus risināt CAPTCHA-testus, uzlabo gan šo Tjūringa testu ģenerēšanas līmeni, gan arī algoritmus, kas noder, lai datoru padarītu cilvēcīgāku. Grafiskā CAPTCHA-testa gadījumā, tie būtu OCR algoritmi, bet nākotnē tie varētu būt, piemēram, sejas vai priekšmetu atpazīšanas algoritmi.

Šī cīņa viennozīmīgi novedīs zinātņi pie jaunām idejām un sasniegumiem visā pasaulē, iespējams, arī citās jomās, tuvinot “cilvēciski domājošu datoru” realitātei.

PATEICĪBAS

Vēlos izteikt pateicību cilvēkiem bez kuriem šī darba tapšanas process būtu bijis daudz smagāks un ilgāks.

Paldies manai Universitātei, kura mani ved pretī zvaigznēm un apskaidrībai, bet īpaši manam darba vadītājam un pasniedzējam doc. Dr. mat. Paulim Ķikustam, par saprotošo attieksmi, par to, ka palīdzēja un ka sniedza skaidras un saprotams atbildes uz jautājumiem, kas man radās šī darba izstrādes sakarā!

Paldies maniem kolēģiem par to, ka darbā spēja kādu laiciņu iztikt bez manis!

Draugi! Piedodiet, ka šo mēnesi un īpaši pēdējo nedēļu biju tik aizņemts, ka izlikos, ka man jūsu nav. Jūs esat! Un paldies jums par to!

IZMANTOTĀ LITERATŪRA UN AVOTI

1. Standarti un normatīvie akti

1. *Letter symbols to be used in electrical technology: Telecommunications and electronics. IEC 60027-2.*
2. *Simple Mail Transfer Protocol. RFC 2821.*
3. *Internet Relay Chat Protocol. RFC 1459.*
4. *Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1. RFC 4408.*
5. *Kārtība, kādā Elektronisko sakaru direkcija pārvalda numerāciju, izveidojot un uzturot numerācijas datubāzi. MK 2008. gada 11. augusta noteikumi Nr. 656.*

2. Disertācijas, maģistru un bakalauru darbi

6. **Ahn, L., Blum, M., Hopper, N. J., Langford, J.** *CAPTCHA: Using Hard AI Problems For Security*. Pittsburgh: Carnegie Mellon University, 2003. 18 p.
7. **Naor M.** *Verification of a human in the loop or Identification via the Turing Test*. Rehovot: Weizmann Institute of Science, 1996. 6 p.
8. **Liu, Y., Cen, R., Zhang, M., Ma S., Ru L.** *Identifying Web Spam with User Behavior Analysis*. Beijing: Tsinghua University, 2008. 8 p.

3. Elektroniskie informācijas avoti

9. *The Official CAPTCHA Site* [tiešsaiste] – [atsauce 14.05.2009.] Pieejams: <http://captcha.net/>
10. *What is reCAPTCHA?* [tiešsaiste] – [atsauce 16.05.2009.] Pieejams: <http://recaptcha.net/learnmore.html>
11. *Inaccessibility of CAPTCHA* [tiešsaiste] – [atsauce 18.04.2009.] Pieejams: <http://www.w3.org/TR/turingtest/>
12. *Send Message(SMS) to mobile phones and pagers in Latvia* [tiešsaiste] – [atsauce 10.05.2009.] Pieejams: <http://web.archive.org/web/19990823152619/http://gobbo.caves.lv/~cikis/phone/all/>
13. *SMS sender* [tiešsaiste] – [atsauce 11.05.2009.] Pieejams: <http://web.archive.org/web/20030208053112/www.ps.id.lv/mobile/sms.php>

14. *Weather forecast + namedays* [tiešsaiste] – [atsauce 11.05.2009.] Pieejams:
<http://web.archive.org/web/20030622093049/www.ps.id.lv/mobile/news.php>
15. *Amigo karte* [tiešsaiste] – [atsauce 14.05.2009.] Pieejams:
<http://amigo.lv/priekshapmaksas-karte/sarunu-karte-amigo/>
16. *Sabiedrisko pakalpojumu regulēšanas komisija* [tiešsaiste] – [atsauce 14.05.2009.] Pieejams: <http://www.sprk.gov.lv/>
17. *E-pasta adrese un tās identifikators* [tiešsaiste] – [atsauce 17.05.2009.] Pieejams: http://www.lmt.lv/lat/abonentiem/pakalpojumi/izsinu_pakalpojumi/sms_no_interneta/sms_no_interneta_identifikators
18. *OKteilis* [tiešsaiste] – [atsauce 14.05.2009.] Pieejams:
http://www.okarte.lv/noflash.php/lv/oriki/citi_oriki/okteilis
19. *Proxy Lists* [tiešsaiste] – [atsauce 17.05.2009.] Pieejams:
<http://www.samair.ru/proxy/>
20. *“sms.lv – Bez maksas SMS sūtīšana”* [tiešsaiste] – [atsauce 18.05.2009.] Pieejams: <http://sms.lv/>
21. *“esms.lv – tavs personīgais īsziņu centrs”* [tiešsaiste] – [atsauce 18.05.2009.] Pieejams: <http://esms.lv/>
22. *“esms.lv” darāmo darbu saraksts* [tiešsaiste] – [atsauce 18.05.2009.] Pieejams: <http://esms.lv/ToDoList.txt>
23. *“smsbox.lv – SMS par brīvu”* [tiešsaiste] – [atsauce 25.05.2009.] Pieejams:
<http://smsbox.lv/>
24. *“bezmaksassms.com – nosūtīt bezmaksas īsziņas uz Latviju”* [tiešsaiste] – [atsauce 25.05.2009.] Pieejams: <http://bezmaksassms.com/>
25. *Numerācijas plāns* [tiešsaiste] – [atsauce 16.05.2009.] Pieejams:
<http://www.esd.lv/inner.php?left=516&left2=519>
26. *E-pasts “Bite plus”* [tiešsaiste] – [atsauce 16.05.2009.] Pieejams:
<http://www.bite.lv/lv/pc/mobilemail/plus>
27. *moot wins, Time Inc. loses* [tiešsaiste] – [atsauce 16.05.2009.] Pieejams:
<http://musicmachinery.com/2009/04/27/moot-wins-time-inc-loses/>
28. *reCAPTCHA rule set for distinguishing between real and fake words* [tiešsaiste] – [atsauce 16.05.2009.] Pieejams:
<http://musicmachinery.files.wordpress.com/2009/04/fakecaptcha2.pdf>

29. *Listing every captcha you can think of* [tiešsaiste] – [atsauce 21.05.2009.]
Pieejams: <http://captcha.org/>

30. *Automated CAPTCHA Solver* [tiešsaiste] – [atsauce 24.05.2009.] Pieejams:
<http://www.captchakiller.com/>

PIELIKUMU SARAKSTS

Bakalaura darbam ir sekojoši pielikumi:

- Pielikums nr. 1: “Vārtejas pieslēguma skripts”
 - Pielikums satur programmas paraugu valodā PHP, ar kura palīdzību iespējams pieslēgties sistēmas vārtejai.
 - Šo paraugu var izmantot, lai uzrakstītu analogisku pieslēguma programmu jebkurā citā programmēšanas valodā, kas atbalsta vai spēj realizēt HTTP.
- Pielikums nr. 2: “Statistikas rīka pirmkods”
 - Pielikums satur statistikas rīka *smsstat.php* pirmkodu.
 - Ar šī statistikas rīka palīdzību iespējams redzēt pēdējā mēneša laikā nosūtīto īsziņu, nenosūtīto īsziņu, apmeklētāju un pat uzlikto automātisko konfigurācijas bloku skaitu.
- Pielikums nr. 3: “Sistēmas testēšanas rezultāti”
 - Pielikums satur sistēmas testu rezultātus, t.i. reālus notikumu protokolus, kas pārbauda, kā sistēma reaģē dažādās parastās un neparedzētās situācijās.
 - Testēšana veikta no lietotāja saskarnēm.
- Pielikums nr. 4: “Lietošanas statistika”
 - Pielikums satur lietošanas statistiku par dažādiem sistēmas aspektiem.
 - Ņemot vērā lietotāju skaita svārstības atkarībā no nedēļas dienas, statistikai tiek izmantoti divi nedēļu gari periodi – viens pirms, bet otrs pēc bakalaura darba.

PIELIKUMS NR. 1: “VĀRTEJAS PIESLĒGUMA SKRIPTS”

Izmantojot failu *paraugs.php*, kas rakstīts valodā PHP, iespējams pieslēgties sistēmas “sms.id.lv” vārtējam un sūtīt īsziņas. No sistēmas īpašnieka iepriekš gan jāsaņem pilnvarojuma kods, ko jānorāda pirmajās rindās, kas sākas ar *\$username* un *\$password*. Rīka pirmkods:

```
<?
function sendSMS($phone,$sms) {
    $username='youruser'; // replace w/ your access parameters
    $password='yourpass';
    $ntext='';
    $sms=substr($sms,0,200);
    for($ir=0;$ir<strlen($sms);$ir++) {
        $dump=ord($sms[$ir]);
        while(strlen($dump)<3)$dump='0'.$dump;
        $ntext.=$dump;
    }
    $to=urlencode(substr($phone,0,20));
    $username=urlencode($username);
    $password=urlencode($password);
    $tosend="user=$username&pass=$password&operator=1&to=$to&sms=$ntext&lang=lv";
    /*
    * POSSIBLE PARAMETERS:
    * (2nd parameter is the DEFAULT - if omitted)
    * lang=lv OR lang=en
    * operator=1 OR operator=0 //return operator name
    * digit=1 OR digit=0 //return answer code
    * nodesc=1 OR nodesc=0 //DO NOT return textual description
    * noqueue=1 OR noqueue=0 //DO NOT queue sms for sending later if target gate down
    * mail=1 OR mail=0 //send the sms via external mail services if available
    */
    $conlen=strlen($tosend);
    $connect='';
    $fp=@fsockopen('85.254.196.155',80,$temp,$temp,5) or
        $connect='no';
    if($connect=='no')
        return 'Nevar pieslēgties centrālajai vārtējai.';
    $data="POST /sendsms.php HTTP/1.0\n";
    $data.="Content-Type: application/x-www-form-urlencoded\n";
    $data.="User-Agent: KirOS SMS script | ".$_SERVER['HTTP_USER_AGENT']."\n";
    $data.="X-Forwarded-For: ".$_SERVER['REMOTE_ADDR'].
    ($_SERVER['HTTP_X_FORWARDED_FOR']?"-".$_SERVER['HTTP_X_FORWARDED_FOR'].":'")."\n";
    $data.="Host: your.gate.sms.id.lv\n";
    $data.="Accept-Encoding: identity\n";
    $data.="Content-Length: $conlen\n";
    $data.="Connection: Close\n";
    $data.="Cache-Control: no-cache\n\n";
    $data.=$tosend;
    @fputs ($fp, $data);
    while($er=@fgets($fp,4096))
        if(trim($er)=='')break;
    // return priority (order) - code, description, operator
    $result=trim(fgets($fp,4096));
    $operator=trim(fgets($fp,4096));
    fclose($fp);
    if($result>'')
        return "$operator: $result";
    else
        return 'Nekas netika atgriezts, droši vien vārtu parametri nav derīgi.';
}
?>
```


PIELIKUMS NR. 2: “STATISTIKAS RĪKA PIRMKODS”

Statistikas rīks *smsstat.php* rakstīts valodā PHP. Rīka pirmkods:

```
<?
require "ssl.php";
require "sms.php";

$delim=10;

switch($t){
  case 'robot': $clause=" realip IS NULL "; break;
  case 'human': $clause="NOT ( realip IS NULL )"; break;
  default: $clause="1"; break;
}

$o=mysql_query("SELECT count(*), count(distinct
realip),sum(status=1),sum(status=-25),sum(status=2),substring(whensent,1,$delim) as
dat,(select count(*) from
blockids where substring(setdate,1,$delim)=`dat` and reason like 'autoban:%') FROM
`processed` WHERE ($clause) and (whensent<DATE(NOW())) and (whensent>0)
GROUP BY dat ORDER BY dat ASC");

$r=array();
echo(mysql_error());
while ( $r[]=mysql_fetch_row($o));

array_pop($r);

for($i=0;$i<sizeof($r);$i++)
  $r[$i][2]=$r[$i][0]-$r[$i][2]-$r[$i][4]-$r[$i][3]; //!!!!/

$textm=array(
  'isziņas, no kurām',
  'dažādi cilvēki sūtīja',
  'netika nosūtītas;',
  'reizes sūtīšana atteikta dēļ bloka. Uzlikti ',
  'tika apsoliņš nosūtīt vēlāk,',
  'laikā',
  'automātiskie bani.',
);
echo("<table>");
foreach($r as $newn){
  echo("<tr>");
  foreach(array(5,1,0,4,2,3,6) as $item)
    echo("<td><b>{$newn[$item]}</b></td><td>{$textm[$item]}</td>");
  echo("</tr>\n");
}
echo("</table>");
?>
```

PIELIKUMS NR. 3: “SISTĒMAS TESTĒŠANAS REZULTĀTI”

Lai pārlicinātos par padarītā kvalitāti, gan no lietojamības, gan drošības viedokļa, tika veikti trīs netriviāli sistēmas testi – katrs, izmantojot savu saskarni.

Kā pirmā tika pārbaudīta populārākā – tīmekļa saskarne “sms.id.lv”. Testi (skatīt sekojošo tabulu) tika izpildīti noteiktajā secībā viens pēc otra.

Adresāts	Teksts	Sistēmas atbilde	Notikums
26036916	sveiki! sveiki!	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 2 sekunžu laikā
26036916	ka iet?	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā
26036916	ka iet?	Filtrs: plūdi - šāda īsziņa jau tikko ir nosūtīta.	Īsziņa netiek izsūtīta
~kirils	<i>(tiek iekopēts, nevis ierakstīts)</i> Sūtītājs darbojas testa režīmā. Šobrīd var sūtīt SMS uz LMT, TELE2 un Triatel tīkliem. Atnāc vēlāk un varbūt būs iespēja sūtīt arī kur citur. :	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā, taču sūtītājam uzrādās, ka īsziņa tiek sūtīta desmit sekundes
26036916	kaa iet?	Filtrs: plūdi - šāda īsziņa jau tikko ir nosūtīta.	Īsziņa netiek izsūtīta
22114441	bezmaksas reklaama! atvaino!	Īsziņu sūtīšana uz šo operatoru šobrīd nav pieejama.	Īsziņa netiek izsūtīta
26036915	bezmaksas reklaama! atvaino!	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta
26799999	bezmaksas reklaama! atvaino!	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta
26799991	bezmaksas reklaama! atvaino!	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta
<i>(tiek pagaidīts 2 minūtes)</i>			
26036916	sveiks, vai shodien aiziesi uz tirgu? vajag iegaadaaties to un sho un veel arii neaizmirsti to. ok?	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta, jo sistēma atceras lietotāju, kā potenciālo masveida ziņojumu izsūtītāju, tāpēc nelaiž cauri viņa ziņojumus. šis ir tikai pagaidu bloks. ja šajā brīdī lietotājs rimsies uz 10 minūtēm, tad viss atkal būs kārtībā
<i>(tiek pagaidīts 5 minūtes)</i>			
26036916	sveiks, vai shodien aiziesi uz tirgu? vajag iegaadaaties to un sho un veel arii neaizmirsti to. ok?	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 3 sekunžu laikā
26343802	cau	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā

26343802	ko dari?	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā
26343802	ku ku!	Filtrs: plūdi - šāda īsziņa jau tikko ir nosūtīta.	Īsziņa netiek izsūtīta
26343802	eu, nu, kapec man nekas nesutas???	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā
26343802	hehee, re kaa piechakareeju!	Filtrs: dalīšana - daudzu īsu ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta
26036916	eu, nu, kapec man nekas nesutas???	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 4 sekunžu laikā
26912345	eu, nu, kapec man nekas nesutas???	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā
26912346	eu, nu, kapec man nekas nesutas???	Īsziņa nosūtīta veiksmīgi.	Īsziņa pienāk 5 sekunžu laikā; sistēma pieļauj līdz pat 4 vienādu īsziņu izsūtīšanu īsā laikā, piemēram, lai savāktu kopā draugu bariņu, bet lielāks īsziņu skaits pieļauts netiek
26726900	eu, nu, kapec man nekas nesutas???	Filtrs: mēstule - nesaskaņota masveida ziņojumu izsūtīšana.	Īsziņa netiek izsūtīta
(tukšs)	tt	Bet kam sūtīt?	Īsziņa pat nenokļūst līdz sistēmas kodolam un netiek sūtīta
159	DOD 11111	Šis nav vietējs publiskā mobilā telefonu tīkla numurs.	Īsziņa netiek izsūtīta
(iepriekšējā rinda vēl 4 reizes)			
159	DOD 11111	Filtrs: dauzīšanās - nepārtraukta sūtīšanas pogas spaidīšana.	Īsziņa netiek izsūtīta. tiek izveidots automātiskais konfigurācijas bloks, kas bloķē īsziņas, kas nāk no sūtītāja IP adreses
22222222	sveiks, jāni!	Filtrs: Šāda īsziņas konfigurācija ir bloķēta. Īsziņa netiks nosūtīta. Lai vienotos par bloka noņemšanu, sazinieties ar prohibit [@at] ctrl.sms.id.lv pa e-pastu. (blocked {by GATE sms.id.lv at 2009-05-31 14:50:37} - #15324: [autoban: dumb hammering].)	Īsziņa netiek izsūtīta
(tiek pagaidīts 30 minūtes vai pat ilgāk)			
26036916	vai biji uz to tirgu, vai nee?	Filtrs: Šāda īsziņas konfigurācija ir bloķēta. Īsziņa netiks nosūtīta. Lai vienotos par bloka noņemšanu, sazinieties ar prohibit [@at] ctrl.sms.id.lv pa e-pastu. (blocked {by GATE sms.id.lv at 2009-05-31 14:50:37} - #15324: [autoban: dumb hammering].)	Īsziņa netiek izsūtīta

No šī testa var secināt, ka lietotājs ir ieguvis rīku, kas kvalitatīvi pilda savas funkcijas, kas nenoslogo lietotāju ar liekām darbībām, nosūtot katru īsziņu, taču vienlaicīgi arī aizsargā saņēmējus no nelietderīgām vai traucējošām īsziņām.

Kā nākošā tika testēta SMTP saskarne. Testēšanai tika izmantota programma “telnet”, lai pārlicinātos, ka SMTP servisam nav iespējami protokola līmeņa uzbrukumi.

Pirmā testēšanas sesija:

```
220 gate.sms.id.lv SMTP SMS Mail Server v1.9
FEAT
500 Command not implemented "FEAT". Donate some $$$ for further development.
HELO ps.id.lv
250 sms.id.lv Hello ps.id.lv [89.248.84.150], pleased to meet you
MAIL FROM: obama@whitehouse.gov
501 You are fake. Go away!
MAIL FROM: dns@dns.dn
501 Your host has no MX record defined! Get one.
MAIL FROM: abrakadabra@[1.1.1.1]
501 You are fake. Go away!
MAIL FROM: test@ps.id.lv
250 OK, you're test@ps.id.lv, i'll trust you this time.
RCPT TO: 26036916@sms.id.lv
250 OK, shall send to 26036916
DATA
354 Enter SMS(143), end with "." on a line by itself. I'll listen now. You speak.
421 Sorry, I'm getting bored, bye. Next time please say something in at least 10
seconds.
```

Šī rezultātā īsziņa pat netika mēģināts sūtīt.

Otrā testēšanas sesija:

```
220 gate.sms.id.lv SMTP SMS Mail Server v1.9
MAIL FROM: admin@ps.id.lv
250 OK, you're admin@ps.id.lv, i'll trust you this time.
RCPT TO: @...
501 And who the hell is that?
RCPT TO: 22114567
250 OK, shall send to 22114567
DATA
354 Enter SMS(143), end with "." on a line by itself. I'll listen now. You speak.
Sveiki!
.
501 SMS NOT sent: Operator currently not supported.
221 OK, I have to go now. See you later!
```

Šī rezultātā savukārt īsziņa caur vārteju nokļūst līdz sistēmas kodolam, kur tiek noskaidrots numura operators un īsziņa noraidīta ar kodu -14.

Trešā testēšanas sesija:

```
220 gate.sms.id.lv SMTP SMS Mail Server v1.9
HELO `cat /etc/passwd`
250 sms.id.lv Hello `cat /etc/passwd` [89.248.84.150], pleased to meet you
MAIL FROM: hacker@`cat /etc/shadow`
501 You are invalid... Sorry.
MAIL FROM: hacker@|cat /etc/passwd
501 You are invalid... Sorry.
MAIL FROM: hacker@a b
501 You are invalid... Sorry.
MAIL FROM: hacker@localhost
```

```
501 Your host has no MX record defined! Get one.
MAIL FROM: billy@microsoft.com
501 You are fake. Go away!
MAIL FROM: lovelyuser@ps.id.lv
250 OK, you're lovelyuser@ps.id.lv, i'll trust you this time.
rcpt to: nosaukums@sms.id.lv
250 OK, shall send to nosaukums
DATA
354 Enter SMS(143), end with "." on a line by itself. I'll listen now. You speak.
atvainojos
.
501 SMS NOT sent: User has not enabled SMS receival from web.
221 OK, I have to go now. See you later!
```

Arī šī testa rezultātā īsziņa caur vārteju nokļūst līdz sistēmas kodolam, kur tā pat tiek sūtīta operatoram, taču operators īsziņu noraida, jo šāds LMT identifikators nav reģistrēts.

Pēdējā testēšanas sesija:

```
220 gate.sms.id.lv SMTP SMS Mail Server v1.9
HELO spam
250 sms.id.lv Hello spam [89.248.84.150], pleased to meet you
MAIL FROM: evil.spammer@ps.id.lv
250 OK, you're evil.spammer@ps.id.lv, i'll trust you this time.
RCPT TO: 26036916
250 OK, shall send to 26036916
RCPT TO: 26036917
550 Only one sms recipient can be specified, sorry, we're out of [food] resources.
DATA
354 Enter SMS(143), end with "." on a line by itself. I'll listen now. You speak.
Subject: vienkārshi SMS
llalallalalalalalalalal lal all :))) hee
un veel viena rinda
.
250 Sent successfully.
221 OK, I have to go now. See you later!
```

Šajā gadījumā SMS tiek piegādāta telefona aparātā 6 sekunžu laikā.

Kā pēdējā tiek testēta WAP saskarne. Ņemot vērā to, ka šī saskarne ir būtība tīmekļa saskarnes “sms.id.lv” ierobežota versija, tests sastāv no divām vienādām īsziņām ar vienādiem adresātiem. Uz pirmo WAP saskarne atbild “Your SMS was sent successfully: Sent successfully.”, bet uz otro “Your SMS was NOT sent successfully: Filter: flood.”

Līdz ar to arī šeit ir nodrošināts maksimāls ērtums, neietekmējot drošību.

PIELIKUMS NR. 4: “LIETOŠANAS STATISTIKA”

Tā kā sistēmas lietotāju skaits un līdz ar to arī slodze un citi parametri ir atkarīgi no nedēļas dienas, tad iespēju robežās, kur tas ir būtiski, tika analizēti dati par nedēļas periodu. Sākumā redzami dati kvantitatīvi par periodu pirms pirms ilgstošā sistēmas darbības pārtraukuma un bakalaura darba uzsākšanas. Dati jāizvērtē ar lielu piesardzību, jo jāņem vērā gan tas, ka neilgi pirms sistēmas darbības pārtraukšanas tā bija ļoti nestabila, gan arī tas, ka bakalaura darba izstrādes laikā vēl nav paspējusi nostabilizēties lietotāju bāze.

Datums	Lietotāji	Īsziņas	Nosūtīts	Aizturēts	Nav nosūtīts	Bloķēts	IRC	WAP	SMTP
2008-08-11	648	1633	929	0	610	94	0	24	0
2008-08-12	721	1826	1049	0	676	101	0	12	0
2008-08-13	694	1810	1013	1	691	105	0	30	0
2008-08-14	608	1619	964	0	559	96	0	19	0
2008-08-15	594	1484	897	0	508	79	0	22	0
2008-08-16	486	1192	650	0	470	72	0	36	0
2008-08-17	471	1226	637	0	488	101	0	32	0
SUMMA	4222	10790	6139	1	4002	648	0	175	0

Šajā statistikā gan IRC, gan SMTP serviss izpelnījies 0 īsziņas pa visu nedēļu. Jāņem vērā, ka IRC serviss šajā periodā funkcionēja, bet SMTP – nefunkcionēja.

Un tagad kvantitatīvā statistika par laika periodu pēc sistēmas “sms.id.lv” darbības atjaunošanas. Jāņem vērā, ka svētdienas dati ņemti no citas nedēļas, jo rakstīšanas brīdi vēl nav datu par attiecīgās nedēļas svētdienu.

Datums	Lietotāji	Īsziņas	Nosūtīts	Aizturēts	Nav nosūtīts	Bloķēts	IRC	WAP	SMTP
2009-05-25	341	1244	804	3	407	30	-	2	39
2009-05-26	302	1202	838	2	335	27	-	3	34
2009-05-27	283	1105	753	14	306	32	-	1	28
2009-05-28	306	1103	752	3	323	25	-	1	41
2009-05-29	314	1248	773	1	394	80	-	2	53
2009-05-30	255	9244	686	1	219	18	-	1	28
2009-05-24	1535	3918	1813	4	1725	376	-	2	26
SUMMA	3336	19064	6419	28	3709	588	-	12	249

Zemāk ir salīdzināta sistēmas darbība vienā un otrā laika periodā. Nākamajā tabulā ir laiks sekundēs, cik ilgi lietotājam bija jāgaida, lai saņemtu katru no kļūdu paziņojumiem.

Kļ.	2	1	0	-1	-3	-4	-5	-9	-10	-11	-13	-14	-16	-17	-21	-22	-23	-24	-25	avg
08	13,04	0,57	0,32	0,50	-	-	0,05	-	3,24	0,01	0,32	2,56	0,33	0,48	0,29	0,31	0,46	0,47	0,47	1,18
09	13,20	0,83	13,22	0,29	16,96	0,68	0,04	7,98	2,53	0,00	0,21	1,44	6,45	0,43	0,16	0,17	0,16	0,31	0,31	1,03

Pēdējā kolonnā redzams vidējais svērtais gaidīšanas laiks. Redzams, ka kopumā īsziņu sūtīšanas ātrums nokrities līdz 1,03 sekundēm. Lietotājam ātrāk tiek sniegtas negatīvās atbildes, kas ļauj viņam ātrāk uzsākt savas kļūdas labošanu.

Nedaudz par OCR kvalitāti, kopš uzsākts bakalaura darbs:

Operators	Atrisināts pareizi	Atrisināts kļūdaini	Kļūdas procents (%)
LMT	3674	82	2,18
TELE2	5712	68	1,18
BITE	0	5	100

Jāņem vērā, ka, saņemot negatīvu atbildi no operatora, sistēmas kodols mēģinās bez lietotāja papildus iejaukšanās vēl maksimums četras reizes automātiski pārslēgties uz citu sesiju un ievadīt citu kodu. Reāla iespējamība, ka lietotājs saņems atteikumu dēļ OCR nepilnības tāpat ir $0.0218^5 = 1$ no 203 miljoniem gadījumiem (LMT) un $0.0118^5 = 1$ no 4.3 miljardiem gadījumu (TELE2). Tas pierādās arī praksē, un, lai gan OCR pārraudzības rīks uzrāda kļūdaini atpazītus OCR, lietotāja atteikums dēļ sliktas kvalitātes OCR kopš sistēmas uzlabojumu pabeigšanas vēl nav konstatēts.

Automātisko konfigurāciju bloķēšanas rezultātā pētāmās nedēļās laikā ir uzlikti 60 dažādu konfigurāciju bloki:

Iemesls	Skaits	Skaidrojums
autoban: dumb hammering	10	Bloķēti lietotāji, kas ir ignorējuši sistēmas kļūdu un vienkārši spieduši pogu "Sūtīt" atkal un atkal. Spilgts piemērs ir lietotāji, kas cenšas nosūtīt īsziņas uz pakalpojumu kodu 159, lai gan viņiem tiek paziņots par neatbilstošu numuru.
autoban: pre-process hammering	10	Bloķēti lietotāji, kas ir ignorējuši sistēmas nosacījumu, kas tiek pārbaudīts vēl pirms kārtīgas īsziņas apstrādes. Parasti tas ir -5 kļūdas kods "īsziņa pārāk gara". Šo izraisa vai nu tas, ka lietotāja pārlūks nesaprot JavaScript, vai arī tas, ka lietotājs neraksta latīņu alfabēta burtiem un tie aizņem vairāk vietas, nekā redzams uz ekrāna. Visi šie bloki tika noņemti.
autoban: flood	1	Bloķēts lietotājs, kas centies numuru 22388___ pārpludināt ar īsziņām ar tekstu "a".
autoban: complex flood	12	Bloķēti lietotāji, kas centušies kādu konkrētu numuru pārpludināt ar bezjēdzīgām īsziņām, cenšoties apmānīt sistēmu, nedaudz pamainot to saturu. Biežākais teksta piemērs: "cxcxc". Šiem lietotājiem vidēji izdevies nosūtīt 2 līdz 3 kaitnieciskās īsziņas līdz sistēma viņus bloķējusi.
autoban: blind spam	15	Bloķēti lietotāji, kas centušies izsūtīt tekstus uz patvaļīgiem numuriem, nemaz nezinot, vai tādi eksistē. Biežākais teksta piemērs: "caw".
autoban: splitting	12	Bloķēti lietotāji, kas sūtījuši daudz īsas īsziņas uz vienu numuru, tā vietā, lai aizsūtītu vienu vai vairākas garākas. Lietotāji vidēji bloķēti pēc apmēram 3 vai 4 īsziņām, kuru vidējais garums svārstījās no 4 līdz pat 30 simboliem.

DOKUMENTĀRĀ LAPA

Bakalaura darbs “Īsziņu sūtīšanas sistēma “sms.id.lv”” izstrādāts LU Fizikas un matemātikas fakultātē.

Ar savu parakstu apliecinu, ka bakalaura darbs veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: Kirils Solovjovs

Rekomendēju darbu aizstāvēšanai.

Vadītājs: docents Dr. mat. Paulis Ķikusts

Recenzents: Rihards Rūmnieks, augstāka izglītība, pielīdzināta maģistra grādam

Darbs iesniegts Datorikas nodaļā

Metodiķe:

Darbs aizstāvēts bakalaura gala pārbaudījuma komisijas sēdē

_____ prot. Nr. __, vērtējums

Komisijas sekretārs: