ISSN-P 2407-2192

Jurnal Teknik Informatika Politeknik Sekayu (TIPS)
Volume II, No. 1, April 2015, h. 41 - 52

MINI TOOL FORENSIK UNTUK *SMARTPHONE ANDROID* MENENTUKAN PROBABILITAS SPAM PADA PESAN PENDEK

Zaid Romegar Mair
NUPN : 9902702271
Teknik Informatika Politeknik Sekayu
Email : romegardm@gmail.com

ABSTRAK

Tekno*log*i merupakan alat bantu yang diperlukan manusia. Salah satunya adalah ponsel, diantara *Smartphone* yang familiar yang digunakan oleh masyarakat adalah *smartphone* berbasis *android*. Kecanggihan fitur layanan yang diberikannya membuat oknum-oknum tertentu menyalahgunakan pemanfaatannya untuk merencanakan sebuah kejahatan. Sehingga dalam penelitian ini dibuatlah sebuah *mini tool* forensik untuk *smartphone android*. Objek kasus yang dianalisis berupa *smartphone* samsung S III mini, Asus T00J dan *Samsung galaxy young*. Berdasarkan penelitian, dilakukan pengujian terhadap memori internal dengan mengikuti langkah-langkah forensik dasar dalam melakukan penanganan kasus digital.sehingga didapatlah hasil dari analisa tersebut bahwa Samsung S III mini yang memiliki 581 SMS, jumlah spam hasil analisis awal 0. Jumlah spam hasil analisis *Naive Bayesian Filtering* Versi 1 : 388. Jumlah spam hasil analisis Naive *Bayesian* Filter Versi 2 : 381. Asus T00J yang memiliki 176 SMS, jumlah spam hasil analisis awal 0. Jumlah spam hasil analisis *Naive Bayesian Filtering* versi 1 : 155. Jumlah spam hasil analisis *Naive Bayesian Filter* Versi 2 : 147. *Samsung galaxy young* yang memiliki 979 SMS, jumlah spam hasil analisis awal 23. Jumlah spam hasil analisis *Naive Bayesian Filtering* versi 1 : 534. Jumlah spam hasil analisis *Naive Bayesian Filter* Versi 2 :357.

Kata Kunci : *Naive Bayesian Filtering*, Forensik, *Smartphone*.

I. PENDAHULUAN

1.1 Latar Belakang

Tekno*log*i merupakan alat bantu yang diperlukan manusia. Salah satunya adalah ponsel, diantara *Smartphone* yang familiar digunakan oleh masyarakat adalah *smartphone* berbasis *android*. Kecanggihan fitur layanan yang diberikannya membuat oknum-oknum tertentu menyalahgunakan pemanfaatannya untuk merencanakan sebuah kejahatan. Maraknya aksi kejahatan yang sering terjadi adalah dengan melakukan pengiriman pesan singkat (SMS). SMS tersebut membuat para pengguna *smartphone* semakin tergangu atas ketidak nyamanan yang dilakukan oleh pihak-pihak tersebut.

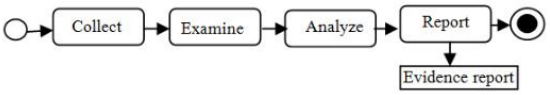
1.2 Perumusan Masalah

Berdasarkan uraian latar belakang dapat diidentifikasi permasalahan sebagai berikut :

1. Bagaimana membangun *mini tool* forensik untuk pengidentifikasian pesan kedalam kategori spam atau bukan?
2. Bagaimana cara melakukan filtering terhadap setiap frase menggunakan *Naïve Bayesian Filter*?
3. Bagaimana melakukan analisis spam terhadap beberapa jenis *smartphone* yang memiliki sistem operasi sama?
4. Bagaimana cara melakukan probabilitas pesan spamming untuk mendapatkan persentase maksimal menggunakan *Naïve Bayesian Filter*?

1.3 Metode Penelitian

Metode yang peneliti gunakan untuk menyelesaikan permasalahan ini adalah metode forensik dasar yang terdiri dari 4 fase penanganan kasus digital. Menurut Kent, Chevalier, Grance and Dang [2006:800-86] yaitu :



Gambar 1.1 Model Proses Forensik

1. Tahap pengumpulan *(Collection)*: pada tahap pertama ini peneliti melakukan pengumpulan bukti *(evidence)* berupa *smartphone* berbasis *android*. Pada *smartphone* tersebut dilakukan instalasi aplikasi berupa *Android Forensic Logical* OSE (AF*Log*ical OSE). Aplikasi tersebut berfungsi mengcapture data-data yang diperlukan untuk dianalisis. Pengambilan data dengan ekstensi *.csv *(Comma Separated Values)* menggunakan transmisi media kabel dari handphone ke komputer.
2. Tahap pemeriksaan *(Examination)*: Pada tahap ini, dilakukan pemilihan terhadap menu yang disediakan *(available providers)* untuk melakukan ekstraksi data oleh aplikasi AF*Log*ical OSE, diantaranya adalah *CallLog Calls*, *Contactc Phones*, MMS, MMSParts, SMS.
3. Tahap analisis *(Analysis)*: Pada tahap analisis ini, peneliti melakukan *filtering* terhadap kata tertentu dengan memasukkan frase dan disimpan pada sebuah *database* spam, memiliki 3 tabel berupa tabel *knowledge_base*, tabel *phrase* dan tabel sms. Tahap ini menentukan apakah pesan tersebut termasuk kedalam kategori spam atau bukan dengan menggunakan algoritma *Bayesian filter*.
4. Tahap laporan *(Reporting):* pada tahap ini bisa dilihat bentuk laporannya dengan memilih menu laporan pada aplikasi yang dibuat. Laporan tersebut menghasilkan kategori pesan spam dan non spam dengan melakukan perbandingan terhadap *frase* dan *Naive Bayesian* Penyaringan algoritma versi 1 dan versi 2.

II. TINJAUAN PUSTAKA

3.1. Forensik dan Turunan Digital Forensik

Menurut K. Franke (2010), Ilmu forensik merupakan metodo*log*i aplikasi yang benar dari spectrum yang luas dari disiplin ilmu, untuk menjawab pertanyaan yang signifikan terhadap sistem hukum. Metode forensik terdiri dari pendekatan untuk melakukan tugas seperti (1) menyelidiki tempat kejadian, (2) mengumpulkan dan menganalisis jejak bukti yang ditemukan, (3) mengidentifikasi, mengklasifikasi, menghitung masing-masing orang, objek, proses, (4) menetapkan hubungan, asosiasi dan rekonstruksi dan (5) menggunakan temuan dalam proses penuntutan atau pembelaan di pengadilan hukum. Forensik sebagian besar menangani kejahatan yang dilakukan sebelumnya, fokusnya untuk mencegah kejahatan di masa depan. Menurut Gavin, P., Robinson, L.A. dan Ellis, R., (2010) berikut adalah cabang dari digital forensik:

Gambar 2. 1 Cabang Digital Forensik

2.2.1. Forensik Komputer dan Forensik Jaringan

Menurut E. Winarno dan A.Zaki (2012) Komputer forensik adalah sebuah proses investigasi peranti komputer atau peranti simpannya, baik berupa komputer pribadi, laptop, *server*, PC kantoran atau media removable untuk menentukan apakah komputer atau peralatan ini digunakan untuk keperluan ilegal, tidak sah atau tidak biasa. Pada prinsipnya kegiatan forensik ini merupakan bagian dari *hacking* hingga akan ditindak lanjuti dengan pemindaian *vulnerability* dan eksploitasi.

2.2.2. Forensik *Database*

Forensik *database* adalah cabang ilmu forensik digital yang berkaitan dengan studi forensik *database* dan metadata. Disiplin ilmu forensik *database* mirip dengan forensik komputer, mengikuti proses forensik normal dan menerapkan teknik-teknik investigasi untuk isi *database* dan metadata. Pemeriksaan forensik *database* berhubungan dengan waktu *update* dari baris dalam tabel relasional yang diperiksa dan diuji validitas untuk verifikasi tindakan dari pengguna *database*. Pemeriksaan forensik dapat fokuskan pada mengidentifikasi transaksi dalam sistem *database* atau aplikasi yang menunjukkan bukti melakukan hal yang salah, seperti penipuan. Perangkat lunak pihak ketiga yang menyediakan lingkungan read-only dapat digunakan untuk memanipulasi dan menganalisis data. Alat-alat ini juga menyediakan kemampuan audit *logging* yang memberikan bukti dokumentasi atau analisis pemeriksa forensik yang dilakukan pada *database.*

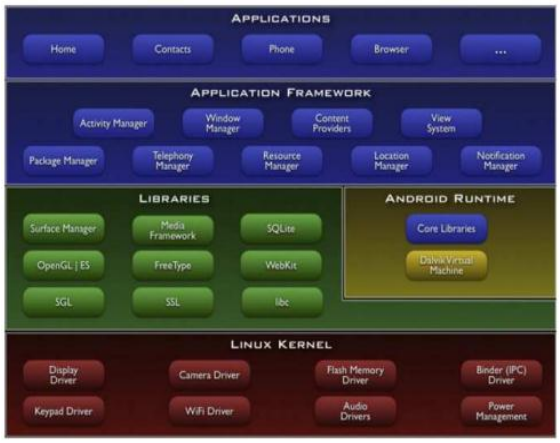
2.2.3. Forensik Perangkat Mobile

Ilmu forensik perangkat *mobile* adalah bidang studi yang relatif baru, berasal dari awal tahun 2000-an. Forensik perangkat *mobile* merupakan cabang sub-forensik digital yang berkaitan dengan pemulihan bukti digital atau data dari perangkat *mobile*. Menurut A. Hoog (2010) Mulai Oktober 2010, perangkat *smartphone* berbasis *android* mewakili 22% dari pasar *smartphone* di Amerika Serikat. Menurut survei yang dilakukan oleh Nielsen (perusahaan *information global* yang berpusat di New York dan belanda) April 2011, 36 persen konsumen *smartphone* memiliki perangkat *Android*, 26 persen untuk Apple IOS *smartphone* (iPhone) dan 23 persen untuk RIM Blackberry. Hal ini diterima secara luas bahwa *android* akan melampaui iPhone dan mungkin pada akhirnya akan menjadi *platform smartphone* yang paling populer.

Menurut E. casey (2004) Investigasi biasanya fokus pada data sederhana seperti data panggilan dan komunikasi (*SMS / Email*) dari pada mendalam pemulihan data yang dihapus. Perangkat *mobile* juga berguna untuk memberikan informasi lokasi, baik dari *gps inbuilt* / lokasi pelacakan atau melalui situs sel *log*, yang melacak perangkat dalam jangkauan mereka (Casey, 2004).

3.2. Arsitektur *Android*

Sistem operasi *android* memiliki 4 lapisan *(layer)* yang merupakan komponen sistem *android*. Gambar 2.2 merupakan lapisan arsitektur sistem operasi *android*:



Gambar 2.2 Arsitektur Sistem Operasi *Android*

Berikut penjelasan gambar arsitektur sistem operasi *android* :

1. *Applications*

Di lapisan teratas akan menemukan fungsi-fungsi dasar *smartphone* seperti menelepon dan mengirim pesan singkat, menjalankan web browser, mengakses daftar kontak, dan lain-lain. Bagi rata-rata pengguna, lapisan inilah yang paling sering mereka akses. Mereka mengakses fungsi-fungsi dasar tersebut melalui user interface.

2. *Application Framework*

Application framework merupakan serangkaian *tool* dasar seperti aplikasi telepon, pergantian antar – proses atau program, dan pelacakan lokasi fisik telepon.

3. *Libraries*

Bertempat di level yang sama dengan *android runtime* adalah *libraries*. *Android* menyertakan satu *set library-library* dalam bahasa C/C++ yang digunakan oleh berbagai komponen yang ada pada sistem *Android*. Kemampuan ini dapat diakses oleh *programmer* melewati *android application framework*. Sebagai contoh *Android* mendukung pemutaran format audio, video, dan gambar.

4. *Android Run Time*

Lapisan setelah kernel linux adalah *android runtime*. *Android runtime* ini berisi *core libraries* dan *dalvik virtual machine*. *Android* terdiri dari satu set core libraries yang menyediakan sebagian besar fungsi yang tersedia dalam *core libraries* dari bahasa pemrograman Java. *Dalvik virtual machine* adalah Aspek unik dari *android* dan komponen penting dalam forensik *smartphone android*, yang dikembangkan oleh Google untuk membuat efisien untuk mengoptimalkan telepon seluler dan aman di lingkungan aplikasi mobile. *Java virtual machine* yang memberi kekuatan pada sistem *android*. Setiap aplikasi yang berjalan pada *android* berjalan pada prosesnya sendiri, dengan instance dari dalvik virtual machine. Dalvik telah dibuat sehingga sebuah piranti yang memakainya dapat menjalankan multi Virtual Machine dengan efisien**,** Untuk mencapai efisiensi, aplikasi yang berjalan di Dalvik VM memiliki Format khusus yang disebut executable Dalvik (. dex) file. Dalvik VM dapat mengeksekusi file dengan format Dalvik executable (.dex) yang telah dioptimasi untuk menggunakan minimal *memory footprint*. Virtual Machine ini register-based, dan menjalankan class-class yang dicompile menggunakan compiler Java yang kemudian ditransformasi menjadi format .dex menggunakan "dx" tool yang telah disertakan. *Dalvik Virtual Machine* (VM) menggunakan kernel Linux untuk menjalankan fungsi-fungsi seperti threading dan *low-level memory management*.

5. *Linux Kernel*

Smartphone android secara keseluruhan bukanlah linux, karena dalam *android* tidak terdapat paket standar yang dimiliki oleh linux lainnya. Tumpukan paling bawah pada arsitektur *Android* ini adalah kernel. yang mencakup memory management, security setting, power management, dan beberapa driver hardware. Kernel berperan sebagai *abstraction layer* antara *hardware* dan keseluruhan *software*. Kernel *Android* terdapat *driver* kamera yang memungkinkan pengguna mengirimkan perintah kepada *hardware* kamera. Linux merupakan sistem operasi terbuka yang handal dalam manajemen memori dan proses. Oleh karenanya pada *android* hanya terdapat beberapa servis yang diperlukan seperti keamanan, manajemen memori, manajemen proses, jaringan dan *driver*. Kernel linux menyediakan *driver* layar, kamera, *keypad*, *Wifi*, *Flash Memory*, audio, dan IPC

(*Inter Process Communication)* untuk mengatur aplikasi dan lubang keamanan.

3.3. Universal Serial Bus (USB) Debugging dan AF*Log*ical-OSE

USB debugging adalah modus sambungan pada perangkat *android* untuk mentransfer data antara komputer dan perangkat *smartphone*, membaca data *log* dari perangkat, dan menggunakan perintah *debugging* saat pengambilan data dari *smartphone* untuk dianalisis. Pembuatan file ekstrak dengan cara mengcapture pesan SMS yaitu dengan menggunakan AF*Log*ical-OSE. AF*Log*ical-OSE merupakan *android* forensik yang digunakan untuk mengekstrak data SD *card* dari *device.* AF*Log*ical OSE dirilis pada bulan Desember 2011 dan kini *host* di GitHub. Aplikasi ini memberikan kerangka dasar untuk mengekstraksi data dari perangkat *Android* menggunakan *Content Provider* dan kemudian menyimpan data ke SD *Card. File* ekstrak berupa kontak, *Log* Panggilan, SMS, MMS, MMS *Parts* dan Info perangkat.

3.4. Short Message Service (SMS) Spam

Tidak ada definisi yang baku secara international dan konstitusi mengenai arti dari SMS *spam*. Beberapa negara memiliki definisi sendiri mengenai arti dari SMS spam. Namun intinya, SMS *spam* adalah SMS yang tidak diinginkan oleh pihak penerima SMS. Contoh dari SMS spam adalah SMS penipuan, SMS ancaman, SMS promosi, dan lain-lain. Menurut I.I Androulidakis (2012) diperkirakan 72% dari semua ponsel pelanggan di seluruh dunia adalah pengguna aktif SMS. Ini merupakan masalah utama secara global, karena semua penggunanya merupakan target bagi spam, spoofing, dan SMS lainnya yang terkait dengan penipuan. Dampak dari spam SMS dapat menurunkan kinerja jaringan dan kualitas layanan (QoS), mengakibatkan hilangnya pendapatan dan meningkatkan kekhawatiran pelanggan.

3.5. Bayes

Bayesian filter merupakan metode terbaru yang digunakan untuk mendeteksi *spam mail*. Algoritma ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas dimasa depan berdasarkan pengalaman di masa sebelumnya. Dua kelompok peneliti, satu oleh Pantel dan Lin, dan yang lain oleh Microsoft Research memperkenalkan metode statistik *Bayesian* ini pada tekno*log*i anti *spam filter*. Tetapi yang membuat algoritma *Bayesian* filtering ini popular adalah pendekatan yang dilakukan oleh Paul Graham. *Bayesian* filter mendeteksi *spam* dengan cara menghitung probabilitas dari suatu pesan berdasarkan isinya. Probabilitas ini dapat dihitung dengan terlebih dahulu membuat suatu database *spam-SMS* dan database *non spam-SMS*. Kemudian dengan suatu *metode training*, *software* anti spam yang menggunakan algoritma *Bayesian* dapat dilatih untuk melihat kata-kata yang sering digunakan pada *spam-mail*, sehingga pada akhirnya dihasilkan filter anti spam yang akurat dengan sesedikit mungkin *false positives*. *False positives* adalah e-mail legal yang ditujukan kepada penerima, tetapi karena kesalahan dari filter anti spam, dikategorikan menjadi *spam SMS.*

3.6. Perhitungan Probabilitas Berdasarkan Algoritma *Bayesian*

Pada awalnya, *Bayesian* filter ini harus di-training terlebih dahulu menggunakan sejumlah spam dan sejumlah ham. *Bayesian* filter akan menghitung probabilitas lokal dari suatu kata,

misalnya kata “adult”, untuk muncul di kelompok spam mail. Probabilitas lokal ini dapat dirumuskan sebagai berikut :

$$P_{local-spam} = N_{spam} / (N_{spam} + N_{non-spam})$$

Keterangan :

$P_{local-spam}$ = probabilitas suatu kata “x” terdapat pada spam – mail

N_{spam} = jumlah spam dengan kata “x” didalamnya

$N_{non-spam}$ = jumlah non-spam dengan kata “x”didalamnya

III. ANALISA DAN PERANCANGAN

3.1 Analisa Permasalahan

Short Message Service (SMS) merupakan media komunikasi yang paling sering digunakan. Kejahatan yang muncul dengan tekno*log*i inipun beragam, sehingga untuk menghadapinya, peneliti mencoba untuk mengusulkan pembuatan *mini tool* forensik untuk mengkategorikan pesan pendek pada *smartphone* dengan memakai bahasa pemrograman PHP, Dreamweaver sebagai editor, dan MySQL sebagai databasenya. Aplikasi ini diharapkan nantinya dapat membantu dalam mengidentifikasi pesan pendek pada *smartphone.*

3.2 Analisis Sistem

Perangkat *smartphone* berbasis *Android* mempunyai fasilitas *messaging* (perpesanan) sebagaimana *handphone* pada umumnya. Dengan dukungan sistem operasi yang bersifat *open source*, maka pengguna lebih leluasa dalam mengelola data, baik yang tersimpan dalam memori internal maupun eksternal tanpa memerlukan peralatan tambahan. Hal ini memungkinkan pengguna untuk mengelola data yang terdapat dalam *smartphone Android* menggunakan berbagai perangkat IT untuk keperluan yang lebih bermanfaat.

3.3 Spesifikasi Alat dan Bahan

Dalam melakukan penelitian ini, Peneliti menggunakan beberapa peralatan yang menunjang kegiatan penelitian, yaitu;

a. Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan terdiri dari *Personal Computer* (PC) Processor Intel Core 2 Duo 1,86 Ghz, RAM 4 GB, VGA NVIDIA Geforce 1 GB, Harddisc 500 GB, dan Monitor 19”

b. Perangkat Lunak (*Software*)

Perangkat lunak yang digunakan dari Sistem Operasi Windows 7 *Ultimate* 64 Bit, *Microsoft Office* 2010 dan *Vertrigoserv.*

3.4 Rancangan Sistem

Rancangan antarmuka aplikasi untuk *mobile* berbasiskan web ditunjukkan pada Gambar 3.1.

Gambar 3.1 Rancangan Sistem

Pada perancangan menu utama sistem identifikasi spamming SMS di atas terdapat beberapa menu, yaitu dari beranda, data, penenentuan jenis SMS, analisis, cek spam, cek non spam, hasil test dan laporan. Data memuat

frase spam dan *import SMS*. *Frase spam* berisikan kumpulan kata atau frase yang merupakan bagian dari SMS, sedangkan *import SMS* berfungsi untuk *import* data SMS yang telah diekstrak dari *smartphone Android* dengan menggunakan tool AF*Log*ical-OSE yang berekstensikan .csv. Penentuan jenis SMS merupakan fasilitas untuk memberikan asumsi atau justifikasi awal dari isi suatu SMS menggunakan notasi "ham" sebagai non spam, "spam" sebagai spam dan "0" untuk isi SMS yang tidak terdefinisi.

IV. IMPLEMENTASI

4.1 Tahap Pengumpulan (*Collection*)

Pengumpulan barang bukti berupa data SMS (baik mengandung *spam* atau tidak) yang dikirimkan melalui nomor provider ke *smartphone android*. Data tersebut diekstrak menggunakan tool AF*Log*ical-OSE, kemudian ditranfer ke computer menggunakan media kabel data. Pesan yang ada pada *smartphone* tersebut berupa iklan, promosi, transaksi jual beli, hadiah dan pesan yang sifatnya umum.

Gambar 4.1 Ekstraksi SMS dengan AF*Log*ical OSE

4.2 Persiapan Web *Server*

Webserver yang digunakan dalam penelitian ini adalah *apache* dan database *server* yang digunakan yaitu *mysql*.

4.3 Halaman Utama

Halaman utama merupakan tampilan awal pada saat sistem pertama kali dibuka. Tampilan utama memiliki tiga menu utama yang berada dibagian atas terdiri dari branda, data (frase spam dan data sma) dan laporan *(reporting)*. Sedangkan dibagian sidebar kiri terdiri dari menu yang digunakan untuk menganalisis data *smartphone android* yang terdiri dari tentukan status sms, analisis, hasil analisis sms, cek probabilitas frase spam dan cek probabilitas frase non spam. Objek penelitian yang dianalisis berupa *smartphone android* samsung S III mini, Asus T00J dan *Samsung galaxy young*.

Gambar 4.2 Halaman Utama

4.4 Membuat dan Mengimport Database

Langkah ini merupakan persiapan untuk menjalankan aplikasi analisis spam. Dalam tahap ini, diperlukan pembuatan database SQL dan mengimport format tabel ke dalam database. Proses ini juga untuk mempermudah menganalisis data dengan menggunakan perintah-perintah sql maupun php. Langkah-langkah menbuat database yaitu pertama menjalankan VertrigoServ. Pastikan servis apache dan mysql berjalan dengan normal. Buka web browser, ketikkan localhost/phpmyadmin di address bar. Isikan username: *root* dan *password*: vertrigo. Isikan nama *database*: spam pada kotak

create new database kemudian klik tombol *create*.

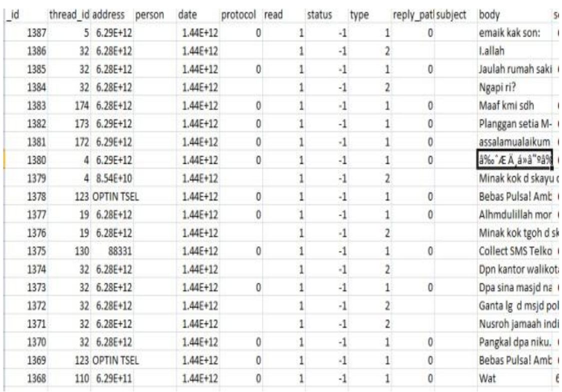
Sedangkan untuk melakukan impor database langkahnya adalah database *spam* sudah dibuat. Pilih database dengan nama *spam* kemudian klik untuk membuka dan melihat daftar tabelnya. Klik Browse dan pilih file *spam.sql* kemudian pilih Import. Hasil dari import database spam mempunyai 3 tabel.

Tabel 4.1 Tabel Database Spam

No	Nama Tabel	Field	Keterangan
1	Phrase	Id	Id frase
		Phrase	Kata atau frase yang dicurigai sebagai spam
		Date_added	Tanggal simpan
		prob	Nilai probabilitas kata atau frase
2	Sms	Id	Id sms
		Pengirim	No pengirim pesan
		Waktu	Waktu pesan dikirim
		Pesan	Isi pesan
		State	Penentuan status berdasakan isi pesan
3	Knowledge_base	Ngram	N-grammer, potongan perkata
		Belong	Status potongan kata
		Repite	Jumlah potongan kata ditemukan dalam pesan
		percent	Prosentase probabilitas

4.4.1. Membaca SMS dari Perangkat Mobile

Pembacaan SMS dilakukan pada perangkat *smartphone* Samsung Galaxy S III mini dengan menggunakan aplikasi AF*Log*ical-OSE yang menghasilkan file berektensi csv yang berisi record pesan pada memori internal *handphone*.



_id	thread_id	address	person	date	protocol	read	status	type	reply_patl	subject	body
1387	5	6.29E+12		1.44E+12	0	1	-1	1	0		emaik kak son:
1386	32	6.28E+12		1.44E+12		1	-1	2			I.allah
1385	32	6.28E+12		1.44E+12	0	1	-1	1	0		Jaulah rumah saki
1384	32	6.28E+12		1.44E+12		1	-1	2			Ngapi ri?
1383	174	6.28E+12		1.44E+12	0	1	-1	1	0		Maaf kmi sdh
1382	173	6.29E+12		1.44E+12	0	1	-1	1	0		Planggan setia M-
1381	172	6.29E+12		1.44E+12	0	1	-1	1	0		assalamualaikum
1380	4	6.29E+12		1.44E+12	0	1	-1	1	0		[illegible]
1379	4	8.54E+10		1.44E+12		1	-1	2			Minak kok d skayu
1378	123	OPTIN TSEL		1.44E+12	0	1	-1	1	0		Bebas Pulsa! Amb
1377	19	6.28E+12		1.44E+12	0	1	-1	1	0		Alhmdulillah mor
1376	19	6.28E+12		1.44E+12		1	-1	2			Minak kok tgoh d sk
1375	130	88331		1.44E+12	0	1	-1	1	0		Collect SMS Telko
1374	32	6.28E+12		1.44E+12		1	-1	2			Dpn kantor walikot
1373	32	6.28E+12		1.44E+12	0	1	-1	1	0		Dpa sina masjd na
1372	32	6.28E+12		1.44E+12		1	-1	2			Ganta lg d msjd pol
1371	32	6.28E+12		1.44E+12		1	-1	2			Nusroh jamaah indi
1370	32	6.28E+12		1.44E+12	0	1	-1	1	0		Pangkal dpa niku.
1369	123	OPTIN TSEL		1.44E+12	0	1	-1	1	0		Bebas Pulsa! Amb
1368	110	6.29E+11		1.44E+12	0	1	-1	1	0		Wat

Gambar 4.3 Hasil Capture AF*Log*ical OSE (dibaca dengan Microsoft Excel)

Gambar 4.3 melukiskan hasil ekspor data SMS dari memori internal handphone. Uraian diatas meliputi _id, thread_id, address, person, date, protocol, read, status, type, replay_path, subject, body, service_center, locked, error_code, seen, deletable dan delivery_date.

4.4.2. Import Data ke Database Spam Tabel SMS

Import data SMS merupakan tahap memasukkan data hasil dari *capture* yang disimpan dengan nama SMS.csv ke dalam database untuk memudahkan pengujian isi pesan dan mendapatkan informasi detail SMS (lihat gambar 4.4).

Gambar 4.4 Import Data ke Database

Hasil import data yang dilukiskan pada gambar 4.4 disajikan dalam bentuk tabel menggunakan perintah dasar html dan skrip php untuk menampilkan data perbaris agar mudah dibaca dan diidentifikasi oleh pengguna. Pengguna juga diberi wewenang untuk menghapus data sms yang tidak dibutuhkan dalam proses analisis.

Data yang diambil adalah data-data yang dibutuhkan dalam proses analisis yaitu _id, address, date dan body dan disimpan dalam field yang berbeda yaitu id, nomor, tanggal dan pesan. Beberapa SMS yang digunakan dalam penelitian ini adalah dengan rincian sesuai pada tabel 4.2.

Tabel 4.2 Beberapa SMS yang Digunakan dalam Penelitian

ID	Nomor	Waktu Kirim	Isi Pesan
28357	8999	2015-08-03 07:57:45	**simPATI Social MAX Bebas Akses BBM,Whatsapp, LINE, Path,& Waze+Kuota Internet sd 2GB Hub:*999# download aplikasi & info : http://tsel.me/simPATI**
27893	8999	2015-07-21 19:23:33	**Anda mendapatkan 2 kupon dlm program Undian Hadiah HUT 20 Telkomsel, cek jumlah total kupon 3hr dr skrg di *123*20#. Info lengkap hub 188**

Tabel 4.2 Beberapa SMS yang Digunakan dalam Penelitian (Lanjutan)

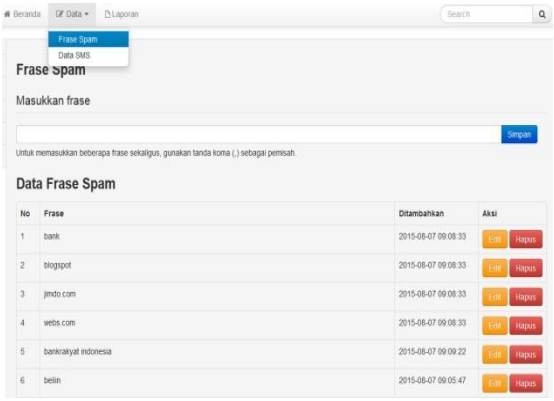
ID	Nomor	Waktu Kirim	Isi Pesan
1244	+6282307816501	2015-07-07 16:37:55	**Slamat!!! Nmr anda 0812155xxxxx Terpilih Men-dpt Hadiah Langsung Rp.50jt.Dari PT.TELKOMSEL No.pin anda.S277DB U/Info klik:www.HadiahTelkomselHut20.blogspot.com**
1078	+6282341933612	2015-06-04 23:57:50	**Plg yth, No.anda M,dpt CEK Rp.75jt & (No.PIN B49C75) dri POIN M-KIOS U/info ketik:NM#ALMT#NO,PIN sms ke 082313512444 /Klik.www.bonusisiulang-mkios.blogspot.com**
022	+6282388387856	2015-05-26 23:35:12	**Maaf tadi sy tlp susah,masalah kontrakan yg bpk/ibu tmpat,in skrang yg puya adik sy Rahmat. no 082386899191 klu ada perlu atau mau perpanjang kabarin aja dia**
776	+6282243956721	2015-04-24 19:39:31	**Surat Keputusan TELKOMSEL NO.VII/04/2015/ pin:HDR7JN7 Menyatakan anda mendpt hadiah ke4 dr TELKOMSELpoin Lihat hadiah di: www.undian-poin-telkomsel.blogspot.com**
590	+6282311939096	2015-03-28 04:26:56	**Nasabah BRI Yth. anda terpilih sebagai pemenang**
			BRI-2015 Pin anda (ijh76k79) info klik www.daftar-pemenang-bri2015.blogspot.com

4.5. Mengelola Data Frase Spam

Kata atau Frase spam diperlukan sebagai penguji apakah sebuah teks pesan berisi spam atau tidak. Pada tahap ini, pengguna mendapatkan wewenang untuk menambahkan, mengubah dan menghapus frase spam. Kata atau frase diambil dari potongan SMS yang dianggap secara umum sebagai spam. Kata atau frase ini selanjutnya dihitung probabilitasnya masing-masing dengan cara menghitung banyaknya kata atau frase yang muncul dalam SMS sehingga diperoleh nilai probabilitasnya.

4.5.1. Menambah Frase Spam

Frase spam yang digunakan dalam penelitian ini berjumlah 66 kata dan frase. Implementasi untuk menambahkan kata/ frase yang dianggap sebagai spam.



Gambar 4.5 Tambah Ubah dan Hapus Frase Spam

4.5.2. Mengubah Frase Spam

Pengguna memiliki wewenang untuk mengubah frase spam yang telah dimasukkan ke dalam database. Hal ini dikarenakan ada keterkaitan antara pesan dan frase spam, sehingga apabila data pesan berubah, maka frase nya juga berubah.

4.5.3. Menghapus Frase Spam

Frase spam yang tidak digunakan dapat dihapus oleh pengguna, sehingga apabila suatu kata atau frase sudah tidak lazim dikategorikan sebagai identifier spam dapat dihapus.

4.6. Hasil Analisis Spam

Tahap akhir dari penelitian ini adalah hasil analisis pesan berserta probabilitasnya. Tahap ini dimulai dengan pengambilan data dari tabel sms, kemudian menghitung probabilitas menggunakan algoritma Naïve Bayessian Filtering versi 1. Apabila hasil yang diperoleh dari versi 1 lebih dari 0.85 (85%) maka disimpulkan bahwa isi SMS tersebut diidentifikasi sebagai spam. Sebaliknya apabila hasil perhitungan probabilitas versi 1 diperoleh hasil kurang dari atau sama dengan 0.85 (85%), maka isi SMS tersebut diidentifikasi sebagai ham (non spam). Tabel 4.3 berikut merupakan beberapa hasil identifikasi isi sms.

Tabel 4.3 Beberapa Hasil dari Identifikasi Isi SMS

No	Pengirim	Pesan	Awal	Bayes v.1
1	88331	**Collect SMS Telkomsel: nomor +6281226122015 gagal mengirim SMS ke Anda. Jika berkenan membayar biaya SMS Rp 165, balas YES dan NO jika tidak. Info: CS 133/188**	Spam	spam (100.00 of spam)%
2	+6282307 816501	**Slamat!!! Nmr anda 0812155xxxxx Terpilih Men-dpt Hadiah Langsung Rp.50jt.Dari PT.TELKOMSEL No.pin anda.S277DB U/Info klik:www.HadiahTelkom selHut20.b*log*spot.com**	Spam	spam (100.00 of spam)%
3	+6282341 933612	**Plg yth, No.anda M,dpt CEK Rp.75jt & (No.PIN B49C75) dri POIN M-KIOS U/info ketik:NM#ALMT#NO,PI N sms ke 082313512444 /Klik.www.bonusisiulang -mkios.b*log*spot.com**	Spam	spam (100.00 of spam)%
4	+6282243 956721	**Surat Keputusan TELKOMSEL NO.VII/04/2015/ pin:HDR7JN7 Menyatakan anda mendpt hadiah ke4 dr TELKOMSELpoin Lihat hadiah di: www.undian-poin-telkomsel.b*log*spot.com**	Spam	spam (100.00 of spam)%
5	+6282311 939096	**Nasabah BRI Yth. anda terpilih sebagai pemenang BRI-2015 Pin anda (ijh76k79) info klik www.daftar-pemenang-bri2015.b*log*spot.com**	Spam	spam (100.00 of spam)%

4.7. Laporan *(reporting)* Hasil Identifikasi Spamming SMS

Spamming SMS dilakukan untuk mengantisipasi terhadap penipuan yang marak terjadi di masyarakat. Pencarian jejak dari tindakan penipuan melalui spamming sms diperoleh dari file .csv. File tersebut diperoleh dari memori internal *smartphone Android* menggunakan tools AF*Log*ical OSE. File .csv yang diperoleh tersebut diimport ke dalam database kemudian dianalisis untuk mengetahui untuk mengetahui probabilitas sebagai spam atau non spam dari isi pesan tersebut. Untuk menganalisis data sms diperlukan adanya frase atau kata yang dicurigai sebagai spam dan berfungsi sebagai pembanding isi pesan. Jika hasil perbandingan diperoleh nilai probabilitas dengan threshold (nilai ambang) tertentu, maka sms tersebut dikategorikan sebagai spam. Kemudian untuk mendapatkan hasil yang maksimal, metode Naïve Bayessian Filtering menawarkan pemecahan isi pesan menjadi n-grammer sehingga mendapatkan hasil yang akurat. Dengan adanya penelitian spamming sms ini diharapkan dapat membantu masyarakat dalam menanggapi sms yang mengandung unsur penipuan maupun sms promosi yang bersifat menggangu sehingga dapat meminimalisir terjadinya tindakankriminal terutama penipuan

yang memanfaatkan fasilitas Short Messaging Service (SMS).

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan pembahasan dan uraian pada bab-bab sebelumnya mengenai pembuatan sistem aplikasi *mini tool* untuk *smartphone android*, maka dapat dsimpulkan sebagai berikut :

1. Sistem aplikasi yang dibangun dibuat dengan menggunakan bahasa pemrograman Predefine Hypertext Preprocessor (PHP).
2. Frase spam yang digunakan dalam penelitian ini berjumlah 66 kata dan frase yang diambil dari potongan SMS yang dianggap secara umum sebagai spam.kata atau frase tersebut dihitung probabilitasnya masing-masing, dengan cara menghitung banyaknya kata atau frase yang muncul dalam SMS.
3. Penelitian dilakukan terhadap tiga jenis *smartphone* berbasis *android* yaitu :
   a. Samsung S III mini yang memiliki 581 SMS, jumlah spam hasil analisis awal 0. Jumlah spam hasil analisis *Naive Bayesian Filtering* Versi 1 : 388. Jumlah spam hasil analisis *Naive Bayesian Filter* Versi 2 : 381.
   b. *Samsung galaxy young* yang memiliki 979 SMS, jumlah spam hasil analisis awal 23. Jumlah spam hasil analisis *Naive Bayesian Filtering* versi 1 : 534. Jumlah spam hasil analisis *Naive Bayesian Filter* Versi 2 : 357.
   c. Asus T00J yang memiliki 176 SMS, jumlah spam hasil analisis awal 0. Jumlah spam hasil analisis *Naive Bayesian Filtering* versi 1 : 155. Jumlah spam hasil analisis *Naive Bayesian Filter* Versi 2 : 147.
4. Proses analisis pesan dimulai dari penentuan n-grammer yang dipecah dengan panjang string (huruf atau kata) 3 sampai 5 kata.

5.2. Saran

Berdasarkan kesimpulan diatas, maka disarankan untuk penelitian selanjutnya supaya :

1. Membangun tool sendiri untuk melakukan ekstraksi data pada *smartphone*.
2. Melakukan cara lain dalam transmisi data dari *smartphone* ke komputer dengan untuk Menganalisa data-data yang diperlukan.
3. Mengembangkan aplikasi supaya dapat berjalan pada *smartphone*. Pesan yang dikategorikan kedalam spam, tanpa pesan tersebut tersimpan pada kotak masuk pesan, yang disesuaikan dengan prinsip investigasi pada mobile forensik.

DAFTAR PUSTAKA

Al-Fedaghi,S dan Al-Babtain,B. 2012. Modeling the Forensics process, Computer Engineering Department, Kuwait University.

Androulidakis, I.I. 2012. *Mobile Phone Security and Forensics*, Springer, New York.

Casey, E. 2004. *Digital Evidence and Computer Crime, Second Edition*, Elsevier, ISBN 0-12-163104-4.

Gavin, P., Robinson, L.A. dan Ellis, R. 2010. *Digital Forensic*

Hoog, A. 2011. *Android Forensics Investigation, Analysis and Mobile Security for Google Android*, United State of America.

K. Kent, S. Chevalier, T. Grance dan H. Dang. 2006. "*Guide to Integrating Forensics into Incident Response",* Special Publication 800-86, Computer Security Division Information Techno*log*y Laboratory.

National Institute of Standards and Techno*log*y, Gaithersburg, MD. **(2006)**. http://csrc.nist.gov/publications/nistpubs/
800-86/SP800-86.pdf.

Winarno, E. dan Zaki, A. 2012. *Windows Forensic*, Jakarta : PT Elex Media Komputindo.

Putri, R.U., 2012, Network Forensic Analysis Case Study SQL Injection Attack on *Server* Gadjah Mada University, Gadjah Mada University, Yogyakarta.

Yadi, I. Z dan Kunang Y. N. 2014. Analisis Forensik pada Platform *Android*, Universitas Bidarma.