BAB 4
Hasil dan Pembahasan

4.1 Konsep 4W1H

Konsep 4W1H (*Who*, *What*, *When*, *Where*, *Why*, dan *How*) adalah kumpulan rumusan pertanyaan yang umumnya digunakan untuk memecahkan suatu masalah. Penggunaan konsep 4W1H dalam dunia forensik ini dapat mempermudah proses investigasi karena dengan menerapkan metode ini, suatu kasus atau masalah dapat terpecahkan. Saat pertanyaan-pertanyaan 4W1H dapat terjawab, hal ini membantu investigator untuk menemukan titik terang, bukti yang kuat, ataupun petunjuk yang merujuk ke bukti selanjutnya dari sebuah kasus yang sedang diinvestigasi.

Penerapan konsep 4W1H dalam forensik email dapat menjelaskan dengan detail terkait kasus yang sedang diinvestigasi dan email menjadi barang bukti dalam kasus tersebut. Pada setiap email terdapat header yang di dalamnya terkandung metadata yang berisi informasi penting terkait email tersebut. Header ini adalah artefak digital yang paling mudah didapatkan dari sebuah email. Informasi pada header email ini sulit dibaca jika dilihat secara langsung, sehingga dibutuhkan *tool* yang dapat menterjemahkan informasi pada header email ini yang kemudian dipetakan menjadi jawaban dari pertanyaan 4W1H. *Tool* Mail Header Extractor dirancang untuk mempermudah investigator dalam melakukan aktivitas forensik email, karena *tool* ini dapat mengekstraksi dan memetakan informasi dari *header* email dengan cepat serta mudah untuk dibaca. Selain itu akan muncul tanda jika *header email* yang diujikan terindikasi sebagai *email fraud* atau *email spoofing*. Dari hasil ekstraksi tersebut, investigator dapat dengan mudah menemukan informasi mengenai siapa pengirim email, siapa penerima email, kapan email dikirim oleh pengirim, kapan email diterima oleh penerima dan informasi lainnya.

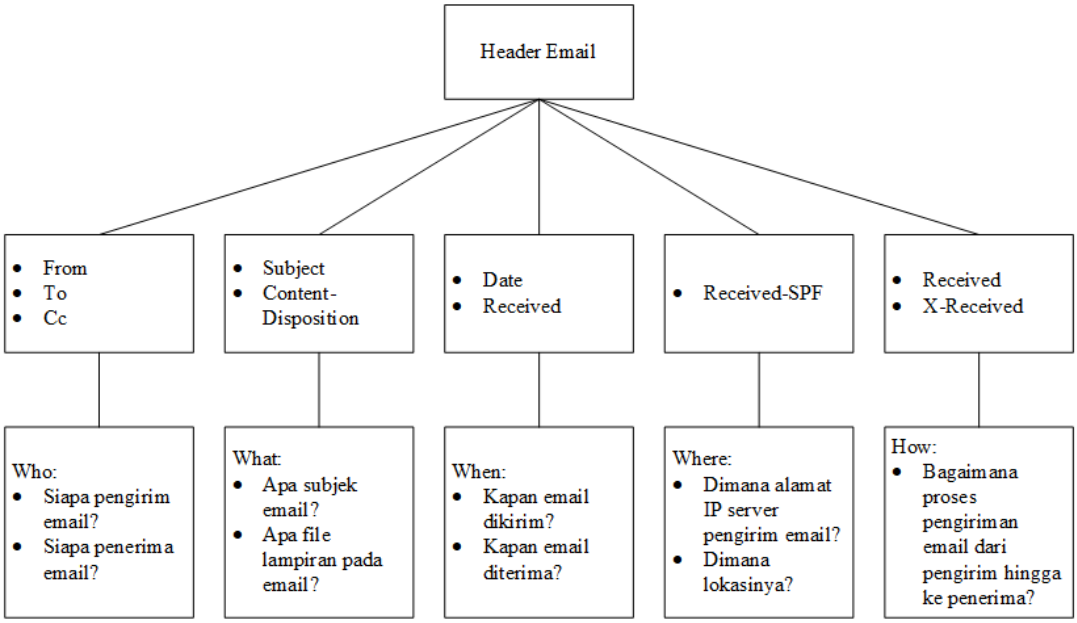
Pertanyaan *Who* ini menjelaskan siapa saja yang terlibat dalam email tersebut, siapa yang mengirim email dan siapa yang menerima email. Pada *header email*, informasi mengenai *who* ini didapatkan dari *field From* (pengirim), *To* (penerima), dan *Cc* (penerima). Dari informasi ini, jika nama alias dan alamat email dari pengirim tidak sama atau tidak mirip dan terlihat mencurigakan, dapat menjadi catatan bagi investigator untuk diselidiki lebih lanjut. *Field Message-ID* juga diambil informasinya untuk mengecek apakah nama domain pengirim email yang tercantum pada *From* cocok dengan nama domain pengirim

email yang tercantum pada *Message-ID*. Email yang benar adalah email yang nama domain di *From* dengan nama domain di *Message-ID* sama.

Pertanyaan *What* menunjukkan subjek dari email dan file lampiran yang ada pada email. Subjek ini dapat menjadi gambaran mengenai isi email tersebut. Informasi mengenai subjek dan file lampiran didapatkan dari *field Subject* dan *Content-Disposition* yang terdapat pada *header email*. Jika subjek dari email terlihat seperti mengintimidasi penerimanya, email dapat dicurigai sebagai *phising* karena biasanya berisi permintaan kepada penerima email untuk mengklik suatu tautan yang mengarah pada website yang telah dihacking dan memancing penerima email untuk mengisikan informasi pribadinya yang kemudian dapat dimanfaatkan oleh pelaku. Kemungkinan lainnya, isi dari email adalah pemerasan kepada penerima email agar akunnya kembali. Selain itu, pertanyaan *What* juga dapat memberikan informasi file lampiran yang ada pada email. Ada kemungkinan file lampirannya merupakan *malware* yang dapat menginfeksi komputer penerima atau pembuka email.

Pertanyaan *When* dapat menunjukkan waktu kapan email dikirim oleh pengirim dan waktu email diterima oleh penerima. Waktu pengiriman dan penerimaan email ini menjadi penting untuk diperhatikan karena pemalsuan waktu pengiriman email bisa digunakan untuk mengelabui penerima email. Hal ini bisa terjadi jika email yang dikirim bertujuan untuk menciptakan adanya selisih waktu dari yang seharusnya, misalnya email dikirim seolah-olah sebelum batas waktu yang ditentukan padahal sebenarnya email dikirim setelah melewati batas waktu. Informasi mengenai waktu pengiriman dan penerimaan email didapatkan dari *field Date* (waktu pengiriman) dan *Received* (waktu penerimaan) pada *header*. Waktu pengiriman email dapat dipalsukan dengan cara mengubah waktu pada komputer yang digunakan untuk mengirim email, karena waktu yang tercantum pada *Date* ini merujuk pada waktu di komputer. Sementara waktu penerimaan pesan adalah waktu riil yang tercatat pada *server* penerima email dan tidak dapat dimanipulasi.

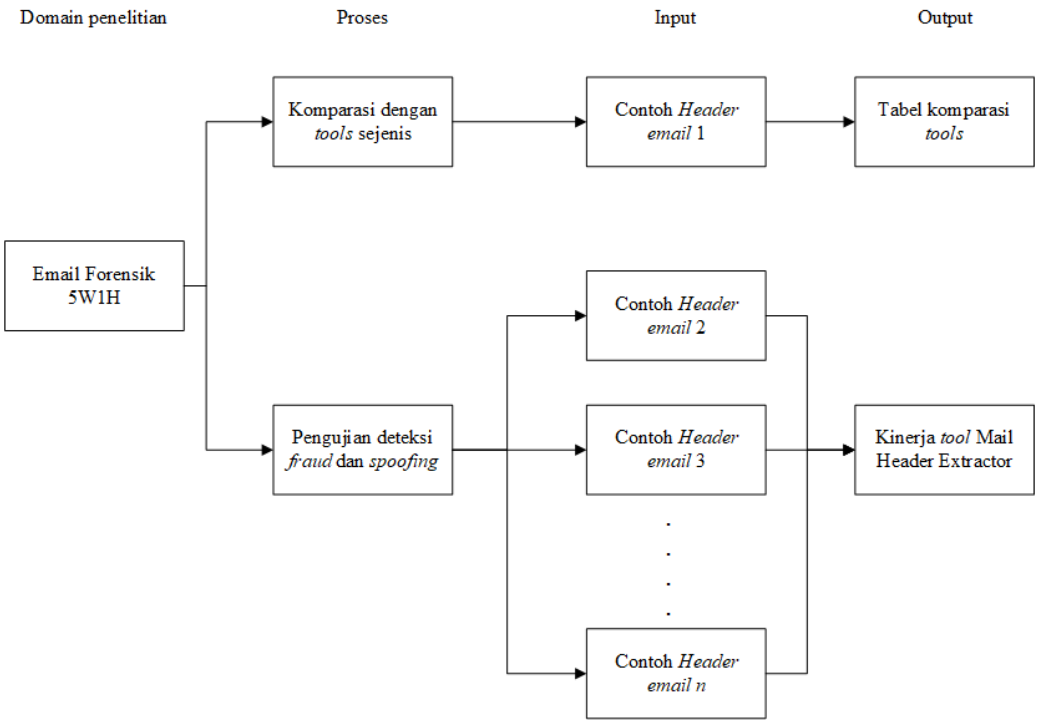
Pertanyaan *Where* dapat menunjukkan alamat IP dari *server* pengirim email beserta lokasinya. Alamat IP ini didapatkan dari *field Received-SPF* pada *header email*. Sementara pertanyaan *How* menunjukkan proses perjalanan email dari pengirim ke penerima, *server* yang dilewati oleh email dan protokol yang digunakan selama di perjalanan. Informasi untuk menjawab pertanyaan *How* diambil dari *field Received* dan *X-Received*. Konsep pemetaan informasi dari *header email* ke dalam konsep 4W1H ditunjukkan pada Gambar 4.1.



Gambar 4.1 Peta Konsep 4W1H dalam Forensik Email

Gambar 4.2 adalah peta konsep penelitian yang dilakukan dalam penelitian ini. Sebelum membuat *tool* Mail Header Extractor, peneliti terlebih dahulu melakukan komparasi dengan sejumlah *tools* email forensik yang menganalisis *header email*. Proses komparasi ini dilakukan untuk mengetahui informasi apa saja yang dapat diekstraksi dari *header email* dari setiap *tool*. Poin yang dilihat dalam melakukan komparasi ini adalah apakah informasi hasil ekstraksi *header email* yang ditampilkan oleh masing-masing *tools* telah dapat menjawab pertanyaan konsep 4W1H (*Who*, *What*, *When*, *Where*, dan *How*). Ada 14 *tools* yang dikomparasikan dengan menggunakan 1 *header email* yang sama (terlampir pada Lampiran), perbedaan tampilan informasi dari masing-masing *tool* dapat dilihat pada Tabel 2.2.

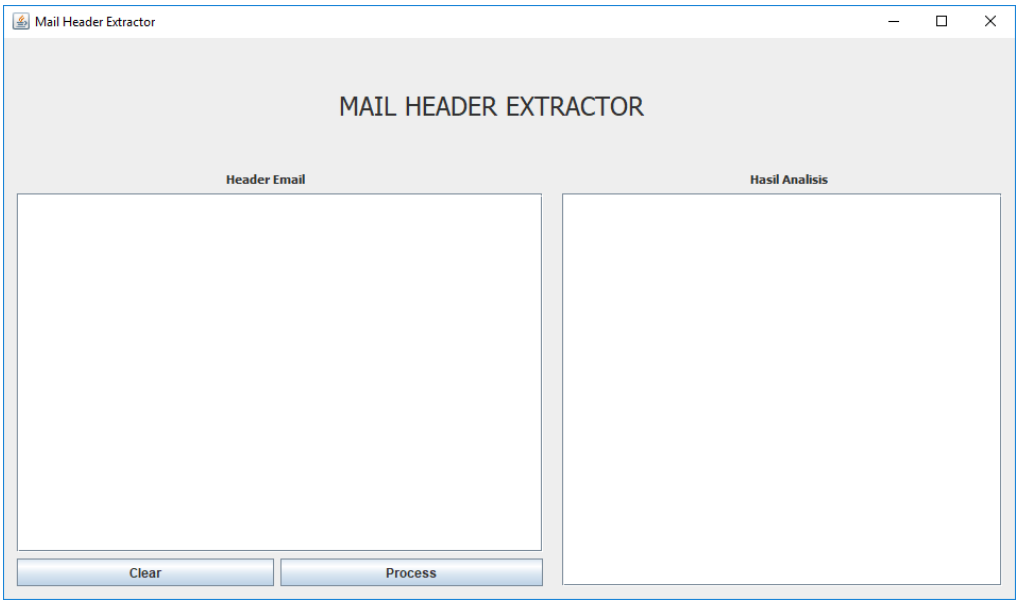
Kemudian *tool* Mail Header Extractor dibuat dengan kemampuan untuk memetakan informasi dari *header email* ke dalam konsep 4W1H (*Who*, *What*, *When*, *Where*, dan *How*) serta dilengkapi dengan kemampuan memberikan indikasi terhadap *email fraud* dan *email spoofing*. Setelah *tool* jadi, dilakukan pengujian dengan menggunakan 15 *header email*, mengevaluasi kinerja *tool* dalam proses menerjemahkan dan memetakan informasi dari *header email* serta proses deteksi *email fraud* dan *email spoofing*.



Gambar 4.2 Peta Konsep Penelitian

4.2 *Tool* Mail Header Extractor

Tool Mail Header Extractor dibuat dengan menggunakan bahasa pemrograman Java dengan compiler Netbeans 8.0. Tampilan dari *tool* Mail Header Extractor ditunjukkan oleh Gambar 4.3.

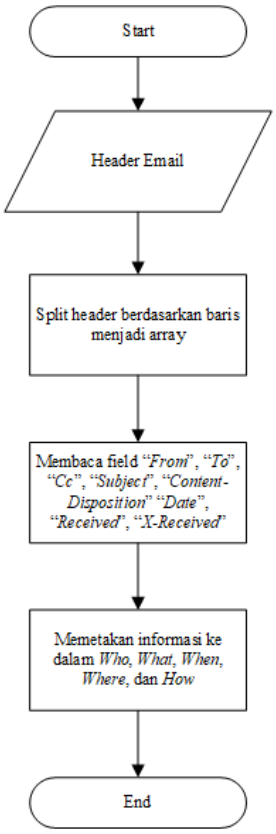


Gambar 4.3 Tampilan *Tool* Mail Header Extractor

Tool Mail Header Extractor dapat menerjemahkan *header email* dan memetakan informasinya ke dalam konsep 4W1H (*Who*, *What*, *When*, *Where*, dan *How*). Sebelum itu, perlu diidentifikasi terlebih dahulu *field* mana saja yang menunjukkan *Who*, *What*, *When*, *Where*, dan *How*. Setelah mengidentifikasi *field* yang dibutuhkan, selanjutnya dilakukan teknik *parsing* untuk mencari *field* tersebut kemudian diambil informasinya. *Field* yang dicari adalah sebagai berikut:

a. *From, To, Cc* untuk pertanyaan *Who*. *From* untuk mendapatkan informasi siapa pengirim email dan *To* dan *Cc* untuk mendapatkan informasi siapa penerima email.
b. *Subject*, *Content-Disposition* untuk pertanyaan *What*. *Subject* untuk mendapatkan informasi mengenai subjek email dan *Content-Disposition* untuk mendapatkan informasi mengenai berkas lampiran yang ada pada email.
c. *Date*, *X-Received* untuk pertanyaan *When*. *Date* untuk mendapatkan informasi kapan email dikirim dan *X-Received* untuk mendapatkan informasi kapan email diterima. Untuk *X-Received* ini, yang diambil adalah yang paling atas.
d. *Received-SPF* untuk menjawab *Where*. *Received-SPF* memuat informasi alamat IP dari *server* pengirim email. Selanjutnya dari informasi alamat IP ini, dicari lokasi alamat IP-nya dengan cara men-*generate* alamat IP tersebut ke website www.ipapi.co.
e. *Received* dan *X-Received* untuk menjawab *How*. Semua *Received* dan *X-Received* yang ada pada header email diekstraksi untuk mendapatkan informasi mengenai perjalanan email dari server pengirim hingga ke penerima beserta informasi protokol yang digunakan.

Proses ekstraksi informasi dari *header email* dimulai dengan memasukkan *header email*. Kemudian *header email* akan diuraikan berdasarkan baris menjadi sebuah array. Proses selanjutnya adalah membaca *header email* dari awal dan mencari *field* kata kunci yang dibutuhkan, yaitu *From, To, Cc, Subject, Content-Disposition, Date, X-Received, Received-SPF*, dan *Received*. Kemudian isi dari *field* tersebut dipetakan ke *Who, What, When, Where,* dan *How*. Diagram alir proses ekstraksi informasi tersebut ditampilkan pada Gambar 4.4.



Gambar 4.4 Diagram alir proses ekstraksi informasi dari header email

Tool Mail Header Extractor memiliki fitur yang berbeda dengan *tools* lainnya yaitu adanya kemampuan untuk mendeteksi *email fraud* (penipuan) dan *email spoofing* (palsu). Fitur ini tidak ditemukan pada semua *tool* yang telah dikomparasi sebelumnya, sehingga menjadi keunggulan dari *tool* Mail Header Extractor. Proses deteksi *email fraud* dilakukan dengan dua acara, yaitu berdasarkan subjek email dan berdasarkan isi email. Deteksi *email fraud* berdasarkan subjek email dilakukan dengan membandingkan kata pada subjek email yang diujikan dengan kata pada kamus. Kamus ini disusun berdasarkan dataset Nigerian Fraud Letters yang didapatkan dari www.kaggle.com. Dataset tersebut memuat 4291 email yang dikumpulkan dari tahun 1998 hingga 2007. Subjek email dari 4291 email tersebut dikumpulkan kemudian dibuat pemeringkatan berdasarkan kata yang paling sering muncul pada subjek tersebut. Berdasarkan hasil pemeringkatan, diambil 10 kata teratas yang paling sering muncul dengan frekuensi kemunculan di atas 200 kali. Kata-kata tersebut yaitu: *urgent* (802 kali), *from* (705 kali), *your* (370 kali), *assist* (358 kali), *please* (319 kali), *need* (299 kali), *business* (295 kali), *assistance* (279 kali), *reply* (273 kali), dan *proposal* (209 kali). Jika ada kata pada subjek email yang cocok dengan kata pada kamus, maka email tersebut diindikasikan sebagai *email fraud*. Selanjutnya ditampilkan pula persentase tingkat

indikasi *email fraud*-nya yang angkanya didapatkan dari perbandingan jumlah banyak kata pada subjek yang cocok dengan kata pada kamus.

Deteksi *email fraud* berdasarkan isi email dilakukan dengan cara yang hampir sama dengan deteksi berdasarkan subjek email. Namun hal yang dibandingkan adalah kata pada isi email yang diujikan dengan kata pada kamus. Isi kamus disusun berdasarkan isi dari 25 sampel email yang ada pada dataset Nigerian Fraud Letters. Selanjutnya dari isi email ini dibuat pemeringkatan berdasarkan kata yang paling sering muncul. Setelah dibuat pemeringkatan, diambil 5 kata dengan frekuensi kemunculan di atas 50 kali, yaitu *transaction* (60 kali), *account* (58 kali), *money* (55 kali), *security* (53 kali), dan *family* (50 kali). Apabila terdapat kata pada isi email yang cocok dengan kata pada kamus, email akan diindikasikan sebagai *email fraud* dengan persentase yang didapatkan berdasarkan perbandingan antara jumlah kata pada isi yang cocok dengan jumlah kata pada kamus.

Deteksi *email address spoofing* dilakukan dengan membandingkan nama domain yang ada pada *field Message-ID* pada *header email* yang diujikan dengan nama domain yang ada pada kamus domain. Data pada kamus domain ini didapatkan dari hasil simulasi pengiriman *email spoofing* dari berbagai situs *fake mailer* yang tersedia di internet. Situs *fake mailer* yang digunakan untuk mengirimkan *email spoofing* pada penelitian ini yaitu: emkei.cz, anonymailer.net, anonymousemail.me, 5ymail.com, dan guerillamail.com. Setelah dilakukan simulasi pengiriman email dari 5 situs *fake mailer* tersebut, selanjutnya diambil nama domain yang ada pada *field Message-ID* dari masing-masing email tersebut. Jika nama domain pada *field Message-ID* dari *header email* yang diujikan cocok dengan salah satu nama domain yang ada di kamus domain, maka email terindikasi sebagai *email address spoofing*.

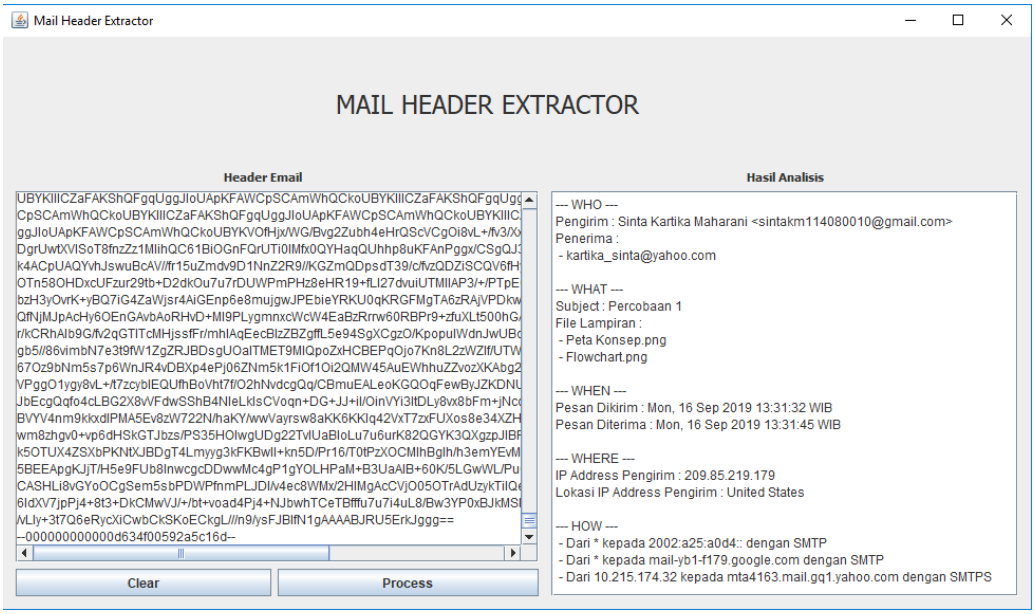
Proses deteksi *email time spoofing* dilakukan dengan membandingkan waktu pengiriman email (didapatkan dari *field Date*) dan waktu penerimaan email (didapatkan dari *field Received* atau *X-Received* yang teratas). Jika selisih waktu antara waktu pengiriman dan penerimaan email lebih dari 5 hari, maka terindikasi email tersebut adalah *email time spoofing*. Waktu yang tercantum pada *field Received* atau *X-Received* yang teratas adalah waktu riil diterimanya email oleh server penerima. Selisih waktu 5 hari dianggap sebagai batas waktu maksimal email terkirim ke penerima. Terjadinya penundaan ini bersifat sementara dan umumnya disebabkan karena sistem dari target atau koneksi rusak (Klensin, 2008).

4.3 Pengujian dan Hasil Analisis Tool Email Header Extractor

Proses pengujian *tool* dilakukan dengan menetapkan beberapa skenario dan kemudian dinilai apakah proses ekstraksi yang dilakukan oleh *tool* telah berhasil dilakukan. Berikut adalah skenario pengujian yang dilakukan:

4.3.1 Email normal

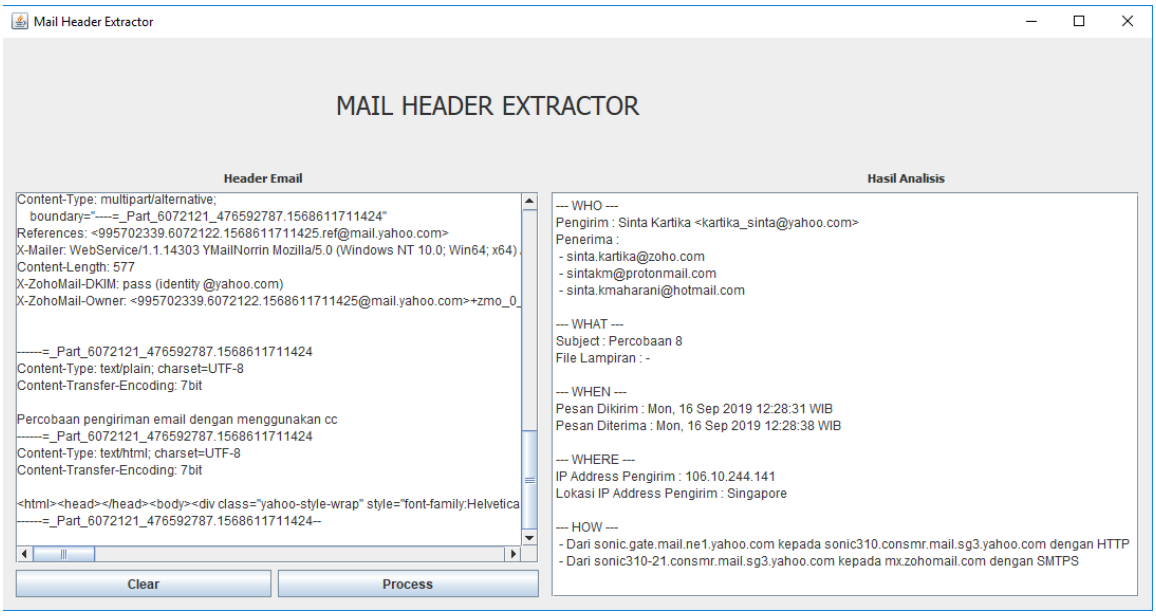
Skenario ini menggunakan *header email* yang berasal dari domain yahoo.com, zoho.com, dan gmail.com. Gambar 4.5 merupakan hasil pengujian dari *header email* domain yahoo.com.



Gambar 4.5 Hasil pengujian *header email* dari domain Yahoo

Gambar 4.5 menunjukkan bahwa *header email* yang diujikan merupakan email yang dikirimkan dari sintakm114080010@gmail.com kepada kartika_sinta@yahoo.com pada tanggal 16 September 2019 jam 13:31:32 WIB dengan subjek “Percobaan 1” dan ada 2 file lampiran di dalamnya, yaitu “Peta Konsep.png” dan “Flowchart.png”. Email tersebut dikirimkan dari alamat IP 209.85.219.179 (alamat IP *server* Google) yang letaknya di Amerika Serikat.

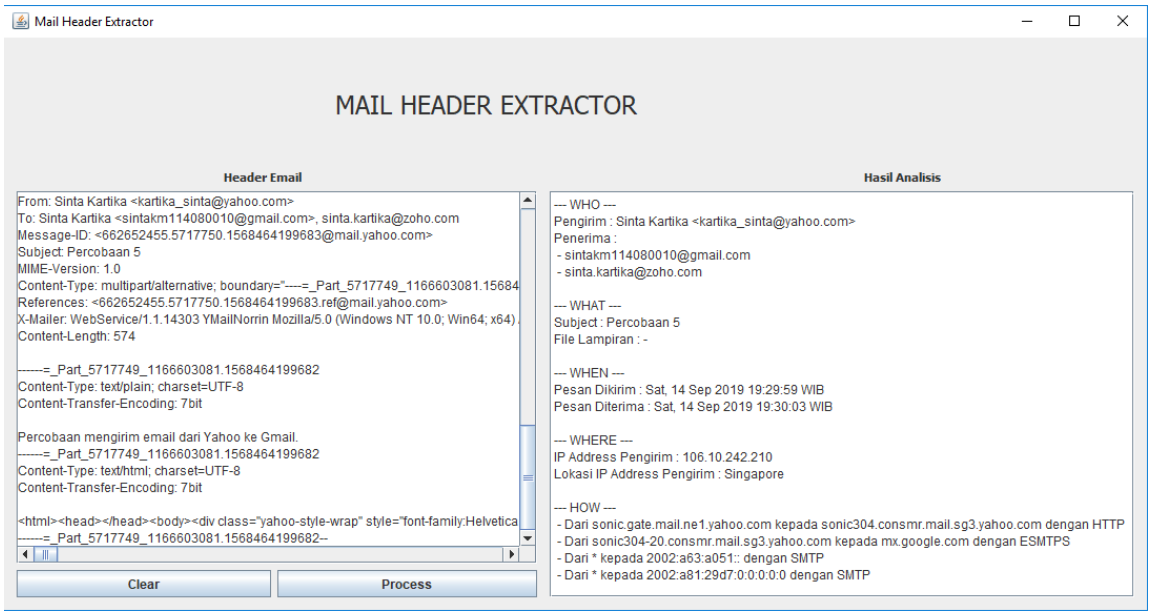
Gambar 4.6 merupakan hasil pengujian dari *header email* domain zoho.com.



Gambar 4.6 Hasil pengujian *header email* dari domain Zoho

Berdasarkan Gambar 4.6, *header email* yang diujikan adalah email dari kartika_sinta@yahoo.com kepada sinta.kartika@zoho.com dan ada cc kepada sintakm@protonmail.com dan sinta.kmaharani@hotmail.com dengan subjek "Percobaan 8" dan dikirimkan pada 16 September 2019 jam 12:28:31 WIB. Email di atas dikirimkan dari alamat IP 106.10.244.141 yang merupakan alamat IP dari *server* Yahoo yang berada di Singapura.

Gambar 4.7 merupakan hasil pengujian dari *header email* domain gmail.com.



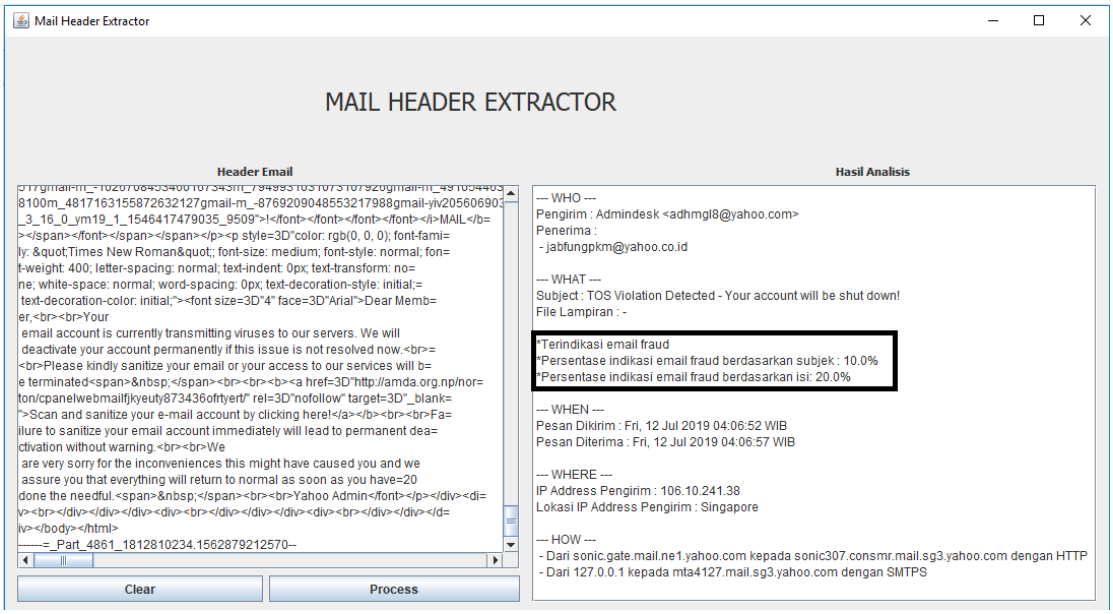
Gambar 4. 7 Hasil pengujian *header email* dari domain Gmail

Gambar 4.7 memperlihatkan bahwa *header email* yang diujikan adalah email dari kartika_sinta@yahoo.com kepada sintakm114080010@gmail.com dan sinta.kartika@zoho.com dengan subjek "Percobaan 5" yang dikirimkan pada 14 September 2019 jam 10:29:50 WIB. Email dikirimkan dari alamat IP 106.10.242.210 yang merupakan alamat IP *server* Yahoo yang letaknya di Singapura.

Berdasarkan ketiga contoh email normal yang diujikan di atas menunjukkan bahwa proses ekstraksi informasi dari *header email* dan pemetaannya ke 4W1H telah berhasil dilakukan.

4.3.2 Email yang subjeknya diindikasikan *fraud*

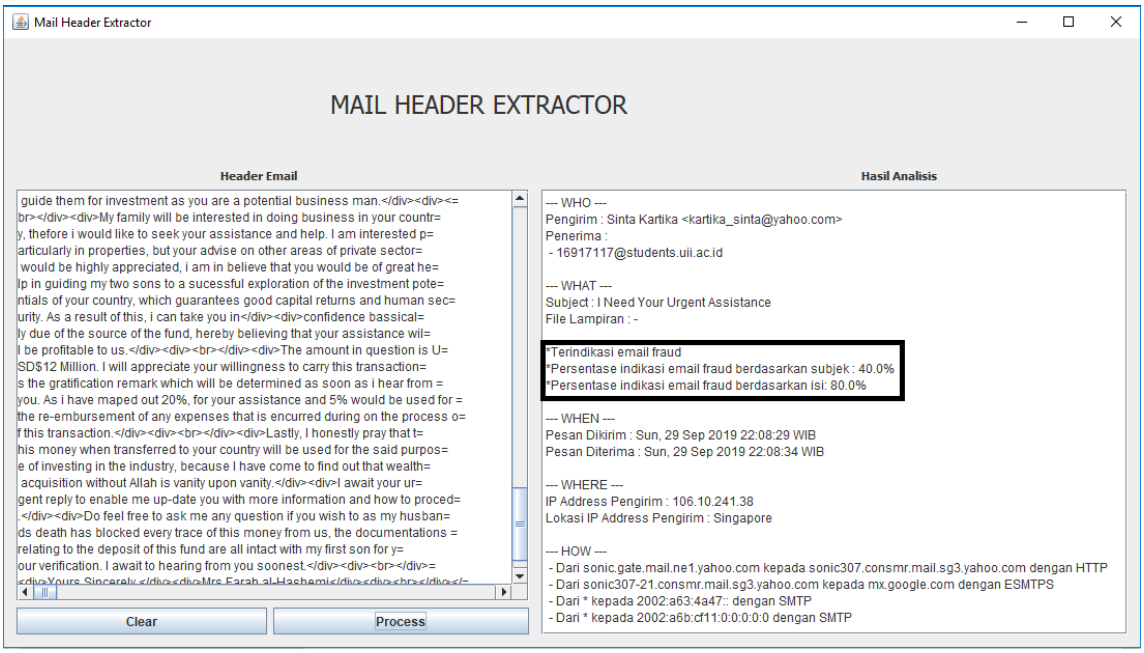
Skenario pengujian email yang subjek diindikasikan sebagai *email fraud* ini dilakukan untuk menguji apakah *tool* Mail Header Extractor telah berhasil memberikan indikasi terhadap email yang diduga *fraud* berdasarkan subjek dan isi dari email. Setiap *header email* yang diujikan akan dibaca subjek dan isinya. Jika ada kata pada subjek dan isi yang cocok dengan kata pada kamus, maka email tersebut diindikasikan sebagai *email fraud*. Hasil dari skenario ini ditunjukkan oleh Gambar 4.8 dan Gambar 4.9.



Gambar 4.8 Hasil pengujian *header email* yang subjeknya terindikasi *fraud*

Gambar 4.8 menunjukkan hasil pengujian *header* dari email yang dikirimkan oleh adhmg18@yahoo.com kepada jabfungpkm@yahoo.co.id dengan subjek "TOS Violation Detected – Your account will be shut down!" yang dikirimkan pada 12 Juli 2019 jam 04:06:52 WIB dari alamat IP 106.10.241.38 yang merupakan alamat IP *server* Yahoo yang

letaknya di Singapura. Email ini diindikasikan sebagai *email fraud* karena pada subjek terdapat kata "*your*" yang termasuk pada kamus pembanding subjek. Persentase indikasi *email fraud* berdasarkan subjek menunjukkan tingkat kemungkinan bahwa email tersebut benar *fraud* adalah 10% yang didapatkan dari jumlah kata pada subjek yang cocok dengan kata pada kamus. Sedangkan persentase indikasi *email fraud* berdasarkan isi menunjukkan bahwa kemungkinan email benar *fraud* sebesar 20%, yang artinya ada 1 kata pada isi email yang cocok dari 5 kata yang terdapat pada kamus pembanding isi.



Gambar 4.9 Hasil pengujian *header email* yang subjeknya terindikasi *fraud*

Gambar 4.9 menunjukkan hasil analisis *header email* yang dikirimkan oleh kartika_sinta@yahoo.com kepada 16917117@students.uii.ac.id yang subjeknya "I Need Your Urgent Assistance" dikirimkan pada 29 September 2019 jam 22:08:29 WIB dari alamat IP 106.10.241.38 yang terletak di Singapura. Email tersebut diindikasikan sebagai *email fraud* karena pada subjeknya terdapat 4 kata yang cocok dengan kata yang ada pada kamus, yaitu "*need*", "*your*", "*urgent*", "*assistance*", sehingga persentasenya adalah 40%. Sementara berdasarkan isi email, persentase indikasi *email fraud* adalah 80%.

Gambar 4.10 merupakan *source code* dari proses deteksi *email fraud* berdasarkan subjeknya. Terlihat bahwa proses awal adalah membaca kamus yang formatnya berupa csv. Subjek email yang tadi sudah didapatkan informasinya di *What* diuraikan per kata kemudian dibandingkan dengan isi kamus.csv.

```
            String file = "C:\\mailheaderparsing\\kamus.csv";

            List<String[]> content = new ArrayList<>();
            try(BufferedReader br = new BufferedReader(new FileReader(file))) {
                String line = "";
                int i=0;
                while ((line = br.readLine()) != null) {
                    content.add(line.split(","));
                    //System.out.println(content.get(i)[0].toString());
                    i++;
                }
            } catch (FileNotFoundException e) {
              //Some error logging
            }

    String[] kata_subject = subject_result.split(" ");
                Boolean statusFraud = false;
                double jml_kata = 10;
                double jml_fraud = 0;
                for(int i=0;i<kata_subject.length;i++){
                    String isFraud="";
                    for (int j=1;j<content.size();j++) { //karena 0 baca header csv

    //System.out.println(content.get(j)[0].toString().toLowerCase()+"=="+kata_subject[i].
    toLowerCase());
                        if
    (content.get(j)[0].toString().toLowerCase().equals(kata_subject[i].toLowerCase())) {
                            if (j<=10){
                                isFraud="(FRAUD)";
                                statusFraud = true;
                                jml_fraud = jml_fraud + 1;
                            }
                        }
                    }
                    //hasil += ""+kata_subject[i]+" "+isFraud+"\n";
                }

                if (statusFraud == true) {
                    hasil += "*Terindikasi email fraud\n";
                    double persentase_fraud = (jml_fraud / jml_kata) * 100;
                    hasil += "*Persentase indikasi email fraud berdasarkan subjek :
    "+String.valueOf(persentase_fraud)+"%\n";
                }
```

Gambar 4.10 *Source code* deteksi *email fraud* berdasarkan subjek

Gambar 4.11 menunjukkan *source code* dari proses deteksi *email fraud* berdasarkan isi email. Proses awalnya adalah dengan mendapatkan daftar kata kunci dari kamus contain.csv. Selanjutnya dilakukan penguraian kata dari isi email dan kemudian membandingkannya dengan isi kamus contain.csv.

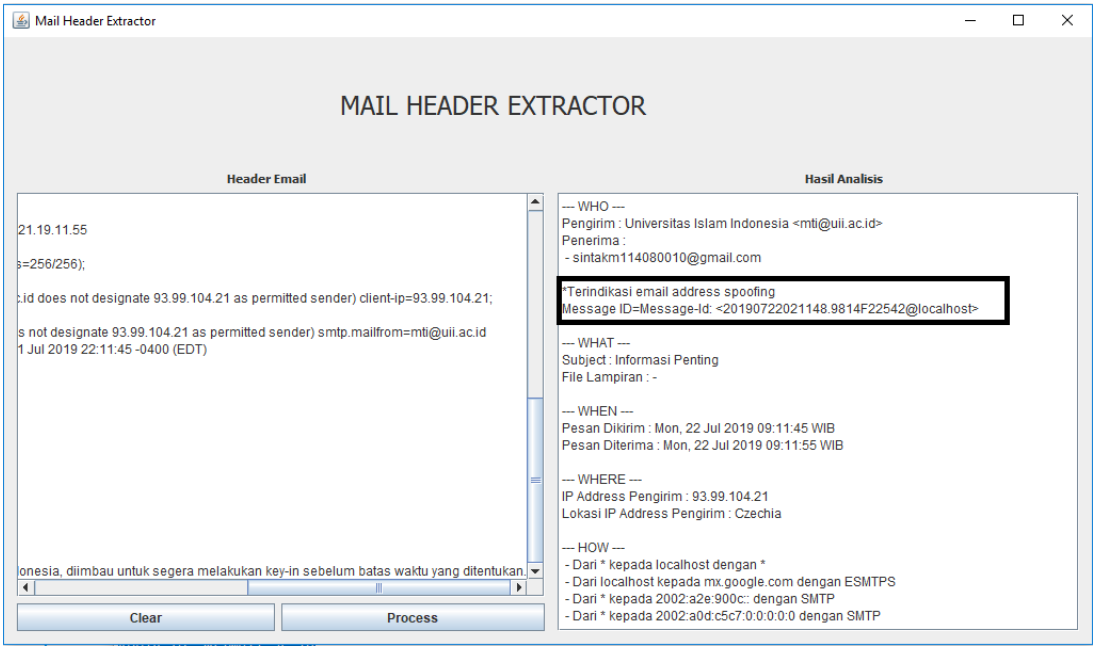
```
    public static String analisis(String input){
        String hasil = "";
        double count = 0;
        double isi = 5;
        List<String> keywordList = getKeywordList("C:\\mailheaderparsing\\contain.csv");
        String extracted = extract(input);
        String[] words_from_mail_body_arr = extracted.split("\\s+");
        Set<String> mail_words = new HashSet<String>();
    for (String word:words from mail body arr){
            if (word.endsWith(".")) word = word.replace(".","");
            if (word.endsWith(",")) word = word.replace(",","");
            mail_words.add(word.toLowerCase());
    }

        for (String word:mail words){
            if (containsIgnoreCase(keywordList, word)) {
                ++count;
            }
    }

        float skor = (float)count/(float)keywordList.size()*100;
        hasil = ""+skor+"%";
        return hasil;
    }
    hasil = hasil += "*Persentase indikasi email fraud berdasarkan isi: "+
   analisis(edtMailHeader.getText())+"\n\n";
```

Gambar 4.11 *Source code* deteksi *email fraud* berdasarkan isi

Berdasarkan contoh pada Gambar 4.8 dan Gambar 4.9 di atas, menunjukkan bahwa *tool* Mail Header Extractor telah berhasil menunjukkan indikasi *email fraud* dengan melakukan perbandingan kata pada subjek dan isi email dengan kata pada kamus.

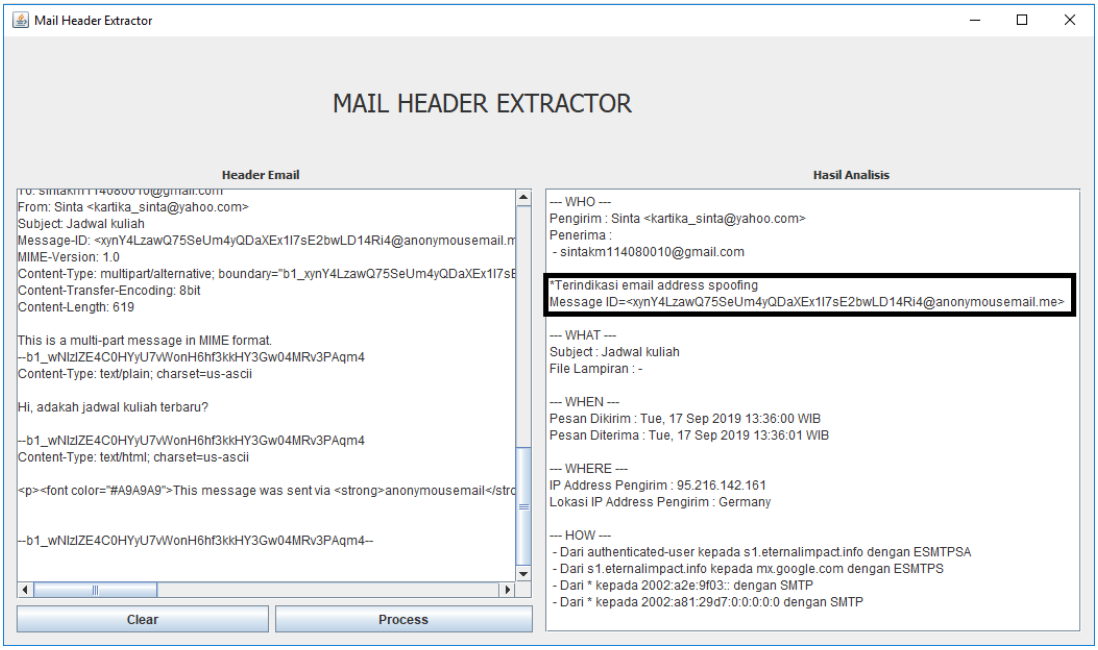
4.3.3 Email yang telah dipalsukan alamat pengirimnya

Skenario pengujian email yang telah dipalsukan alamat pengirimnya dilakukan untuk menguji apakah *tool* Mail Header Extractor telah berhasil memberikan indikasi terhadap email yang diduga *email address spoofing* berdasarkan nama domain yang tercantum di *field Message-ID* dari *header email*. Pada dasarnya setiap pengujian yang dilakukan oleh *tool* akan mengecek *Message-ID*, jika nama domain pada *Message-ID* cocok dengan nama domain yang ada pada kamus domain, maka akan diindikasikan sebagai *email address spoofing*. Hasil dari skenario ini ditunjukkan pada Gambar 4.12 dan Gambar 4.13.



Gambar 4.12 Hasil pengujian *header email* yang terindikasi *email address spoofing*

Header email yang diujikan pada Gambar 4.12 adalah email yang dikirimkan dari mti@uii.ac.id kepada sinta114080010@gmail.com dengan subjek "Informasi Penting" dan dikirimkan pada 22 Juli 2019 jam 09:11:45 WIB. Email dikirimkan dari *server* dengan alamat IP 93.99.104.21 yang terletak di Ceko. Namun email tersebut diindikasikan sebagai *email address spoofing* karena nama domain email pengirimnya adalah "localhost".



Gambar 4.13 Hasil pengujian *header email* yang terindikasi *email address spoofing*

Gambar 4.13 menunjukkan *header email* yang diujikan merupakan email yang dikirimkan dari kartika_sinta@yahoo.com kepada sintakm114080010@gmail.com dengan subjek "Jadwal kuliah" yang dikirimkan pada 17 September 2019 jam 13:36:00 WIB. Tetapi email yang dikirim tersebut terindikasi sebagai *email address spoofing* karena ternyata tidak dikirimkan dari domain Yahoo, melainkan dari domain "anonymousemail.me".

Gambar 4.14 merupakan *source code* dari proses deteksi *email address spoofing*.

```
    String filedomain = "C:\\mailheaderparsing\\domain.csv";

         List<String[]> content = new ArrayList<>();
         try(BufferedReader br = new BufferedReader(new FileReader(file))) {
             String line = "";
             int i=0;
             while ((line = br.readLine()) != null) {
                 content.add(line.split(","));
                 //System.out.println(content.get(i)[0].toString());
                 i++;
             }
         } catch (FileNotFoundException e) {
           //Some error logging
         }
         List<String[]> contentdomain = new ArrayList<>();
         try(BufferedReader br = new BufferedReader(new FileReader(filedomain))) {
             String line = "";
             int i=0;
             while ((line = br.readLine()) != null) {
                 contentdomain.add(line.split(","));
                 //System.out.println(content.get(i)[0].toString());
                 i++;
             }
         } catch (FileNotFoundException e) {
           //Some error logging
         }

if(message_id_pos>=0){
             String message id text = arrList.get(message id pos);
             message_id_text = message_id_text.replaceAll(message_id_key, "");
             int at_pos = message_id_text.indexOf("@");
             int last_close_pos = message_id_text.indexOf(">");
             String host_message_id = message_id_text.substring(at_pos+1,
last_close_pos);
             //hasil = hasil + "\nHost ID"+"="+host message id;
             Boolean statusDomainFraud = false;
             for (int j=1;j<contentdomain.size();j++) { //karena 0 baca header csv

                 if
(contentdomain.get(j)[0].toString().toLowerCase().equals(host_message_id)) {
                     statusDomainFraud = true;
                 }
             }
             if (statusDomainFraud == true) {
                 //hasil = hasil + "\n--- MESSAGE ID ---"+"";
                 hasil +=  "*Terindikasi email address spoofing"+"\n";
                 hasil = hasil + "Message ID"+"="+message_id_text+"\n\n";
             }
         }
```

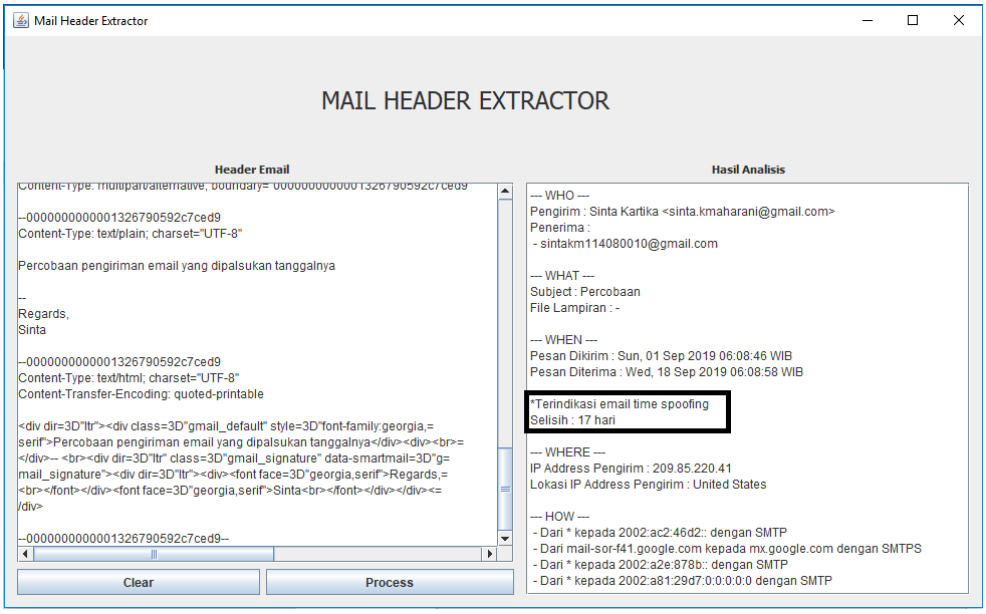
Gambar 4.14 *Source code* deteksi *email address spoofing*

Gambar 4.14 memperlihatkan proses deteksi *email address spoofing*. Informasi nama domain dari *server* pengirim email didapatkan dari *field Message-ID*. Kemudian nama domain ini dibandingkan dengan nama domain dari situs *fake mailer* yang tersimpan pada *file* domain.csv.

Berdasarkan hasil analisis *header email* pada Gambar 4.12 dan Gambar 4.13 terlihat bahwa *tool* Mail Header Extractor telah mampu menunjukkan indikasi *email address spoofing* dengan membandingkan nama domain pada *field Message-ID* dengan nama domain pada kamus domain.

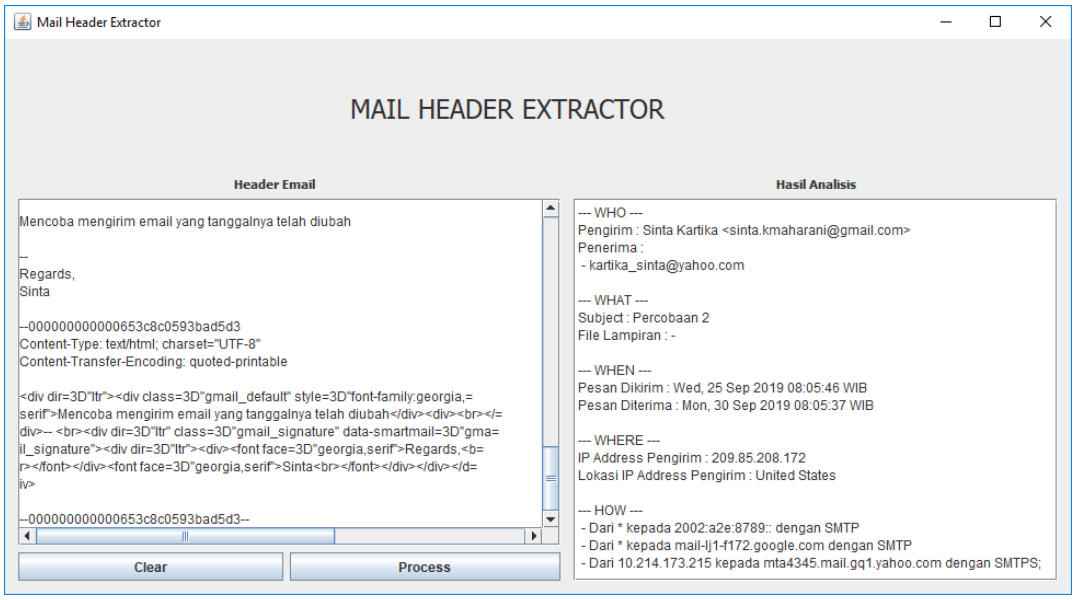
4.3.4 Email yang telah dipalsukan waktu pengirimannya

Skenario pengujian email yang telah dipalsukan tanggal pengirimannya bertujuan untuk menguji apakah *tool* Mail Header Extractor dapat memberikan indikasi terhadap email yang diduga *email time spoofing* berdasarkan selisih antara waktu pengiriman (tercantum pada *field Date*) dan waktu penerimaan (tercantum pada *field Received* atau *X-Received* yang teratas). Jika selisih waktunya lebih dari 5 hari, maka email diindikasikan sebagai *email time spoofing*. Hasil dari skenario ini ditunjukkan pada Gambar 4.15 dan Gambar 4.14.



Gambar 4.15 Hasil pengujian *header email* yang terindikasi *email time spoofing*

Hasil analisis *header email* pada Gambar 4.15 menunjukkan bahwa email dikirimkan dari sinta.kmaharani@gmail.com kepada sintakm114080010@gmail.com dengan subjek "Percobaan" yang dikirimkan pada tanggal 1 September 2019, namun diterima pada 18 September 2019. Karena adanya selisih waktu email diterima oleh penerima yang mencapai 17 hari, sehingga *tool* mengindikasikan email tersebut sebagai *email time spoofing*. Hal ini telah melebihi batas waktu maksimal proses pengiriman pesan dan kemungkinan disebabkan oleh pengirim mengubah tanggal pada komputernya. Sehingga saat pengirim mengirimkan email muncul rentang waktu pengiriman dan penerimaan yang cukup lama.



Gambar 4.16 Hasil pengujian *header email* yang terindikasi *email time spoofing*

Gambar 4.16 menunjukkan bahwa *header email* yang diujikan ke *tool* merupakan email yang dikirimkan dari sinta.kmaharani@gmail.com kepada kartika_sinta@yahoo.com dengan subjek "Percobaan 2" yang dikirimkan pada tanggal 25 September 2019 dan diterima pada tanggal 30 September 2019. Ada selisih waktu 5 hari, namun tidak diindikasikan sebagai *email time spoofing* karena selisih waktu 5 hari ini masih dianggap wajar dan merupakan batas maksimal penundaan pengiriman email oleh *server*.

Gambar 4.17 memperlihatkan *source code* dari proses deteksi *email time spoofing*. Proses deteksi dilakukan dengan membandingkan tanggal email diterima dikurangi dengan tanggal email dikirim. Selisih waktu yang ditetapkan adalah 5 hari, sehingga jika selisih tanggal email dikirim dengan tanggal email diterima lebih dari 5 hari, akan diindikasikan sebagai *email time spoofing*.

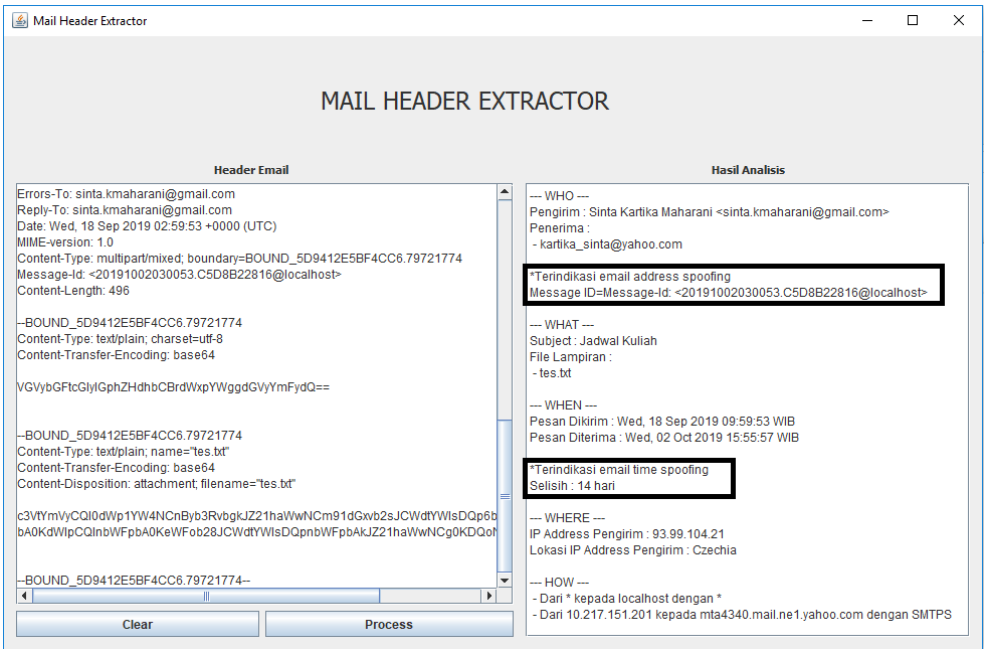
```
Date awal = converToTime(pesan_dikirim_result, from_kirim);
Date akhir = converToTime(pesan_diterima_result, from_terima);
long diff = akhir.getTime() - awal.getTime();
//String selisih = "Difference between  " + awal + " and "+ akhir+" is " + (diff /
(1000 * 60 * 60 * 24)) + " days.";
String selisih = "" + (diff / (1000 * 60 * 60 * 24)) + " hari";

pesan_diterima_result = converTime(pesan_diterima_result,from_terima,to);
System.out.println("pesan_diterima_result_convert :" + pesan_diterima_result);

if ((diff / (1000 * 60 * 60 * 24)) > 5) {
     hasil += "\n*Terindikasi email time spoofing\n";
     hasil += "Selisih : "+selisih+"\n";
}
```

Gambar 4.17 *Source code* proses deteksi *email time spoofing*

Berdasarkan Gambar 4.15 dan Gambar 4.16 terlihat bahwa *tool* Mail Header Extractor telah mampu memberikan indikasi jika terjadi *email time spoofing* dari *header email* yang diujikan dengan membandingkan waktu pengiriman pesan (informasi ini didapatkan dari *field Date*) dan waktu penerimaan pesan (informasi ini didapatkan dari *field Received* atau *X-Received* yang paling atas).

4.3.5 Email yang telah dipalsukan alamat pengirim dan waktu pengirimannya

Skenario ini dilakukan untuk menunjukkan bahwa *tool* Mail Header Extractor dapat memberikan indikasi jika *header email* yang diujikan telah dipalsukan alamat pengirim dan waktu pengirimannya. Hal tersebut ditunjukkan pada Gambar 4.14.

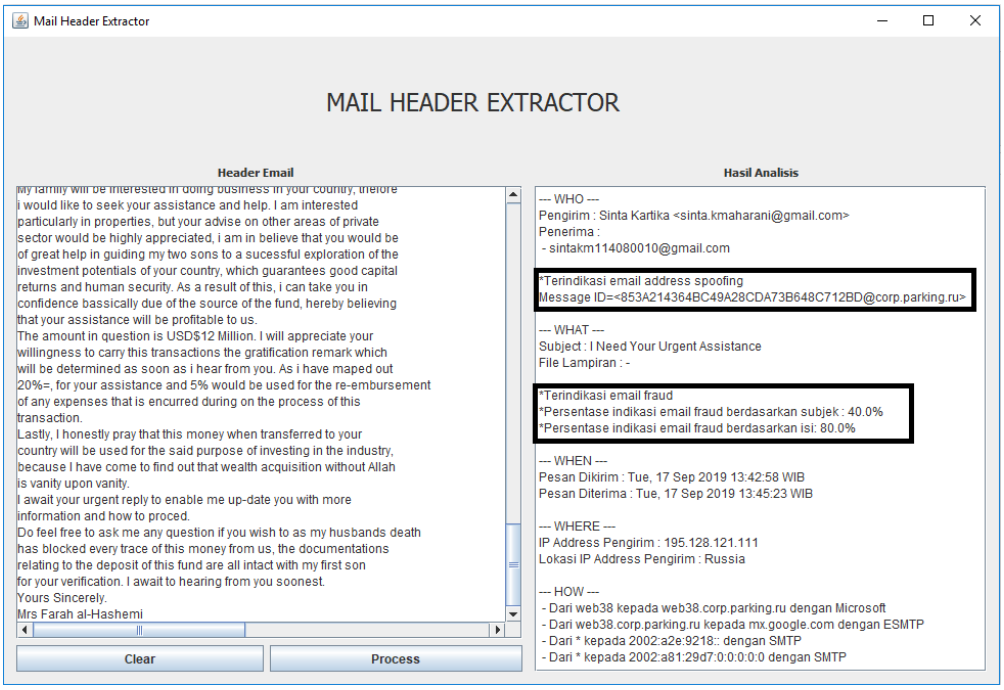


Gambar 4.18 Hasil pengujian *header email* yang terindikasi *email address spoofing* dan *email time spoofing*

Header email yang diujikan pada Gambar 4.18 merupakan *header* dari email yang dikirimkan dari situs *fake mailer* dan fitur untuk memodifikasi tanggal pengiriman emailnya telah diaktifkan. Dari gambar di atas terlihat bahwa email dikirimkan dari sinta.kmaharani@gmail.com, namun sebetulnya dari situs *fake mailer* dengan domain "localhost" kepada kartika_sinta@yahoo.com dengan subjek "Jadwal Kuliah". Karena domain ini termasuk dalam kamus domain, maka email diindikasikan sebagai *email address spoofing*. Email tercatat dikirimkan pada 18 September 2019 namun baru diterima di tanggal 2 Oktober 2019, sehingga diindikasikan *email time spoofing* karena ada selisih 14 hari untuk pengiriman emailnya.

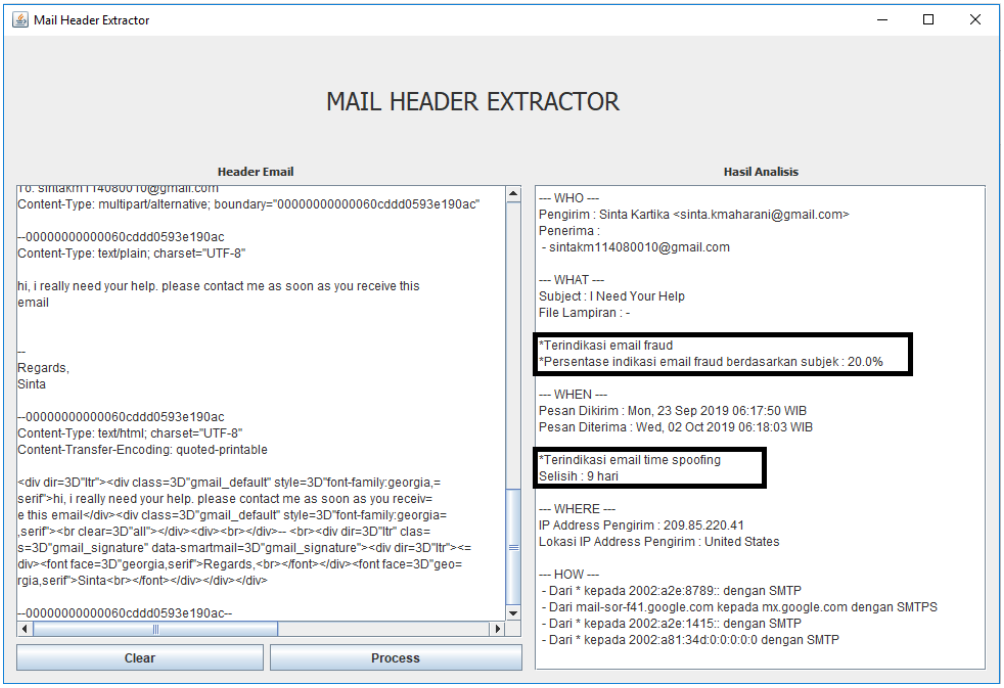
4.3.6 Email yang subjeknya berupa *fraud* serta telah dipalsukan alamat pengirim dan tanggal pengirimannya.

Hasil dari skenario ini ditunjukkan pada Gambar 4.19, Gambar 4.20, dan Gambar 4.21.



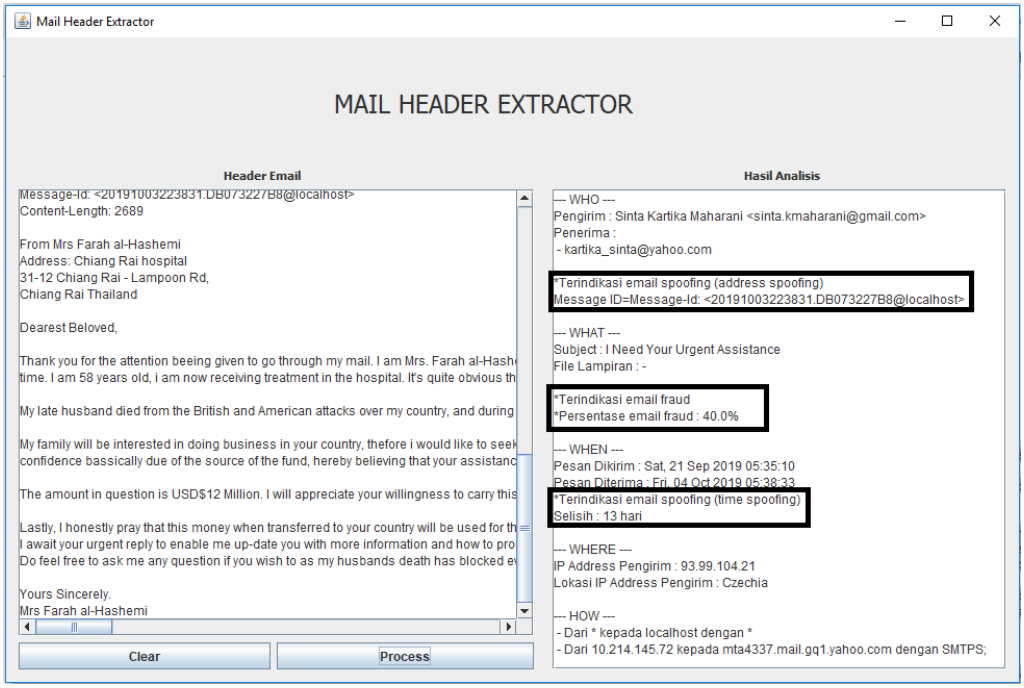
Gambar 4.19 Hasil pengujian *header email* yang terindikasi *email fraud* dan *email address spoofing*

Gambar 4.19 menunjukkan bahwa *header email* yang dikirim dari sinta.kmaharani@gmail.com kepada sintakm114080010@gmail.com diindikasikan sebagai *email fraud* karena subjeknya berjudul "I Need Your Urgent Assistance". Ada beberapa kata pada subjek cocok dengan kata pada kamus, yaitu "*need*", "*your*", "*urgent*", dan "*assistance*", sehingga persentasenya 40%. Selain itu, berdasarkan isinya, persentase indikasi *email fraud*-nya mencapai 80%. Angka 80% ini menunjukkan bahwa ada 4 kata pada isi email yang cocok dengan 5 kata yang ada pada kamus pembanding isi email. Email juga diindikasikan sebagai *email address spoofing* karena dikirimkan dari domain "corp.parking.ru", alih-alih dari domain gmail.



Gambar 4.20 Hasil pengujian *header email* yang terindikasi *email fraud* dan *email time spoofing*

Gambar 4.20 di atas menunjukkan bahwa *header email* yang dirikimkan dari sinta.kmaharani@gmail.com kepada sintakm114080010@gmail.com diindikasikan sebagai *email fraud* karena subjeknya berjudul "I Need Your Help". Pada subjek terdapat kata yang cocok dengan kata pada kamus, yaitu "*need*" dan "*your*", sehingga persentasenya menunjukkan 20%. Namun tidak muncul persentase indikasi *email fraud* berdasarkan isi, hal ini berarti pada isi emailnya tidak ditemukan kata yang cocok dengan kamus yang digunakan untuk deteksi *email fraud* berdasarkan isi. Email juga diindikasikan sebagai *email time spoofing* karena email dikirimkan pada 23 September 2019 dan diterima pada 2 Oktober 2019, selisih waktu pengirimannya mencapai 9 hari.



Gambar 4.21 Hasil pengujian *header email* yang terindikasi *email fraud, email address spoofing,* dan *email time spoofing*

Gambar 4.21 menunjukkan bahwa *header email* yang dirikimkan dari sinta.kmaharani@gmail.com kepada kartika_sinta@yahoo.com diindikasikan sebagai *email fraud* karena subjeknya berjudul "I Need Your Urgent Assistance". Terdapat kecocokan antara kata pada subjek dengan kata pada kamus, yaitu "*need*", "*your*", "*urgent*", dan "*assistance*" sehingga persentasenya menunjukkan 40%. Selain itu, berdasakan isinya, email juga diindikasikan sebagai *fraud* dengan persentase 80%, yang artinya ditemukan 4 kata pada isinya yang cocok dari 5 kata pada kamus. Email juga diindikasikan sebagai *email time spoofing* karena email dikirimkan pada 21 September 2019 dan diterima pada 4 Oktober 2019, sehingga terdapat selisih waktu pengiriman mencapai 13 hari. Selain itu, email diindikasikan sebagai *email address spoofing* karena ternyata email dikirimkan dari domain "localhost", bukan dari domain gmail seperti alamat email seharusnya.

Berdasarkan hasil pengujian yang telah dilakukan, *tool* Mail Header Extractor telah berhasil melakukan ekstraksi dari *header email* kemudian memetakan informasinya ke dalam konsep 4W1H. *Who* dapat menjawab siapa pengirim dan penerima email, *What* dapat menjawab apa subjek dan file lampiran pada email, *When* dapat menjawab kapan waktu email dikirim oleh pengirim dan diterima oleh penerima, *Where* dapat menjawab berapa alamat IP *server* pengirim email beserta di mana lokasinya, dan *How* dapat menjawab

bagaimana proses pengiriman email dari pengirim ke penerima, *server* apa saja yang dilewati beserta protokolnya.

Tool ini juga dapat memberikan indikasi *email fraud* dan *email spoofing* jika *header email* yang diuji diduga sebagai *email fraud* dan *email spoofing*. Fitur deteksi *email fraud* ini merupakan salah satu keunggulan dari *tool* Mail Header Extractor dibandingkan *tool* yang telah ada sebelumnya. Namun pada penelitian ini fitur deteksi *email fraud* masih perlu dioptimasi dalam penerapannya. Salah satu metode yang dapat dikembangkan untuk optimasi deteksi *email fraud* adalah dengan melakukan analisis bentuk kalimat. Analisis bentuk kalimat dapat dilakukan dengan membuat sistem cerdas yang kemudian diintegrasikan pada *tool*. Algoritma yang disusun untuk membentuk sistem cerdas tersebut harus mampu mendeteksi susunan dan bentuk kalimat yang mengindikasikan *email fraud*.