Pengembangan Protokol *Single Sign-On* SAML dengan Kombinasi *Speech* dan *Speaker Recognition*

Patrick Telnoni
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung
Bandung, Indonesia
ptelnoni89@gmail.com

Rinaldi Munir
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung
Bandung, Indonesia
rinaldi@informatika.org

Yusep Rosmansyah
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung
Bandung, Indonesia
yusep@stei.itb.ac.id

ABSTRAK

Single Sign-On (SSO) adalah sistem otentikasi terpusat yang memampukan seseorang memperoleh akses pada banyak layanan dengan satu kali proses otentikasi. Akan tetapi, *Single Sign-On* mempunyai satu kelemahan fatal. Jika sebuah *credential* berhasil diambil oleh seorang *imposter* maka *imposter* dapat mengakses berbagai layanan dengan satu kali otentikasi. Protokol SSO SAML yang cukup banyak diteliti, juga tidak lepas dari kerentanan keamanan ini. Penelitian ini dilakukan untuk memberikan solusi terhadap kerentanan tersebut menggunakan kombinasi *speech* dan *speaker recognition*.

Kata Kunci

Single sign-on, biometric, SAML*, speech recognition, speaker recognition*

1. PENDAHULUAN

Meningkatnya jumlah layanan yang tersedia di Internet yang memerlukan registrasi membuat jumlah *credential* yang dimiliki seorang *user* ikut meningkat. Untuk mengatasi masalah ini, maka diperlukan sistem otentikasi terpusat. Sistem ini biasa disebut dengan *Single Sign-On* (SSO) [5,11].

Masalah utama pada SSO adalah jika sebuah *credential* yang dipakai untuk otentikasi berhasil dicuri, maka pencuri dapat mengakses berbagai layanan dengan satu kali otentikasi. *Credential* yang digunakan umumnya berupa *username* dan *password*. Jenis *credential* ini sangat mudah untuk dicuri. Oleh karena itu, penelitian ini menambahkan *biometric* pada SSO untuk mencegah penyerang terotentikasi pada SSO. Dengan mempertimbangkan harga mikrofon dan semakin banyaknya perangkat yang mengggunakan mikrofon sejak 2006 [17], maka penelitian ini menggunakan *biometric* suara.

2. LANDASAN TEORI

2.1 SAML

SAML [2, 14] adalah portokol SSO yang berbasiskan XML, di mana pesan otentikasi dan protokol-protokol terkait dimuat dalam format XML. SAML melakukan proses *protocol binding* yang memuat struktur pesan SAML ke dalam struktur pesan yang lain untuk mengirim pesan. Sebagai contoh, SAML menggunakan *Simple Object Access Protocol* (SOAP) menggunakan *HTTP Binding* [2]. *Profile* ini mengandung alur kerja protokol dan *security constraint* untuk penggunaan SAML. Contoh pesan dapat dilihat pada gambar 1.

```
<saml:Subject>
    <saml:NameID  Format="urn:oasis:
names:tc:SAML:1.1:nameid-ormat:
unspecified">
patrick.telnoni@students.itb.ac.id
   </saml:NameID>
    <saml:SubjectConfirmation
Method="urn:oasis:names:tc:SAML:2.0:cm:bearer">
        <saml:SubjectConfirmationData
NotOnOrAfter="20014-05-30T02:44:24.173Z"
Recipient="http://localhost:8080"/>
    </saml:SubjectConfirmation>
</saml:Subject>
```

Gambar 1 Contoh SAML *Assertion*

SAML dikembangkan atas pertimbangan sebagai berikut:

1. Untuk kebutuhan SSO yang disebutkan sebelumnya.
2. *Federated identity*, yang berarti bahwa seorang *user* mempunyai satu identitas yang tetap. Hal ini dikarenakan dalam satu *security domain* atau yang lebih umum dikenal dengan situs, *user* A di *Identity Provider* B telah terdaftar di *security domain* C tanpa mendaftarkan diri melalui B.
3. Modularitas. Dibutuhkan modularitas untuk dapat dijalankan pada pada berbagai *platform*. SAML dapat dijalankan pada berbagai *framework*, *web service* dan lain sebagainya.

Berbeda dengan OpenID di mana semua pihak dapat menjadi *ID Provider*, SAML mengharuskan *ID Provider* dan *Service Provider* dipasangkan. SAML menggunakan HTTP, SMTP, FTP dan SOAP untuk mengirim pesan. SAML memiliki komponen sebagai berikut [14]:

1. *Assertions*

   SAML memungkinkan salah satu pihak untuk menyampaikan informasi mengenai sebuah subyek. Contoh dari gambar 1 dapat dilihat bahwa pesan SAML dapat menyatakan subyek bernama "Patrick Telnoni", mempunyai alamat *email* patrick.telnoni@s.itb.ac.id, dan merupakan mahasiswa STEI ITB.

2. *Protocol*

   SAML has memiliki beberapa protokol yang menangani *request* dan *response*.

3. *Binding*

SAML *Binding* memetakan pertukaran pesan pesan *request-response* ke dalam pesan standar atau protokol komunikasi lain. Contohnya SAML mengangkut pesan melalui *transport protocol* pada *layer* bawah.

4. *Profile*

   SAML *Profile* mendefinisikan bagaimana SAML *assertion*, *protocol* dan *binding* dikombinasikan dan ditentukan batasannya untuk memperoleh interoperabilitas dalam beberapa skenario penggunaan.

Selain itu, SAMLjuga memiliki fitur lain yaitu sertifikat X509 dan tanda tangan digital dengan kriptografi asimetris. Tanda tangan digital pada SAML mirip dengan tanda tangan digital pada XML. Lebih lanjut mengenai fitur ini akan dijelaskan pada poin IV.

2.2 *Biometric*

Biometric [4] adalah sebuah cara untuk mengenali seseorang menggunakan seseorang dengan menggunakan karakter dan kelakuan biologisnya. *Biometric* digunakan dengan asumsi bahwa seseorang unik secara fisik dan tingkah laku. *Biometric* banyak digunakan untuk membuka dan membatasi akses ke ruangan, informasi, layanan dan bahkan untuk melewati batas negara

Biometric memiliki beberapa komponen umum, yaitu [8]:

1. *Capture*

   Komponen untuk mengambil sampel *biometric* seseorang. Biasanya disebut sebagai sensor.

2. *Database* referensi

   Database yang menyimpan data mengenai user yang terdaftar, termasuk *biometric* yang terekam.

3. *Matcher*

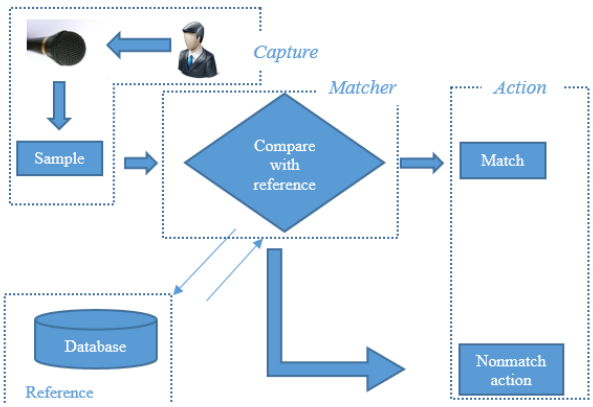
   Membandingkan sampel *biometric* yang diambil dengan sampel yang terekam pada *database*.

4. *Action*

   Keputusan yang diambil terkait hasil yang dihasilkan oleh *matcher*.

2.3 *Speaker Recognition*

Speaker recognition [10] adalah metode untuk mengenali pemilik suara berdasarkan pola suara yang disimpan. *Biometric* suara yang dikeluarkan dalam riset ini erat kaitanya dengan *speech processing*



Gambar 2 Proses *speaker recognition*

Dalam *speaker* recognition, proses-proses yang dilakukan adalah menangkap suara, melakukan *feature extraction* pada suara, melakukan *pattern matching* dan memberikan keputusan terkait hasil proses *matching*. Gambar 2 menunjukan proses *speaker recognition* secara umum.

2.4 *Speech Recognition*

Speech recognition [13] adalah metode untuk mengenali kata yang diucapkan oleh seorang *user*. *Speech recognition* cukup mirip dengan *speaker recognition*. Yang membedakan keduanya adalah, pada *speech recognition*, *feature extraction* dicocokan pada *acoustic model*. *Accoustic model* adalah rekam data yang memuat representasi statistik dengan transkrip teks kata. Contoh *accoustic model* dapat dilihat pada Gambar 3.

```
1720 n_state_map
602 n_tied_state
102 n_tied_ci_state
34 n_tied_tmat
#
# Columns definitions
#base lft  rt p attrib tmat      ... state id's ...
AX_one   -   - -    n/a    0      0      1      2 N
AY_five  -   - -    n/a    1      3      4      5 N
AY_nine  -   - -    n/a    2      6      7      8 N
EH_seven -   - -    n/a    3      9     10     11 N
EY_eight -   - -    n/a    4     12     13     14 N
E_seven  -   - -    n/a    5     15     16     17 N
F_five   -   - -    n/a    6     18     19     20 N
F_four   -   - -    n/a    7     21     22     23 N
II_three -   - -    n/a    8     24     25     26 N
II_zero  -   - -    n/a    9     27     28     29 N
I_six    -   - -    n/a   10     30     31     32 N
K_six    -   - -    n/a   11     33     34     35 N
N_nine   -   - -    n/a   12     36     37     38 N
N_nine_2 -   - -    n/a   13     39     40     41 N
N_one    -   - -    n/a   14     42     43     44 N
N_seven  -   - -    n/a   15     45     46     47 N
```

Gambar 3 Contoh *accoustic model*

Selain itu, *speech recognition* melakukan pencocokan dengan *language model* untuk memprediksi kata yang akan diucapkan oleh *user* [12]. *Language model* adalah rekaman statistic yang menghitung probabilitas terhadap beberapa uruan kata. Contoh *language model* dapat dilihat pada Gambar 4. Terdapat dua acara untuk membuat *language model* yaitu dengan membuat membuat *grammar* sendiri dan dengan membuat model statistic. Gambar 4 menunjukan *language model* yang berupa *grammar* yang dibuat sendiri.

```
#JSGF V1.0;

/**
 * JSGF Grammar for Hello World example
 */

grammar hello;

public <greet> = (Good morning | Hello | Goodbye) ( Bhiksha | Evandro |
```

Gambar 4 Contoh *language model* dalam bentuk *grammar*

Secara umum, proses *speech recognition* dapat dilihat pada gambar 6.

3. RISET TERKAIT

Ada banyak riset terkait yang menggunakan *biometric* untuk melakukan otentikasi, namun yang digunakan pada SSO tidak dapat kami temukan. Oleh karena itu, kami mempelajari riset terkait tentang SAML dan *biometric*.

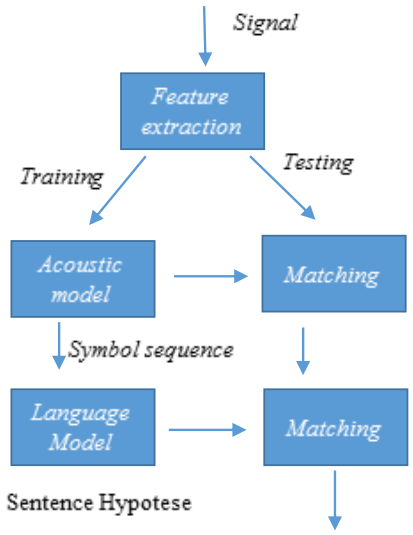
Riset terkait yang mengembangkan SAML dengan *two way authentication* ditemui dalam [16]. Dalam riset tersebut, SAML dikombinasikan dengan menggunakan XACML. XACML adalah standar *web service* berbasis XML untuk komunikasi kebijakan akses kontrol antar layanan. Namun, karena XACML merupakan standar yang berbasis XML, XACML rentan terhadap serangan *XML Signature Wrapping* (XSW) [3].

Seperti yang disebutkan sebelumnya, SAML memiliki tanda tangan digital. Tanda tangan ditial ini menggunakan sertifikat X509. Contoh tanda tangan digital dapat dilihat pada Gambar 5.

```
<ds:Signature
xmlns:ds="http://www.w3.org/2000/09/xmldsig#">
        <ds:SignedInfo>
                        <ds:CanonicalizationMethod
Algorithm="http://www.w3.org/2001/10/xml-exc-
c14n#WithComments" />
                                <ds:SignatureMethod
Algorithm="http://www.w3.org/2000/09/xmlds#rsa-
sha1" />
                <ds:Reference  URI="#ID_ab0392ef-
b557-4453-95a8-a7e168da8ac5">
                    <ds:Transforms>
                                        <ds:Transform
Algorithm="http://www.w3.org/2000/09/xmlds#enve
loped-signature" />
                                        <ds:Transform
Algorithm="http://www.w3.org/2001/10/xml-exc-
c14n#" />
                    </ds:Transforms>
                                    <ds:DigestMethod
Algorithm="http://www.w3.org/2000/09/xmlds#sha1
" />
                    <ds:DigestValue>0Y9QM5c5qCShz5U
WmbFzBmbuTus=</ds:DigestValue>
                </ds:Reference>
            </ds:SignedInfo>
            <ds:SignatureValue>
                se/flQ2htUQ0IUYieVkXNn9cfjnfgv6H99n
FarsTNTpRI9xuSlw5OTai/2PYdZI2Va9+QzzBf99m
                VFyigfFdfrqug6aKFhF0lsujzlFfPfmXBbD
RiTFX+4SkBeV71uuy7rOUI/jRiitEA0QrKqs0e/pV
                \+C8PoaariisK96Mtt7A=
            </ds:SignatureValue>
            <ds:KeyInfo>
                <ds:KeyValue>
                    <ds:RSAKeyValue>
                        <ds:Modulus>
                            suGIyhVTbFvDwZdx8Av62zm
P+aGOlsBN8WUE3eEEcDtOIZgO78SImMQGwB2C0eIVMhiLRz
VPqoW1
                            dCPAveTm653zHOmubaps1fY
0lLJDSZbTbhjeYhoQmmaBro/tDpVw5lKJwspqVnMuRK19ju
2dxpKw
                            lYGGtrP5VQv00dfNPbs=
                        </ds:Modulus>
                        <ds:Exponent>AQAB</ds:Expon
ent>
                    </ds:RSAKeyValue>
                </ds:KeyValue>
            </ds:KeyInfo>
    </ds:Signature>
```

Gambar 5 Contoh tanda tangan digital pada SAML

XSW adalah serangan pada XML dengan menyisipikan elemen XML yang serupa dengan elemen yang didefinisikan dalam tanda tangan digital pada pesan SAML. Di dalam tanda tangan digital XML, Dengan menyisipkan konten yang serupa, diharapkan penerima pesan SAML akan membaca isi konten yang disisipkan tadi. Namun, dalam riset [3] juga diberikan beberapa *countermeasure* teknis yang mudah diimplementasikan.



Gambar 6 Proses *speech recognition*

Riset terkait tentang *biometric* juga dilakukan untuk menguji berbagai *biometric*. Dalam [4], dilakukan uji coba terhadap beberapa *biometric*, yaitu *fingerprint*, iris, wajah, suara, tulisan tangan dan pola pembuluh darah tangan. Dari riset [4], *biometric* yang memberikan hasil paling baik adalah suara dan tulisan tangan.

Dalam riset [6], *biometric* yang dicoba adalah pola pembuluh darah pada tangan. Dalam riset ini, tes dilakukan pada 100 *user*. Dari tes tersebut, diperoleh 0,1% *false acceptance rate* untuk 98,5% *acceptance rate*. Hanya saja, mengingat mahalnya harga kamera termal infra merah, kami memutuskan untuk tidak menggunakan pola pembuluh darah tangan.

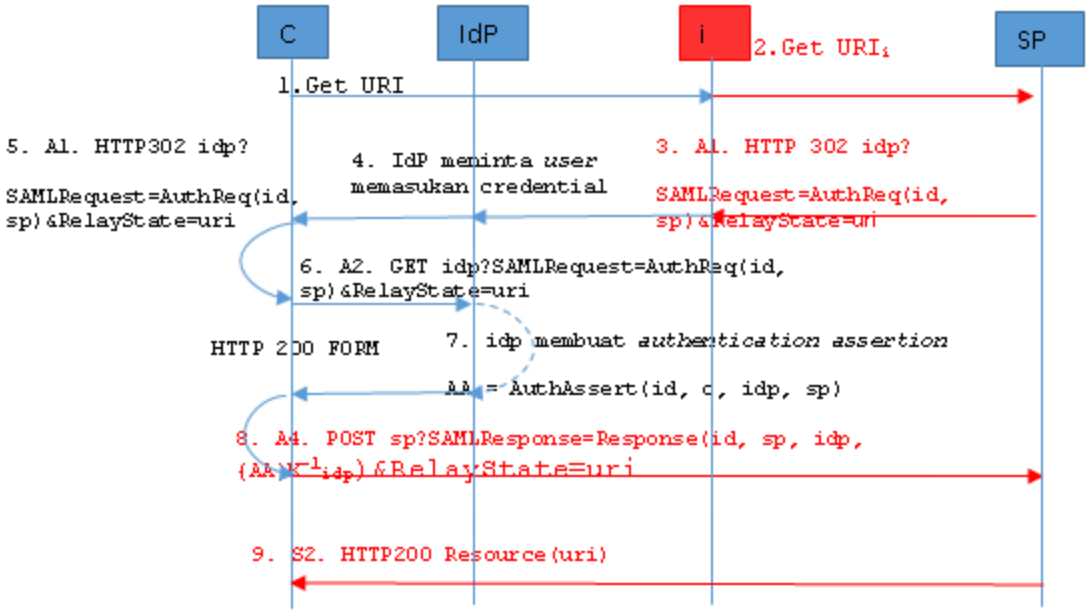
Riset dalam [7] mendesain *Biometric Authentication as a Service* (BioAAS). Dalam riset ini, SAML tidak menjadi fokus utama, namun hanya menjadi komponen teknis dalam membangun BioAAS. BioAAS dalam riset ini dibangun dengan memperhatikan regulasi keamanan data Jerman dan Eropa.

Riset pada [12] menggunakan *keystroke* pada perangkat *keyboard* untuk otentikasi. *Keystroke* adalah irama/ritme suara yang dikeluarkan oleh perangkat *keyboard* saat seseorang sedang mengetik. *Keystroke* pada tiap-tiap manusia memiliki ritme yang berbeda-beda. Hanya saja, *keystroke* kurang dikembangkan menjadi *library* untuk bahasa pemrograman dibandingkan dengan *biometric* lain seperti suara dan wajah.

Dari penelusuran kami, tidak ditemukan riset terkait SAML menggunakan *biometric*. Oleh karena itu, kami mengembangkan SAML dengan menggunakan *biometric* untuk menanggulangi masalah pencurian hak akses. Kami menggunakan *biometric* suara dengan mempertimbangkan aspek ekonomi dan performa seperti yang ditunjukan dalam riset [4,12].

4. DESAIN DAN ANALISIS

SAML cukup banyak diteliti dibandingkan dengan OpenID maupun OAuth, Dalam riset [1], ditemukan celah keamanan pada SAML seperti yang ditunjukan oleh gambar 7. Dalam gambar 7, komunikasi antara *ID Provider* (IdP) dengan *Service Provider* (SP) dapat dicegat oleh penyerang/interceptor (i). Akibatnya, konten berbahaya dapat disisipkan oleh penyerang.



Gambar 7 *Vulnerability* pada SAML

Dalam gambar 7, bagian yang ditandai dengan warna merah menunjukan tindakan dan hasil yang dilakukan oleh penyerang. Celah keamanan ini berpotensi mengakibatkan terjadinya serangan sebagai berikut [1]:

1. *Cross site scripting* (XSS),
2. *Cross Site Resource Forgery* (CSRF),
3. *Redirects* ke situs berbahaya

Ketiga serangan tersebut tercatat dalam OWASP Top 10 Vulnerability 2013 [13]. Serangan nomor 1 dan 3 dapat berujung pada pencurian hak akses dan *credential*. Sebagaimana disebutkan sebelumnya bahwa pada SSO, ketika satu *credential* berhasil dicuri, maka penyerang dapat memperoleh akses pada banyak layanan.

Di dalam [9], disebutkan bahwa untuk meningkatkan level keamanan pada proses otentikasi, maka perlu dikombinasikan dua atau lebih jenis *credential*. Cara ini disebut dengan *two way authentication*. Beberapa jenis *credential* seperti yang disebutkan dalam [9] ada 3 macam, yaitu:

1. *Something you know*, meruapakan tipe *credential* yang paling banyak digunakan. Umumnya berupa *username* dan *password*;
2. *Something you have*, tipe *credential* yang menggunakan barang yang dimiliki *user* seperti *smartcard* dan ATM;
3. *Something you are*, menggunakan sesuatu yang menjadi bawaan pada seseorang secara lahiriah, biasanya *biometric*.

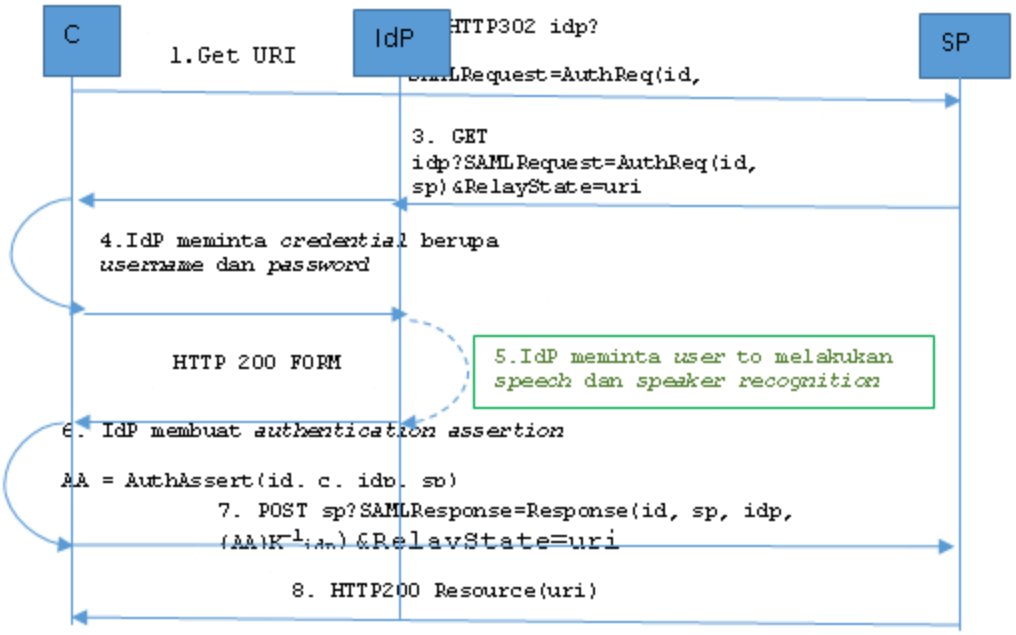
Tipe (1) dapat diperoleh dengan mudah menggunakan berbagai cara seperti *social engineering* dan *sniffing*. Tipe (2) akan membutuhkan banyak biaya untuk produksi dan infrastruktur seperti sensor dibandingkan dengan tipe (3).

Dengan demikian, kami memilih menggunakan tipe (3) dengan biaya sensor yang murah.

Berdasarkan analisis tersebut, *biometric* ditambahkan untuk proses otentikasi pada *ID Provider*. SSO yang dibangun akan menggunakan *speaker recognition* untuk mencegah penyerang terontentikasi pada *ID Provider*. Pemilihan *biometric* sura didasarkan pada beberapa pertimbangan sebagai berikut:

1. Mikrofon yang merupakan sensor *biometric* suara telah menjadi perangkat bawaan dalam kompute dan *smartphone*;
2. Harga mikrofon sangat murah dibandingkan dengan sensor *biometric* lain;
3. Dibandingkan dengan *biometric* wajah yang sensornya juga murah, *biometric* suara lebih sulit untuk ditembus. Dewasa ini, *printer* 3D semakin canggih, sehingga dapat mencetak wajah seseorang mendekati aslinya, sehingga bisa digunakan untuk pengenalan wajah. Walaupun begitu, *printer* 3D masih tergolong sangat mahal.

Dengan menambahkan *biometric*, gambar 7 dimodifikasi menjadi gambar 8.



Gambar 8 Solusi yang diusulkan

Bagian yang ditandai dengan warna hijau merupakan solusi yang ditambahkan. Proses yang lebih rinci dapat dilihat pada gambar 9.

Namun, terdapat isu lain yang terkait dengan proses pengenalan suara. Proses *speaker recognition* dapat ditembus dengan mudah menggunakan rekaman suara dengan kualitas *high definition* (HD) ditambah lagi dengan *smartphone* sekarang yang mampu memberikan suara dengan kualitas HD. Untuk menanggulangi masalah itu, kami menambahkan *speech recognition* untuk mencegah skenario dimana suara *user* berhasil direkam dengan kualitas HD.

Speech recognition akan berperan sebagai *password* suara. Sistem akan membuat sebuah kalimat yang terdiri dari beberapa kata secara acak. Kalimat tadi kita sebut sebagai *keywords*. Dengan demikian, saat *user* melakukan otentikasi, *user* diminta untuk mengucapkan *keywords* yang diminta oleh sistem. Saat *user* mengucapkan *keywords*, suara *user* direkam.

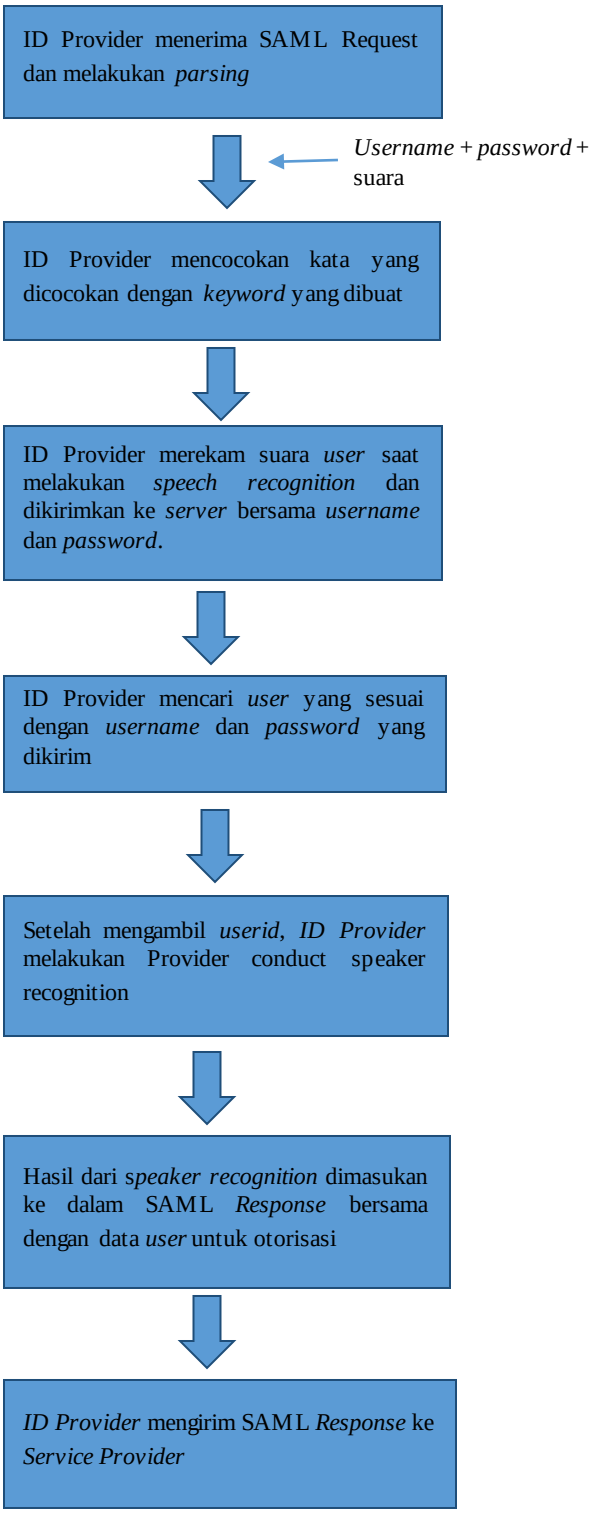
Jika kata yang diucapkan sesuai dengan apa yang diharapkan oleh sistem, informasi dari *form* otentikasi (*username*, *passwords* dan suara) akan dikirim ke *server* untuk diproses. Jika kalimat yang diucapkan tidak sesuai dengan *keywords*, maka *user* akan diminta untuk melakukan *speech recognition* dengan *keywords* yang berbeda. Pada *server*, akan dilakukan otentikasi dengan *username* dan *password* serta *speaker recognition*.

Terkait data suara milik *user*, data ini akan disimpan dalam bentuk rekaman berkekstensi `wav`. Penyimpanan file rekaman ini tidak menggunakan enkripsi maupun *password*. Pertimbangan ini didasari karena sistem SSO yang dikembangkan menggunakan *speech recognition*. Apabila suara rekaman *user* berhasil dicuri, suara rekaman tersebut tidak dapat digunakan untuk keperluan otentikasi, sekalipun menggunakan rekaman *high definition*. Hal ini karena seperti yang telah dijelaskan sebelumnya, *user* akan diminta mengucapkan kalimat *keywords* yang dibuat secara berbeda-beda setiap kali *form* otentikasi dimuat.

Hal ini karena pada saat *speech recognition*, *user* harus mengucapkan kata yang diminta oleh sistem. Hasil *speech recognition* harus sesuai dengan *keyword* yang diminta oleh sistem.

Dari solusi yang diusulkan, kami mendefinisikan tiga *use case*, yaitu:

1. Pendaftaran, mendefinisikan proses pendaftaran *user* yang di dalamnya terdapat proses perekaman suara *user* untuk *data training*.
2. *Login*, mendefinisikan proses otentikasi dengan *username*, *password* dan *biometric* suara.
3. *Logout*, mendefinisikan bagaimana *user* melakukan *logout* dari *session* yang sedang aktif.



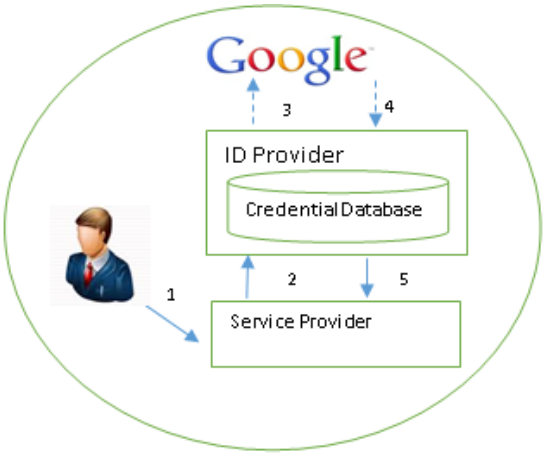
Gambar 9 Usulan solusi secara detil

5. IMPLEMENTASI

Pada tahap implementasi, kami menggunakan HTML5 untuk melakukan perekaman suara dan *speech recognition*. Kami memilih HTML5 sebagai component *front-end* karena HTML5 menawarkan banyak fungsionalitas seperti mengakses mikrofon pada laptop atau *smartphone*. Proses *speech recognition* sendiri dilakukan dalam *server* milik Google. Untuk *speaker recognition*, kami menggunakan *Modular Audio Recognition Framework* (MARF) yang sering digunakan untuk keperluan penelitian.

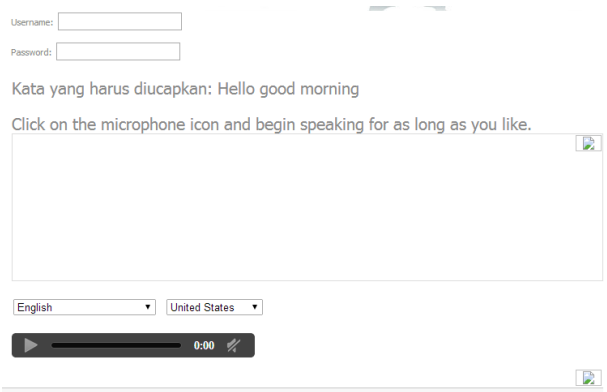
Gambar 10 menunjukan alur kerja SSO yang didesain. Langkah-langkah yang dilakukan adalah sebagai berikut:

1. *User* meminta akses ke *service provider*;
2. *Service Provider* mengarahkan *user* ke *ID Provider*. *ID Provider* meminta *user* untuk memasukan *username* dan *password*. Dalam tahap ini, *user* juga diminta untuk mengucapkan *keyword* yang diminta;
3. *ID Provider* meminta Google untuk melakukan *speech recognition*;
4. Google mengirim hasil *speech recognition*;
5. *ID Provider* melakukan proses *login* menggunakan *username* dan *password*. Setelah proses otentikasi dan *speaker recognition* berhasil, *ID Provider* mengarahkan *user* ke *services provider*;



Gambar 10 Alur proses sistem yang didesain

Pada gambar 11 terdapat contoh *form* otentikasi pada *ID Provider* yang terdapat *field* untuk *username*, *password* dan *keywords* untuk keperluan *speech recognition*. Kalimat yang ditampilkan pada gambar 11 akan berubah-ubah setiap kali *user* akan melakukan otentikasi.



Gambar 11. *Form* othentikasi pada *ID Provider*

MARF dikonfigurasikan sesuai dengan [15], di mana *Endpoint* (sebagai *Preprocessing method*), *Linear Predictive Coding* (sebagai *Feature Extraction method*), *Chebysev Distance* (sebagai *Classification method*) memberikan *acceptance rate* paling tinggi (82.76%). Namun, karena versi MARF yang digunakan masih dalam tahap *beta* maka komponen di dalamnya belum lengkap. Oleh karena itu, kami menggunakan *Raw* (sebagai *Preprocessing method*), *Fast Fourier Transform* (sebagai *Feature Extraction method*), *Eucledian Distance* (sebagai *Classification method*) yang memberikan 75.86% *acceptance rate*.

Spesifikasi lingkungan unutk melakukan tes adalah sebagai berikut:

1. Spesifikasi lingkungan *hardware* terdiri dari:
   a. Processor: Intel® Core™ i5-2430M CPU @ 2.30 GHz
   b. Memory: 4 GB
   c. 16 bit mikrofon pada ASUS A43SA
2. Spesifikasi lingkungan perangkat lunak yang terdiri dari:
   a. Sistem operasi: Windows 8
   b. IDE: Eclipse Kepler
   c. Bahasa pemrograman: Java 1.7.0 menggunakan Servlet
   d. Apache Tomcat 7.0
   e. Google Chrome *Web browser* versi 34.0.1847.131 m

Untuk lingkungan infrastruktur, kami menggunakan koneksi internet ADSL dengan kecepatan 2 Mbps.

Jumlah *user* untuk melakukan tes adalah 29 *user*. Setiap *user* mempunyai minimal 2 sampel suara dan paling banyak 15 sampel yang didaftarkan. Tabel 1 menunjukan pemetaan jumlah sampel suara dengan jumlah *user* pemilik sampel. Secara keseluruhan, terdapat 277 sampel suara sebagai *data training*.

Tabel 1 Pemetaan jumlah sampel dengan *user*

No	Jumlah sampel	Jumlah *user* pemilik sampel
1	2	3
2	3	5
3	7	1
4	8	2
5	9	1
6	12	2
7	14	5
8	15	10

Kami membagi lingkungan percobaan menjadi dua kondisi, yaitu kondisi tenang dan kondisi bising/ramai. Dalam kondisi tenang, kami mengatur agar ruangan tempat melakukan percobaan setenang mungkin. Sedangkan untuk kondisi bising/ramai, percobaan dilakukan pada saat sedang jam kantor di mana bayak terjadi diskusi dan interaksi antar personal. Untuk setiap kondisi, kami melakukan 10 kali tes untuk setiap *use case*. Jumlah *user* yang digunakan untuk pengujian adalah 10 *user*.

Berdasarkan hasil tes, hampir semua *use case* berjalan dengan baik. Namun, fokus riset ini adalah untuk mencegah penyerang terotentikasi menggunakan *biometric*. Oleh karena itu, analisis akan difokuskan pada *use case login*.

Kami memperoleh hasil 100% (10 dari 10) otentikasi yang berhasil pada lingkungan yang hening dan 90% (9 dari 10) otentikasi yang berhasil dalam lingkungan yang bising. Kegagalan pada proses *login* dijelaskan sebagai berikut:

1. Kegagalan pada proses *speech recognition*

   Deskripsi:

   User mengucapkan *keyword* yang sesuai namun hasil yang diberikan berbeda.

   Analisis:

   a. Bahasa yang digunakan unutk melakukan *speech recognition* adalah bahasa inggris. Bahasa ibu dari para *user* adalah bahasa Indonesia. Karena hal ini, terjadi kesalahan *spelling* pada *keyword* yang berakibat pada hasil *speech recognition*.
   b. *User* mengucapkan *keyword* dalam tempo yang cepat.
   c. *Noise* dari lingkungan sekitar *user* ikut terangkut ke *server* Google untuk proses *speech recognition*.

2. Kegagalan pada proses *speaker recognition*

   Deskripsi:

   Suara yang dikirimkan dari *form* otentikasi merupakan suara asli *user* yang sesuai dengan *username* dan *password*. akan tetapi, *speaker recognition* memberikan hasil yang berbeda dari yang diharapkan.

   Analisis:

   a. Saat *ID Provider* merekam suara *user* lingkungan di sekitar *user* cukup berisik. *Filter* yang digunakan dalam metode *preprocessing* adalah *Raw*. Ini berarti suara yang masuk tidak diberikan *filter* apapun.
   b. *Keyword* hanya terdiri dari tida kata dan membuat durasi rekaman menjadi lebih pendek. Akibatnya, pola suara yang diambil juga sedikit.
   c. Format audio suara yang terekam oleh HTML5 berbeda dengan suara yang terkirim dan diproses di server *ID Provider*. Karena perbedaan format inilah terjadi perbedaan kualitas suara yang direkam.

6. KESIMPULAN

Dari penelitian ini, diambil beberapa kesimpulan, antara lain:

1. Penggunaan *biometric* pada SSO cukup efektif untuk mencegah penyerang terotentikasi pada SSO.
2. Faktor penting penggunaan *biometric* untuk keperluan keamanan adalah harga sensor dan *social acceptance*.
3. Faktor penting dalam proses *speech* dan *speaker recognition* adalah lingkungan sekitar *user*. Dibutuhkan kondisi ideal untuk mencepai hasil yang optimal.

7. SARAN RISET BERIKUTNYA

Untuk riset ke depannya, kami memberikan saran sebagai berikut:

1. Kembangkan *server* khusus untuk proses *speech recognition*. Karena *server* Google dapat dimatikan sewaktu-waktu kapanpun.
2. Gunakan *biometric* wajah daripada *biometric* suara. Karena pada penelitian ini, *biometric* suara melakukan dua kali proses komputasi untuk *speech* dan *speaker recognition*. Sedangkan dengan

biometric wajah hanya melakukan satu kalo proses komputasi untuk mengenali wajah. Meskipun dalam poin IV dijelaskan mengenai kemampuan printer 3D untuk membuat replika wajah, namun dengan pertimbangan efisiensi, *biometric* wajah dapat dicoba.

3. Gunakan enkripsi atau perlindungan terhadap suara yang disimpan pada *server*, karena meskipun *speech recognition* dapat menanggulangi masalah *speaker recognition* yang dapat ditembus menggunakan rekaman suara HD, bukan tidak mungkin ke depannya akan muncul teknologi untuk membuat membuat suara yang mengucapkan kalimat atau kata tertentu menggunakan sampel suara sesesorang.

8. REFERENSI

[1] A. Armando, R. Carbone, L. Compagna, J.Cuellar, G. Pellegrino and A. Sorniotti. "An authentication Flaw in Browser-based Single Sign-On Protocols: Impact and Remediations." *Computers & Security*, vol. 33, pp.41-58, Mar.2013.

[2] T. Grob. "Security Analysis of The SAML Single Sign-on Browser/Artifact Protocol." Computer Security Applications Conference, 2003. Proceedings. 19th Annual , 2003, pp. 298 - 307.

[3] J. Somorovsky, A. Mayer, J. Schwenk, M. Kampmann and M.Jensen.. "On Breaking SAML: Be Whoever You Want to Be", *21st USENIX Security*, vol. 33, pp. 397-412, Aug. 2012.

[4] D.Liu, Z.J. Zhang, N.Zhang. "A biometrics-based SSO authentication scheme in Telematics,". *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discover*, 2012, pp. 191-194.

[5] V. Radhaa & D.H. Reddy. "A Survey on Single Sign-On Techniques". *Procedia Technology*, 2012, pp. 134-139.

[6] A. Kumar, M. Hanmandlu and H. M. Gupta. "Online Biometric Authentication Using Hand Vein Patterns". *Computational Intelligence in Security and Defense Applications*, 2009.

[7] C.Senk and F.Dosler. "Biometric Authentication as a Service for Enterprise Identity Management Deployment: A Data Protection Perspective." Availability, Reliability and Security (ARES) Sixth International Conference, 2011, pp.43-50.

[8] J.N. Pato, J, L.I. Millet. *Biometric Recognition: Challenges and Opportunities*. Washington, D.C.: The National Academies Press, 2010, pp.1-8.

[9] N.Daswani, C. Kern, C. and A. Kesavan. "Foundations of Security What Every Programmer Needs to Know," 1st ed., New York: Appress , 2007, pp.7-22.

[10] J.P. Campbell. "Speaker Recognition: A Tutorial". *Proceedings of The IEEE*, vol. 85, no. 9, pp. 1437-1462, 1997.

[11] J.D. Clerq. "Single Sign On," in *Windows Server 2003 Security Infrastructures: Core Security Features*, 1st ed., USA: Digital Press, 2004, pp.533-579.

[12] J. Roth, X.Liu, A.Ross and D.Metaxas. (2013). "Biometric Authentication via Keystroke Sound". Biometrics (ICB), 2013 International Conference on *IEEE*, 2013, pp.1-8.

[13] K. Samudravijaya. "Speech and Speaker Recognition: A Tutorial". *Tata Institute of Fundamental Research*., 2001.

[14] OASIS. (2008). "Security Assertion Markup Language (SAML) V2.0 Technical Overview. OASIS.

[15] I. Clement, S. Mokhov, D. Nicolacopoulos and S. Sinclair. "Experimentation Results" in Modular Audio Recognition Framework v.0.3.0.6 (0.3.0 final) and its Applications," The MARF Research and Development Group, Montreal, Rep. Qc, Dec, 2007.

[16] S. Fugkeaw, P. Manpanpanich and S. Juntapremjitt, "Adding SAML To Two-Factor Authentication And Single Sign-On Model For Dynamic Access Control," in Information, Communications & Signal Processing, 2007 6th International Conference, Singapore., 2007, pp. 1-5.

[17] A. Kounoudes, V. Kekatos, and S. Mavromoustakos, "Voice Biometric Authentication For Enhancing Internet Service Security," *Information and Communication Technologies*, 2006, Damascus, pp. 1020-1025