Multimodal Biometrik pada Keystroke User-Adaptive Feature dan Mahalanobis Distance

1st Ardityo Cahyo Putro Hutomo
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
ardityoc@students.telkomuniversity.ac.id

2nd Prasti Eko Yunanto
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
gppras@telkomuniversity.ac.id

3rd Febryanti Sthevanie
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
sthevanie@telkomuniversity.ac.id

***Abstrak* — Penelitian ini menganalisis efektivitas kombinasi metode User-Adaptive dan Mahalanobis Distance dalam sistem autentikasi *keystroke biometrics*. Menggunakan Biomey Keystroke Dataset dengan 40 responden, studi ini bertujuan meningkatkan akurasi dan keandalan autentikasi berbasis KD. Sistem yang dikembangkan terdiri dari tahap *enrollment* dan *authentication*, dengan User-Adaptive sebagai metode *feature extraction* dan Mahalanobis Distance untuk *feature matching*. Teknik *decision level fusion* diterapkan untuk mengintegrasikan hasil dari berbagai fitur *keystroke*. Hasil yang diperoleh menunjukkan bahwa teknik *fusion* dengan Mahalanobis Distance menunjukkan hasil yang lebih baik dibandingkan dengan fitur non-*fusion* dengan rata-rata penurunan nilai error sebesar 8,73%. Panjang vektor optimal (Fn) ditemukan pada n = 5 dengan nilai error 12,07%. Pencarian threshold terbaik menghasilkan FAR 15,6% dan FRR 6% pada n sebanyak 5. Hasil yang diperoleh pada penelitian ini menunjukkan nilai *error rate* lebih rendah dengan rata-rata penurunan nilai error sebesar 9,9% dengan penelitian sebelumnya. Penelitian ini membuktikan potensi Mahalanobis Distance dan teknik *fusion* dalam meningkatkan akurasi sistem autentikasi *keystroke biometrics*, membuka peluang pengembangan sistem keamanan yang lebih handal. Studi lebih lanjut disarankan untuk mengeksplorasi *pattern* tertentu pada layar sentuh dan penggunaan dataset yang lebih bervariatif serta menggunakan data *testing real-time*.**

***Kata kunci*— autentikasi, biometrik, digraf, keystroke dynamic, user-adaptive, mahalanobis distance**

I. PENDAHULUAN

Autentikasi sangatlah penting bagi sistem informasi, seiring dengan transformasi digital data, baik itu data sensitif maupun pribadi, membuat perlindungan data pengguna menjadi suatu hal yang sangatlah penting. Saat ini, masih banyak sekali sistem yang menggunakan keamanan dengan mekanisme tradisional berbasis pengetahuan, seperti PIN, Kata Sandi maupun berbasis Token [1][2]. Kekurangan dalam metode tersebut adalah pengguna cenderung memilih kata sandi yang pendek dan mudah diingat, menjadikan kata sandi tersebut mudah ditebak, selain itu kecenderungan penggunaan kata sandi yang sama untuk beberapa akun, serangan *bruteforce*, serangan *keylogger*, dan lain sebagainya [3][4][5]. Pendekatan alternatif yang bisa digunakan yaitu Biometrik, media autentikasi yang menggunakan ciri fisik dan perilaku dari pengguna, sehingga pengguna tidak perlu mengingat ataupun membawa media autentikasi [1]. Ciri fisik memiliki kelemahan terhadap serangan *spoofing* atau kebutuhan perangkat yang mahal [6], sementara ciri perilaku memiliki kekurangan karena sifat dinamis pada adanya adaptasi dan ketidakstabilan dari beberapa faktor temporer sesuai dengan kondisi emosional pengguna, sehingga berdampak pada rendahnya hasil autentikasi [7]. Meski begitu ciri perilaku berpotensi menjadi suatu kelebihan atau kekuatan bahwa ciri biometrik ini susah untuk ditiru karena karena sifatnya yang dinamis. Hal ini akan tercapai apabila hasil ciri yang diekstraksi merupakan ciri yang dinamis, bukan tentang apa yang pengguna ketik, melainkan bagaimana pengguna mengetik [8]. *Keystroke dynamic* (KD) merupakan salah satu metode autentikasi pengguna berbasis perilaku seseorang pada saat pengetikan suatu teks tertentu [7]. Secara umum, ciri perilaku dapat dikenali dari ritme jari saat menekan dan melepas tombol pada perangkat mobile atau komputer. KD lebih hemat biaya, nyaman digunakan, dan tidak memerlukan perangkat tambahan karena sudah tersedia di perangkat tersebut. [7][9]. Penelitian terkait KD terdapat pada [10], metode User-Adaptive diperkenalkan sebagai metode ekstraksi fitur dalam sistem KD yang memanfaatkan kecepatan dan ciri perilaku mengetik pengguna (digraf). Penelitian tersebut menunjukkan bahwa metode User-Adaptive dapat meningkatkan akurasi sebesar 44% lebih baik dari penelitian sebelumnya.

Autentikasi dilakukan dengan membandingkan antara fitur hasil ciri ekstraksi yang diperoleh dengan fitur ciri ekstraksi yang terdaftar dalam sistem. Perbandingan tersebut akan diperoleh nilai jarak yang mana penghitungan jarak menggunakan *distance-based*, penerapan tersebut telah dilakukan pada sejumlah penelitian sebelumnya. Salah satu *distance-based* yang digunakan terdapat pada penelitian [11], Mahalanobis Distance digunakan untuk *generate threshold* yang nantinya diolah untuk penerapan empat klasifikasi yang berbeda diantaranya Random Forest, Isolation Forest, Gradiant Boosting, dan SVM. Akurasi terbaik yang didapat

TABEL 1.
Skenario *Fusion* Multimodal Biometrik [14]

Skenario *Fusion*	Keterangan
Satu jenis Biometrik, Banyak sensor	Menggunakan pengambilan data berbagai sensor dari biometrik yang sama, contohnya autentikasi wajah dengan menggunakan sensor 2D dan sensor 3D.
Satu jenis biometrik, Banyak model klasifikasi	Menggunakan satu sensor dan diolah menggunakan berbagai model klasifikasi, contohnya menggunakan ketiga model klasifikasi fingerprint yang berbeda dalam melakukan *matching score*.
Satu jenis biometrik, Banyak fitur	Menggunakan satu sensor dan dilakukannya ekstraksi fitur agar mendapatkan beberapa fitur untuk diolah, contohnya mendapatkan ciri fisiologi wajah dari segala arah.
Lebih dari satu jenis Biometrik dan Sensor yang digunakan	Menggunakan berbagai sensor sesuai dengan ciri biometrik yang digunakan, contohnya menggunakan ciri fisiologi wajah dan suara pengguna untuk autentikasi.

adalah 90% dari model klasifikasi Isolated Forest. Penerapan autentikasi dengan enam metode *distance-based* yang berbeda telah dilakukan pada penelitian [12], beberapa diantaranya ITAD, KDE, Manhattan Distance, Scaled Manhattan Distance, Mahalanobis Distance dan Transformed Mahalanobis Distance dengan menggunakan Clarskson II Dataset. Penelitian tersebut menerapkan teknik *fusion* pada fitur waktu dan diperoleh Equal Error Rate (EER) terbaik sebesar 12,3% pada metode ITAD, dan Mahalanobis Distance memperoleh Equal Error Rate (EER) sebesar 22%.

Tujuan Penelitian ini adalah mengeksplorasi lebih lanjut mengenai Mahalanobis Distance, tidak hanya untuk mendapatkan threshold tetapi juga untuk autentikasi, selain itu dengan menggunakan metode User-Adaptive sebagai *feature extraction* dapat membuka potensi bagi Mahalanobis Distance sebagai *feature matching* untuk mendapatkan hasil yang lebih baik dengan menggunakan Biomey Biometrics Dataset.

II. KAJIAN TEORI

A. Biometrik

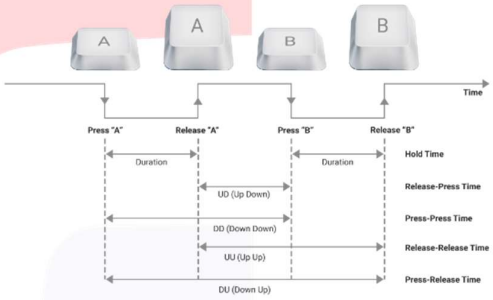
Biometrik menawarkan solusi alami dalam autentikasi berdasarkan ciri fisik atau perilaku yang dimiliki oleh pengguna [1][10]. Keuntungan penggunaan biometrik adalah pengguna tidak perlu membawa atau mengingat, mengingat bahwa biometric sudah terdapat pada tiap individu [10][13]. Biometrik dibagi menjadi 2 (dua) bagian, yakni berdasarkan ciri fisik (fisiologis) yang diperoleh dari sidik jari, wajah, iris, geometri tangan, dan lainnya; dan ciri perilaku (*behavioral*) salah satunya adalah *keystroke dynamics* [8].

B. Keystroke Dynamics

Keystroke dynamics merupakan salah satu bentuk autentikasi berbasis biometrik ciri perilaku berdasarkan pada ritme pengetikan pengguna [7]. Dikarenakan sifatnya yang dinamis, ekstraksi fitur pada biometrik ini menggunakan n-graf yang bertujuan untuk mengukur kecepatan waktu antar penekanan tombol. Pada studi sebelumnya, umumnya n-graf yang digunakan adalah monograf dan digraf [3]. Monograf hanya memiliki 2 fitur waktu, yakni *press* dan *release*, sedangkan fitur waktu digraf terdiri dari 5 (lima) fitur waktu berdasarkan Basic Keystroke Time beberapa diantaranya DD, UD, UU, DU, dan Duration [10]. Kelima fitur ini didapatkan dari Gambar 1 dan (1), (2), (3), (4), dan (5).

TABEL 2.
Perbedaan monograf dengan digraf

Input	"Lorem Ipsum"
Monograf	[l], [o], [r], [e], [m], [SPACE], [i], [p], [s], [u], [m]
Digraf	[l,o], [o,r], [r,e], [e,m], [m,SPACE], [SPACE, i], [i,p], [p,s], [s,u], [u,m]



GAMBAR 1.
Basic Keystroke Time

Istilah D dan U diangkat dari kata *Down* yang mewakili kata *press time*/ waktu saat menekan tombol dan *Up* yang mewakili kata *release time*/ waktu saat melepas tombol.

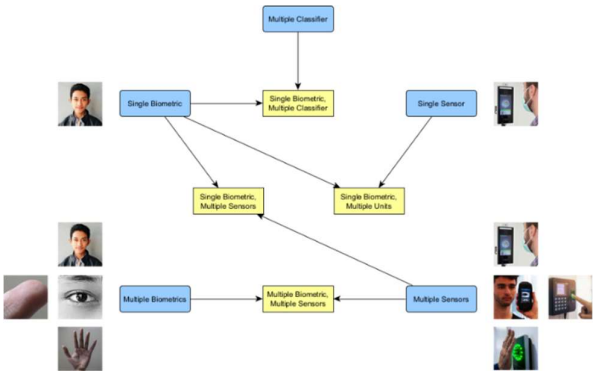
$$DD = press_{time\ i+1} - press_{time\ i} \quad (1)$$
$$UD = press\_time_{i+1} - release\_time_i \quad (2)$$
$$UU = release\_time_{i+1} - release\_time_i \quad (3)$$
$$DU = release\_time_{i+1} - press\_time_i \quad (4)$$
$$Duration = DU_i - UD_i \quad (5)$$

Kelebihan biometrik adalah pengguna hanya perlu perangkat input dasar seperti keyboard, yang sudah ada di setiap perangkat mobile atau komputer. Selain itu, pengguna tidak perlu mengingat atau membawa sesuatu untuk autentikasi. [7][9].

C. Multimodal Biometrik

Penggunaan ciri biometrik yang bergantung pada satu sumber dapat menimbulkan berbagai keterbatasan, namun dapat diatasi dengan penerapan multimodal biometrik dengan menggabungkan berbagai ciri fitur untuk autentikasi untuk menutupi kekurangan saat autentikasi.

Multimodal Biometrik memiliki beberapa skenario *fusion*, diantaranya terdapat pada Tabel 1 [6][14].



GAMBAR 2.
Skenario *Fusion* Multimodal Biometrik

D. User-Adaptive Feature Extraction

Pada penelitian [10], telah memperkenalkan User-Adaptive, sebuah metode ekstraksi fitur keystroke dynamic yang memanfaatkan kecepatan dan perilaku pengetikan khas seseorang dengan mengelompokkan fitur waktu digraf pengguna sebanyak vektor n dimensi. Kecepatan pengetikan tersebut diolah dengan mengurutkan waktu dan melakukan rerata pada tiap waktu digraf yang duplikat, yang nantinya dibentuk berdasarkan urutan *ascending* dan dikelompokkan menjadi vektor n dimensi. Proses tersebut dihitung sebanyak keseluruhan Fitur Basic Keystroke Time.

Start	End	Time
l	o	121
o	r	177
r	e	91
e	m	64
m	_	88
_	i	113
i	p	105
p	s	100
s	u	40
u	m	197
m	_	286
_	d	222
d	o	77
o	l	175
l	o	184
o	r	250
r	_	148
_	s	25
s	i	200
i	t	108
t	_	226
_	a	293
a	m	144
m	e	106
e	t	136

Start	End	Time
_	s	25
s	u	40
e	m	64
d	o	77
r	e	91
p	s	100
i	p	105
m	e	106
i	t	108
_	i	113
e	t	136
a	m	144
r	_	148
l	o	152,5
o	l	175
m	_	187
u	m	197
s	i	200
o	r	213,5
_	d	222
t	_	226
_	a	293

GAMBAR 3.
Pengurutan Digraf pada kalimat "Lorem ipsum dolor sit amet"

F1	F2	F3	F4	F5	F6	F7	F8
_,s	d,o	i,p	_,i	r,_	m,_	o,r	t,_
s,u	r,e	m,e	e,t	l,o	u,m	_,d	_,a
e,m	p,s	i,t	a,m	o,l	s,i		

F1	F2	F3	F4	F5	F6	F7	F8
25	77	105	113	148	187	213,5	226
40	91	106	136	152,5	197	222	293
64	100	108	144	175	200		

F1	F2	F3	F4	F5	F6	F7	F8
43	89,33	106,3	131	158,5	194,7	217,8	259,5

GAMBAR 4.
Penerapan Metode User-Adaptive pada kalimat "Lorem ipsum dolor sit amet"

E. Mahalanobis Distance

Mahalanobis Distance biasa digunakan untuk melakukan *feature matching* dalam *keystroke biometrics* [15]. Tidak seperti metode *distance* lain yang menganggap fitur sebagai independen, metode ini mempertimbangkan korelasi antar fitur sehingga sesuai untuk data yang bersifat *multivariate*. Metode ini dapat menemukan *outlier multivariate* yang mana dapat membedakan sifat unik pengetikan tiap pengguna pada *keystroke biometrics* [16].

$$D_{Mahalanobis}(\vec{x},\vec{y}) = \sqrt{(\vec{x}-\vec{y})^T S^{-1}(\vec{x}-\vec{y})} \quad (6)$$

Persamaan Mahalanobis terdapat pada (6), di mana x merupakan vektor fitur dari *testing*, y merupakan fitur *training* yang telah dilakukan rata-rata, dan S−1 merupakan matriks inverse covariace dari fitur *training* tersebut.

F. Biometrics Error Rate

Performansi yang paling umun digunakan dalam biometrik yakni FAR, FRR, dan EER. FAR (False Acceptence Rate) merupakan persentase dimana sistem salah mengenali pengguna yang tidak sah sebagai pengguna yang sah [1][14].

$$FAR = \frac{NFA}{NIRA} \times 100\% \quad (7)$$

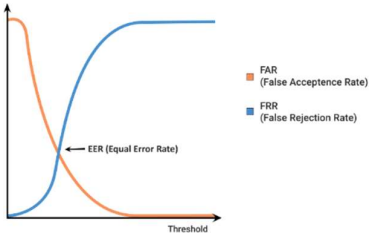
Dimana NFA (*Number of False Acceptence*) mewakili banyaknya percobaan pengguna tidak sah yang diterima, sedangkan NIRA (*Number of Imposter Recognition Attempts*) mewakili total percobaan pengguna tidak sah.

FRR (False Reject Rate) merupakan persentase dimana sistem salah mengenali pengguna yang sah sebagai pengguna tidak sah.

$$FRR = \frac{NFR}{NGRA} \times 100\% \quad (8)$$

Dimana NFR (*Number of False Rejected*) mewakili banyaknya percobaan pengguna sah yang ditolak, sedangkan NGRA (*Number of Genuine Recognition Attempts*) mewakili total percobaan pengguna sah.

EER (Equal Error Rate) merupakan titik perpotongan dimana FAR sama dengan FRR, semakin rendah nilai EER yang didapat menunjukkan kinerja yang lebih baik. Jika nilai FAR tinggi menunjukkan sistem yang kurang aman (pengguna tidak sah lebih mudah untuk memperoleh akses), sedangkan nilai FRR tinggi menunjukkan sistem terlalu ketat (pengguna yang sah sulit untuk memperoleh akses). Maka dari itu EER menyeimbangkan kedua titik ini [16].



GAMBAR 5.
Kurva Relasi antara FAR, FRR, dan EER

TABEL 3.
Penerapan User-Adaptive pada tiap fitur dalam bentuk key (DD)

F1	F2	F3	F4	F5	F6	F7	F8
(g,g)	(t,i)	(n,g)	(i,n)	(a,r)	(m,a)	(a,n)	(SPACE,j)
(u,n)	(n.t)	(g,u)	(g,SPACE)	(j,a)	(r,i)		

TABEL 4.
Hasil Ekstraksi Fitur yang digunakan untuk learning algoritma autentikasi (DD)

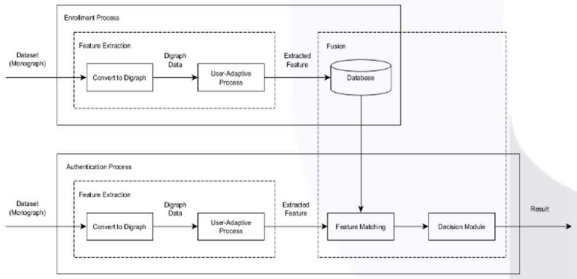
F1	F2	F3	F4	F5	F6	F7	F8
315,5	474,5	640	773,25	983	1257	2240	2295

TABEL 5.
Hasil Ekstraksi Fitur dengan User-Adaptive tiap Fitur waktu digraf (Basic Keystroke Time)

Features	F1	F2	F3	F4	F5	F6	F7	F8
DD	315,5	474,5	640	773,25	983	1257	2240	2295
UD	211,5	379	547,83	689,5	878,5	1187	2129	2231
UU	295	463,5	646,33	770	979,5	1295	2251	2317
DU	383	575	738,5	853,75	1080	1369	2362	2381
Duration	148	156,5	175,25	185,17	193	205,5	233	235

III. METODE

Proses *enrollment* dan *authentication* merupakan 2 (dua) tahap utama yang dilakukan pada penelitian ini. Proses *enrollment* dilakukan untuk menyimpan pola *keystroke* yang telah melalui tahap ekstraksi fitur dan autentikasi merupakan tahap akhir yang bertujuan untuk menentukan (atau memverifikasi) klaim pengguna adalah sah atau tidak. Gambaran umum terdapat pada Gambar 6.



GAMBAR 6.
Sistem Autentikasi dengan Fitur User-Adaptive dan Mahalanobis Distance

A. Dataset

Pada penelitian ini digunakannya Biomey Biometric Dataset sebagai sumber data valid. Dengan melibatkan pengguna sebanyak 40 Responden, Biomey Biometric Dataset dibuat dengan menggunakan aplikasi berbasis android berdasarkan pada teks acak dari *corpus* yang telah ditentukan (enam kata terdiri dari *alphabetic* dan spasi), dimana data tersebut berbentuk monograf terdiri dari 30 sesi dan 10 kalimat unik di tiap responden [17].

B. Enrollment

Pada tahap *enrollment*, dilakukannnya *feature extraction* yang terdiri dari beberapa tahapan, yakni Pembentukan fitur untuk mengubah Biomey Biometric Dataset yang masih berupa monograf menjadi digraf, dan mengimplementasikan metode User-Adaptive untuk memperoleh fitur *keystroke* dengan jumlah yang lebih sedikit. Proses ekstraksi ciri ini akan menghasilkan fitur yang lebih kaya dan merepresentasikan ciri dari *keystroke*.

Pada metode User-Adaptive, implementasi diawali dengan pembentukan fitur waktu digraf dengan contoh kalimat "manggung jarinting".

Adapun tahap selanjutnya yakni melakukan penghitungan rerata tiap digraf duplikat pada *session* (SID) tersebut, dilanjutkan dengan pengurutan secara *ascending* berdasarkan time (DD), sehingga 10 (sepuluh) kalimat yang ada pada *session* tersebut akan menjadi suatu kesatuan (fitur user-adaptive). Blok warna pada Tabel 6 merupakan 2 (dua) contoh kecil mewakili digraf duplikat yang telah dirata-ratakan pada *session* tersebut. Warna ungu mewakili digraf (n, g) dan warna biru mewakili (i, n).

TABEL 6.
Digraf DD yang telah diurutkan pada kalimat "manggung jarinting".

index	uid	sid	start	end	DD
0	94022	1	g	g	182
1	94022	1	u	n	449
2	94022	1	t	i	451
3	94022	1	n	t	498
4	94022	1	n	g	625
5	94022	1	g	u	655
6	94022	1	i	n	728,5
7	94022	1	g	SPACE	818
8	94022	1	a	r	944
9	94022	1	j	a	1022
10	94022	1	m	a	1026
11	94022	1	r	i	1488
12	94022	1	a	n	2240
13	94022	1	SPACE	j	2295

TABEL 7.
Teknik Pembentukan Fitur *Fusion*

UID	Fitur					Hasil Prediksi (*Fusion*)
	DD	UD	UU	DU	Duration	
94022	Sah	Sah	Tidak Sah	Sah	Tidak Sah	Sah
93349	Tidak Sah	Sah	Sah	Tidak Sah	Sah	Sah
93749	Tidak Sah	Tidak Sah	Tidak Sah	Sah	Sah	Tidak Sah
98439	Sah	Tidak Sah	Sah	Tidak Sah	Tidak Sah	Tidak Sah
93272	Sah	Sah	Sah	Sah	Tidak Sah	Sah

TABEL 8.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (DD & UD) dari Mahalanobis Distance

No	UID	DD				UD			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	4,8	0,1846	0,2	0,1923	4,3	0,2744	0,28	0,2772
2	94022	4,4	0,3128	0,36	0,3364	4,1	0,3487	0,36	0,3544
3	80154	5,2	0,1359	0,12	0,1279	5,6	0,118	0,12	0,119
4	84101	5,2	0,1949	0,16	0,1774	5,4	0,1897	0,2	0,1949
…	…	…	…	…	…	…	…	…	…
37	24376	4,2	0,3462	0,32	0,3331	4	0,2385	0,24	0,2392
38	13036	5,4	0,1615	0,16	0,1608	4,9	0,1564	0,16	0,1582
39	64089	4,6	0,3026	0,32	0,3113	4,3	0,2385	0,24	0,2392
40	21016	4,1	0,2436	0,2	0,2218	4,1	0,4487	0,48	0,4644

C. Authentication

Autentikasi memiliki beberapa tahap yang seragam dengan *enrollment*, namun yang membedakan adalah tujuannya. Berbeda dengan *enrollment*, *authentication* memiliki tahap *fusion* yang terdiri dari *feature matching* dan *decision making*. Mahalanobis Distance digunakan dalam *feature matching* dan Majority Voting (MV) digunakan dalam Decision Making. Hasil dari prediksi *voting* kelima fitur waktu tersebut digunakan sebagai acuan fitur *fusion*.

IV. HASIL DAN PEMBAHASAN

Terdapat 3 (tiga) skenario pengujian yang dilakukan pada penelitian ini, diantaranya Skenario Pertama yakni pengujian performansi pada Basic Keystroke Time dan *fusion*, dengan membandingkan nilai *error rate* antara unimodal dengan multimodal dan membandingkan hasil performansi antara Mahalanobis Distance dan Euclidean Distance. Skenario Kedua, Pengujian performansi pada pengaruh banyak Fn pada User-Adaptive dengan membandingkan nilai *error rate* dengan beragam nilai panjang vektor n pada 3, 5, 7, dan 9. Skenario Ketiga, Pengujian performansi pada Basic Keystroke Time dan *fusion* dari *threshold* rata-rata terbaik yang telah didapat. Dengan membandingkan nilai *error rate* antara Panjang vektor n pada 5 dan 8.

Pada skenario pertama dan kedua, pengujian dilakukan pada fitur waktu dari *keystroke* menggunakan 25 dari 30 data *session* responden, dengan sisa data selisihnya sebanyak 5 digunakan untuk pengujian pada skenario ketiga. *feature extraction* menggunakan User-Adaptive, dengan n = 8 untuk skenario pertama dan percobaan dengan n = 3, 5, 7, dan 9 pada skenario kedua. Cross-validation dilakukan dengan 5 *fold*, membagi data menjadi 20 data *training* dan 5 data *testing*, dengan penambahan 2 fitur dari pengguna lain per fold, sehingga setiap fold memiliki komposisi 20 data *training* dan 83 data *testing*. Pada skenario pertama, feature matching menggunakan Mahalanobis Distance dan Euclidean Distance, dengan *threshold* dan EER diperoleh dari rata-rata FAR dan FRR kelima fold. Skenario kedua hanya menggunakan Mahalanobis Distance untuk feature matching, dan *threshold* dari kedua skenario yang didapat digunakan pada skenario ketiga. Skenario ketiga menggunakan 25 data training dan 200 data testing dari penggabungan 5 data *testing* pengguna dari skenario pertama. Pengujian pada skenario ini menggunakan panjang vektor optimal 5 dari skenario kedua, menghasilkan rata-rata FAR dan FRR per pengguna.

A. Skenario 1

Hasil akhir pengujian skenario ini diperoleh dari rerata EER dari seluruh pengguna (UID). Pengujian dilakukan untuk mengetahui dan membandingkan nilai *error rate* antara non-*fusion* (unimodal) dengan *fusion* (multimodal) serta membandingkan nilai *error rate* Mahalanobis Distance dengan Euclidean Distance yang mana *distance* tersebut umum digunakan terkait *keystroke dynamics*.

Berdasarkan hasil yang diperoleh, nilai terbaik didapatkan pada *feature matching* menggunakan Mahalanobis Distance bahwa nilai performansi menggunakan teknik *fusion* jauh lebih baik dibandingkan dengan non-*fusion*. Sedangkan pada *feature matching* dengan menggunakan Euclidean Distance, nilai terendah didapat dari fitur non-*fusion* yakni Duration. Tabel 8-14

merupakan hasil performansi dengan *feature matching* menggunakan Mahalanobis Distance dan Euclidean Distance sebagai pembanding sehingga *thresholds* dan nilai EER didapatkan dengan merata-ratakan FAR dan FRR dari kelima *fold* untuk tiap pengguna.

TABEL 9.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (UU & DU) dari Mahalanobis Distance

No.	UID	UU				DU			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	4,7	0,2103	0,2	0,2051	4,4	0,1128	0,12	0,1164
2	94022	4,4	0,3718	0,36	0,3659	4,1	0,3103	0,32	0,3151
3	80154	5,7	0,2	0,2	0,2	5,3	0,1821	0,2	0,1910
4	84101	4,7	0,2436	0,24	0,2418	5	0,2128	0,24	0,2264
…	…	…	…	…	…	…	…	…	…
37	24376	4,7	0,2923	0,28	0,2862	4,5	0,3103	0,32	0,3151
38	13036	4,6	0,1538	0,16	0,1569	4,8	0,1641	0,16	0,1621
39	64089	4,5	0,2821	0,28	0,281	4,3	0,1667	0,16	0,1633
40	21016	4,2	0,2538	0,24	0,2469	5,1	0,2385	0,24	0,2392

TABEL 10.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (*Duration* & *Fusion*) dari Mahalanobis Distance

No.	UID	Duration				Fusion			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	4,7	0,1564	0,16	0,1582	4,4	0,0923	0,08	0,0862
2	94022	4,3	0,2846	0,32	0,3023	4,1	0,2103	0,2	0,2051
3	80154	4,6	0,1769	0,2	0,1885	5,1	0,0821	0,12	0,1010
4	84101	5,1	0,2026	0,2	0,2013	5,1	0,1282	0,12	0,1241
…	…	…	…	…	…	…	…	…	…
37	24376	4	0,3385	0,36	0,3492	4,2	0,2308	0,24	0,2354
38	13036	4,2	0,2077	0,24	0,2238	4,5	0,0846	0,08	0,0823
39	64089	4,3	0,2359	0,24	0,2379	4,5	0,1846	0,2	0,1923
40	21016	4,9	0,4462	0,44	0,4431	4,3	0,2308	0,24	0,2354

TABEL 11.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (DD & UD) dari Euclidean Distance

No.	UID	DD				UD			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	3,8	0,2769	0,28	0,2785	4,3	0,3718	0,36	0,3659
2	94022	4,8	0,2154	0,2	0,2077	4,8	0,2897	0,28	0,2849
3	80154	3,8	0,2333	0,2	0,2167	4,8	0,1821	0,16	0,1710
4	84101	5,1	0,2026	0,2	0,2013	5,1	0,1282	0,12	0,1241
…	…	…	…	…	…	…	…	…	…
37	24376	2,9	0,4487	0,48	0,4643	2,9	0,4487	0,48	0,4643
38	13036	3,9	0,2462	0,2	0,2231	4,2	0,2718	0,28	0,2759
39	64089	5,4	0,5308	0,52	0,5254	5,9	0,5308	0,52	0,5254
40	21016	3,2	0,3359	0,28	0,3079	3,4	0,3769	0,4	0,3885

TABEL 12.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (UU & DU) dari Euclidean Distance

No.	UID	UU				DU			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	3,8	0,2872	0,32	0,3036	3,9	0,1949	0,2	0,1974
2	94022	4,3	0,2154	0,2	0,2077	4,3	0,1718	0,16	0,1659
3	80154	3,3	0,2615	0,28	0,2708	3,6	0,2590	0,24	0,2495
4	84101	3,9	0,4436	0,4	0,4218	3,6	0,3513	0,36	0,3556
…	…	…	…	…	…	…	…	…	…
37	24376	3,3	0,4103	0,44	0,4251	3,2	0,4077	0,4	0,4038
38	13036	3,6	0,2872	0,24	0,2636	3,9	0,2641	0,28	0,2721
39	64089	4	0,4821	0,48	0,4810	5,8	0,5359	0,52	0,5279
40	21016	4	0,4128	0,4	0,4064	4,2	0,4154	0,36	0,3877

TABEL 13.
Hasil Rerata FAR & FRR dari 5 Hasil k-Fold dan EER (*Duration* & *Fusion*) dari Euclidean

No.	UID	Duration				Fusion			
		Thresholds	Avg. FAR	Avg. FRR	EER	Thresholds	Avg. FAR	Avg. FRR	EER
1	94026	4,3	0,1154	0,12	0,1177	3,8	0,1795	0,2	0,1897
2	94022	4	0,2205	0,2	0,2103	4,4	0,1795	0,16	0,1697
3	80154	4,7	0,2	0,2	0,2	3,9	0,1462	0,12	0,1331
4	84101	5,9	0,2410	0,2	0,2205	4	0,3179	0,36	0,3390
…	…	…	…	…	…	…	…	…	…
37	24376	3,6	0,3282	0,32	0,3241	3,3	0,3718	0,4	0,3859
38	13036	4	0,2051	0,2	0,2026	4	0,2385	0,2	0,2192
39	64089	4,5	0,2846	0,28	0,2823	5,3	0,4256	0,44	0,4328
40	21016	4,7	0,2538	0,28	0,2669	3,9	0,3256	0,32	0,3228

TABEL 14.
Hasil Rerata EER dari Keseluruhan Pengguna dengan panjang vektor n = 8 menggunakan Mahalanobis Distance dan Euclidean Distance

Distance Method	Features					
	DD	UD	UU	DU	Duration	Fusion
Mahalanobis Distance	0,197891026	0,214044872	0,202762821	0,188429487	0,254711538	0,124307692
Euclidean Distance	0,379589744	0,379935897	0,374628205	0,363051282	0,241365385	0,30475641

B. Skenario 2

Hasil akhir pengujian skenario ini diperoleh dari rerata EER dari seluruh pengguna (UID). Pengujian ini dilakukan untuk mengetahui nilai n optiman dengan membandingkan nilai *error rate* nilai panjang vektor n pada 3, 5, 7, dan 9. Berdasarkan hasil yang diperoleh, menunjukkan bahwa jumlah kelompok User-Adaptive yang digunakan sangat mempengaruhi kinerja sistem autentikasi. Terlihat pada Tabel 16 diperoleh bahwa panjang vektor optimal bernilai 5 dengan nilai error 12,07%, di mana pertambahan panjang vektor tidak berbanding lurus dengan performansi sistem.

TABEL 15.
Rerata *Threshold* dari Keseluruhan Pengguna

n	Avg. Thresholds (40 UID)					
	DD	UD	UU	DU	Duration	Fusion
3	2,4075	2,285	2,3275	2,3675	2,5275	2,395
5	3,4675	3,3875	3,4025	3,5075	3,2675	3,345
7	4,45	4,3825	4,365	4,485	4,1175	4,2825
8	5,0225	4,86	4,845	4,9425	4,58	4,735
9	5,4875	5,51	5,4125	5,55	5,0425	5,305

TABEL 16.
Hasil Rerata EER dari Keseluruhan Pengguna (Mahalanobis Distance)

n	Avg. EER (40 UID)					
	DD	**UD**	**UU**	**DU**	**Duration**	**Fusion**
3	0,202878205	0,212814103	0,20149359	0,20124359	0,189820513	0,142769231
5	0,182923077	0,19774359	0,19249359	0,18775	0,211205128	0,120711538
7	0,191	0,208910256	0,19400641	0,195929487	0,238705128	0,12174359
8	0,197891026	0,214044872	0,202762821	0,188429487	0,254711538	0,124307692
9	0,211480769	0,226429487	0,209134615	0,204160256	0,255967949	0,13974359

C. Skenario 3

Hasil pengujian skenario ini didapatkan dari rata-rata FAR dan FRR tiap pengguna. Pengujian ini dilakukan untuk mengetahui performasi dari pengaruh *threshold* tiap pengguna yang telah diperoleh sebelumnya dan membandingkan pengelompokkan fitur User-Adaptive n sebesar 5 dan 8. Berdasarkan hasil yang diperoleh, menunjukkan bahwa nilai performansi n = 5 lebih baik signifikan dari n = 8 dengan nilai FAR terendah sebesar 15,6% dan FRR sebesar 6%. Namun, hasil FRR terendah dengan nilai *error rate* yang lebih stabil tiap fitur waktu diperoleh pada n = 8 dengan nilai FRR sebesar 4% dengan FAR sebesar 21,1%.

TABEL 17.
Hasil FAR & FRR dengan *Threshold* Terbaik tiap Pengguna (DD, UD & UU) pada n = 5

No.	**UID**	**DD**			**UD**			**UU**		
		Thresholds	**FAR**	**FRR**	**Thresholds**	**FAR**	**FRR**	**Thresholds**	**FAR**	**FRR**
1	94026	3,4	0,174358974	0,2	3,1	0,215384615	0,2	3,2	0,164102564	0
2	94022	3,2	0,276923077	0,2	3,1	0,374358974	0,2	3,4	0,348717949	0,4
3	80154	4,8	0,271794872	0	4,1	0,148717949	0	4,3	0,148717949	0,2
4	84101	4	0,215384615	0	3,7	0,123076923	0	3,1	0,179487179	0,4
…	…	…	…	…	…	…	…	…	…	…
37	24376	3,1	0,246153846	0,4	2,8	0,302564103	0,4	3	0,276923077	0,6
38	13036	3,5	0,205128205	0	3,6	0,18974359	0	3,9	0,225641026	0
39	64089	2,8	0,184615385	0,2	2,8	0,328205128	0,2	2,8	0,353846154	0
40	21016	3	0,261538462	0,2	2,8	0,41025641	0,2	3,2	0,276923077	0,4

TABEL 18.
Hasil FAR & FRR dengan *Threshold* Terbaik tiap Pengguna (DU, *Duration* & *Fusion*) pada n = 5

No.	**UID**	**DU**			**Duration**			**Fusion**		
		Thresholds	**FAR**	**FRR**	**Thresholds**	**FAR**	**FRR**	**Thresholds**	**FAR**	**FRR**
1	94026	3,5	0,123076923	0,4	3,6	0,138461538	0,2	2,9	0,133333333	0,2
2	94022	3,2	0,194871795	0,8	3	0,230769231	0	3,1	0,225641026	0
3	80154	3,6	0,158974359	0	3,3	0,21025641	0	3,7	0,133333333	0
4	84101	3,4	0,215384615	0,2	3,9	0,179487179	0,4	3,7	0,117948718	0
…	…	…	…	…	…	…	…	…	…	…
37	24376	3,1	0,317948718	0,2	3,1	0,364102564	0,2	3	0,271794872	0,4
38	13036	4,2	0,241025641	0,2	3,1	0,230769231	0,2	3,6	0,2	0
39	64089	3,3	0,230769231	0,4	3,6	0,169230769	0	3	0,179487179	0
40	21016	3,2	0,215384615	0,4	3,3	0,374358974	0,2	3	0,251282051	0,2

TABEL 19.
Hasil FAR & FRR dengan *Threshold* Terbaik tiap Pengguna (DD, UD & UU) pada n = 8

No.	UID	DD			UD			UU		
		Thresholds	FAR	FRR	Thresholds	FAR	FRR	Thresholds	FAR	FRR
1	94026	4,8	0,241025641	0,2	4,3	0,405128205	0	4,7	0,292307692	0
2	94022	4,4	0,394871795	0,2	4,1	0,497435897	0,2	4,4	0,523076923	0
3	80154	5,2	0,138461538	0	5,6	0,138461538	0	5,7	0,230769231	0
4	84101	5,2	0,241025641	0,2	5,4	0,230769231	0	4,7	0,271794872	0
…	…	…	…	…	…	…	…	…	…	…
37	24376	4,2	0,333333333	0,2	4	0,246153846	0,8	4,7	0,333333333	0,2
38	13036	5,4	0,251282051	0,2	4,9	0,21025641	0	4,6	0,194871795	0
39	64089	4,6	0,379487179	0	4,3	0,297435897	0	4,5	0,374358974	0,2
40	21016	4,1	0,297435897	0,2	4,1	0,543589744	0,2	4,2	0,374358974	0,2

TABEL 20.
Hasil FAR & FRR dengan *Threshold* Terbaik tiap Pengguna (DU, *Duration* & *Fusion*) pada n = 8

No.	UID	DU			Duration			Fusion		
		Thresholds	FAR	FRR	Thresholds	FAR	FRR	Thresholds	FAR	FRR
1	94026	4,4	0,143589744	0,4	4,7	0,205128205	0	4,4	0,164102564	0
2	94022	4,1	0,451282051	0,4	4,3	0,38974359	0,2	4,1	0,41025641	0
3	80154	5,3	0,271794872	0	4,6	0,251282051	0	5,1	0,148717949	0
4	84101	5	0,287179487	0	5,1	0,317948718	0	5,1	0,18974359	0
…	…	…	…	…	…	…	…	…	…	…
37	24376	4,5	0,374358974	0,4	4	0,405128205	0,4	4,2	0,312820513	0,2
38	13036	4,8	0,220512821	0,4	4,2	0,266666667	0	4,5	0,18974359	0
39	64089	4,3	0,179487179	0,2	4,3	0,261538462	0,2	4,5	0,246153846	0
40	21016	5,1	0,374358974	0	4,9	0,58974359	0	4,3	0,394871795	0

TABEL 21.
Hasil Pengujian dengan menggunakan rata-rata *Threshold* terbaik yang telah diperoleh pada n 5 dan 8.

Features	*n* = 5		*n* = 8	
	Avg. FAR	Avg. FRR	Avg. FAR	Avg. FRR
DD	0,212307692	0,135	0,259358974	0,125
UD	0,223717949	0,14	0,262051282	0,125
UU	0,208205128	0,17	0,250769231	0,12
DU	0,19974359	0,18	0,240769231	0,12
Duration	0,225512821	0,16	0,31474359	0,135
Fusion	0,156410256	0,06	0,211025641	0,04

V. KESIMPULAN

Pada penelitian ini, metode User-Adaptive digunakan pada *keystroke biometrics* dengan Mahalanobis Distance sebagai *feature matching* dan menerapkan decision level *fusion* sebagai teknik *fusion* yang digunakan. Hasil penelitian menunjukkan bahwa penggunaan teknik fusion dengan Mahalanobis Distance menghasilkan performa yang lebih baik dibandingkan dengan fitur non-fusion, dengan rata-rata penurunan nilai error sebesar 8,73%. Selanjutnya panjang vektor (Fn) yang optimal adalah 5 dengan nilai error 12,07%. Berdasarkan pencarian nilai *threshold* terbaik dan penggunaan teknik *fusion* diperoleh nilai FAR terendah sebesar 15,6% dan FRR sebesar 6%. Jika dibandingkan dengan penelitian sebelumnya, hasil yang diperoleh pada penelitian ini menunjukkan nilai *error rate* lebih rendah dengan rata-rata penurunan nilai error sebesar 9,9%. Secara keseluruhan, hasil penelitian menunjukkan bahwa penggunaan Mahalanobis Distance dan Teknik *fusion* memperlihatkan hasil yang menjanjikan pada sistem autentikasi menggunakan *keystroke biometrics.*

Saran untuk penelitian selanjutnya, perlu dilakukan penelitian lebih lanjut, yang mana pola *keystroke* bisa diperoleh dari *pattern* tertentu pada layar sentuh, dan juga

perlu menambahkan jumlah dataset atau menggunakan dataset lain yang lebih bervariatif atau homogen serta menggunakan data *testing real-time*.

REFERENSI

[1] A. K. Jain and A. Ross, "Introduction to Biometrics," in Handbook of Biometrics, A. K. Jain, P. Flynn, and A. A. Ross, Eds., Boston, MA: Springer US, 2008, pp. 1–22. doi: 10.1007/978-0-387-71041-9_1.

[2] B. Ayotte, J. Huang, M. K. Banavar, D. Hou, and S. Schuckers, "Fast Continuous User Authentication Using Distance Metric Fusion of Free-Text Keystroke Data," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2019, pp. 2380–2388. doi: 10.1109/CVPRW.2019.00292.

[3] Y. M. Alginahi and M. N. Kabir, Eds., Authentication Technologies for Cloud Computing, IoT and Big Data. Institution of Engineering and Technology, 2019. doi: 10.1049/PBSE009E.

[4] H. Kalita, E. Maiorana, and P. Campisi, "Keystroke Dynamics for Biometric Recognition in Handheld Devices," in 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Jul. 2020, pp. 410–416. doi: 10.1109/TSP49548.2020.9163524.

[5] A. Rahman et al., "Multimodal EEG and Keystroke Dynamics Based Biometric System Using Machine Learning Algorithms," IEEE Access, vol. 9, pp. 94625–94643, 2021, doi: 10.1109/ACCESS.2021.3092840.

[6] A. Ross and A. K. Jain, "Multimodal biometrics: An overview," in 2004 12th European Signal Processing Conference, Sep. 2004, pp. 1221–1224. Accessed: Jun. 23, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/7080214.

[7] Y. Zhong and Y. Deng, "A Survey on Keystroke Dynamics Biometrics: Approaches, Advances, and Evaluations," 2015, pp. 1–22. doi: 10.15579/gcsr.vol2.ch1.

[8] Y. Amankar, S. Gangurde, A. Khachane, and P. Y. Itankar, "User Authentication using Keystroke Analysis," vol. 07, no. 04, 2020.

[9] B. Ayotte, M. Banavar, D. Hou, and S. Schuckers, "Fast and Accurate Continuous User Authentication by Fusion of Instance-based, Free-text Keystroke Dynamics," 2019 International Conference of the Biometrics Special Interest Group (BIOSIG), Sep. 2019, Accessed: Apr. 24, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Fast-and-Accurate-Continuous-User-Authentication-by-Ayotte-Banavar/66249aca2610910caa80d71154db9396ac54f771.

[10] J. Kim, H. Kim, and P. Kang, "Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection," Applied Soft Computing, vol. 62, pp. 1077–1087, Jan. 2018, doi: 10.1016/j.asoc.2017.09.045.

[11] I. M, I. Karunanithi, and S. B. U, "Enhancing User Authentication through Keystroke Dynamics Analysis using Isolation Forest algorithm," in 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), Feb. 2024, pp. 1–5. doi: 10.1109/ic-ETITE58242.2024.10493648.

[12] B. Ayotte, M. Banavar, D. Hou, and S. Schuckers, "Fast Free-Text Authentication via Instance-Based Keystroke Dynamics," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 2, no. 4, pp. 377–387, Oct. 2020, doi: 10.1109/TBIOM.2020.3003988.

[13] J. Kim and P. Kang, "Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features," Pattern Recognition, vol. 108, p. 107556, Dec. 2020, doi: 10.1016/j.patcog.2020.107556.

[14] Z. Qin, P. Zhao, T. Zhuang, F. Deng, Y. Ding, and T. Chen, "A survey of identity recognition via data fusion and feature learning - ScienceDirect," Information Fusion, vol. 91, pp. 694–712, Mar. 2023, doi: 10.1016/j.inffus.2022.10.032.

[15] A.-C. Iapa and V.-I. Cretu, "Modified Distance Metric That Generates Better Performance For The Authentication Algorithm Based On Free-Text Keystroke Dynamics," in 2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI), May 2021, pp. 000455–000460. doi: 10.1109/SACI51354.2021.9465601.

[16] R. Shadman, A. A. Wahab, M. Manno, M. Lukaszewski, D. Hou, and F. Hussain, "Keystroke Dynamics: Concepts, Techniques, and Applications," Mar. 08, 2023, arXiv: arXiv:2303.04605. doi: 10.48550/arXiv.2303.04605.

[17] I. R. Arrazaan, Sistem Autentikasi Pengguna Berbasis Keystroke Biometrics Dinamis Menggunakan Metode FACT. Universitas Telkom, S1 Informatika, 2023. Accessed: Aug. 23, 2024. [Online]. Available: https://openlibrary.telkomuniversity.ac.id/pustaka/197049/sistem-autentikasi-pengguna-berbasis-keystroke-biometrics-dinamis-menggunakan-metode-fact.html