Forum Statistika dan Komputasi : *Indonesian Journal of Statistics* available online at:
ISSN : 0853-8115 *journal.ipb.ac.id/index.php/statistika*
Vol. 17 No.1, April 2012, p: 21-30

PENERAPAN RANTAI MARKOV PADA PENGEMBANGAN UJI KETERDUGAAN KUNCI

(*Markov Chain Technique in Key Predictability Test Development*)

Sari Agustini Hafman[1)], Anang Kurnia[2)], Agus Buono[3)]
[1]Lembaga Sandi Negara, Republik Indonesia
[2]Departemen Statistika FMIPA – IPB
[3]Departemen Ilmu Komputer FMIPA – IPB
e-mail : [1]hafman76@yahoo.com

Abstract

One Time Key (OTK) system with key from alphabetical sequences is one of symmetric encryption algorithm that used in Indonesia to protect secret information. Alphabetic sequences in OTK system must be cryptographically secure pseudorandom sequences. OTK system in Indonesia only tested by overlapping m-tuple test developed by Marsaglia (2005). Overlapping m-tuple test doesn't check the unpredictability of alphabetical sequences, it just tests distribution form and indpendency of alphabetical sequences. So, an alphabetical sequence in OTK system cannot be used in cryptography application by the reason of unpredictability sequence is unknown. Because some of Pseudorandom Number Generator (PRNG) algorithm based on block cipher algorithm that has markovian properties, markov chain model used to detect predictability alphabetical sequences. Data in this study consists of two data sources i.e. simulation data that generated from four classes PRNG and OTK system keys in 2005 that used in three communication units of foreign ministry. Simulation data is used to develop key predictability test methodology by find predictability threshold value based on characteristic of match level. OTK system keys will be predictability tested by comparing characteristic of match level with threshold value that is obtained from simulation data. The first result of this study shows the alphabetical sequence generated by first, second and fourth PRNG class can't be modeled with first-order markov chain until third-order. The third PRNG class, except PRNG LCG1, LCG2, coveyou, rand and randu, also can't be modeled with first order markov chain until third-order. Sequence generated by LCG2, coveyou, rand and randu are not fit for use in cryptography because it has a high probability to be modeled by high orders of markov chain (above the order of three). The second result obtains predictability threshold value with markov chains based on the minimum and maximum match level on the second-order and third-order. The last result shows the size of training data must be greater than the size of the observation data with the best ratio between the size of training data with observational data is 100: 10. The results of testing using 10 times repeated shows that the match level average of the OTK system key match on the all of three-order less than 4.5 x 10^{-2}, so the OTK system the is feasible to secure information in three communication units.

Keywords: One Time Key (OTK), markov chain, PRNG, probability transition, match level

PENDAHULUAN

Salah satu dampak negatif perkembangan teknologi informasi adalah timbulnya kerawanan dalam komunikasi seperti pemalsuan, penyadapan, perusakan dan pengubahan informasi. Dalam pengamanan informasi terdapat tiga aspek yang harus diperhatikan yaitu pengamanan fisik, administratif dan *logic*. Penggunaan kriptografi merupakan salah satu upaya pengamanan secara *logic*.

Berdasarkan prinsip Kerckhoffs (1883), keamanan sistem kriptografi harus hanya bergantung pada kunci. Kunci umumnya dihasilkan oleh pembangkit bilangan acak nyata (PBAN) atau pembangkit bilangan acaksemu (PBAS). Output dari PBAN atau PBAS ini berupa barisan kunci berbentuk bit atau diubah menjadi bentuk barisan lain bergantung pada kebutuhan sistem kriptografi seperti barisan digit (0-9), barisan bilangan heksadesimal (0-F), barisan karakter (0-255) dan barisan abjad (A-Z).

Terdapat tiga tipe barisan yang dihasilkan oleh PBAN dan PBAS yaitu *pseudo-random sequences* (barisan acaksemu), *cryptographically secure pseudo-random sequences* (barisan acaksemu yang aman secara kriptografis) dan *real random sequences* (barisan yang acak nyata). Barisan dikatakan acaksemu jika secara statistik terlihat acak (berdistribusi seragam dan saling bebas). Barisan dikatakan aman secara kriptografis bila barisan tersebut secara statistik terlihat acak serta *unpredictable* (ketidakterdugaan). Barisan dikatakan acak nyata bila memenuhi tiga syarat yaitu barisan tersebut secara statistik terlihat acak, ketidakterdugaan dan barisan yang sama tidak dapat dihasilkan kembali (Schneier 1996). Hanya barisan acaksemu yang aman secara kriptografis dan barisan acak nyata yang dapat digunakan dalam sistem kriptografi.

Sistem *One Time Key* (OTK) yang menggunakan kunci berupa barisan abjad merupakan salah satu contoh sistem kriptografi yang masih digunakan di Indonesia untuk mengamankan informasi yang bersifat rahasia. Berdasarkan prinsip Kerckhoffs (1883), barisan abjad pada OTK minimal harus berupa barisan acaksemu yang aman secara kriptografis.

Uji statistik untuk menguji bentuk distribusi dari suatu barisan kunci mulai berkembang sejak masa perang dunia I yang dipelopori oleh Kendall dan Smith (1938). Uji ini bertujuan menguji barisan digit dan terdiri atas empat uji yaitu uji frekwensi, uji serial, uji poker dan uji gap. Keempat uji tersebut merupakan pengembangan dari uji kecocokan *chi-square*. Sejak tahun 1938 sampai dengan tahun 2005, uji-uji statistik untuk menguji barisan abjad hanya bertujuan mengetahui bentuk distribusi dari barisan kunci. Marsaglia (2005) mengajukan uji *overlapping m-tuple test* yang merupakan pengembangan dari uji serial yang dikembangkan oleh Beker dan Piper (1982).

Selama ini kunci yang digunakan dalam sistem OTK di Indonesia hanya diuji dengan menggunakan *overlapping m-tuple test* yang dikembangkan oleh Marsaglia (2005). Padahal uji tersebut hanya bertujuan menguji bentuk distribusi dan kesalingbebasan sehingga barisan yang telah lulus *overlapping m-tuple test* belum dapat digunakan sebagai kunci sistem OTK karena ketidakterdugaan barisan tersebut belum diketahui.

Mengingat belum adanya penelitian mengenai ketidakterdugaan maka dilakukan penelitian untuk membahas pengujian terhadap keterdugaan suatu barisan abjad dengan menggunakan pendekatan rantai markov. Penelitian dibatasi pada pemodelan rantai markov karena beberapa algorirtma pembentuk PBAS yaitu DES dalam Lai (1992) serta AES dalam Daemen dan Rijmen (2007) merupakan *markov cipher* yang memiliki sifat markov. Tujuan dari penelitian ini adalah mengembangkan metodologi untuk menguji keterdugaan suatu barisan abjad yang dihasilkan PBAS berdasarkan model rantai markov waktu diskrit

METODOLOGI PENELITIAN

Data

Sumber data pada penelitian ini terdiri atas dua sumber data yaitu data simulasi dan data kunci sistem OTK. Data simulasi digunakan untuk mengembangkan metodologi uji keterdugaan dengan rantai markov. Data simulasi berasal dari PBAS yang masing-masing berukuran satu juta huruf. Data simulasi ini dibangkitkan langsung dari empat kelas PBAS seperti yang diperlihatkan Tabel 1.

Tabel 1 Empat kelas PBAS

Kelas	Basis	Nama PBAS
Satu	Algoritma penyandian blok	PBAS ANSI X9.17 dan ANSI X9.31
Dua	Faktorisasi bilangan bulat	*Blum Blum Shub* (BBS)
Tiga	*Linear Congruential Generator* (LCG)	coveyou, fishman18, fishman20, fishman2x, knuthran, knuthran2, lecuyer21, minstd, LCG1, LCG2, cmrg, mrg, rand,rand48, randu, ran0, ran1, ran2, ran3, gfsr4 dan zuf
Empat	*Linear Feedback Shift-Register* (LFSR)	rand128_bsd, rand128_glibc2, rand128_libc5, rand32_bsd, rand32_glibc2, rand32_glibc2, rand64_bsd, rand64_libc2, mt19937, mt19937_1999 dan mt19937_1998

Data kedua adalah data kunci sistem OTK Tahun 2005 yang digunakan oleh 3 unit komunikasi Departemen Luar Negeri yaitu Canbera, Jenewa dan New York. Data tersebut akan diuji keterdugaannya dengan menerapkan hasil pengembangan metodologi uji keterdugaan dengan rantai markov yang telah diperoleh sebelumnya.

Metode Penelitian

Penelitian ini terdiri atas dua tahap yaitu pengembangan metodologi uji keterdugaan dengan rantai markov dan penerapan hasil pengembangan metodologi tersebut. Langkah-langkah pengembangan metodologi uji keterdugaan dengan rantai markov secara terperinci adalah sebagai berikut :

1. Membangkitkan barisan huruf dari rantai markov waktu diskrit orde satu,dua dan tiga. Langkah-langkah untuk membangkitkan barisan adalah :
   a. Membangkitkan barisan huruf dari keempat kelas masing-masing sebesar satu juta huruf.
   b. Mengelompokkan barisan huruf kedalam tiga tipe gugus data seperti pada Tabel 2.

Pengambilan ketiga tipe gugus data ini dilakukan secara *overlap* (tumpang tindih) dan tanpa *overlap* dengan jumlah huruf yang *overlap* sebanyak 10.000 huruf.

Tabel 2 Tipe gugus data

Tipe	Perbandingan	Jumlah Huruf	
		Data Pelatihan	Data Observasi
Satu	50:50	50.000	50.000
Dua	75:25	75.000	25.000
Tiga	100:10	100.000	10.000

c. Menghitung frekwensi 2-*gram* (AA-ZZ) s.d. 4-*gram* (AAAA-ZZZZ) dari data pelatihan pada ketiga tipe gugus data dengan menggunakan algoritma *sliding window counts*.
d. Menduga peluang matriks transisi orde pertama s.d orde ketiga berdasarkan frekwensi 2-*gram* s.d. 4-*gram* dari data pelatihan
   orde 1 : $P(j|i) = \frac{N(i,j)}{\sum_{l=0}^{25} N(i,l)}, 0 \le i,j < 26$
   direpresentasikan dalam matriks peluang transisi berukuran $26 \cdot 26$ (*state* awal : A-Z, *state* akhir : A-Z)
   orde 2 : $P(k|i,j) = \frac{N(i,j,k)}{\sum_{l=0}^{25} N(ij,l)}$, $0 \le i,j,k < 26$
   direpresentasikan dalam matriks peluang transisi berukuran $26^2 \cdot 26^2$ (*state* awal : AA – ZZ, *state* akhir : AA-ZZ)
   orde 3 : $P(m|i,j,k) = \frac{N(i,j,k,m)}{\sum_{l=0}^{25} N(i,j,m,l)}$, $0 \le i,j,k,m < 26$
   direpresentasikan dalam matriks peluang transisi berukuran $26^3 \cdot 26^3$ (*state* awal : AAA-ZZZ, *state* akhir : AAA-ZZZ)
e. Pembangkitkan huruf sebesar ukuran data observasi ketiga tipe gugus data berdasarkan peluang transisi rantai markov mulai orde pertama s.d. orde ketiga. Langkah tersebut diulang sebanyak 10 kali.
   Berikut ini adalah tahapan dalam pembangkitan huruf dengan menggunakan peluang matriks transisi.
   1) Mengambil huruf sebanyak tingkat orde mulai dari posisi ke-(*n*-tingkat orde) sampai dengan ke-*n* dimana *n* adalah banyaknya huruf pada gugus data pelatihan. Sekumpulan huruf ini selanjutnya dinamakan dengan prefiks.
   2) Mengambil *state* awal peluang matriks transisi pada orde tersebut sesuai prefiks kemudian menghitung frekwensi kumulatif dari *state* tersebut.
   3) Membangkitkan sebuah angka acak U[0,1]. Angka ini digunakan untuk menunjuk posisi *state* akhir dari frekwensi kumulatif peluang matriks transisi
   4) *State* akhir inilah yang merupakan huruf hasil bangkitan dari rantai markov pada orde tersebut.

2. Analisis tingkat kecocokan barisan huruf antara data bangkitan dengan data observasi.
   Pada tahap ini dilakukan langkah-langkah sebagai berikut :
   a. Menghitung tingkat kecocokan dengan cara :
      1) Mencocokan gugus data observasi dengan gugus data hasil bangkitan rantai markov pada berbagai orde untuk mengetahui jumlah huruf yang cocok diantara kedua gugus data tersebut dengan menggunakan operator relasi setara (==) yang terdapat pada bahasa pemrograman matlab.
         Jumlah huruf yang cocok diukur oleh operator relasi == dengan cara menghitung banyaknya huruf-huruf yang sama di posisi yang bersesuaian pada dua barisan dengan panjang sama yang berasal dari kedua gugus data.
      2) Menghitung tingkat kecocokan dengan membandingkan banyaknya huruf yang sama dengan jumlah seluruh huruf dalam gugus data menggunakan persamaan :
         $$tingkat\ kecocokan = \frac{\sum huruf\ yang\ cocok}{\sum huruf\ yang\ diuji}$$
      3) Menghitung rataan tingkat kecocokan dari 10 ulangan gugus data.
   b. Melakukan analisis karakteristik tingkat kecocokan :
      1) Setiap kelas PBAS.
         Tujuan analisis ini untuk memperoleh rekomendasi PBAS yang dapat digunakan dalam sistem kriptografi.
      2) Antara keempat kelas PBAS.
         Hasil analisis ini akan digunakan untuk menentukan nilai ambang (*threshold*) keterdugaan model rantai markov orde 1 s.d. 3.
      3) Ketiga tipe gugus data kelas kesatu baik tanpa *overlap* maupun dengan *overlap*.
         Tujuan analisis ini adalah untuk memperoleh rekomendasi mengenai ukuran perbandingan antara data pelatihan dan data observasi terbaik.

HASIL DAN PEMBAHASAN

Pembahasan mengenai pengembangan metodologi uji keterdugaan dengan rantai markov dibatasi pada pencarian karakteristik tingkat

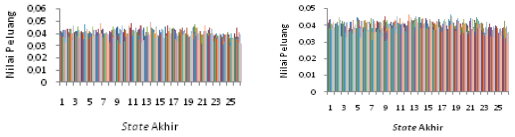
kecocokan dalam setiap kelas PBAS, membandingkan karakteristik tingkat kecocokan diantara keempat kelas PBAS serta melihat karakteristik tingkat kecocokan antara ketiga tipe gugus data tanpa *overlap* maupun dengan *overlap.* Hasil analisis dan pembahasan tersebut diuraikan dibawah ini.

Keempat Kelas PBAS

Untuk memperoleh rekomendasi PBAS yang dapat digunakan dalam sistem kriptografi maka dilakukan analisis terhadap karakteristik tingkat kecocokan pada keempat kelas PBAS. Analisis dilakukan dengan mengamati matriks peluang transisi serta grafik tingkat kecocokan yang dihasilkan oleh gugus data tanpa *overlap* maupun *overlap* dalam kelas tersebut.

1. Kelas Kesatu

Nilai peluang pada matriks transisi berpengaruh terhadap tingkat kecocokan yang akan dicapai oleh suatu gugus data karena ketika *state* awal ke-*i* bertransisi ke semua *state* akhir *j* maka kemungkinan untuk memperoleh huruf yang cocok akan semakin sedikit (peluang = 1/26). Ketika *state* awal ke-*i* hanya bertransisi ke beberapa *state* saja maka kemungkinan untuk memperoleh huruf yang cocok memiliki peluang lebih besar dari 1/26. Identifikasi awal dapat dilihat pada plot nilai peluang matriks transisi setiap PBAS pada ketiga tipe gugus data.



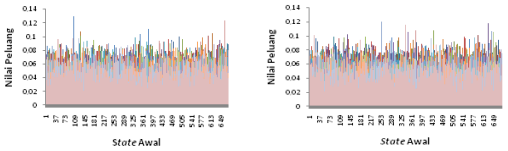
(a) (b)

Gambar 1 Plot nilai peluang matriks transisi orde satu kelas kesatu gugus data tipe 3 tanpa *overlap* (a) PBAS X9.17; (b) PBAS X9.31.

Gambar 1 dan Gambar 2 menunjukkan plot nilai peluang transisi orde satu dan orde dua PBAS X9.17 dan X9.31 pada gugus data tipe ketiga tanpa *overlap*.

Pada Gambar 1 terlihat bahwa nilai peluang transisi orde satu PBAS X9.17 pada ketiga tipe gugus data berada diantara nilai 2.76×10^{-2} s.d. 5.20×10^{-2} sedangkan pada PBAS X9.31 berada diantara 2.12×10^{-2} s.d. 5.36×10^{-2}. Hal ini menyebabkan semua *state* pada matriks peluang transisi X9.17 dan X9.31 dapat bertransisi secara langsung dari satu *state* ke *state* lain sehingga rantai markov yang terbentuk merupakan rantai markov tidak tereduksi dan hanya terdiri atas satu kelas *state* tertutup yaitu {A,B,C,D,E,F, ...,Z}.

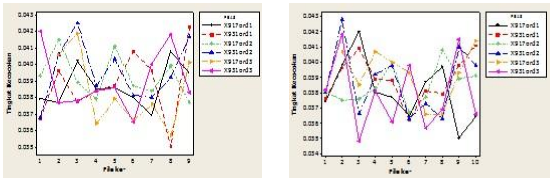
Pada Gambar 2 terlihat bahwa nilai peluang transisi orde dua pada ketiga tipe gugus data X9.17 dan X9.31 mengalami perubahan. Nilai peluang transisi orde dua PBAS X917 berada diantara nilai 0 s.d 1.54×10^{-1} sedangkan pada PBAS X9.31 diantara 0 s.d. 1.62×10^{-1}.



(a) (b)

Gambar 2 Plot nilai peluang matriks transisi orde dua kelas kesatu gugus data tipe 3 tanpa *overlap* (a) PBAS X9.17; (b) PBAS X9.31.

Dari Gambar 3 terlihat bahwa bahwa perubahan nilai peluang matriks transisi orde dua tidak berpengaruh secara signifikan pada perolehan tingkat kecocokan di orde dua. Tingkat kecocokan yang dicapai baik pada data tanpa *overlap* maupun dengan *overlap* pada ketiga orde relatif sama yaitu berada diantara 3.4×10^{-2} s.d. 4.3×10^{-2}. Atau dengan kata lain, barisan huruf yang dihasilkan oleh PBAS X9.17 dan PBAS X9.31 pada ketiga tipe gugus data dengan *overlap* maupun tanpa *overlap*, belum dapat dimodelkan dengan rantai markov orde satu,dua maupun tiga.



(a) (b)

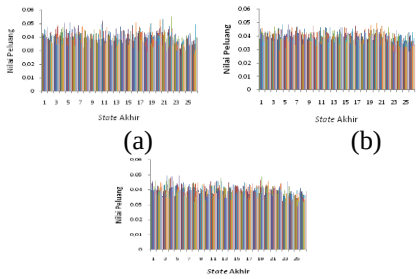
Gambar 3 Plot Tingkat Kecocokan Gugus Data Tipe 3 Kelas Kesatu Orde Satu, Dua dan Tiga (a) tanpa *Overlap* ; (b) dengan *Overlap*

2. Kelas Kedua

Tidak seperti kelas kesatu, kelas kedua hanya terdiri atas satu PBAS yaitu PBAS BBS. Gambar 4 dan Gambar 5 menunjukkan plot nilai peluang transisi orde satu dan orde dua PBAS BBS pada ketiga tipe gugus data tanpa *overlap*.

Pada Gambar 4 terlihat bahwa nilai peluang transisi orde satu PBAS BBS pada ketiga tipe gugus data berada diantara nilai 2.43×10^{-2} s.d. 5.53×10^{-2}. Hal ini menyebabkan semua *state* pada matriks peluang transisi BBS dapat bertransisi secara langsung dari satu *state* ke *state* lain sehingga rantai markov yang terbentuk merupakan rantai markov tidak

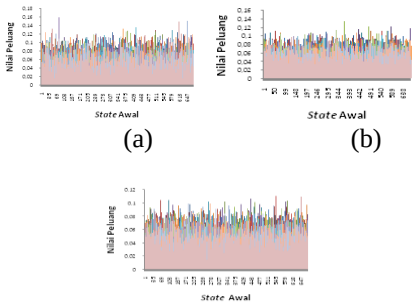
tereduksi dan hanya terdiri atas satu kelas *state* tertutup yaitu {A,B,C,D,E,F, ...,Z}.



(a) (b)

(c)

Gambar 4 Plot nilai peluang matrik transisi orde satu kelas kedua tanpa *overlap* (a) tipe 1; (b) tipe 2; (c) tipe 3.

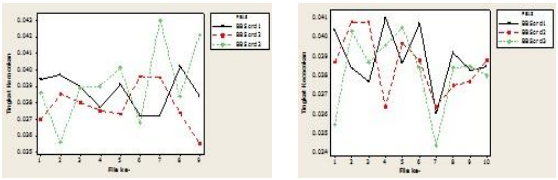


(a) (b)

(c)

Gambar 5 Plot nilai peluang matrik transisi orde dua kelas kedua tanpa *overlap* (a) tipe 1; (b) tipe 2; (c) tipe 3.

Gambar 5 menunjukkan bahwa nilai peluang transisi orde dua pada ketiga tipe gugus data BBS mengalami perubahan. Nilai peluang transisi orde dua berada diantara nilai 0 s.d. 1.62 x 10^{-1}.



(a) (b)

Gambar 6 Plot tingkat kecocokan gugus data tipe 3 kelas kedua (a) tanpa *overlap*; (b) dengan *overlap*.

Dari Gambar 6 terlihat bahwa perubahan nilai peluang matriks transisi orde dua tidak berpengaruh secara signifikan pada perolehan tingkat kecocokan di orde dua. Tingkat kecocokan yang dicapai baik pada data tanpa *overlap* maupun dengan *overlap* pada ketiga orde relatif sama yaitu berada diantara 3.4 x 10^{-2} s.d. 4.3 x 10^{-2}. Atau dengan kata lain, barisan huruf yang dihasilkan oleh PBAS BBS pada ketiga gugus data dengan *overlap* maupun tanpa *overlap*, belum dapat dimodelkan dengan rantai markov orde satu,dua maupun tiga.

3. Kelas Ketiga

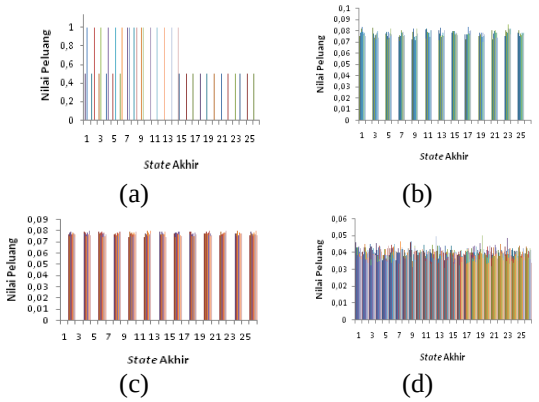
Kelas ketiga terdiri atas dua puluh PBAS berbasis LCG. Disebut berbasis LCG karena algoritma pembangkitan huruf yang digunakan pada kedua puluh PBAS ini pada dasarnya sama yaitu menggunakan persamaan:

$$x_n = ax_{n-1} + b \bmod m, \; n \geq 1,$$

Perbedaannya hanya terletak pada pemilihan nilai parameter *a*, *b*, *m* dan x_{n-1}.

Gambar 7 dan 8 menunjukkan plot nilai peluang transisi orde satu dan dua PBAS LCG1, Coveyou, LCG2 dan gfsr4 pada tipe gugus data tipe ketiga tanpa *overlap*.

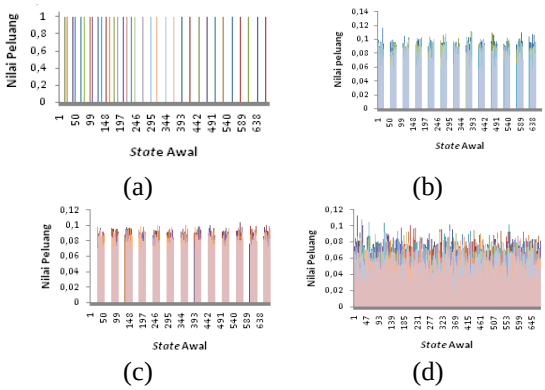
Pada Gambar 7 dan Gambar 8 terlihat bahwa nilai peluang transisi orde satu dan orde dua dari kelas ketiga pada ketiga tipe gugus data terbagi dalam tiga kelompok yaitu (a) kelompok 1 berisi nilai peluang transisi LCG1, (b) kelompok 2 berisi nilai peluang transisi coveyou, LCG2, rand, dan randu, (c) kelompok 3 berisi nilai peluang transisi keenam belas PBAS lain. Pada kelompok 1, nilai peluang transisi orde satu hanya berada pada nilai 0, 4.99 x 10^{-1} , 0.5 dan 1 sedangkan pada orde dua nilai peluangnya hanya bernilai 0 dan 1. Pada kelompok 2, nilai peluang matrik transisi orde satu selain bernilai 0 juga berada pada nilai 6.67 x 10^{-2} s.d. 8.88 x 10^{-2} sedangkan pada orde dua selain bernilai 0 juga berada pada nilai 2.25 x 10^{-2} s.d. 1.47 x 10^{-1}. Pada kelompok 3 nilai peluang matrik transisi orde satu pada ketiga tipe gugus data berada diantara nilai 2.62 x 10^{-2} s.d. 4.99 x 10^{-2} sedangkan pada orde dua selain bernilai 0 juga berada diantara nilai 5.75 x 10^{-3} s.d. 1.58 x 10^{-1}.



(a) (b)

(c) (d)

Gambar 7 Plot nilai peluang matrik transisi orde satu kelas ketiga gugus data tipe 3 tanpa *overlap* (a) PBAS LCG1; (b) PBAS Coveyou; (c) PBAS LCG2; (d) PBAS gfsr4.

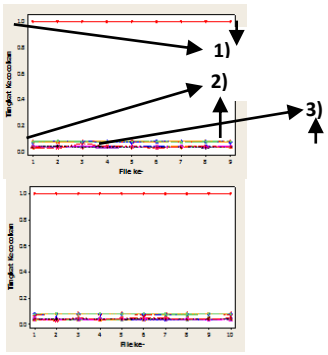


Gambar 8 Plot nilai peluang matrik transisi orde dua kelas ketiga gugus data tipe 3 tanpa *overlap* (a) PBAS LCG1; (b) PBAS Coveyou; (c) PBAS LCG2; (d) PBAS gfsr4.

Penyebaran nilai peluang pada matriks transisi pada ketiga kelompok tersebut disebabkan oleh parameter yang digunakan oleh PBAS tersebut.

Penjelasannya adalah sebagai berikut :

a. parameter yang digunakan dalam LCG1 adalah $a = 1$, $b = 23$, $m = 35$ dan $x_1 = 9$. Periode barisan yang dihasilkan LCG1 maksimal yaitu 34 huruf karena parameter *b* dan *m* relatif prima serta parameter *m* lebih besar daripada nilai parameter *a*, *b* dan x_1. Meskipun demikian, periode tersebut terlalu pendek jika dibandingkan dengan variasi kemunculan bigram yaitu 676. Akibatnya tidak semua variasi bigram muncul dan nilai peluang pada matriks peluang transisi hanya tersebar pada *state-state* tertentu saja.
b. parameter yang digunakan dalam LCG2 adalah $a = 1227$, $b = 0$, $m = 131072$ dan $x_1 = 1$.
Periode barisan yang dihasilkan LCG2 tidak mencapai maksimal yaitu hanya 32.768 huruf karena parameter *b* dan *m* komposit. Meskipun periode LCG2 lebih panjang dari LCG1 tetapi tidak semua variasi bigram muncul sehingga nilai peluang pada matriks peluang transisi hanya tersebar pada *state-state* tertentu saja.
c. penggunaan modulus yang bukan bilangan prima pada rand dan coveyou menyebabkan tidak semua variasi 2-gram, 3-gram muncul dalam barisan sehingga nilai peluang pada matriks peluang transisi hanya tersebar pada *state-state* tertentu saja.
d. pada enam belas PBAS lain yang menggunakan modulus berupa bilangan prima, hampir semua variasi 2-gram, 3-gram muncul dalam barisan sehingga nilai peluang pada matriks peluang transisi menyebar pada seluruh *state*. Hal ini tidak berlaku pada PBAS randu yang juga menggunakan modulus bilangan prima (2^{31}). Dalam barisan yang dihasilkan oleh PBAS randu, tidak semua variasi 2-gram, 3-gram muncul sehingga nilai peluang pada matriks peluang transisi hanya tersebar pada *state-state* tertentu saja.



Ket : 1) PBAS LCG1 orde2 dan orde3, 2) PBAS LCG2, Coveyou, Rand dan Randu, 3)PBAS fishman18, fishman20, fishman2x, knuthran, knuthran2, lecuyer21, minstd, cmrg, mrg, rand48, ran0, ran1, ran2, ran3, gfsr4 dan zuf

Gambar 9 Plot tingkat kecocokan gugus data tipe 3 kelas ketiga orde 2 dan orde 3 (a) tanpa *overlap*; (b) dengan *overlap*.

Perilaku rantai markov orde satu kelompok 1 dan 2 menunjukkan bahwa *state* pada matriks transisi coveyou, LCG1, LCG2, rand dan randu tidak dapat bertransisi secara langsung dari satu *state* ke *state* lain sedangkan *state* dari matriks transisi PBAS lainnya dapat bertransisi secara langsung. Meskipun demikian, rantai markov dari ke-21 PBAS ini tidak tereduksi dan hanya terdiri atas satu kelas *state* yang tertutup yaitu {A,B,C,D,E,F, ...,Z}.

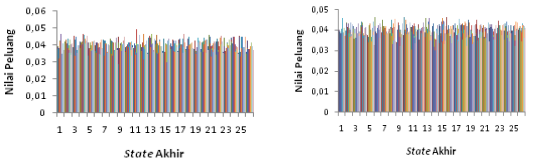
Pada Gambar 9 terlihat bahwa tingkat kecocokan PBAS LCG1 pada ketiga tipe gugus data mulai dari orde dua mencapai 1. Hal ini berarti PBAS LCG1 dapat dimodelkan dengan rantai markov orde dua atau barisan yang dihasilkan oleh LCG1 merupakan barisan yang dapat diduga dengan rantai markov orde dua.

Pada PBAS coveyou, LCG2, rand dan randu, perubahan nilai peluang matrik transisi orde dua dan tiga tidak berpengaruh secara signifikan pada perolehan tingkat kecocokan di orde dua dan tiga. Tingkat kecocokan yang dicapai baik pada data tanpa *overlap* maupun dengan *overlap* pada orde dua dan tiga relatif sama dengan tingkat kecocokan pada orde satu yaitu berada diantara 7.02×10^{-2} s.d. 8.39×10^{-2}. Hal ini pun terjadi pada keenam belas PBAS lain. Perubahan nilai peluang transisi orde dua dan tiga tidak berpengaruh secara signifikan pada perolehan tingkat kecocokan di orde dua

dan tiga. Tingkat kecocokan yang dicapai baik pada data tanpa *overlap* maupun dengan *overlap* pada orde dua dan ketiga relatif sama dengan tingkat kecocokan pada orde satu yaitu berada diantara 3.34×10^{-2} s.d. 4.44×10^{-2}. Atau dengan kata lain, barisan huruf yang dihasilkan oleh ke-20 PBAS pada kelas kedua di ketiga tipe gugus data baik dengan *overlap* maupun tanpa *overlap* tidak dapat dimodelkan dengan rantai markov orde satu,dua maupun tiga.

4. Kelas Keempat

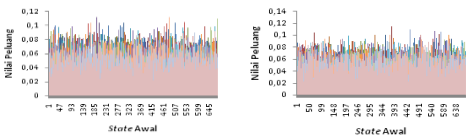
Kelas keempat terdiri atas sebelas PBAS. Gambar 10 dan Gambar 11 menunjukkan plot nilai peluang transisi orde satu dan orde dua PBAS mt19937_1999 dan rand128_bsd pada gugus data tipe ketiga tanpa *overlap*.



(a) (b)

Gambar 10 Plot nilai peluang matriks transisi orde satu kelas keempat gugus data tipe 3 tanpa *overlap* (a) PBAS mt19937_1999; (b) PBAS random128_bsd.

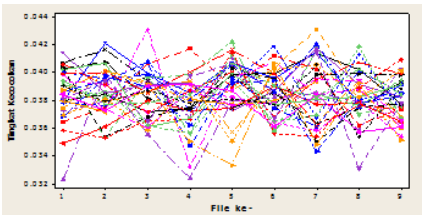
Pada Gambar 10 terlihat bahwa nilai peluang matriks transisi orde satu PBAS mt19937_1999 pada ketiga tipe gugus data berada diantara nilai 2.39×10^{-2} s.d. 5.29×10^{-2} sedangkan pada PBAS rand128_bsd berada diantara 2.42×10^{-2} s.d. 5.31×10^{-2}. Hal ini menyebabkan semua *state* pada matriks peluang transisi PBAS mt19937_1999 dan rand128_bsd dapat bertransisi secara langsung dari satu *state* ke *state* lain sehingga rantai markov yang terbentuk merupakan rantai markov tidak tereduksi dan hanya terdiri atas satu kelas *state* tertutup yaitu {A,B,C,D,E,F, ...,Z}.



(a) (b)

Gambar 11 Plot nilai peluang matrik transisi orde dua kelas keempat gugus data tipe 3 tanpa *overlap* (a) PBAS mt19937_1999; (b) PBAS random128_bsd.

Gambar 11 menunjukkan bahwa nilai peluang matriks transisi orde dua pada ketiga tipe gugus data PBAS mt19937_1999 dan rand128_bsd mengalami perubahan. Nilai peluang matriks transisi orde dua PBAS mt19937_1999 selain bernilai 0 juga berada pada nilai 5.53×10^{-3} s.d. 1.67×10^{-1} sedangkan pada PBAS rand128_bsd selain bernilai 0 juga berada pada nilai 5.53×10^{-3} s.d. 1.88×10^{-1}.



Gambar 12 Plot tingkat kecocokan gugus data tipe 3 pbas kelas keempat tanpa *overlap* pada ketiga orde.

Pada Gambar 12 terlihat bahwa perubahan nilai peluang matriks transisi orde dua dan tiga tidak berpengaruh secara signifikan pada perolehan tingkat kecocokan di orde dua dan tiga. Akibatnya tingkat kecocokan yang dicapai baik pada data tanpa *overlap* maupun dengan *overlap* pada orde dua dan orde tiga relatif sama dengan orde satu yaitu berada diantara 3.23×10^{-2} s.d. 4.09×10^{-2}. Atau dengan kata lain, barisan huruf yang dihasilkan oleh PBAS kelas keempat pada ketiga tipe gugus data dengan *overlap* maupun tanpa *overlap*, belum dapat dimodelkan dengan rantai markov orde satu,dua maupun tiga.

Berdasarkan hasil penelitian *yang* telah dilakukan, dapat disimpulkan bahwa barisan abjad yang dihasilkan oleh PBAS kelas kesatu, kedua, dan keempat tidak dapat dimodelkan dengan rantai markov. Demikian pula dengan PBAS kelas ketiga kecuali barisan yang dihasilkan oleh PBAS LCG1, LCG2, coveyou, rand dan randu. Bila barisan tersebut juga acak secara statistik maka barisan huruf adalah barisan acaksemu yang aman secara kriptografis sehingga layak digunakan dalam kriptografi.

Meskipun barisan yang dihasilkan oleh PBAS LCG2, coveyou, rand dan randu tidak dapat dimodelkan dengan rantai markov orde satu, dua dan tiga tetapi memiliki kemungkinan yang tinggi untuk dapat dimodelkan dengan rantai markov orde-orde tinggi (diatas orde tiga). Oleh karena itu barisan yang dihasilkan oleh keempat PBAS tersebut tidak layak digunakan dalam kriptografi.

Evaluasi Antar Keempat Kelas PBAS

Berdasarkan plot karakteristik tingkat kecocokan, terlihat bahwa tingkat kecocokan pada

keempat kelas cenderung tidak berubah mulai dari orde ke-2. Oleh karena itu penentuan nilai ambang keterdugaan dengan rantai markov dibuat berdasarkan nilai minimum dan maksimum tingkat kecocokan dengan pembulatan ke bawah pada orde ke-2 dan orde ke-3. Nilai ambang tersebut adalah :

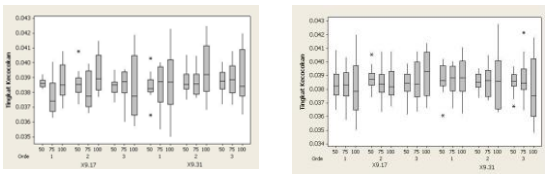
1. Untuk gugus data tanpa *overlap* :
   a. bila tingkat kecocokan dari suatu barisan berada dalam rentang $0 \leq$ tingkat kecocokan $\leq 4.5 \times 10^{-2}$ maka barisan tersebut tidak dapat diduga dengan rantai markov orde satu, dua dan tiga.
   b. bila tingkat kecocokan dari suatu barisan berada dalam rentang $4.5 \times 10^{-2} <$ tingkat kecocokan $\leq 8.2 \times 10^{-2}$ maka masih ada kemungkinan barisan tersebut dapat diduga dengan rantai markov orde satu,dua dan tiga.
   c. bila tingkat kecocokan $8.2 \times 10^{-2} <$ tingkat kecocokan ≤ 1 maka barisan tersebut memiliki kemungkinan yang cukup tinggi dapat diduga dengan rantai markov orde satu,dua dan tiga.

2. Untuk gugus data dengan *overlap* :
   a. bila tingkat kecocokan dari suatu barisan berada dalam rentang $0 \leq$ tingkat kecocokan $\leq 4.4 \times 10^{-2}$ maka barisan tersebut tidak dapat diduga dengan rantai markov orde satu, dua dan tiga.
   b. bila tingkat kecocokan berada dalam rentang $4.4 \times 10^{-2} <$ tingkat kecocokan $\leq 8.3 \times 10^{-2}$ maka masih ada kemungkinan barisan tersebut dapat diduga dengan rantai markov orde satu,dua dan tiga.
   c. bila tingkat kecocokan $8.3 \times 10^{-2} <$ tingkat kecocokan ≤ 1 maka barisan tersebut memiliki kemungkinan yang cukup tinggi dapat diduga dengan rantai markov orde satu,dua dan tiga.

Jika suatu barisan acak secara statistik serta memiliki tingkat kecocokan yang berada dalam rentang $0 \leq$ tingkat kecocokan $\leq 4.5 \times 10^{-2}$ untuk gugus data tanpa *overlap* dan dalam rentang $0 \leq$ tingkat kecocokan $\leq 4.4 \times 10^{-2}$ untuk data dengan *overlap* maka barisan tersebut merupakan barisan acaksemu yang aman secara kriptografis. Atau dengan kata lain barisan tersebut layak digunakan dalam kriptografi.

Evaluasi Antar Ketiga Tipe Gugus Data

Untuk mengetahui ukuran perbandingan antara data pelatihan dan data observasi terbaik maka dilakukan analisis terhadap karakteristik tingkat kecocokan ketiga tipe gugus data kelas kesatu. Analisis dilakukan dengan mengamati diagram kotak serta ringkasan statistik tingkat kecocokan pada kelas kesatu. Kelas kesatu dipilih karena dianggap telah mewakili kelas PBAS lain. Gambar 13 menunjukkan boxplot tingkat kecocokan ketiga tipe gugus data tanpa dan dengan *overlap* mulai dari orde 1 s.d. orde 3 kelas kesatu.



(a) (b)

Gambar 13 Diagram kotak tingkat kecocokan ketiga tipe gugus data orde 1, 2 dan 3 kelas kesatu gugus data (a) tanpa *overlap*; (b) dengan *overlap*.

Gambar 13 memperlihatkan :

1. Gugus data tipe ketiga baik pada data tanpa maupun dengan *overlap* memiliki keragaman tingkat kecocokan tertinggi, diikuti dengan tipe kedua. Sedangkan tipe kesatu memiliki keragaman yang paling rendah. Hal ini karena jumlah data pelatihan yang digunakan untuk menduga matriks peluang transisi lebih besar sehingga lebih banyak diperoleh informasi mengenai bigram (AA-ZZ) dan trigram (AAA-ZZZ).

2. Tingkat kecocokan gugus data tipe ketiga pada ketiga orde tidak simetris, bahkan distribusi tingkat kecocokannya cenderung menjulur ke kanan (*skewness* positif).

Berdasarkan analisis di atas disimpulkan bahwa ukuran data pelatihan harus lebih besar daripada ukuran data observasi dengan perbandingan ukuran data pelatihan dan data observasi terbaik adalah 100:10.

Contoh Penerapan

Hasil pengembangan metodologi uji keterdugaan yang telah diperoleh dalam penelitian ini, selanjutnya diterapkan untuk menguji keterdugaan suatu barisan kunci. Barisan kunci yang akan diuji adalah barisan kunci OTK yang digunakan pada tahun 2005 oleh 3 unit komunikasi Departeman Luar Negeri yaitu unit komunikasi Canbera, Jenewa dan New York yang masing-masing berukuran 10.000 huruf. Barisan kunci ini sebelumnya telah diuji keacakannya dengan *overlapping m-tupple test* dan dinyatakan acak secara statistik. Oleh karena itu untuk mengetahui apakah barisan kunci tersebut layak atau tidak digunakan dalam kriptografi maka akan diuji keterdugaannya terhadap model markov orde satu,dua dan tiga.

Sebelum diuji keterdugaannya, terlebih dahulu barisan kunci OTK dibagi menjadi data pelatihan dan data observasi. Perbandingan ukuran data pelatihan dan data observasi yang digunakan adalah 9: 1 sehingga kedua data tersebut masing-

masing berukuran 9000 huruf dan berukuran 1000 huruf. Selanjutnya akan dibangkitkan 1000 huruf berdasarkan peluang matriks transisi rantai markov orde satu, dua dan tiga. Pembangkitan barisan ini dilakukan sebanyak 10 kali untuk tiap kunci OTK. Proses pembangkitannya seperti yang telah dijelaskan dalam subbaab Metode Analisis.

Langkah selanjutnya adalah menghitung tingkat kecocokan serta rataan tingkat kecocokan dari 10 barisan tersebut. Hasil pengujian berupa rataan tingkat kecocokan disajikan pada Tabel 3.

Tabel 3 Hasil pengujian keterdugaan dengan rantai markov orde satu, dua dan tiga pada barisan kunci OTK

Unit Komunikasi	Rataan Tingkat Kecocokan		
	Orde 1	Orde 2	Orde 3
Canbera	3.71×10^{-2}	4.10×10^{-2}	3.75×10^{-2}
Jenewa	3.62×10^{-2}	3.91×10^{-2}	3.89×10^{-2}
New York	3.98×10^{-2}	3.98×10^{-2}	3.65×10^{-2}

Tabel 3 memperlihatkan bahwa rataan tingkat kecocokan ketiga kunci OTK pada orde satu, dua dan tiga berada dalam rentang nilai ambang 0 s.d. 4.5×10^{-2} sehingga kunci OTK tersebut tidak dapat dimodelkan dengan rantai markov. Karena barisan tersebut acak secara statistik dan tidak dapat dimodelkan dengan rantai markov orde satu, dua dan tiga maka disimpulkan bahwa barisan kunci tersebut layak digunakan untuk mengamankan informasi di unit komunikasi Canbera, Jenewa dan New York.

KESIMPULAN DAN SARAN

Kesimpulan

Simpulan yang diperoleh pada penelitian ini adalah :

1. Barisan abjad yang dihasilkan oleh PBAS kelas kesatu, kedua, dan keempat tidak dapat dimodelkan dengan rantai markov. Demikian pula dengan PBAS kelas ketiga kecuali barisan yang dihasilkan oleh PBAS LCG1, LCG2, coveyou, rand dan randu.
2. Barisan yang dihasilkan oleh PBAS LCG2, coveyou, rand dan randu tidak layak digunakan dalam kriptografi karena memiliki kemungkinan yang tinggi untuk dapat dimodelkan dengan rantai markov orde-orde tinggi (diatas orde tiga).
3. Analisis terhadap karakteristik tingkat kecocokan antara keempat kelas PBAS menghasilkan nilai ambang (*threshold*) keterdugaan dengan rantai markov yang dibuat berdasarkan nilai minimum dan maksimum tingkat kecocokan pada orde 2 dan orde 3.
4. Berdasarkan nilai ambang, suatu barisan yang acak secara statistik serta memiliki tingkat kecocokan dalam rentang $0 \leq \text{tingkat kecocokan} \leq 4.5 \times 10^{-2}$ untuk gugus data tanpa *overlap* dan dalam rentang $0 \leq \text{tingkat kecocokan} \leq 4.4 \times 10^{-2}$ untuk data dengan *overlap*, merupakan barisan acaksemu yang aman secara kriptografis sehingga layak digunakan dalam kriptografi
5. Analisis terhadap karakteristik tingkat kecocokan ketiga tipe gugus data kelas kesatu menunjukkan bahwa ukuran data pelatihan harus lebih besar daripada ukuran data observasi dengan perbandingan ukuran yang terbaik antara data pelatihan dengan data observasi adalah 100: 10.
6. Hasil pengujian terhadap barisan kunci OTK yang digunakan oleh unit komunikasi Canbera, Jenewa dan New York pada Tahun 2005 menunjukkan bahwa rataan tingkat kecocokan kunci OTK pada ketiga orde kurang dari 4.5×10^{-2} sehingga kunci OTK layak digunakan dalam kriptografi untuk mengamankan informasi di 3 unit komunikasi tersebut.

Saran

Dalam penelitian ini, parameter rantai markov hanya diduga dengan menggunakan metode kemungkinan maksimum dengan orde tertinggi yang diteliti adalah tiga. Pada penelitian selanjutnya, perlu dilakukan analisis dan kajian teori mengenai pendugaan parameter rantai markov dengan menggunakan metode Bayes serta efeknya terhadap perolehan tingkat kecocokan. Selain itu, perlu dilakukan kajian lebih lanjut mengenai pengujian ketidakterdugaan dengan menggunakan model *random walk*.

DAFTAR PUSTAKA

Beker H, Piper. 1982. *Cipher System the Protection of Communications*. London: Northwood Books.

Casella G, Berger RL. 2002. *Statistical Inference*. Ed ke-2. California: Duxbury.

Daemen J, Rijmen V. 2007. Probability Distributions of Correlation and Differentials in Block Cipher. *Journal of Mathematical Cryptology* 1:221-241.

Dewdney AK. 1989. Computer Recreations. *Scientific American* 260:122-125

Hadiwibowo. 2006. Informasi Rahasia. http://hadiwibowo.wordpress.com/2006/12/25/informasi-rahasia [17 Januari 2011]

Kendall MG, Smith BB. 1938. Randomness and Random Sampling Numbers. *Journal of the Royal Statistical Society* 101 No.1:147-166.

Kerckhoffs A. 1883. La Cryptographic Militaire. *Journal des Sciences Militaires* IX:5-38

Konheim AG. 2007. *Computer Security and Cryptography*. New Jersey : John Wiley & Sons.

Lai X. 1995. *On the Design and Security of Block Ciphers*. Zurich : Hartung-Gorre Verlag Konstanz.

Lidl R, Pilz G. 1997. *Applied Abstract Algebra*. Ed ke-2. New York : Springer-Verlag.

Mangku IW. 2005. *Proses Stokastik*. Bogor : Departemen Matematika Fakultas MIPA IPB.

Marsaglia G. 2005. Monkeying Goodness of Fit Test. *Journal of Statistics Software* 13 Issue 14

Menezes AJ, Van Oorschot PC, Vanstone SA. 1997. *Handbook of Applied Cryptography*. Florida : CRC Press.

Schneier B. 1996. *Applied Cryptography* : *Protocols, Algorithms and Source Code in C*. Ed ke-2. Canada : John Wiley & Sons.

Shannon CE. 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal* 27:379-423

Stalling W. 1998. *Cryptograhy and Network Securty Principles and Practice*. Ed ke-2. New Jersey : Prentice Hall.

Stinson DR. 1995. *Cryptography Theory and Practice*. Ed ke-3. Florida : CRC Press.

Van Tilborg HCA. 2005. *Encyclopedia of Cryptography and Security*. New York: Springer

Xu X, Tsang WW. 2007. An Empirical Study on the Power of the *Overlap* Serial Test. HKU CS Tech Report TR 09.