# Analisis Pengaruh Penggunaan Manhattan Distance Pada Algoritma Clustering Isodata ( Self-Organizing Data Analysis Technique) Untuk Sistem Deteksi Anomali Trafik

# Analysis Of Manhattan Distance Usage Effects on Isodata Clustering Algorithm (Self-Organizing Data Analysis Technique) For Traffic Anomaly Detection

**A.A.Ngr Wisnu Gautama[1], Yudha Purwanto[2], Tito Waluyo Purboyo[3]**

[1,2,3]Prodi S1 Teknik Komputer, Fakultas Teknik Elektro, Universitas Telkom

[1]wisnugautama@student.telkomuniversity.ac.id, [2]omyudha@telkomuniversity.ac.id, titowaluyo@telkomuniversity@telkomuniversity.ac.id

**Abstrak**

**Ketertarikan masyarakat terhadap berbagai informasi yang mudah didapat menyebabkan meningkatnya penggunaan internet. Seiring banyaknya masyarakat yang mengakses internet menyebabkan adanya fenomena anomali trafik. Fenomena anomali trafik ini dapat berupa serangan DDoS dan *flash crowd*. Dilihat dari dampak negatif yang terdapat anomali trafik tersebut, maka pada penelitian ini dibuat sebuah metode *Intrusion Detection System* (IDS) dengan teknik *unsupervised learning* yang menggunakan algoritma ISODATA *clustering* dengan penambahan metode berbasis pengukuran jarak *Manhattan Distance* dan metode *Dunn Index* untuk menghitung kualitas cluster yang dihasilkan. Hasil dari penelitian ini, sistem yang sudah dibangun menunjukan hasil performansi yang baik dan minimnya dalam kesalahan deteksi dilihat dari sistem yang sudah mampu membedakan trafik anomali dengan trafik normal. Dengan menggunakan metode *Manhattan Distance,* waktu proses yang dibutuhkan lebih singkat dibandingkan dengan metode *Euclidean Distance*.**

**Kata Kunci : Anomali trafik, DDoS, Flash Crowd, Isodata, Clustering, Manhattan Distance, Dunn Index**

**Abstract**

**The Interest of people's to any information's with easily to get has increased the using of internet. As the number of people's which accessing the internet become the phenomenon of traffic anomaly. The phenomenon of traffic anomaly there are DDoS attack and flash crowd. Considering the negative impacts of traffic anomaly, then in this research be made a Intrusion Detection System (IDS) method with unsupervised learning technique by using ISODATA algorithm with the addition of Manhattan Distance based method and Dunn Index to calculate the quality of cluster. The result of this research, the system has been made show the good performance and low value of false detection. The system has been differentiate the anomaly traffic and normal traffic. By using the Manhattan Distance method, the time proccess neede is less than Euclidean Distance.**

**Keywords: Traffic anomaly, DDoS, Flash Crowd, Isodata, Clustering, Manhattan Distance, Dunn Index**

## 1. Pendahuluan

Anomali Trafik merupakan penyimpangan atau keanehan yang terdapat pada suatu trafik jaringan. Anomali Trafik ini meliputi serangan *Distributed Denial of Service* (DDoS) dan *flash crowd*. *Distributed Denial of Service* (DoS) adalah suatu jenis serangan terhadap sebuah komputer atau *server* didalam suatu jaringan internet dengan cara menghabiskan sumber (*Resources*) komputer tersebut sehingga komputer tersebut tidak dapat menjalankan fungsinya dengan baik dan benar sehingga secara tidak langsung mencegah komputer lain mengakses komputer tersebut [1] [2]. Sedangkan *flash crowd* bukanlah sebuah serangan, melainkan sistuasi terjadinya peningkatan trafik yang sangat tinggi secara signifikan dalam suatu jaringan sehingga tidak dapat diakses dalam rentang waktu tertentu [1]. *Flash crowd* ini juga dapat dikatakan sebagai situasi dimana saat banyak user mengakses sebuah *website* dalam waktu yang sama.

Dalam mendeteksi adanya anomaly trafik, informasi 5-*tuple* dari *IP* menjadi karakteristik acuan dalam menganalisis. Dimana diantaranya, *protocol type, source IP address, destination IP address, source port, dan destination port* [1] [2]. Pada penelitian sebelumnya [4], dalam mendeteksi anomaly trafik menggunakan metode *Data Mining* dengan algoritma *K-means*, dikatakan berhasil dalam membedakan trafik normal dengan trafik

anomaly. Namun, algoritma *K-Means* masih sangat sensitif terhadap adanya *outlier* dan *cluster* yang dihasilkan cenderung berbentuk oval [5]. Dari kelemahan algortima *K-Means* tersebut, maka digunakan algoritma ISODATA dalam mengatasi *outlier*. Proses pembelahan yang terdapat pada algoritma ISODATA dapat mengatasi hasil *cluster* yang cenderung berbentuk oval yang dihasilkan oleh algoritma *K-Means* [5]. Algoritma ISODATA ini merupakan perkembangan dari algoritma *K-Means* dimana terdapat proses – proses baru seperti penggabungan *cluster* dan pembelahan *cluster* serta kepadatan suatu *cluster* dapat dikontrol dengan algoritma [4] [6] [7] [8].

Pada penelitian ini akan dibuat sebuah sistem deteksi anomaly trafik yang dapat membedakan antara serangan DDoS dengan *flash crowd* dengan menggunakan algoritma ISODATA dan penambahan metode berbasis pengukuran jarak *Manhattan Distance*. Metode *Dunn Index* juga digunakan dalam menghitung kualitas *cluster* yang dihasilkan.

## 2. Dasar Teori dan Perancangan

### 2.1. Sistem Deteksi Anomali

Dalam sistem deteksi anomlai trafik, terdapat dua istilah yang sering muncul yaitu *Itrusion Detection System* (IDS) dan *Intrusion Prevesion System* (IPS). Pada sistem IDS, sistem harus memonitor atau mengawasi terlebih dahulu aliran trafik dan jika mengalami sererangan sistem akan memberikan tanda yang berupa alarm sehingga sistem akan minindak lanjuti secara manual oleh operator. Sistem IPS merupakan pengembangan dari sistem IDS dengan kemampuan sistem memonitor atau mengawasi alitran trafik dan jika mengalami serangan akan langsung ditindaklanjuti secara otomatis [1] [2].

### 2.2. Distributed-Denial of Service (DDoS)

*Denial of Service* (DoS) adalah salah satu seranag yang ditakutkan di dunia internet. Dalam sebuah serangan *Denial of Service* (DoS) penyerang (*hacker*) akan mencoba untuk mencegah komputer lain atau client untuk mengakses suatu komputer atau jaringan dengan beberapa cara atau teknik , yaitu *traffic flooding* , *request flooding* dan mengganggu komunikasi antara sebuah host dan kliennya yang terdaftar dengan menggunakn banyak cara, termasuk dengan mengubah informasi konfigurasi sistem atau bahkan perusakan fisik terhadap komponen dan server. Sesuai dengan namanya *flooding traffic* dan *request flooding* ini bekerja dengan cara membanjiri suatu jaringan. *Traffic flooding* membanjiri lalu lintas jaringan yang datang dari pengguna yang terdaftar menjadi tidak dapat asuk ke dalam sistem jaringan. Sedangkan *Request flooding* membanjiri dengan banyak request terhadap sebuah layanan jaringan yang disediakan oleh sebuah *host* sehingga request yang dating dari pengguna terdaftar tidak dapat dilayani oleh layanan tersebut [1] [2] [3].

### 2.3. Flashcrowd

*Flash Crowd* bukanlah merupakan suatu serangan seperti *Denial of Service* (DoS) ataupun Distributed *Denial of Service*(DDoS) melainkan dimana situasi terjadinya sebuah peningkatan trafik yang sangat tinggi dalam suatu jaringan sehingga tidak dapat diakses dalam rentang waktu tertentu [1]. Peningkatan trafik ini terjadi karena user yang mengakses sebuah jaringan tersebut sangat banyak. Kejadian *flashcrowd* ini dapat terjadi kapan saja, karena peningkatan akses secara dramatis/tinggi ke suatu server dipengaruhi dari suatu kejadian seperti bencana alam, peluncuran produk, *breaking news*, dll [1].

### 2.4. ISODATA Clustering

Clustering merupakan salah satu teknik pengelompokan data berdasarkan kesamaan karakteristik data. Clustering-based memiliki beberapa tipe penting, diantaranya *Partitional Clustering*. *Partitional Clustering* merupakan pembagian data ke dalam sebuah himpunan data (*cluster*) yang tidak overlap sedemikian setiap data berada dalam satu *cluster* saja. Terdapat beberapa Algoritma dalam *partitional clustering* ini, diantaranya Algoritma ISODATA Algoritma. ISODATA (*Self-Organizing Data Analysis Technique*) deperkenalkan oleh Ball, Hall dkk pada sekitar tahun 1960an ialah clustering berbasis *unsupervised* algoritma yang pengembangan dari algoritma *K-Means* . Pengembangan di lakukan pada pada algoritma ISODATA ini penambahan proses pembagian, penggabungan, dan penghapusan cluster, Algoritma ISODATA mampu mengatur jumlah cluster fleksibel dan mengkontrol kepadatan suatu *cluster* [4] [6] [7] [8]. Algoritma ISODATA dapat membentuk cluster lebih optimal dalam melakukan analisis. Dalam algoritma terdapat masukan sistem yang mendukung kenerja algoritma ini, masukan tersebut ialah jumlah cluster awal (*k*), banyaknya iterasi (*i*), maksimum jarak untuk melakukan pengabungan (*minjar*), maksimum variansi (*var*) dalam melakukan pembelahan, serta minimum anggota sebuah cluster (*minjum*).

Agoritma ISODATA dapat membentuk suatu *cluster* dari sebuah data yang atribut - atributnya berasal dari tipe yang berbeda - beda, dengan cara mengubah atribut atribut tersebut ke dalam indeks *similarity* (kesamaan) atau *dissimilarity* (ketidaksamaan). Dengan algoritma ISODATA melakukan clustering akan lebih spesifik dan efisien [4]. Dalam menngunkan algoritma ISODATA data tidak perlu terdistribusi normal. Jika iterasi yang ditetapkan cukup, algoritma clustering ISODATA ini mudah untuk menemukan cluster yang benar dalam data. Namun, lebih banyak waktu komputasi yang dibutuhkan.

### 2.5. Manhattan Distance

*Manhattan Distance*/City Block Distance, merupakan salah satu teknik yang sering digunakan untuk menentukan kesamaan antara dua buah obyek. Pengukuran ini dihasilkan berdasarkan penjumlahan jarak selisih antara dua buah obyek dan hasil yang didapatkan dari Manhattan Distance bernilai mutlak [9]. Dimana, *Manhattan Distance* melakukan perhitungan jarak dengan cara tegak lurus.

$$\sum \qquad (1)$$

### 2.6. Dunn Index

Pada penelitian sebelumnya [10] [11], *Dunn Index* digunakan untuk validasi dan pelabelan sebuah *cluster*. Parameter yang digunakan dalam proses dunn index nilai jarak *inter cluster* dan nilai jarak *intra cluster*. Validasi *cluster* dikatakan baik jika nilai dari *dunn index* yang didapatkan tinggi. *Dunn index* mempunyai renta nilai dari 0 samapi ∞.

Formulasi dari *dunn index* $\{ \qquad \text{————————} \}$ (2)

### 2.7. Dataset DARPA 1999 dan World Cup 1998

Pada penelitian ini menggunakan *dataset* yang sudah sering digunakan pada peneletian serupa sebelumnya. Untuk pengujian *traffic* DDoS digunakan *dataset* DARPA 1998 [12] dan untuk pengujian *traffic flash crowd* [13] digunakan dataset World Cup 1998, kedua *dataset* tersebut sudah menjalani proses *preprocessing* agar mudah dianalisia. Untuk pengujian metode IDS yang dirancang, dilakukan simulasi menggunakan bahasa pemrograman Java

### 2.8. Masukan Sistem

Penentuan masukan sistem sebelum algoritma memulai prosesnya, ditentukan sesuai dengan beberpa kali melakukan percobaan dengan nialai yang bervariasi dan dengan beberapa jenis dataset dengan fitur – fitur yang berbeda nilainya. Masukan sistem tersebut dapt dilihat pada tabel 1 masukan sistem

Tabel 1 Masukan Sistem

| | Masukan | Keterangan |
|---|---|---|
| 1 | Jumlah *Cluster* (*k*) | Menentukan berapa banyak jumlah *cluster* awal untuk |
| 2 | Iterasi (*i*) | Menentukan berapa banyak jumlah Iterasi |
| 3 | Minimal jumlah data dalam satu | Menentukan jumlah minimal data dalam satu *cluster*, |
| 4 | Minimal jarak antara *Cluster* | Menentukan jarak minimal *centroid* / titik pusat *cluster*, |
| 5 | Minimal Variansi (*var*) | Menentukan angka minimal variansi satu *cluster*, |

### 2.9. Parameter Uji

Beberapa parameter yang digunakan digunakan untuk mengetahui sebereapa akuratnya algoritma isodata clustering dengan menggunakan euclidean distance dalam melakukan pembedaan antara traffik normal dan traffik anomali atau serangan. Beberapa parameter awal yang digunakan dalam mengukur keakuratan algoritma sebagai berikut :

Tabel 2 Matching matrix

| *ACTUAL* | *PREDICTION* | |
|---|---|---|
| | *ATTACK* | *NORMAL* |
| *ATTACK* | *TRUE POSITIVE (TP)* | *FALSE NEGATIVE (FN)* |
| *NORMAL* | *FALSE POSITIVE (FP)* | *TRUE NAGATIVE (TN)* |

*True positive* (TP) adalah kondisi dimana algoritma mendeteksi data sebagai serangan dan kelanjutan sebenarnya memang data tersebut merupakan serangan. *True nagative* (TN) adalah dimana algoritma mendeteksi data sebagai kondisi normal dan kenyataannya memang data tersebut merupakan data normal. *False positive* (FP) adalah dimana kondisi algoritma mendeteksi data dengan kondisi normal tetapi disebut sebuah serangan. *False negative* (FN) adalah kondisi dimana algoritma melakukan salah deteksi yang menyatakan data dengan kondisi serangan disebut segabagi kondisi normal.

**Detection Rate (DR)**

*Detection rate* merupakan presentase yang menyatakan seberapa besar algoritma dapat memberikan true alarm terhadap serangan yang terjadi. Formulasi untuk *detection rate* (DR) sebagai berikut :

( 2 )

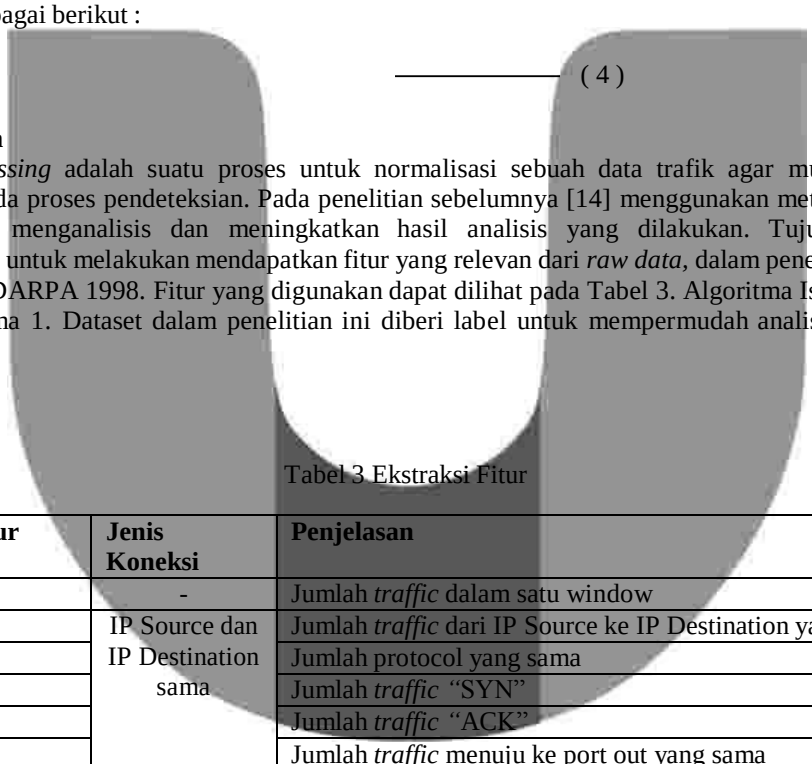
**False Positive Rate (FPR)**

*False positive rate* merupakan presentase yang menyatakan seberapa besar kesalahan algoritma memberikan false alarm, dimana algoritma mendeteksi sebuah kondisi serangan yang sebenarnya adalah kondisi normal. Formulasi untuk *false positive rate* (FPR) sebagai berikut:

( 3 )

**Akurasi (Acc)**

Akurasi merupakan presentase yang menyatakan seberapa benarnya algoritma melakukan pendekteksian, serta seberapa besar memisahkan data normal dan data serangan. Formulasi untuk akurasi algoritma sebagai berikut :



( 4 )

## 3. Pembahasan

*Preprocessing* adalah suatu proses untuk normalisasi sebuah data trafik agar mudah/cocok untuk digunakan pada proses pendeteksian. Pada penelitian sebelumnya [14] menggunakan metode *preprocessing* memudahkan menganalisis dan meningkatkan hasil analisis yang dilakukan. Tujuan dalam proses *preprocessing* untuk melakukan mendapatkan fitur yang relevan dari *raw data*, dalam penelitian ini dilakukan pada dataset DARPA 1998. Fitur yang digunakan dapat dilihat pada Tabel 3. Algoritma Isodata dapat dilihat pada Algoritma 1. Dataset dalam penelitian ini diberi label untuk mempermudah analisa hasil akhir yang didapatkan.

Tabel 3 Ekstraksi Fitur

| **Nama Fitur** | **Jenis Koneksi** | **Penjelasan** |
|---|---|---|
| Count | - | Jumlah *traffic* dalam satu window |
| IP_source | IP Source dan IP Destination sama | Jumlah *traffic* dari IP Source ke IP Destination yang sama |
| Protocol | | Jumlah protocol yang sama |
| SYN | | Jumlah *traffic* "SYN" |
| ACK | | Jumlah *traffic* "ACK" |
| Port_Out | | Jumlah *traffic* menuju ke port out yang sama |
| Length | | Jumlah *traffic* dengan length yang sama |
| Different_Source | IP Destination sama | Jumlah *traffic* dengan IP Source berbeda |
| New_IP | - | Jumlah kemunculan IP baru |

**Algoritma 1** : Proses Clustering dengan Isodata (*extracted dataset, k, i, minjum, minjar, var*)

1: Masukkan *extracted dataset*

2: Masaukan *k* sebagai jumlah *cluster* awal

3: Masukan *i* sebagai jumlah iterasi

4: Masukan *minjum* sebagai minimal jumlah anggota *cluster*

5: Masukan *minjar* sebagai minimal jarak antara centroid

6: Masukan *var* sebagai minimal variansi

7: Bentuk *k-centroid* secara acak sebanyak *k* buah

8: ***for*** 1 to *x* **do**

9: Hitung jarak $x_n$ ke *k-centroid*

10: Tetapkan $x_n$ ke *k-centroid* terdekat

11: ***end for***

12: Hapus $k_{kosong}$

13: ***repeat***

14: ***if*** *i* = ganjil ***do***

15: ***for*** 1 to *k* **do**

16: ***if*** jumlah *x* $k_n$ < *minjum* ***do***

17: Hapus k-centroid

18: *x* anggota *k-centroid* tetapkan ke *cluster* terdekat

19: ***end if***

20: Hitung jarak $k_n$-*centroid* ke $k_n$-*centroid* lainnya

21: ***if*** jarak $k_n$-*centroid* < *minjar* ***do***

22: Gabungkan kedua *k-cluster*

23: ***end if***

24: ***end if***

25: ***end if***

26: ***if*** *i* = genap ***do***

27: ***for*** 1 to *k* **do**

28: ***if*** jumlah *x* $k_n$ < *minjum* ***do***

29: Hapus k-centroid

30: *x* anggota *k-centroid* tetapkan ke *cluster* terdekat

31: ***end if***

32: Hiting variansi $k_n$

33: ***if*** variansi $k_n$ > var ***do***

34: Belah kn menjadi dua buah cluster baru

35: hitung *x* anggota $k_n$ dengan *k* baru

36: Tetapkan *x* anggota $k_n$ ke *cluster* baru yang terdekat

37: ***end if***

38: ***end if***

39: ***end if***

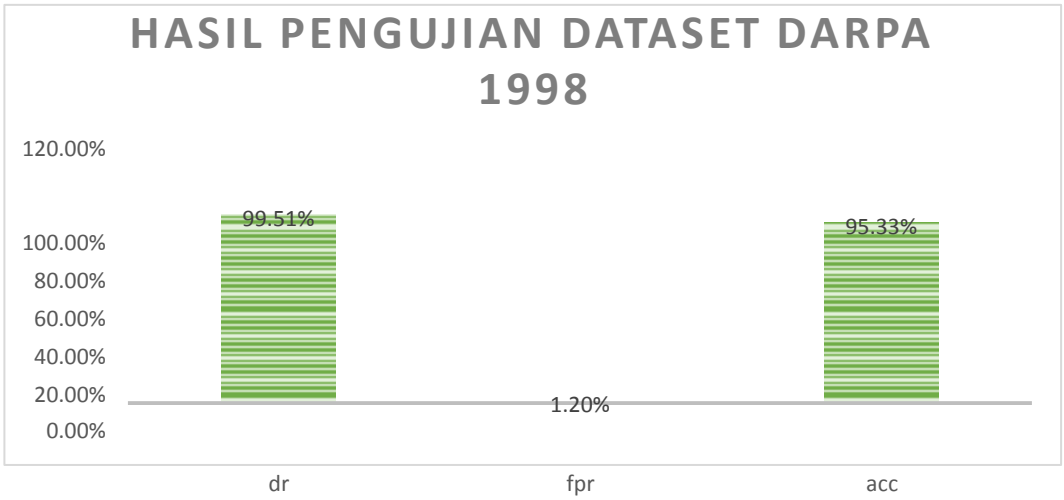
40: ***until iterasi terakhir***

41: Menghitung *DR*, *FPR*, dan Akurasi

42: Menghitung kualitas *cluster* dengan dunn indeks

### 3.1. Pengujian Dataset DDoS

Data dalam dataset DARPA 1998 ini masih berupa *raw data*. Maka sebelum diuji akan dilakukan proses *preprocesseing* terlebih dahulu. Tujuan dari proses *preprocessing* untuk medapatkan karekteristik dari *traffic* DDoS sehingga hasil performansi deteksi yang dihasilkan lebih baik. Pengujian dilakukan pada *dataset* normal dan serangan mendapati hasil yang beragam.hasil pengujian *dataset* DARPA 1999 dapat dilihat pada Gambar 1 dibawah.

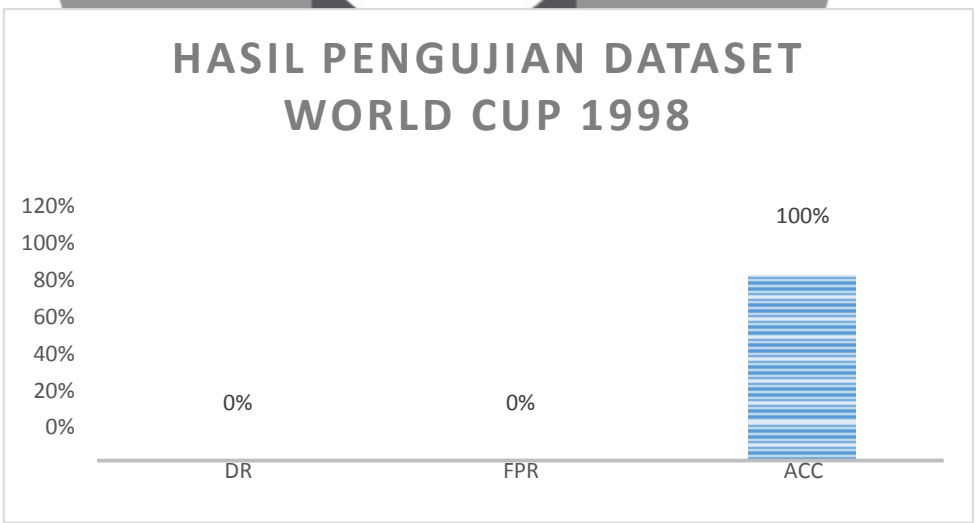



Gambar 1 Hasil Pengujian Dataset DARPA 1998

Pada pengujian *dataset* DARPA 1998, pengujian dilakukan berulang kali dengan masukan sistem yang berbeda. Dilihat dari hasil performansi sistem pada gambar 1 diatas dilihat hasil terbaik didapat dengan percobaan dengan masukan sistem sebagi berikut ,$k$ = 10, $i$ = 4, *minjum* = 1, *minjar* = 5, *var* = 0.5. Masukan sistem yang berbeda mempengaruhi dari hasil pengujian ini, namun hasil yang didapatkan tidak terlalu berbeda jauh. Rata rata waktu pemrosesan dataset DARPA 1998 ialah sekitar 41,283 detik. Dalam hal ini membuktikan sistem deteksi anomali menggunakan algoritma ISODATA dan *Manhattan distance* dapat mendeteksi adanya sebuah serangan dalam dataset dengan waktu yang cukup singkat.

### 3.2. Pengujian Dataset Flash Crowd

Dataset World Cup 1998 memuat memuat dua jenis trafik yaitu, trafik normal dan trafik *flash crowds*. Pelabelan *flash crowds* dilakukan pada saat trafik mulai menunjukan kenaikan jumlahnya. Pada pengujian dataset ini nilai masukan tidak mempengaruhi hasil akhir, akan tetapi mempengaruhi waktu eksekusi dataset oleh sistem deteksi. Dari hasil berbagai percobaan didapat hasil yang digambarkan pada gambar 2 dibawah.



Gambar 2 Hasil Pengujian Dataset World Cup 1998

Dapat dilihat dari gambar 2 diatas, hasil performansi sistem deteksi anomali dengan algoritma ISODATA dan *Manhattan distance* dapat bekerja dengan baik. Dapat dilihat sistem tidak melakukan kesalahan deteksi. *Cluster* yang dihasilkan berjumlah 1 buah, dimana trafik *flash crowd* dan trafik normal berada dalam 1 buah *cluster* karena kedua t*raffic* tersebut bukanlah sebuah serangan.

## 4. Kesimpulan

Sistem deteksi anomali trafik meggunakan algoritma ISODATA dengan metode pengukuran jarak *Manhattan Distance* baik untuk diterapkan, dilihat dari hasil performansi sistem yang sudah baik. Dalam penggunaan metode pengukuran jarak *Manhattan Distance*, waktu proses yang dihasilkan lebih singkat dibandingkan dengan metode *Euclidean Distance* dalam memproses sebuah dataset. Nilai DR, FPR, dan ACC juga dipengaruhi oleh masukan sistem yang bervariasi, dimana nilai yang dihasilkan tidak terlalu berbeda jauh.

Saran untuk penelitian selanjutnya, modifikasi sistem deteksi menggunakan algoritma ISODATA dengan penambahan metode *windowing* untuk pemotongan data agar mudah dalam mendeteksi. Gunakan metode pengukuran jarak lainnya seperti *Mahalanobis Distance* dan *Minkowski Distance*, yang diharapkan mendapatkan hasil performansi yang lebih baik.

## Daftar Pustaka

[1] Y. Purwanto, Kuspriyanto, Hendrawan dan B. Rahardjo, "Traffic Anomaly Detection in DDoS Flooding," *International Conference on Telecommunication Systems Services and Applications (TSSA)*, vol. 8, pp. 313-318, 2014.

[2] K. Ramadhani, M. Yusuf dan H. E. Wahanani, "Pendeteksian Dini Sserangan UDP Flood Berdasarkan Anomali Perubahan Ttraffic Jaringan Berbais Cusum Algoritm," *Computer security*, 2013.

[3] F. Kargl, J. Maier dan M. Weber, "Protecting Web Servers from Distributed Denial of Service Attacks," dalam *Proceedings of the 10th International Conference on World Wide Web*, ACM, 2001, pp. 514-524.

[4] M. Merzougui, M. Nasri dan B. Bouali, "Image Segmentation using Isodata Clustering with Parameters Estimated by Evolutionary Approach: Application to Quality Control," *International Journal of Computer Applications*, vol. 66, pp. 25-30, 2013.

[5] R. Xu dan D. Wunsch, "Survey of Clustering Algorithms," *Neural Networks, IEEE Transactions*, vol. 16, no. 3, pp. 645 - 678, 2005

[6] A. Kohei dan B. XianQianhg, "ISODATA clustering with parameter (threshold for merge and split) estimation based on GA: Genetic Algorithm," *Reports of the Faculty of Science and Engineering, Saga University*, vol. 36, pp. 17-23, 2007.

[7] N. MEMARSADEGHI, D. M. MOUNT, N. S. NETANYAHU dan J. L. MOIGNE, "A FAST IMPLEMENTATION OF THE ISODATA CLUSTERING ALGORITHM," *Int. J. Comput. Geometry Appl*, pp. 71-103, 2007.

[8] Seok-Woo Jang, Gye-Young Kim dan Siwoo Byun, "Clustering-Based Pattern Abnormality Detection in Distributed Sensor Networks," *International Journal of Distributed Sensor Networks*, 2014.

[9] D. Sinwar dan R. Kaushik, "Study of Euclidean and Manhattan Distance Metrics using Simple K-Means Clustering," *NTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)*, vol. 2, no. 5, 2014.

[10] S. Saitta, B. Raphael dan I. F. Smith, "A Bounded Index for Cluster Validity," *MLDM '07 Proceedings of*

*the 5th international conference on Machine Learning and Data Mining in Pattern Recognition,* pp. 174 - 187, 2007.

[11] F. Kovács, C. Legány dan A. Babos, "Cluster Validity Measurement Techniques," *AIKED'06 Proceedings of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases,* pp. 388 - 393 , 2006.

[12] "Cyber System and Technology," Lincoln Laboratory Messachusetts Institute of Technology, 4 December 1998. [Online]. Available: http://www.ll.mit.edu/ideval/data/. [Diakses 23 October 2014].

[13] P. Danzig, J. Mogul, V. Paxson dan M. Schwartz, "WorldCup98," ACM SIGCOMM, [Online]. Available: http://ita.ee.lbl.gov/html/contrib/WorldCup.html.

[14] Made Indra W P, Yudha P dan Fiky Y S, "DDoS Detection Using Modified K-Means Clustering with Chain Initialization Over Landmark Window," *International Conference on Control, Electronics, Renewable Energy, and Communications,* 2015.