



JITE (*Journal of Informatics and Telecommunication Engineering*)

Available online <http://ojs.uma.ac.id/index.php/jite> DOI : 10.31289/jite.v4i1.xxx

Received: 10 Juni 2020

Accepted: 07 Juli 2020

Published: 20 Juli 2020

Path Smoothing With Support Vector Regression

Donni Richasdy^{1)*} & Saiful Akbar²⁾

1) Teknik Informatika, Fakultas Informatika, Universitas Telkom, Indonesia
 2) Teknik Informatika, STEI, Institut Teknologi Bandung, Indonesia

*Coresponding Email:donnir@telkomuniversity.ac.id

Abstrak

Salah satu masalah dalam topik penelitian *moving object* adalah tidak lengkapnya data yang diperoleh oleh alat Geo-tracking. Fenomena ini dapat ditemukan dalam pelacakan pesawat dari darat yang data hilangnya mendekati 5 menit. Masalah ini memerlukan solusi *path smoothing* untuk melengkapi datanya. Salah satu metode *path smoothing* adalah dengan menggunakan fisika gerak, sedangkan *path smoothing* dalam penelitian ini menggunakan *machine learning* yaitu *Support Vector Regression* (SVR). Penelitian ini akan mengoptimalkan parameter konfigurasi SVR seperti *kernel*, *common*, *gamma*, *epsilon* dan *degree*. *Support vector regression* akan memprediksi nilai data yang hilang dari data pelacakan pesawat. Kami menggunakan kombinasi *mean absolute error* (MAE) dan *mean absolute error error* (MAPE) untuk mendapatkan akurasi. MAE akan menjelaskan nilai rata-rata kesalahan yang terjadi, sementara MAPE akan menjelaskan persentase kesalahan pada data. Dalam percobaan, nilai kesalahan terbaik MAE 0,52 dan MAPE 2,07, yang berarti data kesalahan $\pm 0,52$, ini sama dengan 2,07% dari nilai data keseluruhan.

Kata Kunci: *Moving Object, Path Smoothing, Support Vector Regression, MAE.*

Abstract

One of moving object problems is the incomplete data that acquired by Geo-tracking technology. This phenomenon can be found in aircraft ground-based tracking with data loss come near to 5 minutes. It needs path smoothing process to complete the data. One solution of path smoothing is using physics of motion, while this research performs path smoothing process using machine learning algorithm that is Support Vector Regression (SVR). This study will optimize the SVR configuration parameters such as kernel, common, gamma, epsilon and degree. Support Vector Regression will predict value of the data lost from aircraft tracking data. We use combination of mean absolute error (MAE) and mean absolute percentage error (MAPE) to get more accuracy. MAE will explain the average value of error that occurs, while MAPE will explain the error percentage to the data. In the experiment, the best error value MAE 0.52 and MAPE 2.07, which means error data ± 0.52 , this is equal to 2.07% of the overall data value.

Keywords: *Moving Object, Path Smoothing, Support Vector Regression, MAE.*

How to Cite: Richasdy, D. & Akbar, S. (2020). Path Smoothing With Support Vector Regression. *JITE (Journal Of Informatics And Telecommunication Engineering)*. 4 (1): 142-150

I. PENDAHULUAN

Perkembangan teknologi telekomunikasi dan sistem GPS memberikan kemampuan untuk melakukan tracking pada moving object secara terus menerus. Teknologi ini memungkinkan kita untuk mendapatkan data spatio-temporal dalam jumlah yang sangat banyak. Tidak hanya data yang berasal dari pergerakan manusia tetapi juga berasal dari sensor-sensor telemetri yang dibawa oleh hewan-hewan dan gejala-gejala alam seperti pergerakan awan, angin.

Data moving object berpotensi besar sebagai sumber informasi dan pengetahuan dalam mempelajari fenomena sosial seperti ekonomi, antropologi, kesehatan, pangan, transportasi, serta fenomena alam. Informasi seperti pola migrasi hewan, perubahan iklim, dan keseimbangan ekosistem sangat dibutuhkan oleh pihak berkepentingan sebagai dasar pengambilan keputusan untuk masalah seperti kemacetan, bencana alam, kerusakan ekosistem.. (I, Lychkov. 2018) (A, Nishad. 2017) (Z, Wang. 2017)

Untuk mendapatkan informasi dan pengetahuan yang terkandung dari data moving object, para peneliti mengembangkan berbagai metode visual analisis. Visual analisis dilakukan dengan

menggabungkan antara data mining dan visualisasi data sehingga memberikan informasi yang lebih intuitif terhadap cara manusia mengambil keputusan. (J, Lee. 2018) (Y, Yakufu. 2016) (G, Swalaganata. 2018) (K, Wai. 2018)

Masalah yang sering terjadi dalam data moving object adalah tidak lengkapnya data, noise dan rendahnya ruang sampel dari data yang bisa diakuisisi oleh teknologi geo-tracking. Masalah kualitas data akan membuat hambatan yang signifikan dalam ekstraksi informasi dan pengetahuan dari moving object.

Untuk memperbaiki kualitas data pada moving object dilakukan path smoothing. Path smoothing, sering disebut juga sebagai path planning dan path generation, memiliki beberapa pendekatan metode dalam pemecahannya. Beberapa teknik yang digunakan adalah sebagai berikut:

- a. *Machine learning*
- b. *Statistik*
- c. *Fisika benda bergerak*

Data *moving object* pada penelitian ini adalah data penerbangan di Indonesia. Data adalah *timeseries* dengan lima dimensi yaitu *latitude*, *longitude*, *altitude*, kecepatan dan arah. Penelitian akan mengkaji, menerapkan dan mengevaluasi *Support Vector Regression* dalam menyelesaikan masalah *path smoothing*.

II. STUDI PUSTAKA

Analisis path smoothing lebih sering ditemukan dalam penelitian dikomunitas mobile robot dan UAV, sebagai optimasi pergerakan. Path yang kasar akan menyebabkan gerak robot tiba-tiba berhenti karena robot perlu merubah orientasi rodanya kemudian melakukan percepatan lagi (L, Xu. 2017). Solusinya yang berupa proses perlambatan, mengubah orintasi dan percepatan ini akan mengkonsumsi energi yang tidak efisien (G, Swalaganata. 2018) Efek lain dari path yang kasar adalah slips (terpeleset) saat robot sedang bergerak.

Metode yang digunakan dalam menemukan solusi path smoothing adalah sebagai berikut :

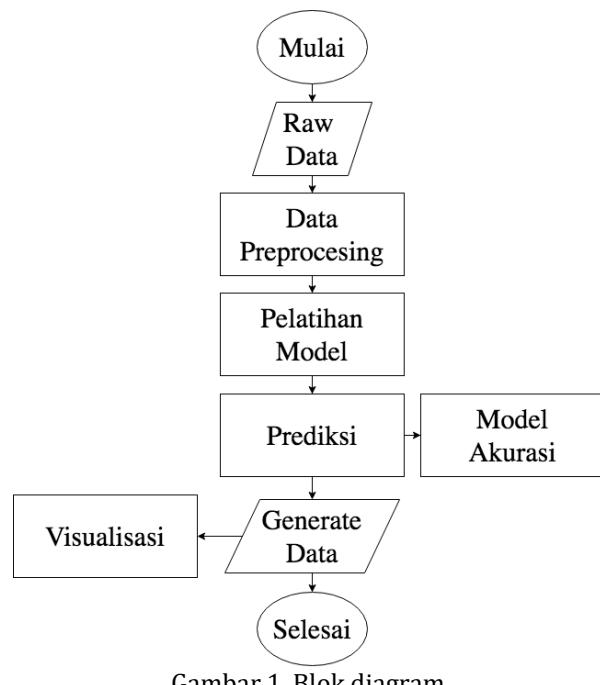
- Continuous-Curvature* (L, Xu. 2017) (R, Cimurs. 2017)
- Metode Geometri (L, Xu. 2017)
- η^3 splines generation (S, Yang. 2017)
- Canonical Curve* (G, Swalaganata. 2018)
- Arcs line and cubic spiral* ((G, Swalaganata. 2018)
- Fuzzy inference* (Y, Yan. 2016) (Q, Wu. 2016)
- Quadratic Programming* (A, Khan. 2016) (Y, Hao. 2017) (W, Jing. 2018)

Algoritma Cubic Bezier Curve dan algoritma line-arcs and cubic spirals adalah penyelesaian path smoothing

dengan menggunakan pendekatan matematika, sedangkan algoritma A* dan algoritma genetika adalah algorima searching. Dari studi literatur, masih belum banyak ditemukan solusi path smoothing dengan menggunakan pendekatan machine learning sehingga perlu dilakukan penelitian lebih lanjut di area ini.

III. METODE PENELITIAN

Metode yang diajukan dalam penelitian terdiri dari langkah proses Gambar. 1



Gambar 1. Blok diagram

Path smoothing terdiri dari empat proses utama, yaitu:

- Data preprocessing*: proses mengubah data mentah menjadi data yang berkualitas.

- b. Pelatihan model: proses pembentukan model dengan menggunakan *data training* sebagai *input*-nya. Model akan digunakan sebagai fungsi untuk membentuk nilai dari *input*-an yang dicari
- c. Prediksi: proses pembentukan *path* yang hilang dari data asli. Prediksi dilakukan untuk *data training* maupun *data testing*.
- d. Model akurasi: proses pengukuran kualitas model yang dibangun berdasarkan ukuran fungsi *error* yang telah ditentukan.
- e. Visualisasi: memproyeksikan data asli dan data hasil *path smoothing*.

A. Preprocessing

Langkah pertama adalah melakukan *parsing* data, menghapus atribut yang tidak diperlukan dan menghilangkan *noise* seperti *null* ataupun data yang tidak *valid*.

Preprocessing akan terlihat dari Tabel 1

Tabel 1. Attribut Dataset

Raw Data	Preprocessing Data
Timestamp	Timestamp
UTC	Latitude
Callsign	Longitude
Position	Altitude
Altitude	Speed
Speed	Direction
Direction	

B. Pelatihan Model

Training adalah proses pembentukan model yang akan digunakan sebagai predictor. Dalam penelitian ini proses

pembentukan model menggunakan algoritma *support vector regression*. Untuk mendapatkan model yang lebih baik dan menghindari *overfit* maka data akan dibagi menjadi data *training* dan *testing* dengan rasio 70:30

C. Prediksi

Prediksi adalah proses pembentukan data yang hilang dari data penerbangan. Berdasarkan analisis data, data yang hilang memiliki jeda sekitar 5 menit. Pada proses *prediction* jeda waktu akan diisi dengan data baru berdasarkan model *training* yang telah dibuat. *Prediction* akan menghasilkan data *moving object* yang memiliki *sample rate* yang lebih tinggi dibandingkan dengan data asli.

Tabel 2. Hasil Path Smoothing

Timestamp	Altitude	Speed	D	Lat	Long
1475763546	0.04572	262.984	99	-7.377182	112.770844
1475763545	0.04826	262.984	99	-7.377096	112.770296
1475763544	0.05080	262.984	99	-7.377010	112.769748
1475763543	0.05334	262.984	99	-7.376925	112.769199
1475763542	0.05588	262.984	99	-7.376839	112.768651
1475763541	0.05842	262.984	99	-7.376753	112.768103
1475763540	0.06096	262.984	99	-7.376667	112.767555
1475763539	0.06350	262.984	99	-7.376581	112.767026
1475763538	0.06604	262.984	99	-7.376496	112.766497
1475763537	0.06858	262.984	99	-7.376411	112.765968
1475763536	0.07112	262.984	98	-7.376326	112.765439
1475763535	0.07366	262.984	98	-7.376240	112.764910
1475763534	0.07620	262.984	98	-7.376155	112.764381

D adalah data *distance*, Lat adalah data *latitude* sedangkan Long adalah data *longitude*. Baris tabel tanpa arsir adalah data asli penerbangan, sedangkan baris

tabel dengan arsir adalah hasil prediksi data.

D. Model Akurasi

Dalam kasus regresi, terdapat dua fungsi umum yang digunakan untuk menentukan akurasi yaitu *loss function* atau *error function* dan *penalty function*. *Error function* adalah biaya yang diperlukan untuk sebuah in-akurasi dari masalah sebuah prediksi. Fungsi *penalty* adalah fungsi untuk mendapatkan nilai optimal suatu *constraint*, contoh seperti optimasi menggunakan *Genetic Algoritm*.

Dalam penelitian ini *digunakan error function* untuk mengidentifikasi akurasi model. Metode yang biasa digunakan dalam menghitung akurasi sebuah prediksi:

1. Mean Absolute Errors (MAE)

MAE adalah ukuran akurasi yang terikat pada data, artinya besaran MAE akan sangat tergantung dengan variansi data, semakin besar variansi data maka nilai MAE akan besar juga.

Untuk mendapatkan akurasi yang lebih general maka MAE saja tidak cukup, sehingga perlu digabungkan dengan ukuran akurasi lain yang bukan dari kelas yang terikat pada data. Ukuran akurasi yang tidak terikat pada data misalnya adalah *error persentase* atau lebih dikenal

dengan *mean absolute tercentage error* (MAPE).

2. Mean Absolute Percentage Error;

MAPE akan menjawab kekurangan MAE yang dianggap sangat bergantung dengan jumlah dan variasi yang dimiliki data. MAPE menambahkan fungsi pembagi dengan jumlah data x_t sehingga membuat MAPE lebih bebas.

E. Visualization

Visualisasi adalah salah satu bentuk model akurasi yang bisa digunakan untuk memverifikasi hasil prediksi data. Seperti pada Gambar. 2 adalah hasil plot untuk atribut *latitude* dan *longitude*.



Gambar 2. Jalur perpindahan

IV. HASIL DAN PEMBAHASAN

Hasil dari percobaan akan dijelaskan pada bab ini, sesuai dengan rancang blok diagram.

A. Skenario Percobaan

Penelitian ini menggunakan data *moving object* yang didapatkan dari flightradar24.com. Data penerbangan yang diambil adalah data penerbangan khusus untuk wilayah Indonesia,

penerbangan semua maskapai Indonesia dengan rute Jakarta-Surabaya. Data penerbangan diambil dari Juli 2016 sampai dengan Desember 2016. Data terdiri dari 8 maskapai, selama 6 bulan, setiap bulan ada 1 penerbangan dalam seharinya, sehingga ada 1440 penerbangan yang disimpan dalam file csv. Masing-masing penerbangan memiliki 200 record, sehingga total data menjadi 288.000 record moving object.

Eksperimen SVR akan dilakukan beberapa kali dengan memvariasikan nilai parameter SVR untuk mengetahui perubahan atau pengaruhnya terhadap performa teknik yang diimplementasikan dalam tool. Berikut ini adalah parameter tersebut:

a. *Common (C)*.

b. *Gamma*

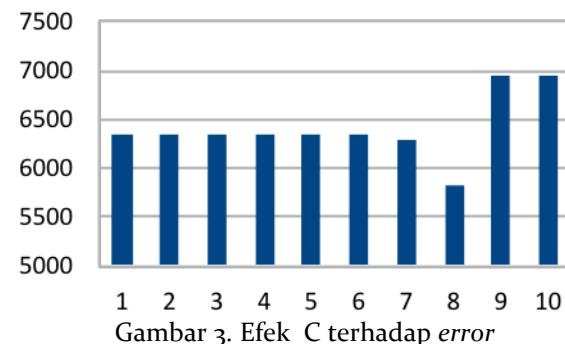
c. *Epsilon*

Percobaan dilakukan untuk melihat pengaruh masing-masing konfigurasi SVR terhadap model yang dibentuk berdasarkan akurasi model yang dihasilkan. Konfigurasi SVR terbaik didapatkan dengan melakukan percobaan tertutup dari parameter konfigurasi

B. *Percobaan Common (C)*

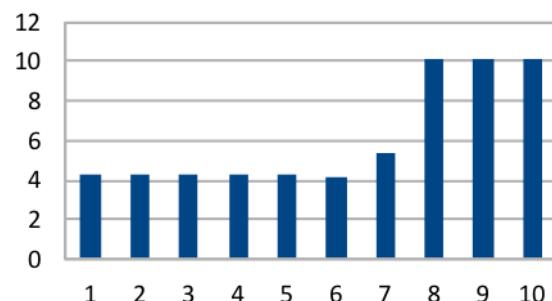
Gambar. 3 adalah pengaruh C terhadap model prediksi yang dihasilkan. Terlihat dari grafik bahwa tren yang dihasilkan memiliki akurasi yang semakin

baik jika nilai C semakin kecil. Percobaan diatas menggunakan nilai C 1e-10 untuk data altitude penerbangan.



C. *Percobaan Gamma*

Gambar. 4 adalah pengaruh *gamma* terhadap model prediksi yang dihasilkan. Terlihat dari grafik bahwa tren yang dihasilkan memiliki akurasi yang semakin baik jika nilai *gamma* semakin kecil. Percobaan diatas menggunakan nilai *Gamma* 1e-10 untuk data *direction* penerbangan

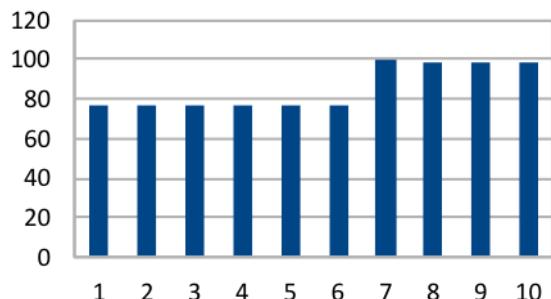


Gambar 4. Efek gamma terhadap error

D. *Percobaan Epsilon*

Gambar. 5 adalah pengaruh *epsilon* terhadap model prediksi yang dihasilkan. Terlihat dari grafik bahwa tren yang dihasilkan memiliki akurasi yang semakin baik jika nilai epsilon semakin kecil.

Percobaan diatas menggunakan nilai C 1e-10 untuk data *speed* penerbangan.



Gambar 5. Efek epsilon terhadap error

Kernel yang umum digunakan oleh *support vector regression* adalah RBF, *polinomial*, *linier*, *sigmoid*. Pemilihan *kernel* sangat tergantung dengan jenis data yang digunakan, misalkan untuk data *linier* kita bisa menggunakan *kernel linier* agar SVR bisa melakukan prediksi yang lebih tepat. Berdasarkan persebaran datanya, SVR lebih sesuai menggunakan *kernel RBF*.

E. Hasil

Tabel 3. Konfigurasi SVR Terbaik untuk MAE

Data	G	K	E	C	1	2	3	4	5
Altitude	1e-10	rbf	1e-10	1e-10	0	6334.07	2533.33	1891.86	1763.32
Direction	1e-10	rbf	1e-10	1e-10	0	4.3	4.67	5.42	5.19
Latitude	1e-10	rbf	1e-10	1e-10	0	0.52	0.13	0.49	0.1
Longitude	1e-10	rbf	1e-10	1e-10	0	2.43	0.72	2.2	0.61
Speed	1e-10	rbf	1e-10	1e-10	0	76.42	50.67	52.08	28.67

Tabel 4. Konfigurasi SVR Terbaik untuk MAPE

Data	G	K	E	C	1	2	3	4	5
Altitude	1e-10	rbf	1e-10	1e-10	0	240.52	89.29	250.45	4.01
Direction	1e-10	rbf	1e-10	1e-10	0	22.51	22.51	15.47	7.77
Latitude	1e-10	rbf	1e-10	1e-10	0	6.98	5.38	6.45	3.33
Longitude	1e-10	rbf	1e-10	1e-10	0	2.07	1.51	1.72	0.92
Speed	1e-10	rbf	1e-10	1e-10	0	34.38	24.9	19.6	5.48

Table 3 dan Table 4, G adalah gamma, K adalah kernel, E adalah epsilon, 1 adalah mean fit time, 2 adalah mean test score, 3 mean training score, 4 standar deviasi test score, 5 standar deviasi training score.

Dari Table 3 dan Table 4 ditemukan bahwa konfigurasi terbaik dari model prediksi pada nilai gamma, epsilon dan C yang paling kecil. Nilai error yang di dapatkan untuk setiap atribut data berbeda dikarenakan masing-masing data memiliki variansi yang berbeda.

V. SIMPULAN

Penelitian ini telah melakukan proses *path smoothing* dengan menggunakan model *Support Vector Regression*. Konfigurasi terbaik dari SVR adalah pada kernel RBF dengan nilai *gamma*, *epsilon*, dan C paling kecil, hal ini memiliki semantik bahwa ; *hyperplane* lebih lurus untuk mencegah *overfit*, semua sampel yang digunakan dalam proses training dan nilai toleransi *error* harus kecil. Nilai akurasi terbaik yang didapatkan adalah :

Tabel 5. Konfigurasi SVR Terbaik

Error	Train time	Test accuracy	Train accuracy
MAE	0	0.52	0.13
MAPE	0	2.07	1.51

Dari Table 5 bahwa nilai error paling baik dihasilkan adalah 0.52 yang berarti 2.07% dari nilai keseluruhan

Support Vector Regression terdiri dari 4 konfigurasi utama yaitu kernel,

epsilon, gamma dan C. Masing-masing konfigurasi memiliki pengaruh besar dalam pembentukan model. Semantik dari konfigurasi adalah ; kernel sangat berpengaruh terhadap distribusi data yang kita gunakan, gamma merupakan fleksibilitas hyperplane yang dibuat, epsilon menentukan jumlah sampel data yang digunakan pada training dan C merupakan toleransi error yang akan diadaptasikan

Untuk rekomendasi penelitian selanjutnya bisa melakukan perbandingan terhadap metode prediksi lain untuk meningkatkan akurasi dan melakukan pelatihan dengan *multi attribut data*.

DAFTAR PUSTAKA

- A. Khan, I. Noreen and Z. Habib, "Coverage Path Planning of Mobile Robots Using Rational Quadratic Bézier Spline," 2016 International Conference on Frontiers of Information Technology (FIT), Islamabad, 2016, pp. 319-323, doi: 10.1109/FIT.2016.065. (A, Khan. 2016)
- A. Nishad and S. Abraham, "Semantic trajectory analysis for identifying locations of interest of moving objects," 2017 International Conference on Networks & Advances in Computational Technologies (NetACT), Thiruvananthapuram, 2017, pp. 257-261, doi: 10.1109/NETACT.2017.8076776. (A, Nishad. 2017)
- G. Swalaganata, Muniri and Y. Affriyenni, "Moving object tracking using hybrid method," 2018 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, 2018, pp. 607-611, doi: 10.1109/ICOIACT.2018.8350740. (G, Swalaganata. 2018)
- I. I. Lychkov, A. N. Alfimtsev and S. A. Sakulin, "Tracking of Moving Objects With Regeneration of Object Feature Points," 2018 Global Smart Industry Conference (GloSIC), Chelyabinsk, 2018, pp. 1-6, doi: 10.1109/GloSIC.2018.8570061. (I, Lychkov. 2018)
- J. Lee, B. Hong, W. C. Kim and A. Gedeon Mengara Mengara, "Generation of Data Set for Tactical Moving Objects," 2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, 2018, pp. 189-192, doi: 10.1109/AIKE.2018.00043. (J, Lee. 2018)
- K. P. Wai and T. Nwe Aung, "Distance-based Clustering of Moving Objects' Trajectories from Spatiotemporal Big Data*," 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), Singapore, 2018, pp. 567-572, doi: 10.1109/ICIS.2018.8466465. (K, Wai. 2018)
- L. Xu, D. Wang, B. Song and M. Cao, "Global smooth path planning for mobile robots based on continuous Bezier curve," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 2081-2085, doi: 10.1109/CAC.2017.8243114. (L, Xu. 2017)
- Q. Wu, Y. Lu and X. Liu, "Real-Time globally optimized path planning in a dynamic environment combing artificial potential field and fuzzy neural network," 2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, 2016, pp. 114-118, doi: 10.1109/ICCWAMTIP.2016.8079817. (Q, Wu. 2016)
- R. Cimurs, J. Hwang and I. H. Suh, "Bezier Curve-Based Smoothing for Path Planner with Curvature Constraint," 2017 First IEEE International Conference on Robotic Computing (IRC), Taichung, 2017, pp. 241-248, doi: 10.1109/IRC.2017.13. (R, Cimurs. 2017)
- S. Yang, Z. Wang and H. Zhang, "Kinematic model based real-time path planning method with guide line for autonomous vehicle," 2017 36th Chinese Control Conference (CCC), Dalian, 2017, pp. 990-994, doi: 10.23919/ChiCC.2017.8027474. (S, Yang. 2017)
- W. Jing, D. Feng, P. Zhang, S. Zhang, S. Lin and B. Tang, "A Multi-Objective Optimization-based Path Planning Method for Parallel Parking of Autonomous Vehicle via Nonlinear Programming," 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore, 2018, pp. 1665-1670, doi: 10.1109/ICARCV.2018.8581195. (W, Jing. 2018)
- Y. Hao, B. Li, L. Shao, Y. Zhang and J. Cui, "Multi-objective path planning for unmanned aerial

- vehicle based on mixed integer programming," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 7035-7039, doi: 10.1109/CAC.2017.8244046. (Y, Hao. 2017)
- Y. Yakufu and C. E. Atay, "A data mining application on moving object data," 2016 Eleventh International Conference on Digital Information Management (ICDIM), Porto, 2016, pp. 91-95, doi: 10.1109/ICDIM.2016.7829757. (Y, Yakufu. 2016)
- Y. Yan and Y. Li, "Mobile robot autonomous path planning based on fuzzy logic and filter smoothing in dynamic environment," 2016 12th World Congress on Intelligent Control and Automation (WCICA), Guilin, 2016, pp. 1479-1484, doi: 10.1109/WCICA.2016.7578767. (Y, Yan. 2016)
- Z. Wang, "Research on Moving Objects Trajectories Collection Based on Data Mining," 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA), Changsha, 2017, pp. 538-541, doi: 10.1109/ICSGEA.2017.8244046. (Z, Wang. 2017)