Dipartimento di Informatica dell'Università di Bologna Corso di Laurea Magistrale in Informatica – Curriculum A

Relazione Finale di Tirocinio

Enhancement Algorithms for Railway Localization based on GNSS with RTK

Svoltosi presso *Sadel S.p.A.* dal *20/06/2023* al *01/08/2023*

Tirocinante: Zhiguang Li

Matricola: 0001029938

Tutor aziendale: Ing. Davide Amato

Tutor Università: Prof. Alessandro Vanelli-Coralli

Tutor operativo Università: Dr.ssa Carla Amatetti

Tutor operativo Università: Dr. Riccardo Campana

Anno Accademico 2022-2023

The host structure and the context

The internship was carried out at Sadel.S.p.A., located in Via M. Serenari, 1 40013 Castel Maggiore (BO), under the supervision of the Academic tutors Prof. Alessandro Vanelli-Coralli, Dr. Carla Amatetti, Dr. Riccardo Campana and the corporate tutor Eng. Davide Amato, in a mixed mode of on-site and remote work.

The aim of the internship is to research and develop a novel sensor fusion algorithm, to integrate the collected IMU data with GNSS and RTK, and to use the algorithm to solve the measurement errors caused by various factors in the case of combining GNSS and RTK. This internship will carry out a preliminary classification of the problems in the data, use machine learning and neural network methods to classify the causes of the dynamic test results using GNSS and RTK sensor fusion on the railway lines of Formigine and Modena.

Internship by analyzing the data results of the dynamic test using GNSS and RTK sensor fusion on the railway lines of Formigine and Modena, comparing the number of satellites, accuracy, latitude and longitude position information of U-center, RTK correction information and road environment conditions of Google Earth, to classify the errors of accuracy in the data. Through the method of artificial intelligence, the system can automatically identify and classify. The classification results of the current research are the following three situations: a. Clear view; b. Multipath correction; c. Tunel - the situation where the train is running in a tunnel or underground.

Based on the research of the paper "*Towards the Future Generation of Railway Localization Exploiting RTK and GNSS*", we have come to the following conclusions after completing the performance tests of GNSS and RTK under static and dynamic conditions:

In static conditions, centimeter-level accuracy can be achieved using RTK if there is a clear view of the sky, while in urban environments multipath has a significant impact on positioning accuracy, limiting or even degrading the achievable accuracy. In the static test of the metal plate with signal shielding, under the premise of keeping the number of GNSS satellite signals during reception constant, we verified that the positioning accuracy decreases when the view of the sky is blocked. The floating range of accuracy is described in detail in the paper.

Under the conditions of the dynamic test, the researchers conducted a dynamic test on the railway line between Formigine and Modena in Emilia-Romagna, Italy, using a custom sensor node to collect GNSS, RTK and IMU measurement data in different environments. At the same time, the horizontal accuracy and satellite coverage of the GNSS data with RTK will be evaluated. In an environment without long bridges and tunnels and with a clear view of the sky, the accuracy of the RTK and GNSS modules is on average within 16 cm. However, with long bridges and tunnels, the number of GNSS satellites changes. Especially in tunnels, the number of available GNSS satellites is close to 0-1.

Using the U-centre to analyze the data set of this dynamic test, when GNSS is not available or the angle of accuracy is poor, the additional RTK sensor brings some help for accurate positioning. At the same time, there are limitations and the latter can even reduce the positioning accuracy compared to GNSS alone in unfavorable environmental conditions.

Therefore, the development of an algorithm that integrates sensors and navigation and positioning systems and combines collected IMU data, RTK data and GNSS data will become the basis for future work.

The programme of this study is a preliminary fusion algorithm, which aims to combine the test data, including the number of GNSS satellites and positioning accuracy, by using neural network and supervised learning algorithms, and classify the data according to the situation. It can be divided into three situations: clear sky, multi-path GNSS positioning and long bridge or tunnel.

The activity performed

During the meeting and continuous communication with Eng. Davide Amato, Dr. Carla Amatetti, Dr. Riccardo Campana, we have defined the internship plan, the report is as follows：

1. Learn "*Towards the Future Generation of Railway Localization Exploiting RTK and GNSS* ".
2. Learn the basics of GNSS and RTK.
3. Use U-center and Google Earth to analyze the different scenarios of GNSS errors.
4. Learn the basics of Machine Learning.
5. Develop the algorithm (in Python programming language).
6. Train the model and evaluate the algorithm and do a presentation.

It took us about 2 weeks to complete steps 1, 2 and 3, during which we met each other with Eng. Davide Amato 2-3 times a week at SADEL. I spent 1 week studying the Machine Learning course in Google Developer. I spent 2 weeks developing the algorithm of the project in Python and training the model with final evaluation of the algorithm. I spent the final week writing the report.

We have four meetings online and face to face with Dr. Carla Amatetti and Dr.Riccardo Campagna. Otherwise, we discuss questions with MS Teams.

GNSS

GNSS (Global Navigation Satellite System) It is a technology system that utilizes satellite signals to provide global positioning and navigation services. It can achieve accurate positioning and navigation for location on Earth by involving satellite constellations and receiver devices working together.

GNSS constellations list:

1. GPS (Global Positioning System) is developed and operated by the US government. (1970-now), has 24 satellites.
2. The BeiDou Compass System. is developed and operated by China. (2000-2018), has 35 satellites.
3. GLONASS is developed and operated by Russia. (1970-now) , has 24 satellites.
4. Galileo is developed and operated by ESA. Europe (2000-2020) , has 22 satellites.

Accuracy for GPS, DGPS, GPS with RTK-Correction

1. DGPS stands for differential GPS, and represents use of ground reference situations to send differential correction data to correct data from GPS.
2. GPS with RTK-Correction stands use real time kinematic technico to correct the data from GPS.
3. The CEP accuracy of standard GPS positioning in general is 5-10 meters.
4. The CEP accuracy of standard DGPS positioning in general is 0.5-3 meters.
5. The CEP accuracy of GPS with RTK is 1-5 centimeters.

The way to improve the accuracy of GNSS localization

1. the smaller Geometric Dilution of Precision (DoP) value and the higher accuracy

2. receiver performance: the better performance of the chip, and the higher accuracy.

3. the method of measure: combine other sensors like RTK, IMU.

RTK

Real-time kinematic positioning (RTK) is the application of surveying to correct for common errors in current satellite navigation (GNSS) systems. It uses measurements of the phase of the signal's *carrier wave* in addition to the information content of the signal and relies on a *single reference station* or *interpolated virtual station* to provide real-time corrections, providing up to centimeter-level accuracy.

RTK uses a fixed base station and a rover to reduce the rover's position error. The base station transmits correction data to the rover.

What is the improvement with the RTK correction method to the standard GNSS measure?

1. Improved positioning accuracy.Standard GNSS positioning accuracy is generally 0.5m - 10m. (Depending which satellite constellations and which device we used. Standard GPS generally is 5-10 meters, DGPS generally is 0.5 - 3 meters. When we use RTK positioning accuracy is centimeter.)

2. High real-time performance. RTK can obtain high-precision positioning results in real time.

3. Enhanced anti-interference capability. RTK technology has a strong ability to correct atmospheric effects, satellite clock errors.

4. Low cost. GNSS receivers can get position centimeters and need at least 5 satellites. And the cost of the device is more expensive. We can use RTK through software and algorithms to improve accuracy.

5. Continuous measurement possible. RTK can use the moving environment, even if the atmospheric effects signal of GNSS, and RTK still run normally.

Accuracy

When will there be poor satellites?

There are a few situations where GPS satellites may provide poor coverage:

1. High latitudes

2. Urban canyons

3. Signal blockage: Objects such as tall buildings, trees and tunnels can block satellite signals, resulting in fewer satellites available.

Poor satellites on the railway, there are some situations like trees and tunnels and under the bridge.

What is multipath interference?

In radio communication, multipath is the propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths. When the same signal is received over more than one path, it can create interference and phase shifting of the signal. Destructive interference causes fading; this may cause a radio signal to become too weak in certain areas to be received adequately. For this reason, this effect is also known as multipath interference.

In this picture, the position of the left car will reflect two or three radios to sensors, so it will be an accuracy error.

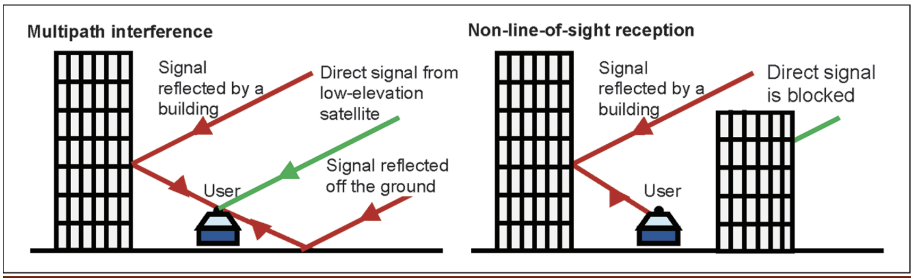


FIGURE 1 Multipath interference and NLOS reception

U-centre

u-center is highly intuitive GNSS evaluation software that is easy to use, personalized, and compatible with leading u-blox technologies.

Some parameters of the dataset

Longitude - The east-west position in degrees (-180 to 180)

Latitude - The north-south position in degrees (-90 to 90)

Altitude - The height above mean sea level in meters

Altitude ellipsoidal - the elevation above a mathematical model that approximates the shape of the earth.

TTFF - Time to First Fix - The time taken to get the initial position fix

Fix Mode - The type of position fix. RTK Fixed or No Fixed, FLOAT-RTK FLOAT

3D Acc. - The 3D position accuracy in meters

2D Acc. - The 2D position (horizontal) accuracy in meters

PDOP - Position Dilution of Precision - A measure of satellite geometry

HDOP - Horizontal Dilution of Precision

Satellites - The number of satellites used in the position fix

Code Analysis

Code on github: https://github.com/davide-li/co-gnss-rtk

1. Data Processing

2. Model Construction

   Construct a simple fully connected neural network with 4 dense layers. Each layer uses the ReLU activation function, and the last layer uses softmax output. The input layer accepts 2 features, and the output layer outputs the probability of 3 categories.

3. Train the model

4. Prediction and Result Processing

   Make predictions on the test set and get the probability that each sample belongs to 3 categories.Convert the probabilities into one-hot encoded predicted labels.Write the prediction results back to the test set DataFrame.Display and export the test set results to CSV file.

   Using row.argmax(), we can get the index of the maximum value of each row, the predicted category. Using this index, create an array pred with all 0s, and set the corresponding index position to 1. Add pred to the list preds, preds contains all one-hot encoding results. Finally, assign preds to the column preds of the test set test_df to complete the conversion.

   The whole process realizes the transition from softmax probability to one-hot encoding.

5. Evaluate the Model

   Use the confusion matrix to evaluate the result of the prediction and the true result.

Model Training

1. Why do I use the ReLU for the dense layer and Softmax for the output layer?

   - ReLU is a linear rectification function, mainly used in the hidden layer of the neural network. It can introduce nonlinearity and help the model learn complex function mappings.

   - Softmax is mainly used in the output layer of classification problems. It converts the outputs values of multiple neurons into probability distributions for multi-classification problems.

   - ReLU is usually used in the hidden layer to extract features.

   - Softmax is usually used in the output layer for classification.

2. Using categorical_crossentropy as the loss function

   - How to determine our loss function:

   First, according to our needs, we can know that our scenario is a classification problem, so we exclude the use of loss functions for other regression problems;

   Second, according to the classification results, we divide the final results into three categories, so we exclude the use of binary_crossentropy, we use categorical_crossentropy;

   - Categorical_crossentropy is a common loss function for multi-classification tasks.

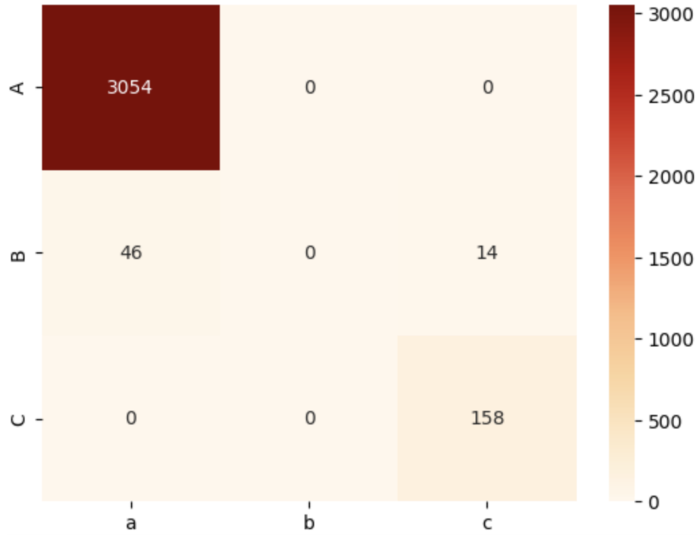
   It is suitable for multi-category problems where the categories are mutually exclusive, i.e., each sample belongs to only one category and does not belong to more than one category at the same time.

3. Using accuracy as an evaluation metric

Evaluate Model

Use the confusion matrix to evaluate the result of the prediction and the true result.

This graph is the result of Evaluate.



The skills acquired

Through this internship, I have mastered the basic knowledge related to GNSS and RTK, and have the ability to directly analyze the causes of corresponding data accuracy problems through data. U-center, Google Earth and other software can be used, and accuracy data can be analyzed on these software.

In addition, it is possible to analyze and optimize GNSS and RTK data sets by using machine learning and neural network methods, and implement preliminary problem classification algorithms. The accuracy of the prediction results can be evaluated using the confusion matrix.

Finally, I would like to thank Eng. Davide for his full help, seriousness, patience and wisdom to help me learn the knowledge related to GNSS, so that I have a solid foundation of the concepts of GNSS. Thanks to Prof. Alessandro for arranging two doctors, Dr. Carla and Dr. Riccardo, to help me with the machine learning algorithm. Special thanks to SAF and Sadel for giving me this internship opportunity.

Data

Firma del tirocinante

Firma del tutor aziendale (per i tirocini in azienda)