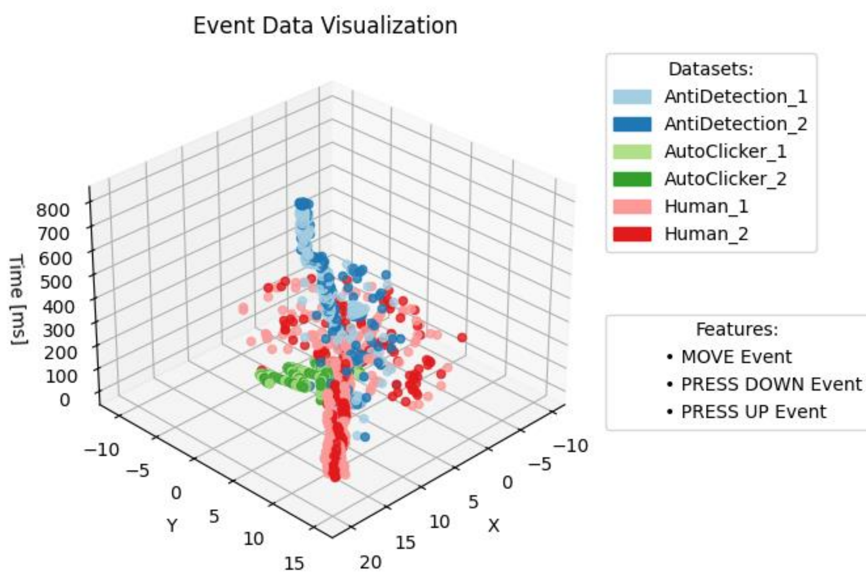
ANALYSIS PROJECT

BOT DETECTION FROM SENSOR DATA



IVAN SIČAJA

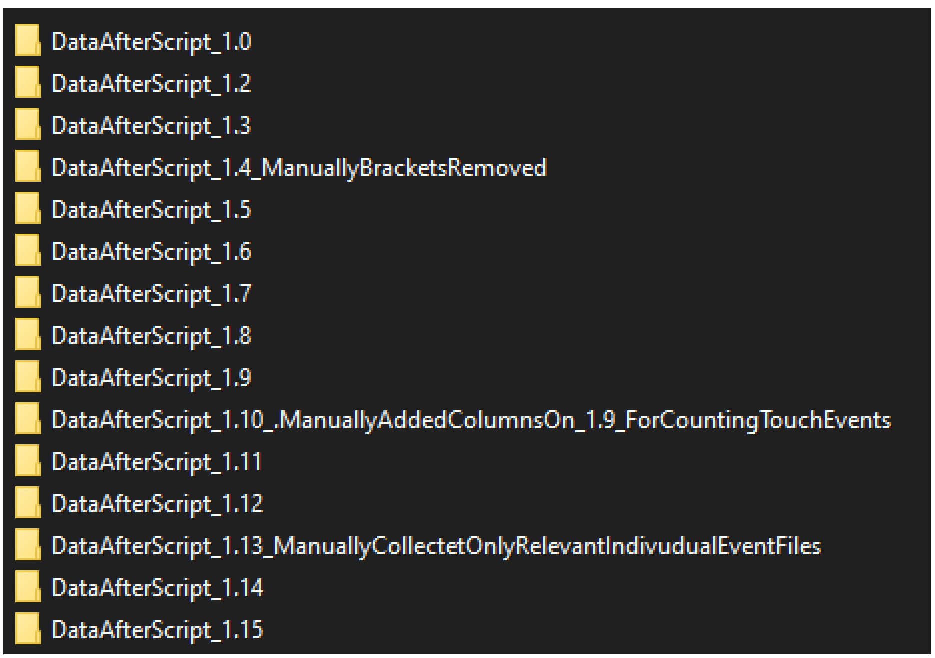
Stuttgart, August 2024.

1 Introduction

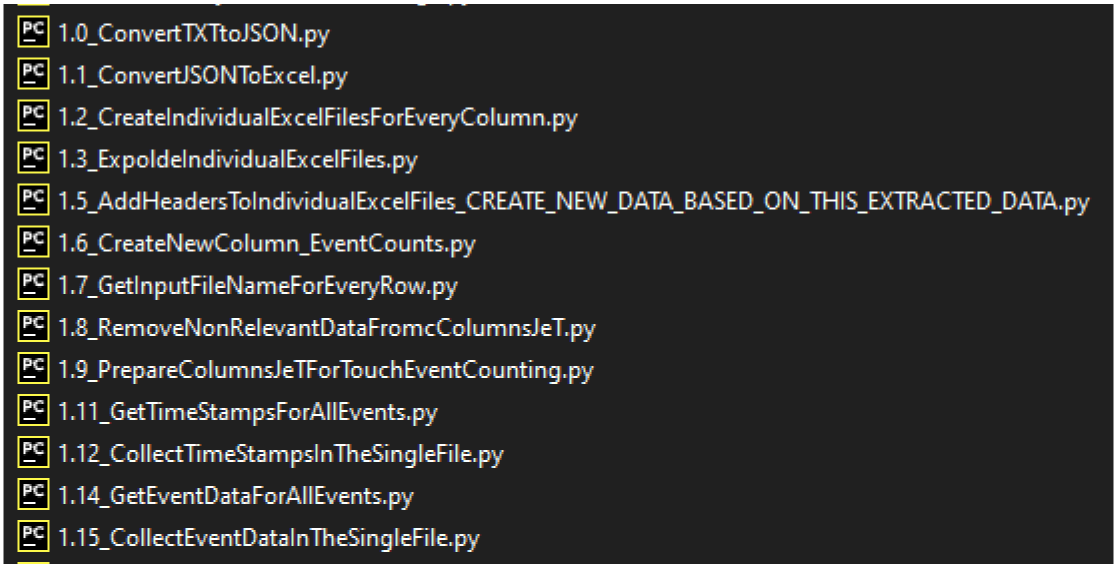
Original data is converted from ".json" format to the row-column data. This approach is chosen because the data is small and extracting relevant data will be possible to do with Python scripts together with the usage of the module "pandas".

All Python scripts are executed locally to achieve time efficiency.

Output data of every script is stored in the folder at the following path: "0_Data/1_Unzipped/1_OutputData", with corresponding folder names as shown in the following image.



The converting .json data to te row-column data is done by the Python scripts displayed in the next image.



- All code files are attached to a directory in order to improve the document's readability.
- All scripts are located in the root directory.
- Tableau workbook is also attached to the directory with the name :
- „Events count.twbx“

2 Tasks

2.1 Event types contribution to bots detection

- Battery Level: Not very useful for bot detection, as it doesn't directly relate to user interaction.
- Accelerometer Data: Human interactions tend to show natural variability in movement, while bots might generate repetitive or less variable patterns.
- Magnetic Data: Similar to accelerometer data, magnetic data can reflect device movement, which might help in detecting non-human patterns.
  (Data is missing!)
- Gyroscope Data: This data provides insights into device orientation changes. Human-induced movements are generally more varied compared to bots.
- Proximity Data: This could be useful if the game interaction involves proximity sensors, indicating how close a user holds the device.
  (Data is missing!)
- Light Sensor Data: Less likely to be useful unless there's a specific game mechanic involving changes in light.
  (Data is missing!)
- Pressure Sensor Data: Might indicate forcefulness of touches, potentially different for bots versus humans.
- Temperature: Unlikely to be useful for bot detection.
  (Data is missing!)
- Step Detector Data: Relevant if the game involves movement that affects the step count.
  (Data is missing!)
- Geomagnetic Rotation: Helps in understanding device orientation, potentially useful similar to gyroscope data.
  (Data is missing!)
- Key Events: (Data is missing!)
- Touch Events: Crucial for detecting bots. Patterns in touch events (like consistent timing or location of touches) can be strong indicators of bot activity.

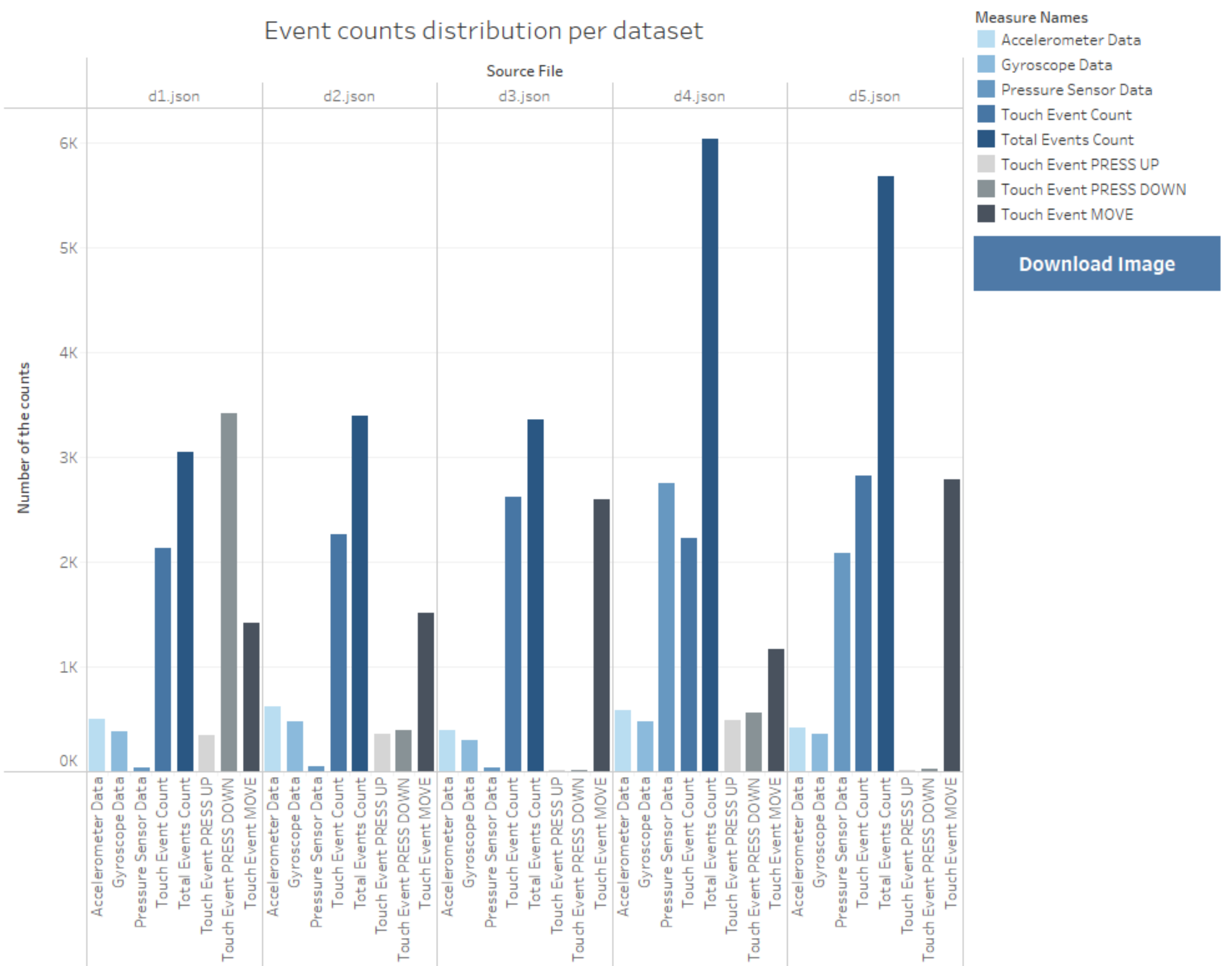
Conclusion:

Only 4 event groups are relevant for the further analysis. The main focus is analyzing the touch events data.

2.2 Event counts distribution per type

After the row-column data is created, it is easy with the python-pandas code, built-in Excel, or Tableau functions to count how many times we have the corresponding event occurred. Generic names for 5 datasets are used: “d1.json”,”d2.json”…

The results image from the Tableau is shown on the following image.



2.3 Plots

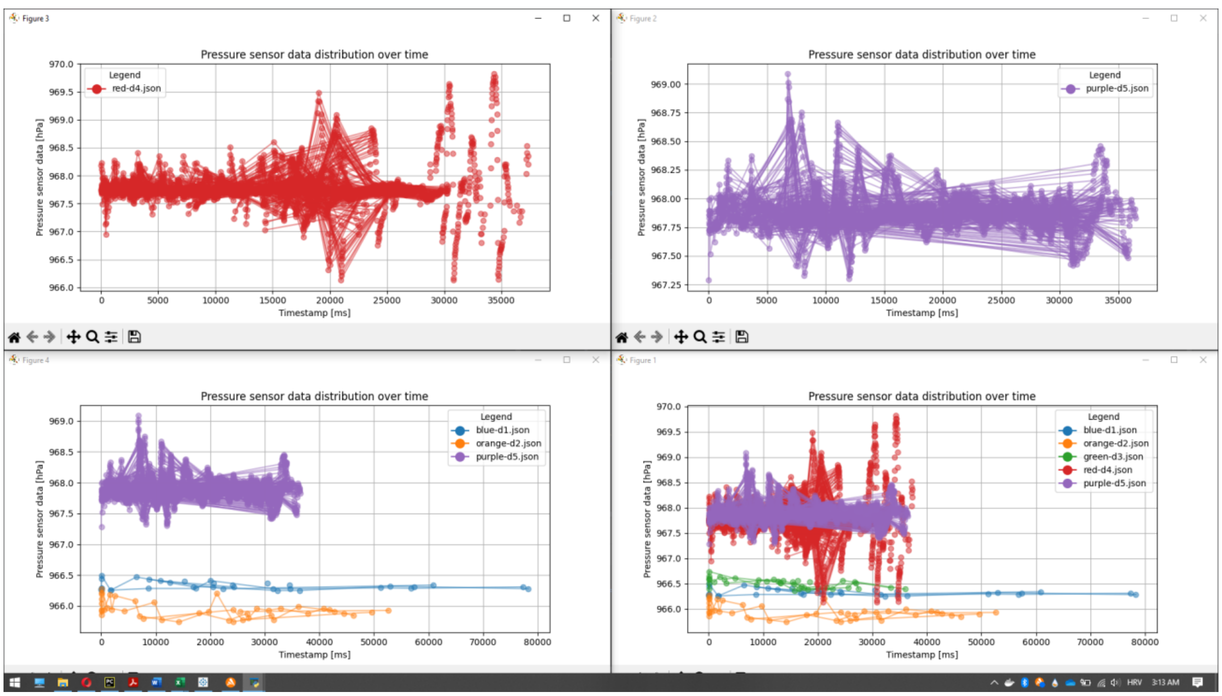
2.3.1 Plotting 1D data - pressure sensor data

The script for plotting 1-dimensional pressure sensor data is:
„2.1.0_Plot_1D_pressure_sensor_data.py“

The script gets points from the database and assigns corresponding timestamps and colors to every point.

Also to get more info every previous point is connected with the line with the next point in order to get info on which points belong to which timestamp.

The result of the script is the following image.



Conclusion:

From the plotting pressure sensor data is possible to see a pretty regular pattern in the red and blue dataset that can indicate non-human behavior.

Nothing can be concluded with the high probability from those plots. A further analysis of other events is done in the next chapters.

2.3.2 Plotting 2D data- touch events

The used script is:

“2.2.5_Plot_2D_touch_events_data_with_filtering_&_synchronous_rotation_with_enter_angle_&_color_filtering.py”

The script is created from scratch with the usage of Python, Matplotlib, and Tkinter GUI.

The script gets points from the database and assigns corresponding timestamps and colors to every point.

X-Y coordinates are positioned on the ground floor of the 3D cube.

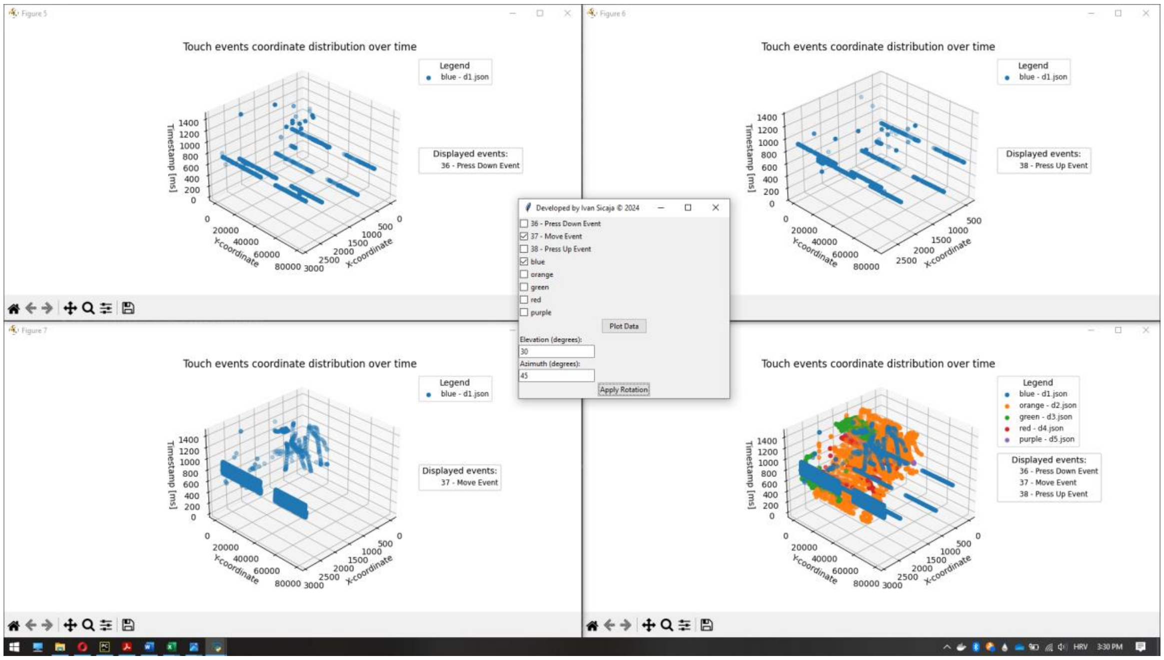
The Z-axis is the timestamp which is positioned vertically to the X-Y plane. In that way, data will be easy to understand.

Filters for specific events are created.

Additionally, filters that allow us to plot only points from the wanted dataset improve readability because reduce distractions from the other dataset’s points overlapping.

A great feature of this script is the possibility to show multiple filtered points at the same time and rotate every filtered point simultaneously, in order to compare multidimensional data.

The example of the plot is shown on the next image.

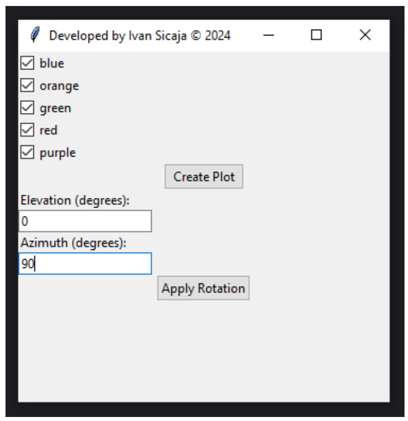


More examples are shown in the next chapter.

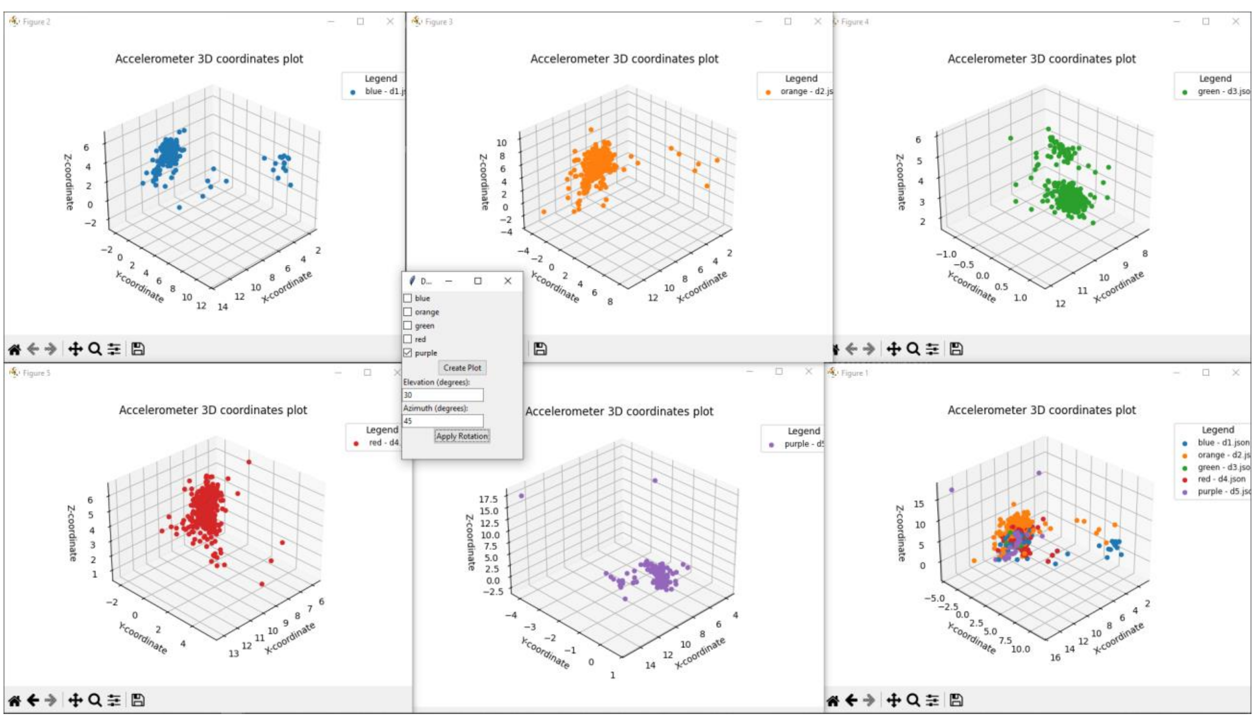
2.3.3 Plotting 3D data- accelerometer, gyroscope

Here only 3D coordinates are used without timestamps because the focus is monitoring the variance of the point groups which can be done without timestamps.

The look of the GUI is shown in the following image.



The accelerometer 3D coordinates plot is shown in the next image.



2.4 Data analiyze

There are three possible classes to which the corresponding dataset can belong:

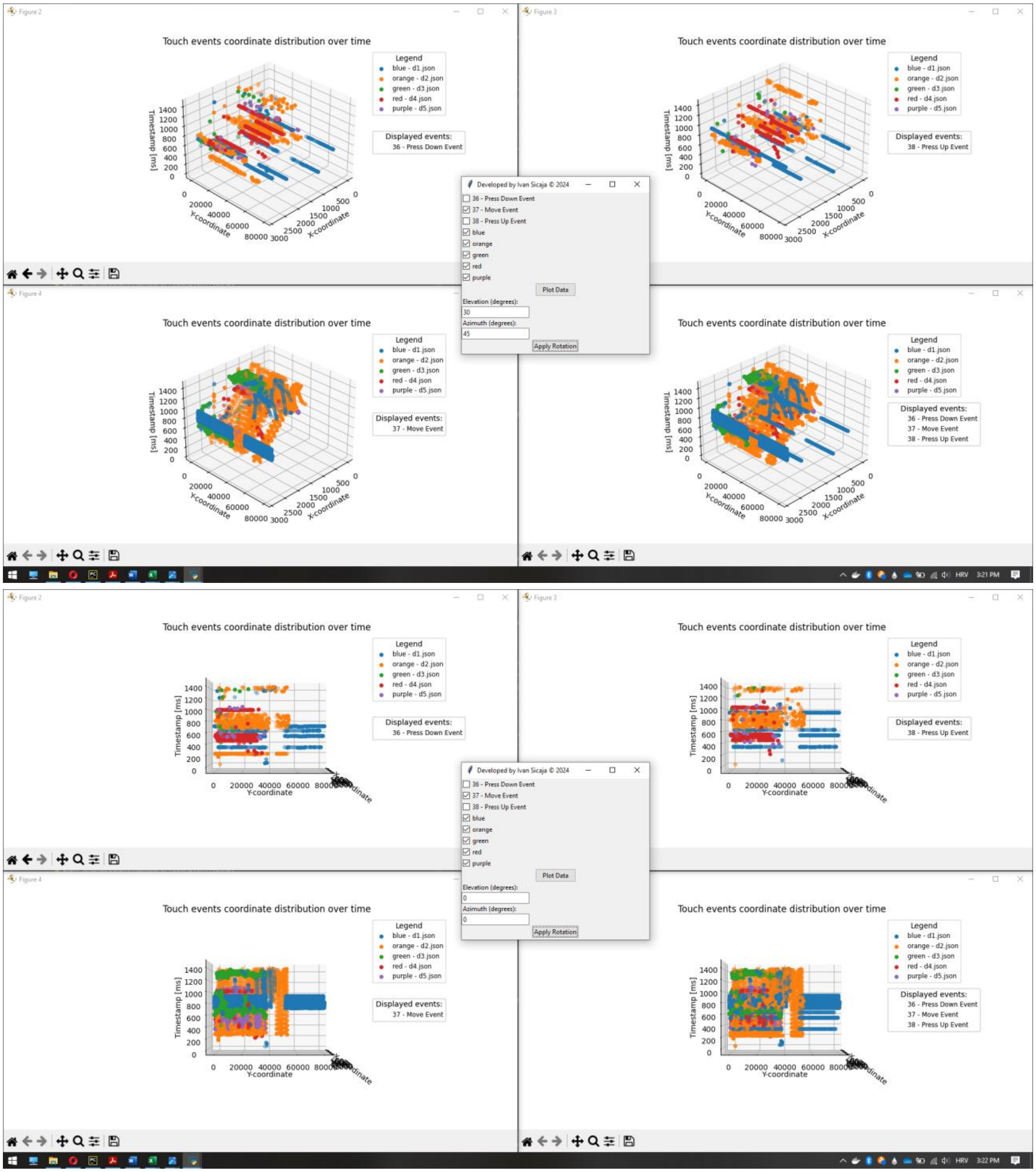
- human-only behavior,
- auto-clicker without anti-detection
- auto-clicker with anti-detection enabled

In order to identify which dataset belongs to which possible classes as spoken earlier main focus is on the “Touch Events” data.

In order to recognize patterns and differences between specific events. A Python script was created that allows displaying all relevant plots simultaneously, with elevation and azimuth rotation applied to all plots at the same time.

Also, it is possible to filter plots by wanted touch event and specific dataset in order to remove distractions produced by the overlapping points from the data which is not in the focus at the moment.

2.4.1 Analyze- touch events



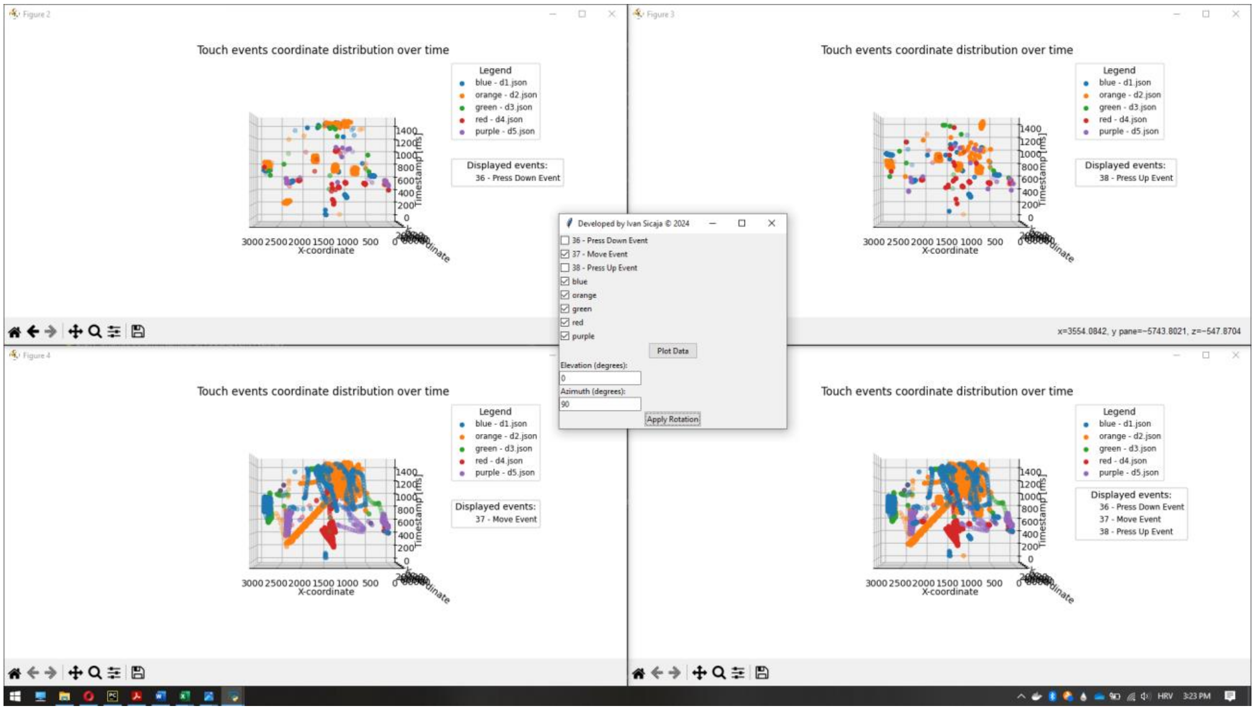


Figure 2
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
X-coordinate
Timestamp [ms]
Figure 3
38 - Press Up Event
Figure 4
37 - Move Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
90
Apply Rotation
x=3554.0842, y pane=-5743.8021, z=-547.8704
3:23 PM

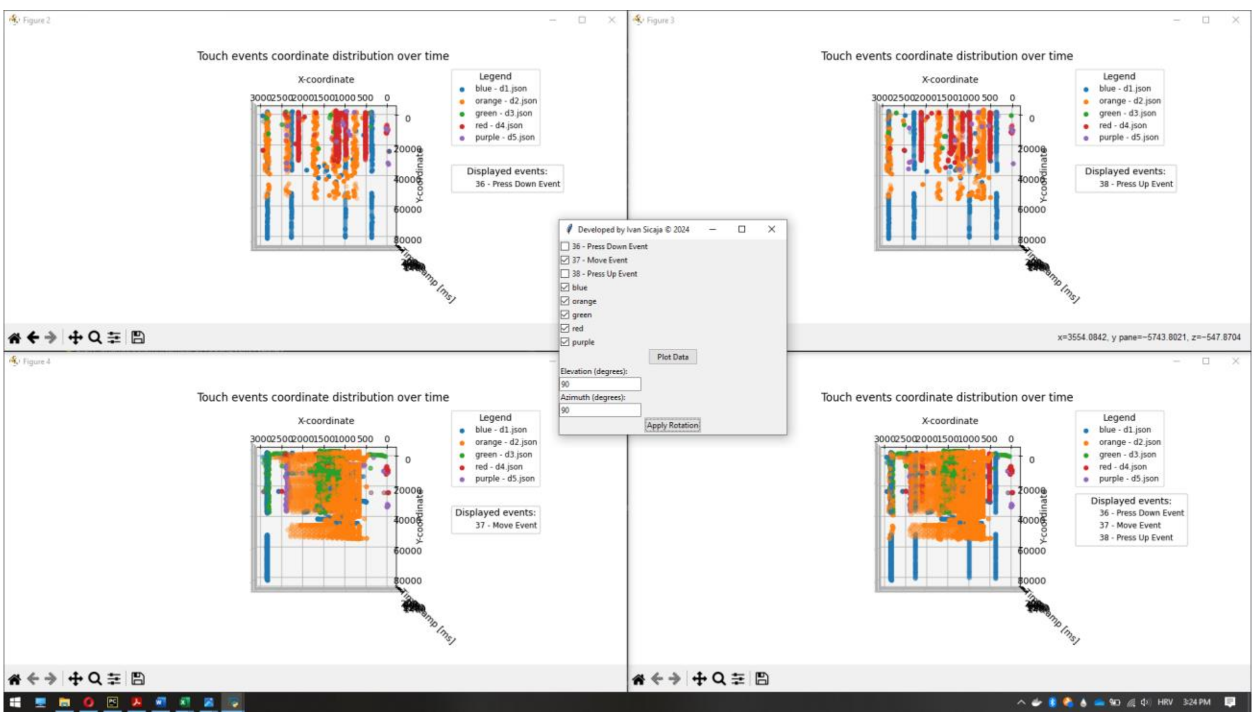


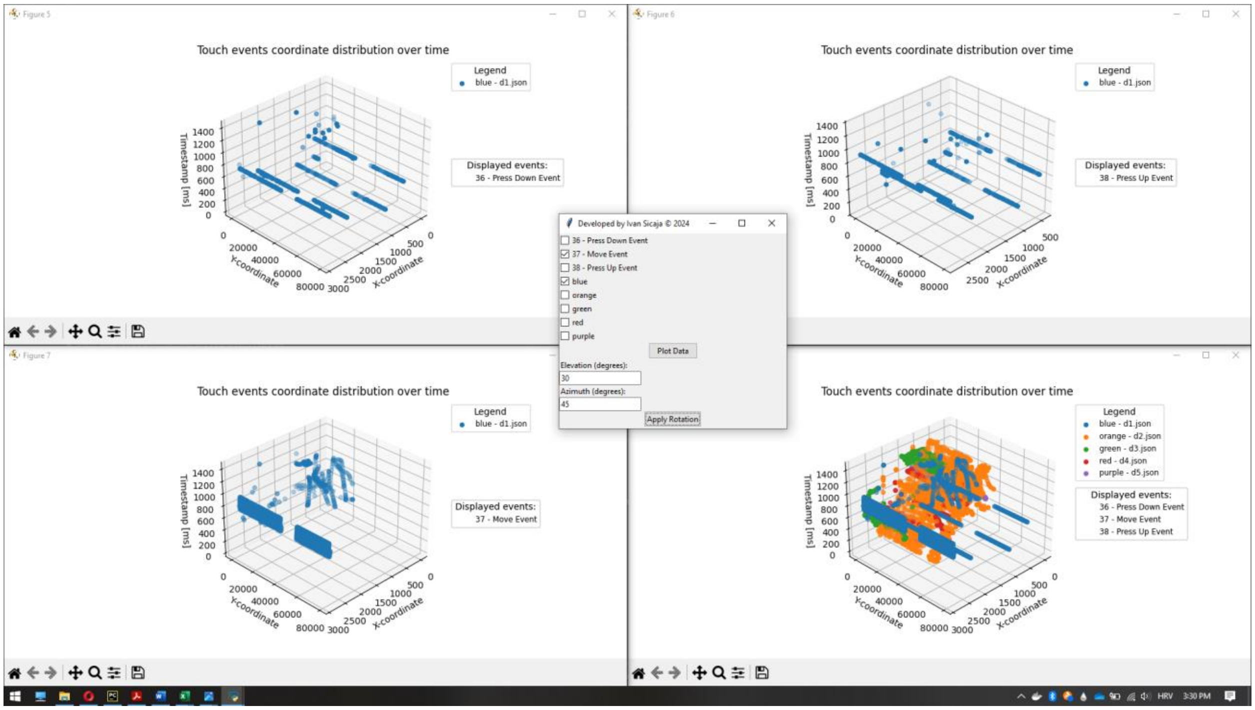
Figure 2
Touch events coordinate distribution over time
X-coordinate
Y-coordinate
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
Figure 3
38 - Press Up Event
Figure 4
37 - Move Event
Developed by Ivan Sicaja © 2024
Plot Data
Elevation (degrees):
90
Azimuth (degrees):
90
Apply Rotation
x=3554.0842, y pane=-5743.8021, z=-547.8704
3:24 PM

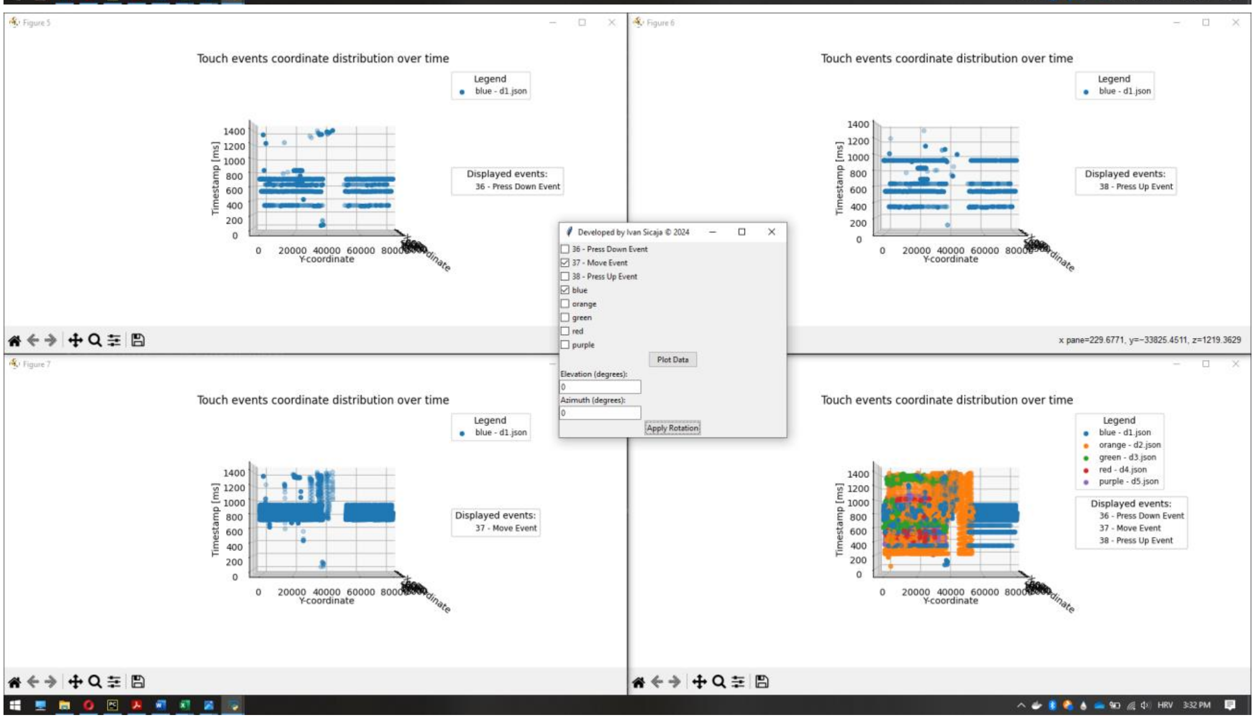
Observations:

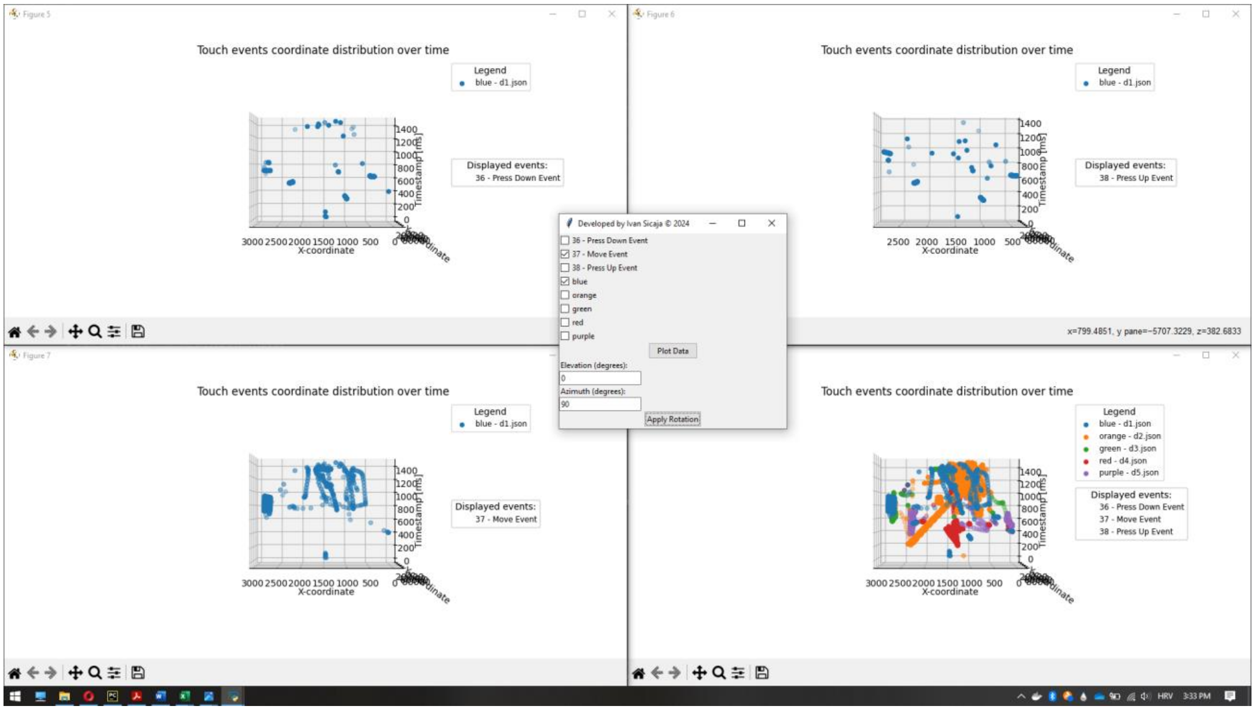
"Blue – d1.json" is categorized as auto-clicker.

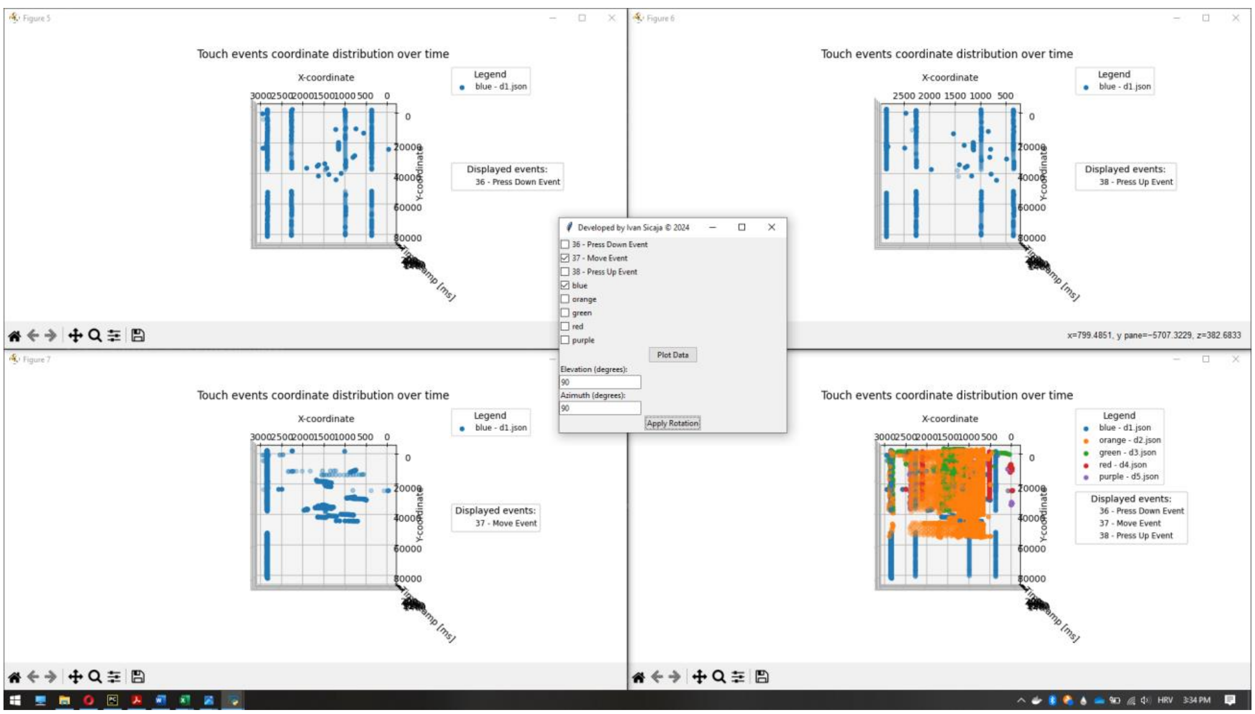
Reasons:

- "Press Down" and "Press Up" events series on the multiple paths between two specific positions.
- The variance during continuous shoots is small, which indicates the high shooting precision. That is a bot's characteristic.
- The "Press Down" and "Press Up" events series are executed fast in comparison with some of the other datasets (low variance visible on the z-axis). That shows significantly high shooting speed. That is a bot's characteristic.
- Movement positions are iterative over time which shows the high movement precision. That is a bot's characteristic.
- Movements at paths between two specific positions are done fast (z-axis small variance), which indicates high moving speed. That is a bot's characteristic.
- The small amount of the "Press Down" and "Press Up" events points are not on the paths between two specific positions.


Touch events coordinate distribution over time
Legend
blue - d1.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Timestamp [ms]
Y-coordinate
X-coordinate
Developed by Ivan Sicaja © 2024
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
30
Azimuth (degrees):
45
Apply Rotation


Touch events coordinate distribution over time
Legend
blue - d1.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Timestamp [ms]
Y-coordinate
Developed by Ivan Sicaja © 2024
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
0
Apply Rotation
x pane=229.6771, y=-33825.4511, z=1219.3629


Touch events coordinate distribution over time
Legend
blue - d1.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
Azimuth (degrees):
Apply Rotation
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
X-coordinate
Timestamp [ms]


Touch events coordinate distribution over time
X-coordinate
Y-coordinate
Legend
blue - d1.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
Developed by Ivan Sicaja © 2024
Plot Data
Elevation (degrees):
Azimuth (degrees):
Apply Rotation
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json

“Orange– d2.json” is categorized as an auto-clicker with anti-detection enabled.

Reasons:

- It is the only dataset that contains the characteristics for both “Human-only” and “Auto-clicker without anti-detection” behavior. Humans can't simulate the precision and the speed of the three mentioned events because of human physics limitations. The “bot” can simulate slow and non-precise events. Speed and high-precision patterns that can not be reproduced by humans are proof of the auto-clicker without anti-detection. Since we also have a lot of events which are executed slower with lower precision, but still with pretty regular iterative patterns it is obvious that this dataset is auto-clicker with anti-detection enabled.

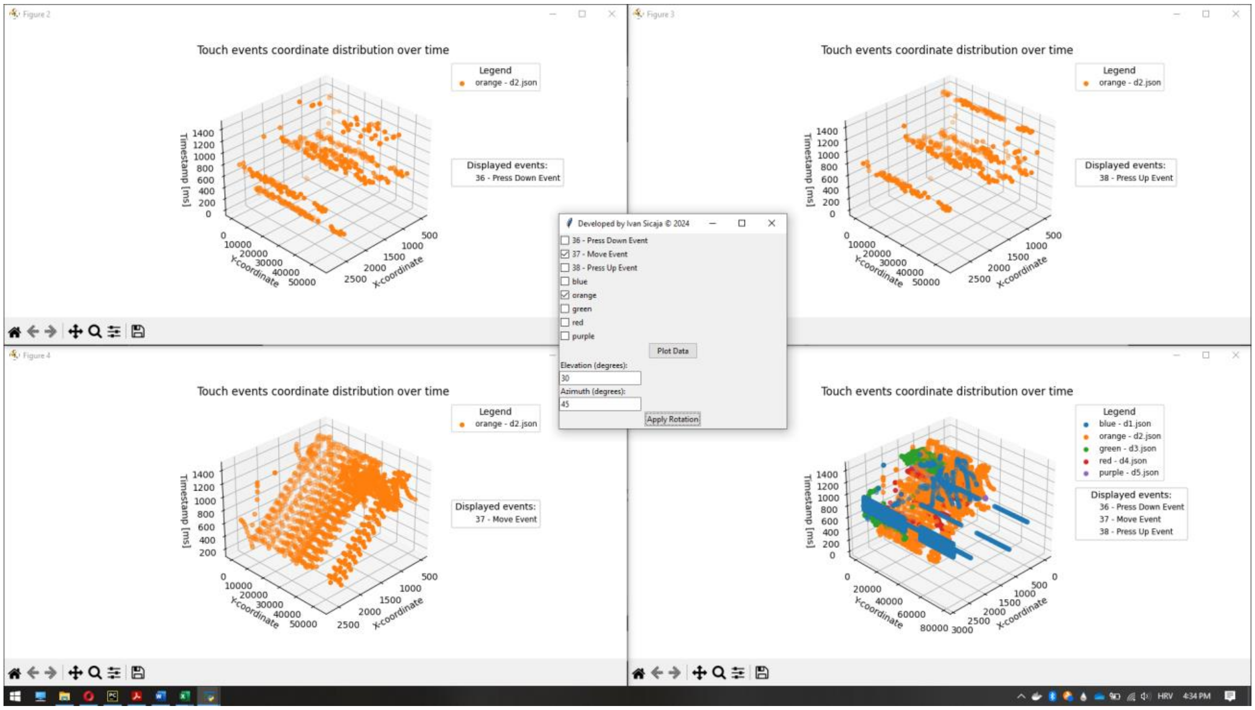


Figure 2
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
36 - Press Down Event
Timestamp [ms]
Y-coordinate
X-coordinate
Figure 3
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
38 - Press Up Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
30
Azimuth (degrees):
45
Apply Rotation
Figure 4
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
4:34 PM

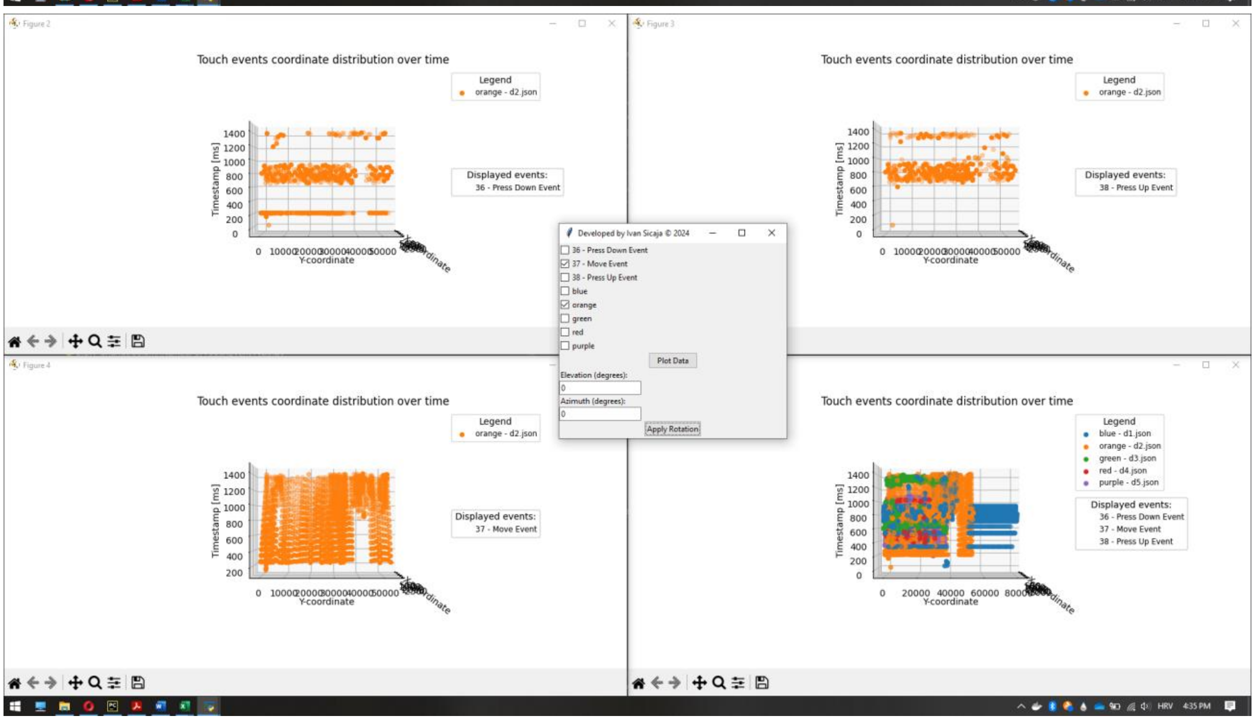
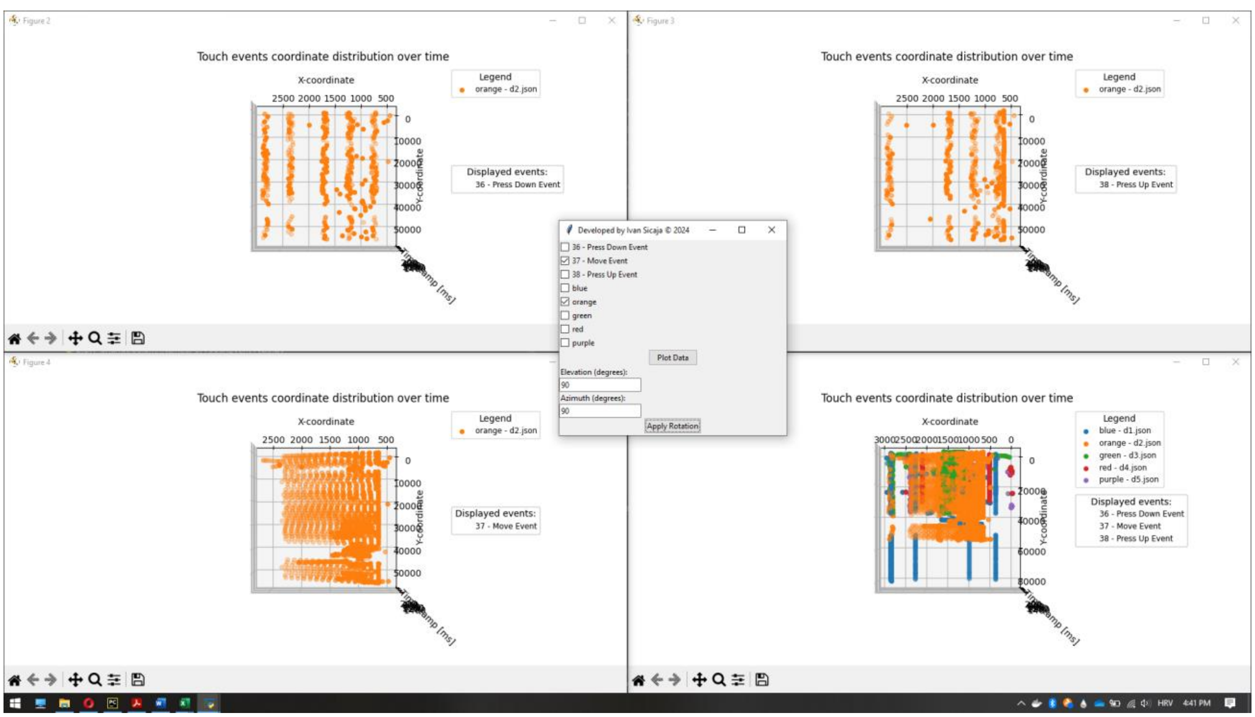



Figure 2
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
36 - Press Down Event
Timestamp [ms]
Y-coordinate
Figure 3
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
38 - Press Up Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
0
Apply Rotation
Figure 4
Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
4:35 PM


Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
green - d3.json
red - d4.json
purple - d5.json
36 - Press Down Event
37 - Move Event
38 - Press Up Event
X-coordinate
Timestamp [ms]
Developed by Ivan Sicaja © 2024
Plot Data
Elevation (degrees):
Azimuth (degrees):
Apply Rotation

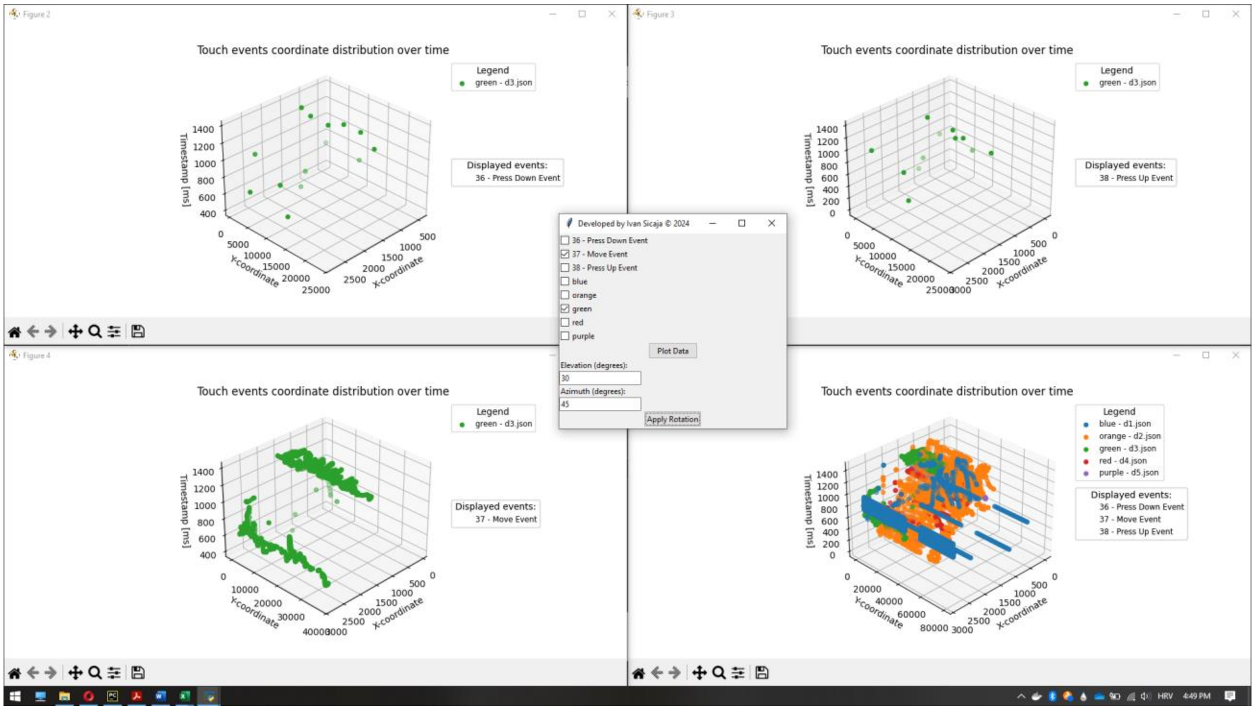

Touch events coordinate distribution over time
Legend
orange - d2.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
green - d3.json
red - d4.json
purple - d5.json
X-coordinate
Y-coordinate
Developed by Ivan Sicaja © 2024
Plot Data
Elevation (degrees):
Azimuth (degrees):
Apply Rotation

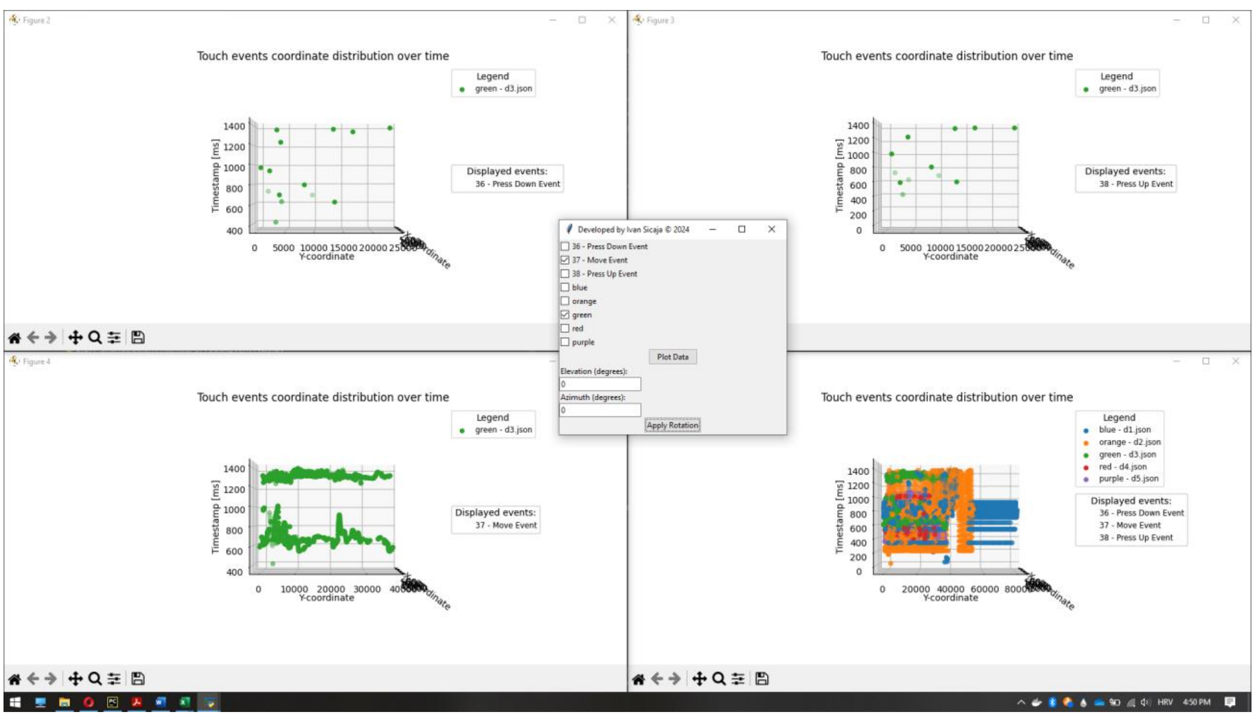
“Green – d3.json” is categorized as human-only behavior.

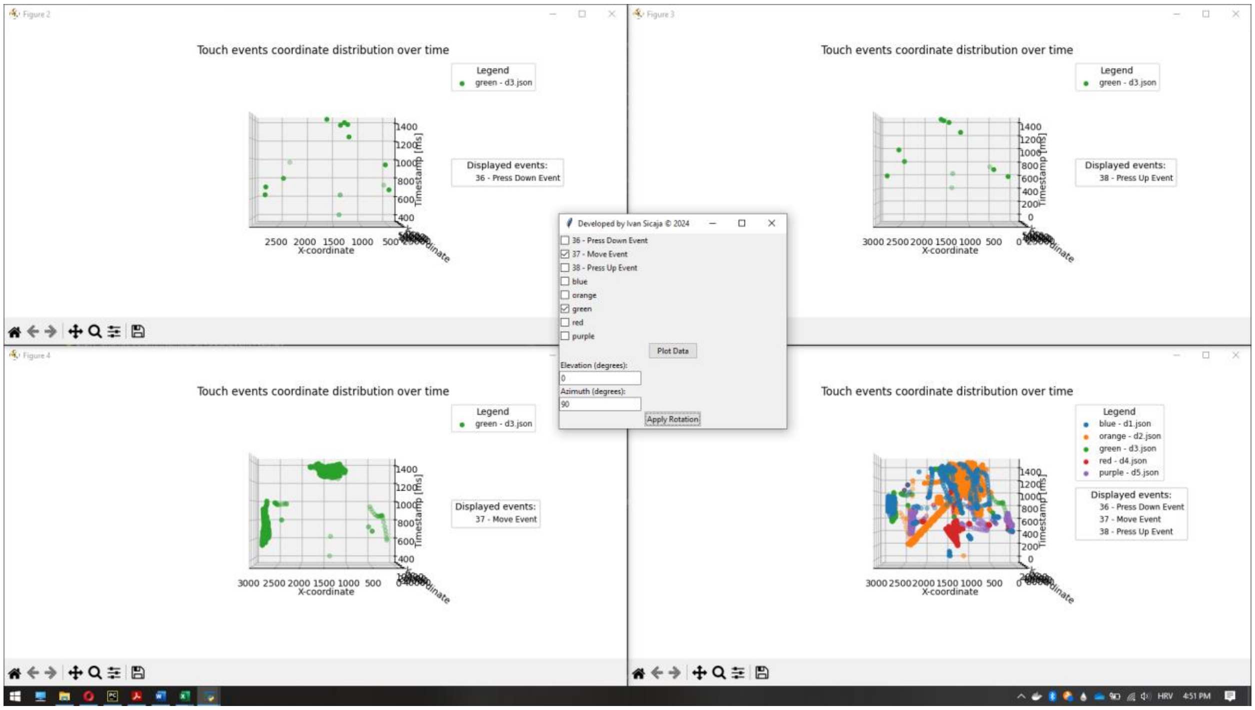
Reasons:

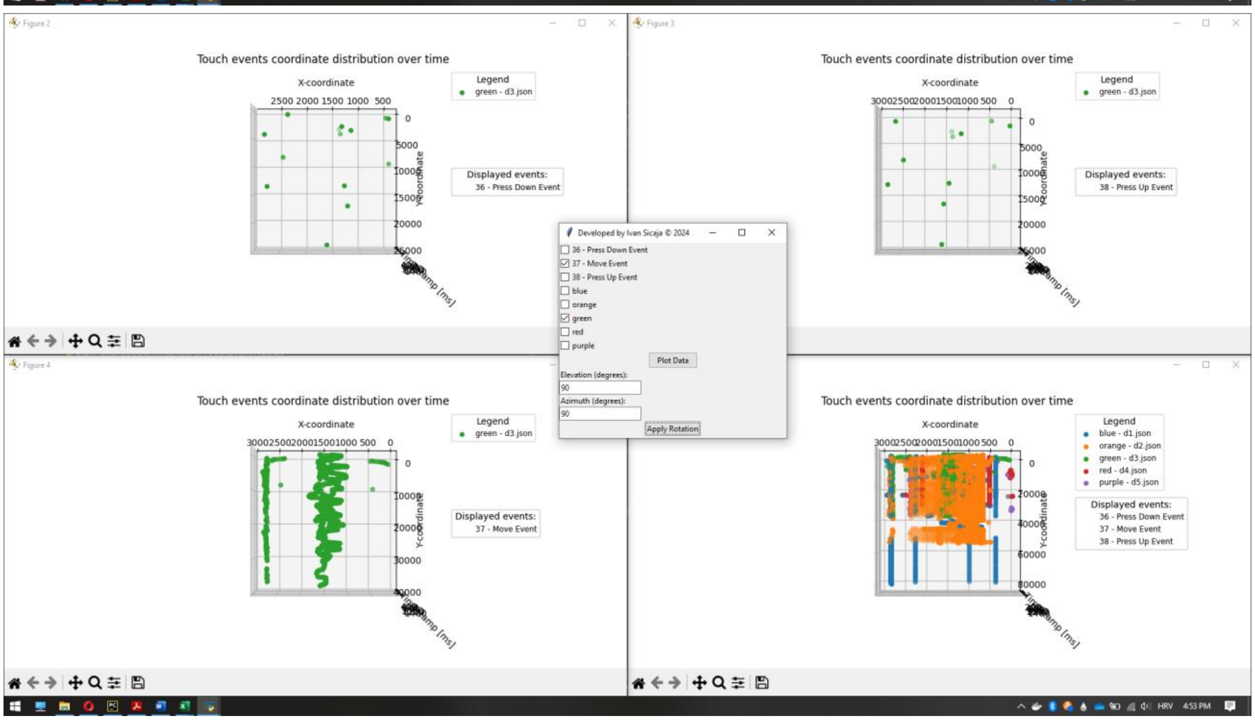
Absence of the following auto-clicker patterns:

- “Press Down” and “Press Up” events series on the multiple paths between two specific positions.
- The variance during continuous shoots is small, which indicates the high shooting precision. That is a bot's characteristic.
- The “Press Down” and “Press Up” events series are executed fast in comparison with some of the other datasets (low variance visible on the z-axis). That shows significantly high shooting speed. That is a bot's characteristic.
- Movement positions are iterative over time which shows the high movement precision. That is a bot's characteristic.
- Movements at paths between two specific positions are done fast (z-axis small variance), which indicates high moving speed. That is a bot's characteristic.
- The small amount of the “Press Down” and “Press Up” events points are not on the paths between two specific positions.


Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
orange - d2.json
red - d4.json
purple - d5.json
Timestamp [ms]
Y-coordinate
X-coordinate
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
30
Azimuth (degrees):
45
Apply Rotation


Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
orange - d2.json
red - d4.json
purple - d5.json
Timestamp [ms]
Y-coordinate
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
0
Apply Rotation


Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
36 - Press Down Event
Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
38 - Press Up Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
90
Apply Rotation
Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
X-coordinate
Timestamp [ms]


Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
36 - Press Down Event
Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
38 - Press Up Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
90
Azimuth (degrees):
90
Apply Rotation
Touch events coordinate distribution over time
Legend
green - d3.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
X-coordinate
Y-coordinate

“Red – d4.json” is categorized as auto-clicker.

Reasons:

- “Press Down” and “Press Up” events series on the multiple paths between two specific positions.
- The variance during continuous shoots is small, which indicates the high shooting precision. That is a bot's characteristic.
- The “Press Down” and “Press Up” events series are executed fast in comparison with some of the other datasets (low variance visible on the z-axis). That shows significantly high shooting speed. That is a bot's characteristic.
- Movement positions are iterative over time which shows the high movement precision. That is a bot's characteristic.
- Movements at paths between two specific positions are done fast (z-axis small variance), which indicates high moving speed. That is a bot's characteristic.
- The small amount of the “Press Down” and “Press Up” events points are not on the paths between two specific positions.

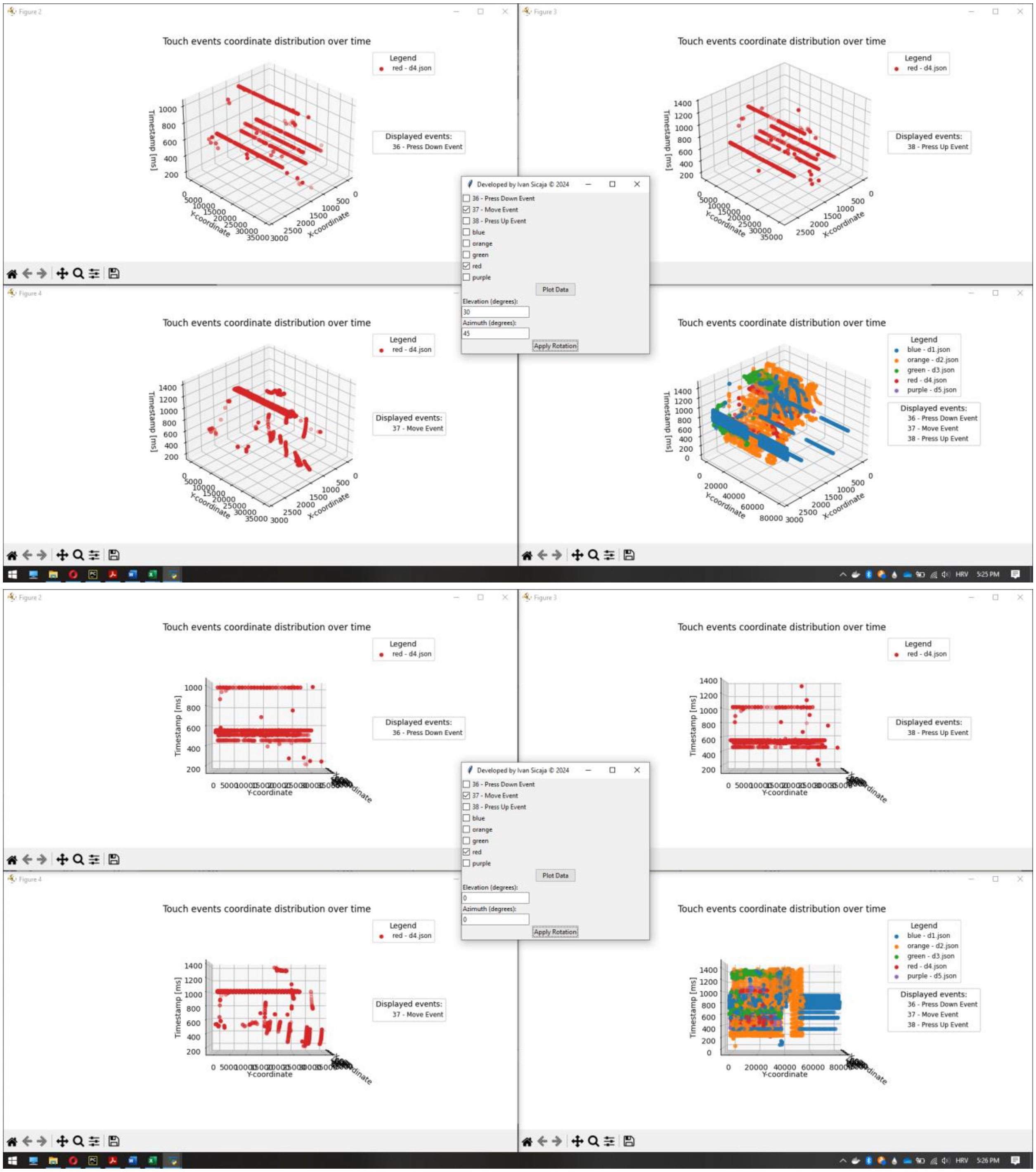


Figure 2
Figure 3
Figure 4
Touch events coordinate distribution over time
Legend
red - d4.json
blue - d1.json
orange - d2.json
green - d3.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
Timestamp [ms]
Y-coordinate
X-coordinate
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
30
Azimuth (degrees):
45
Apply Rotation
Elevation (degrees):
0
Azimuth (degrees):
0
Apply Rotation
HRV
5:25 PM
5:26 PM

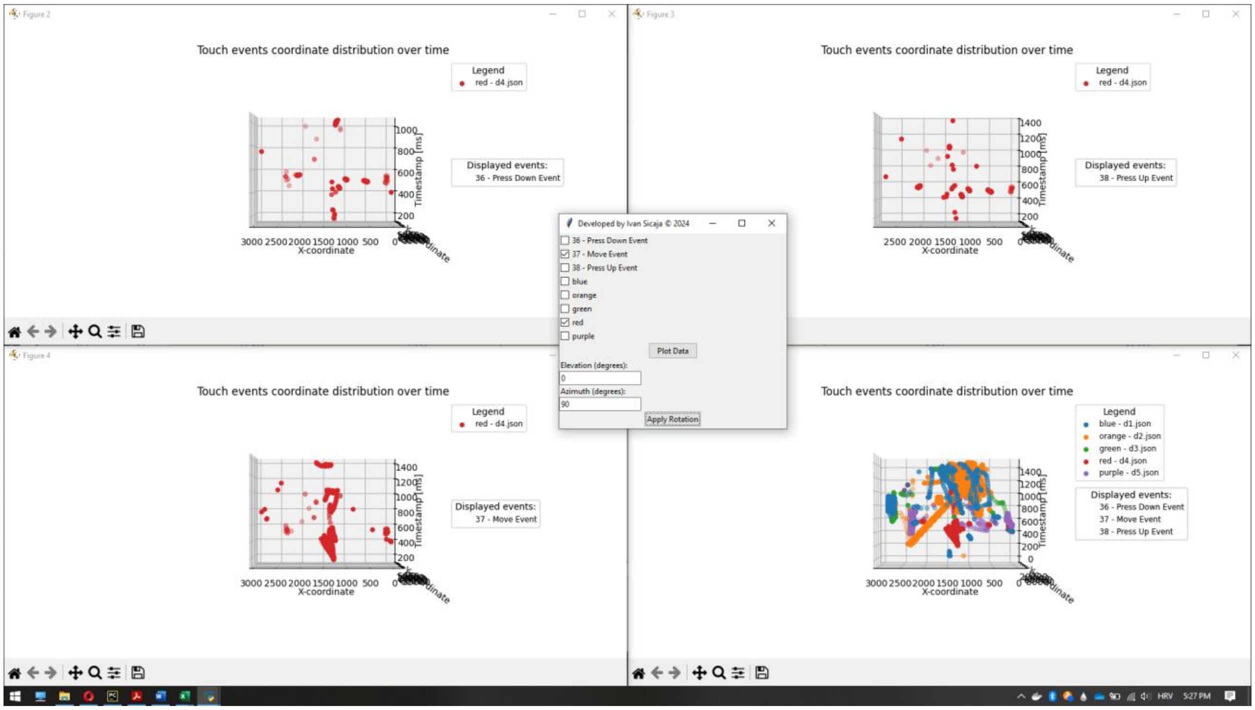


Figure 2
Touch events coordinate distribution over time
Legend
red - d4.json
Displayed events:
36 - Press Down Event
X-coordinate
Timestamp [ms]
Figure 3
Touch events coordinate distribution over time
Legend
red - d4.json
Displayed events:
38 - Press Up Event
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
90
Apply Rotation
Figure 4
Touch events coordinate distribution over time
Legend
red - d4.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
HRV
5:27 PM

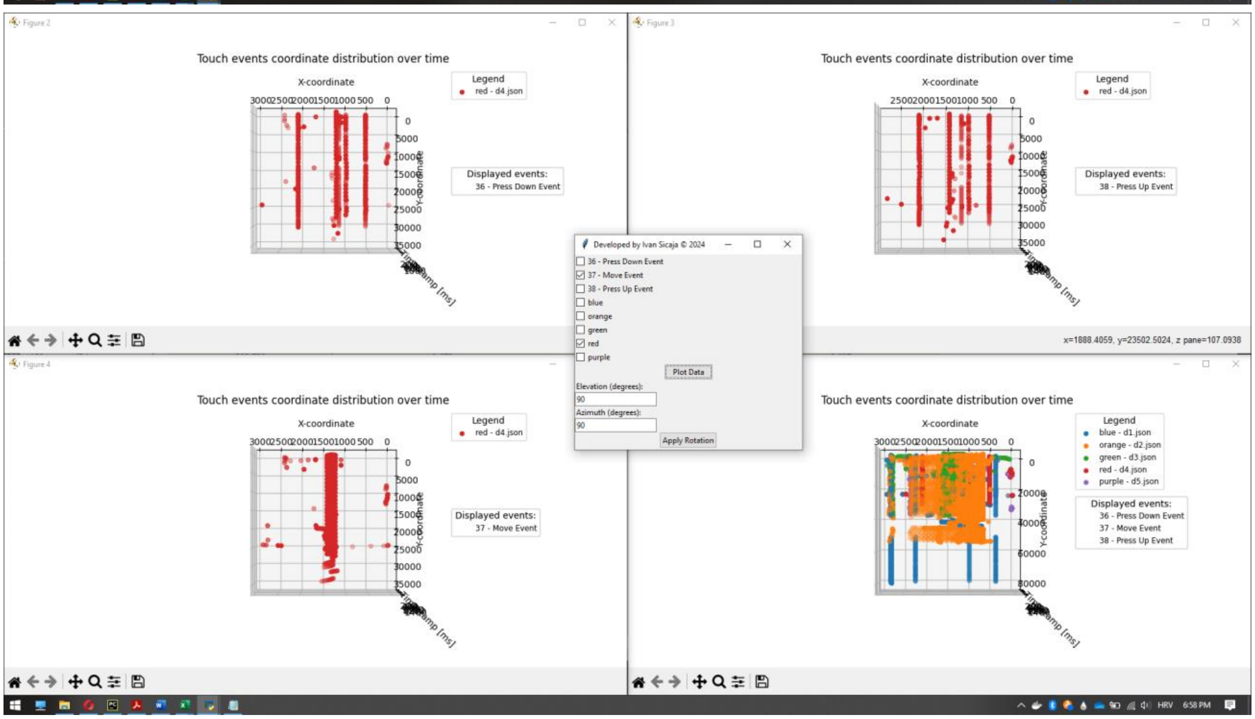


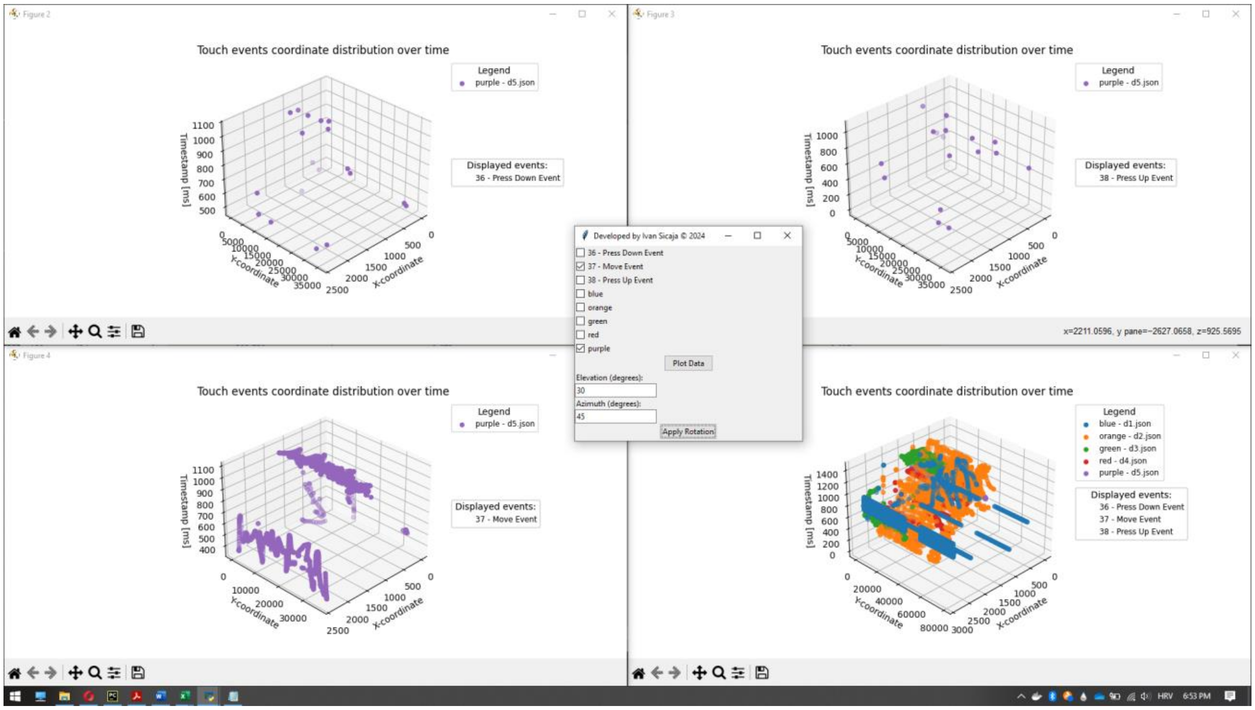
Figure 2
Touch events coordinate distribution over time
X-coordinate
Legend
red - d4.json
Displayed events:
36 - Press Down Event
Y-coordinate
Timestamp [ms]
Figure 3
Touch events coordinate distribution over time
X-coordinate
Legend
red - d4.json
Displayed events:
38 - Press Up Event
x=1888.4059, y=23502.5024, z pane=107.0938
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
90
Azimuth (degrees):
90
Apply Rotation
Figure 4
Touch events coordinate distribution over time
X-coordinate
Legend
red - d4.json
Displayed events:
37 - Move Event
Touch events coordinate distribution over time
X-coordinate
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Displayed events:
36 - Press Down Event
37 - Move Event
38 - Press Up Event
HRV
6:58 PM

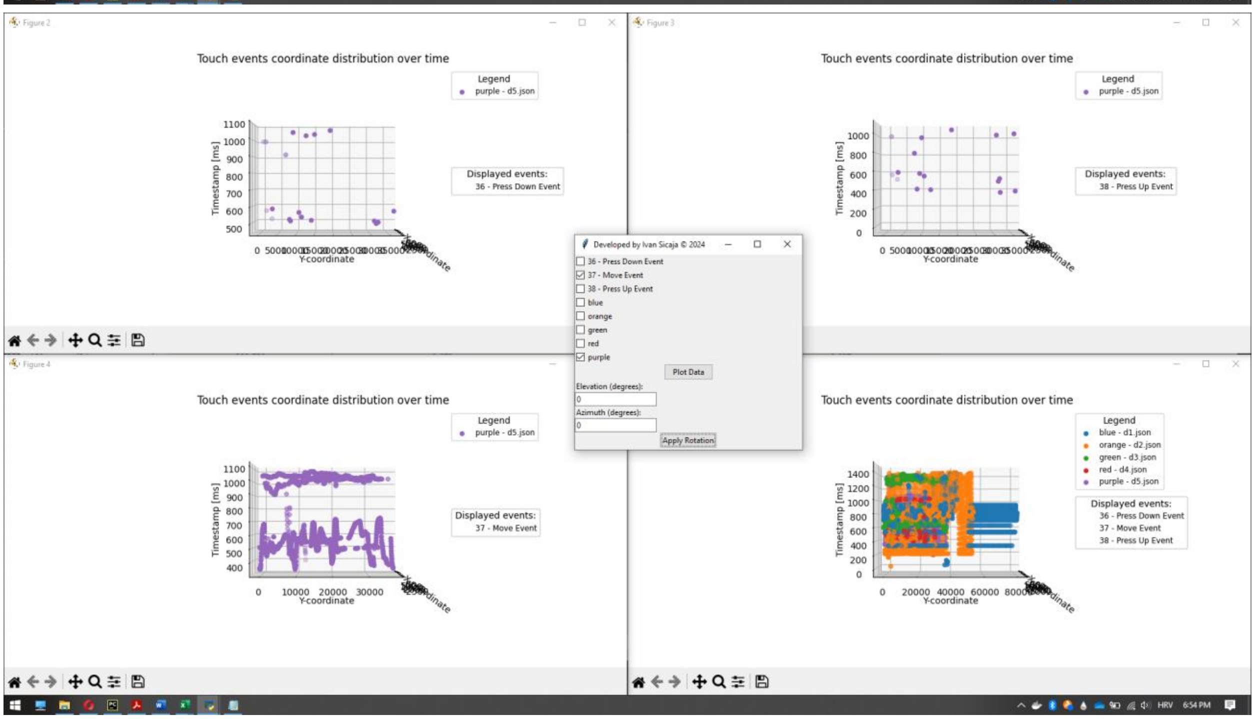
“Purple – d5.json” is categorized as human-only behavior.

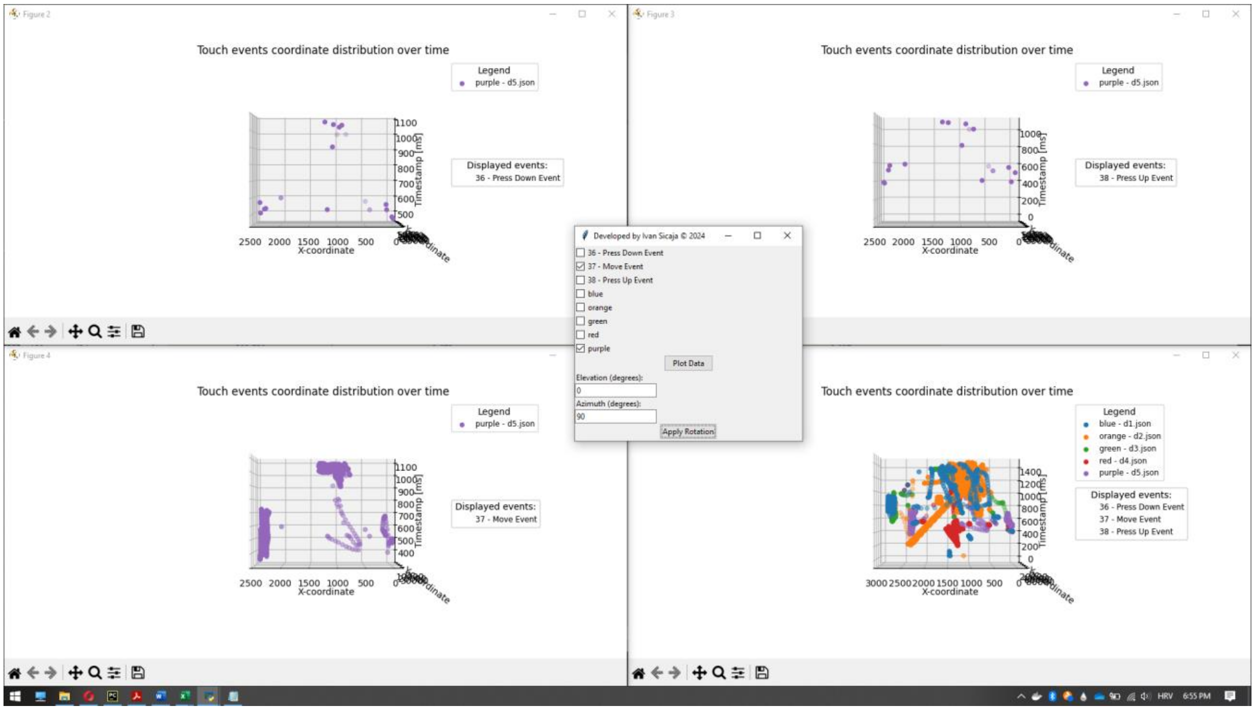
Reasons:

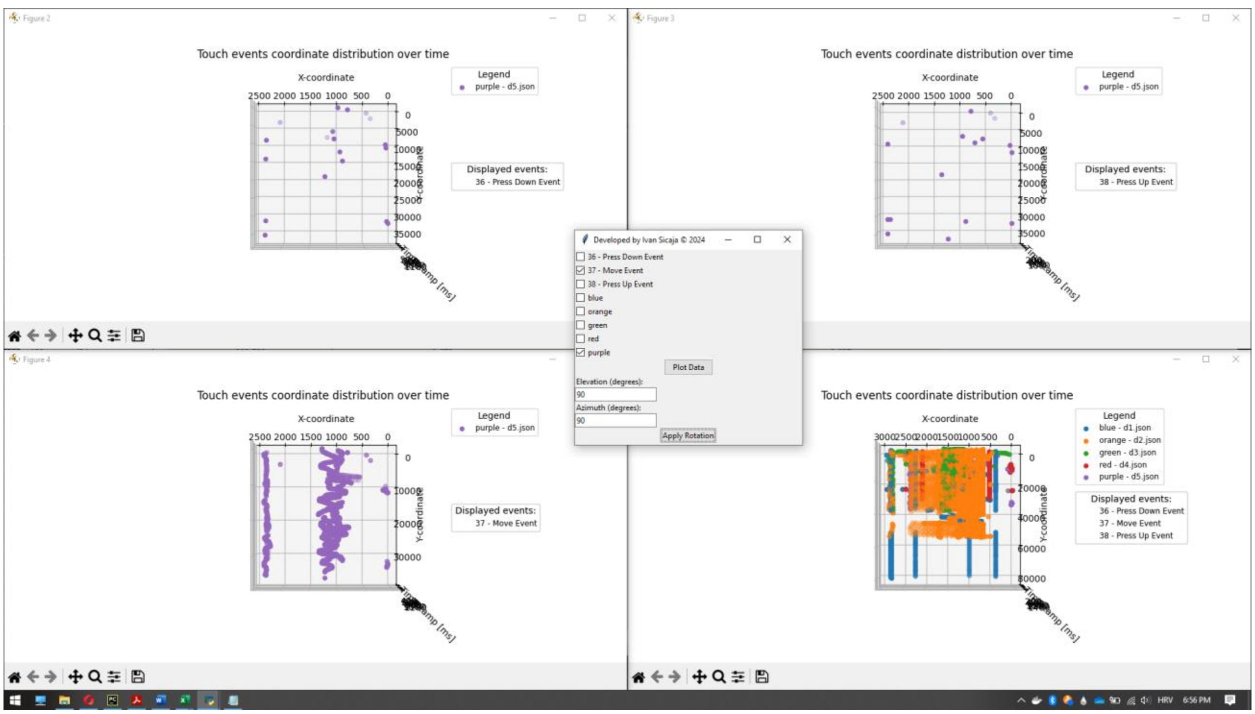
Absence of the following auto-clicker patterns:

- “Press Down” and “Press Up” events series on the multiple paths between two specific positions.
- The variance during continuous shoots is small, which indicates the high shooting precision. That is a bot's characteristic.
- The “Press Down” and “Press Up” events series are executed fast in comparison with some of the other datasets (low variance visible on the z-axis). That shows significantly high shooting speed. That is a bot's characteristic.
- Movement positions are iterative over time which shows the high movement precision. That is a bot's characteristic.
- Movements at paths between two specific positions are done fast (z-axis small variance), which indicates high moving speed. That is a bot's characteristic.
- The small amount of the “Press Down” and “Press Up” events points are not on the paths between two specific positions.

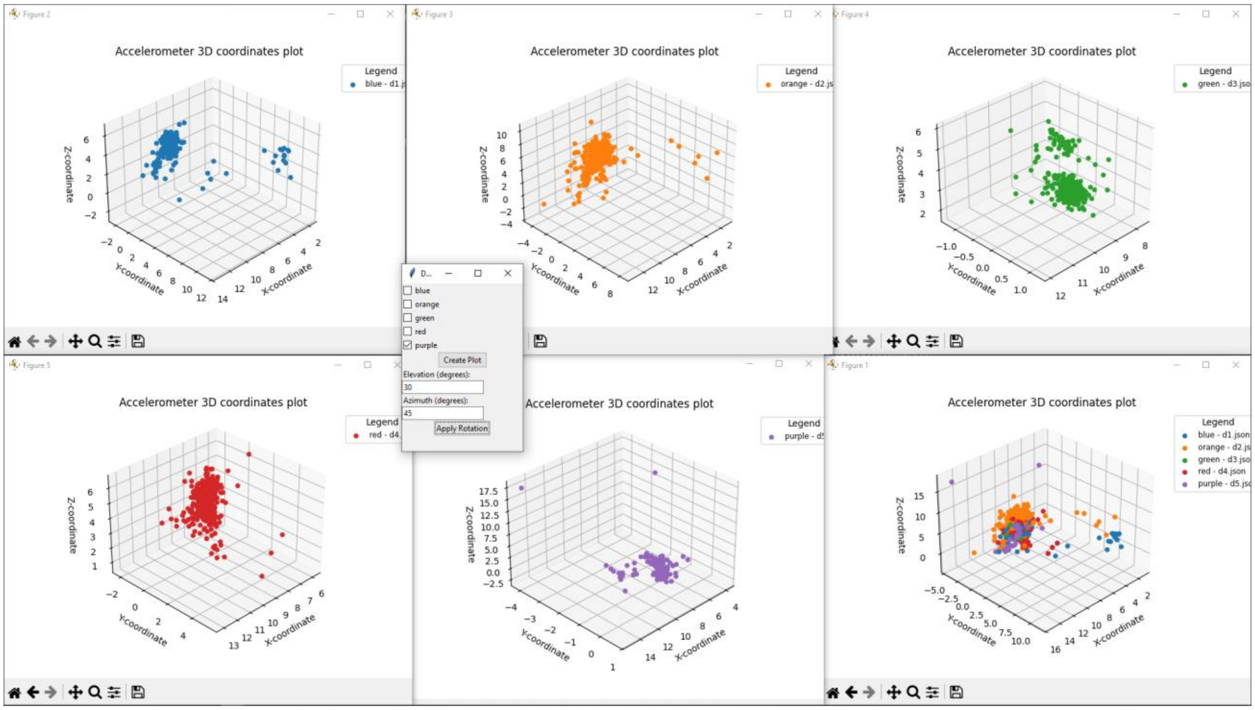

Touch events coordinate distribution over time
Legend
purple - d5.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
Timestamp [ms]
Y-coordinate
X-coordinate
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
30
Azimuth (degrees):
45
Apply Rotation

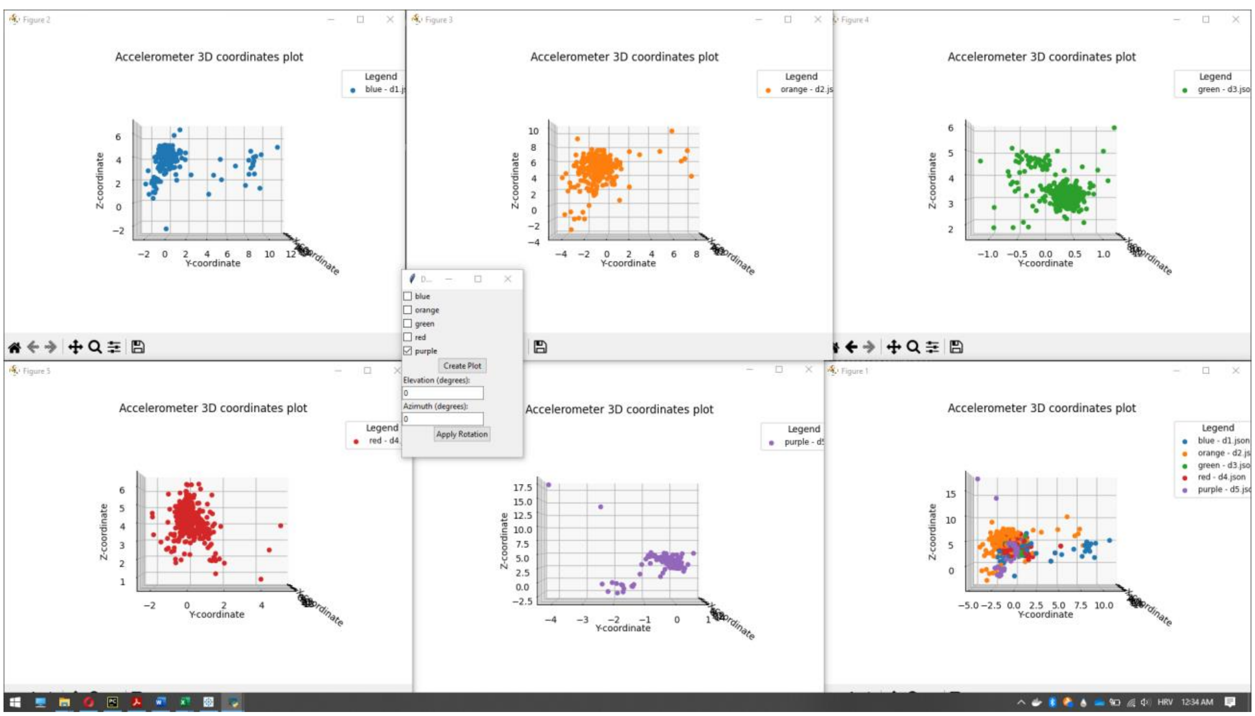

Touch events coordinate distribution over time
Legend
purple - d5.json
Displayed events:
36 - Press Down Event
38 - Press Up Event
37 - Move Event
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
Timestamp [ms]
Y-coordinate
Developed by Ivan Sicaja © 2024
36 - Press Down Event
37 - Move Event
38 - Press Up Event
blue
orange
green
red
purple
Plot Data
Elevation (degrees):
0
Azimuth (degrees):
0
Apply Rotation

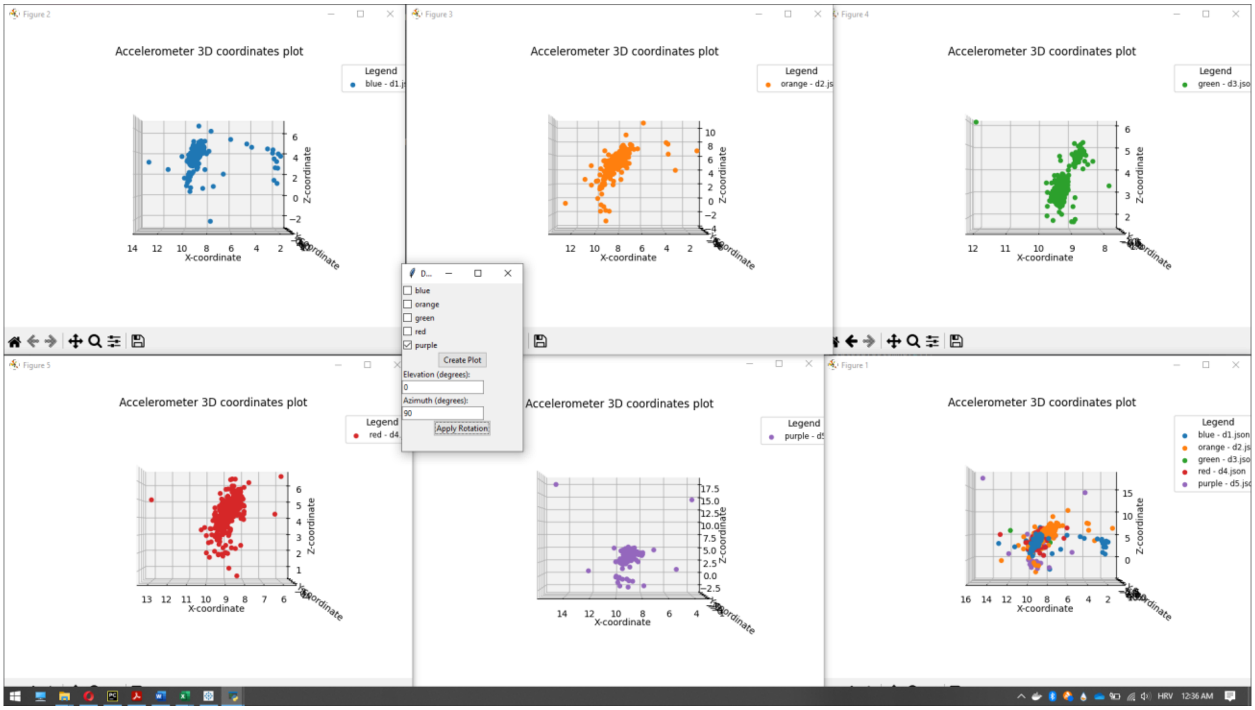


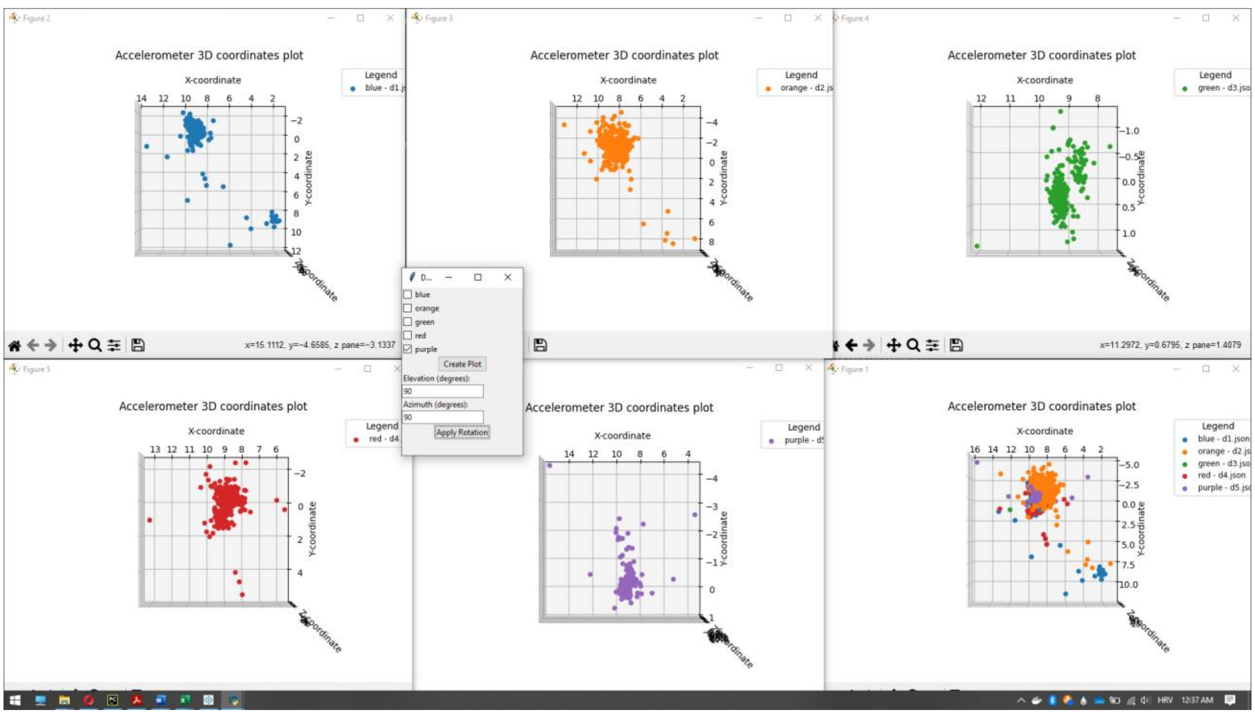


2.4.2 Analyze - accelerometer data







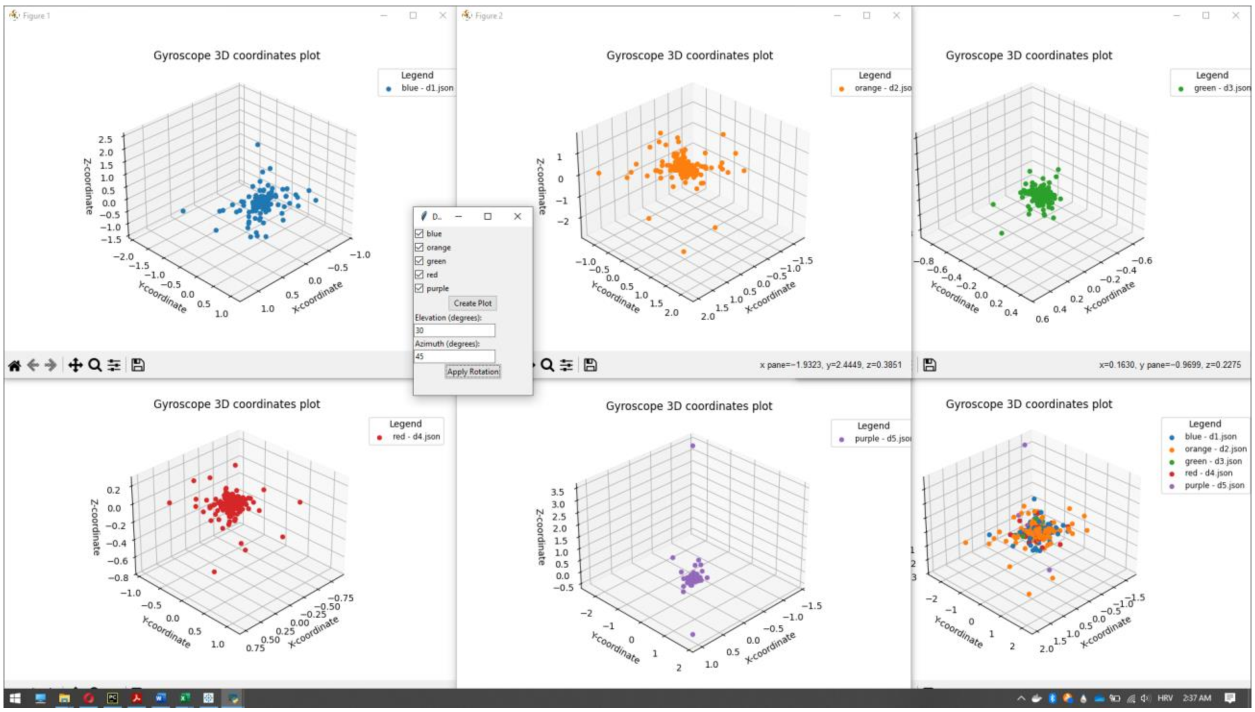


Conclusion:

All dataset except the "Green-d3.json" and "Purple-d5.json" has more points that converge to the midpoint of the dataset which is formed by more regular points dispersion in comparison with the points dispersion of "Green-d3.json" and "Purple-d5.json" dataset.

Accelerometer 3D data doesn't give so much useful information based on which is possible, with a high probability, to categorize provided datasets into three provided classes.

2.4.3 Analyze - gyroscope data



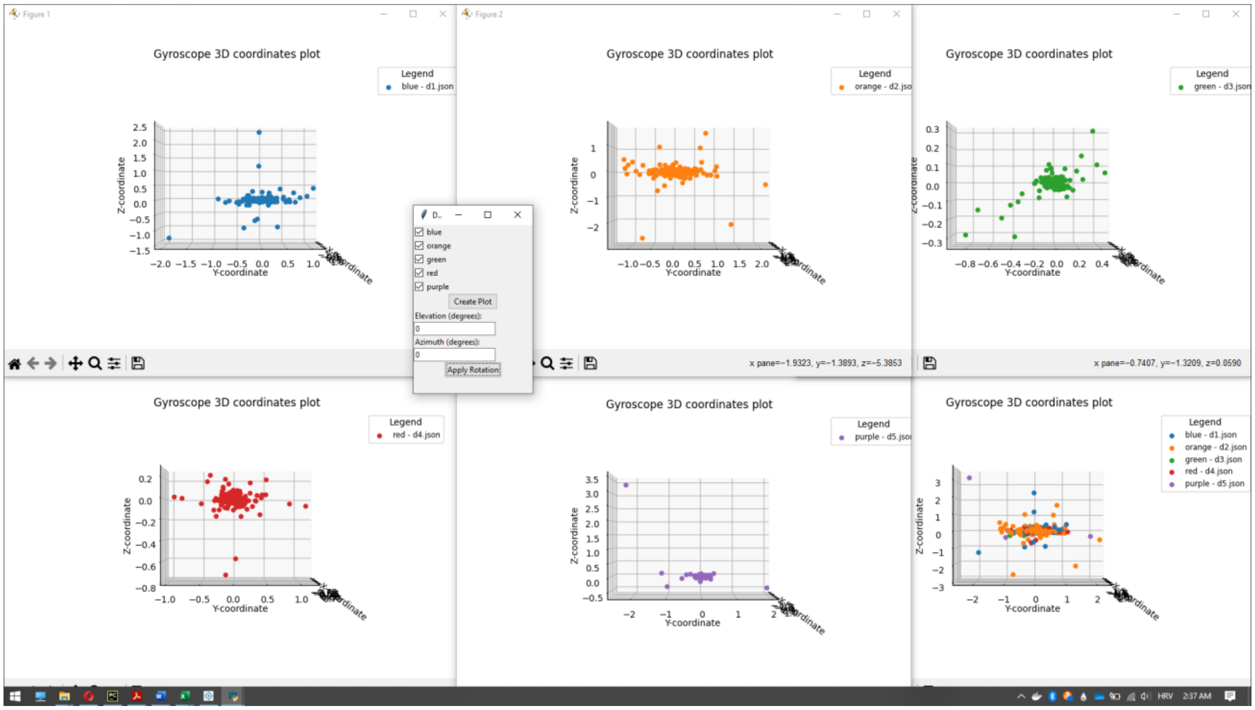


Figure 1
Figure 2
Gyroscope 3D coordinates plot
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
Y-coordinate
Z-coordinate
blue
orange
green
red
purple
Create Plot
Elevation (degrees):
Azimuth (degrees):
Apply Rotation

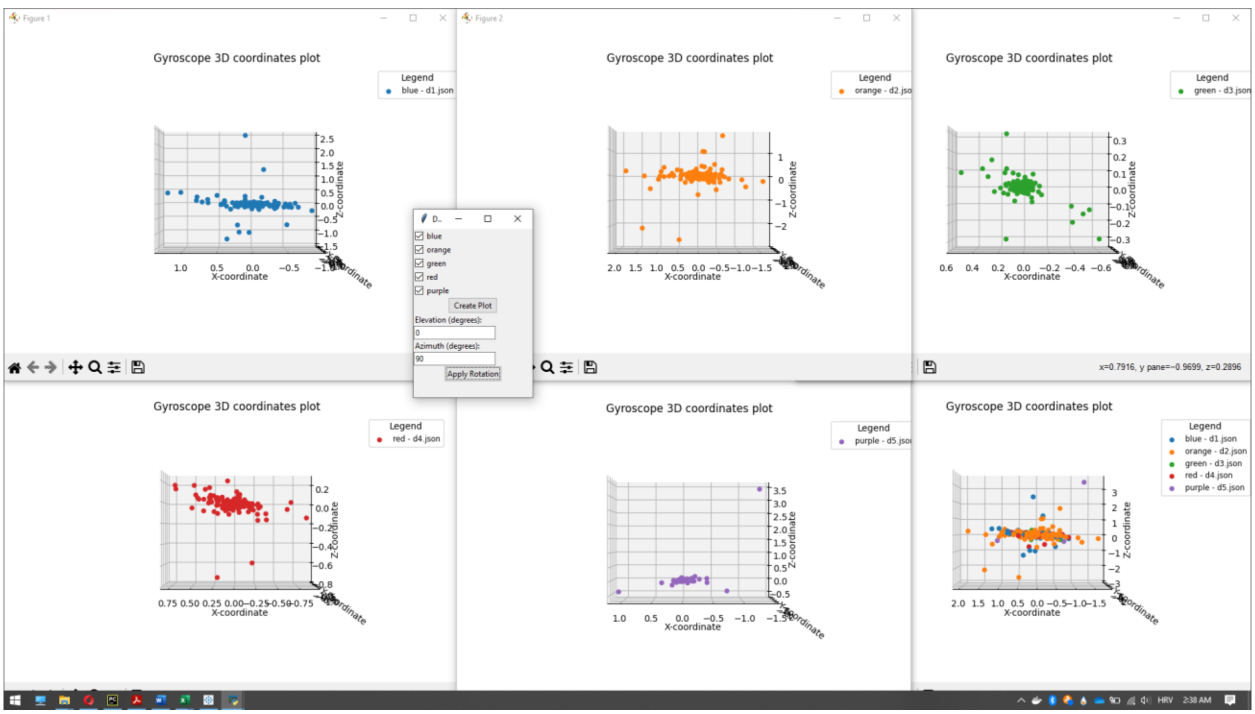
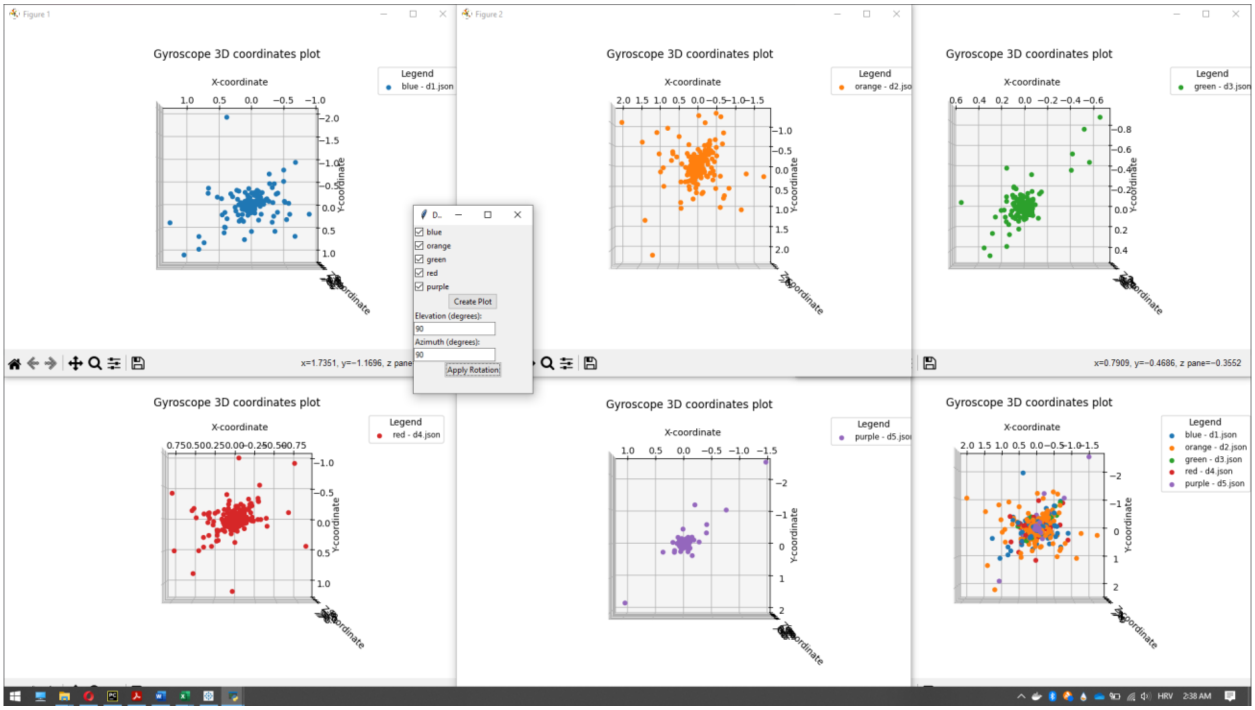


Figure 1
Figure 2
Gyroscope 3D coordinates plot
Legend
blue - d1.json
orange - d2.json
green - d3.json
red - d4.json
purple - d5.json
X-coordinate
Z-coordinate
blue
orange
green
red
purple
Create Plot
Elevation (degrees):
Azimuth (degrees):
Apply Rotation



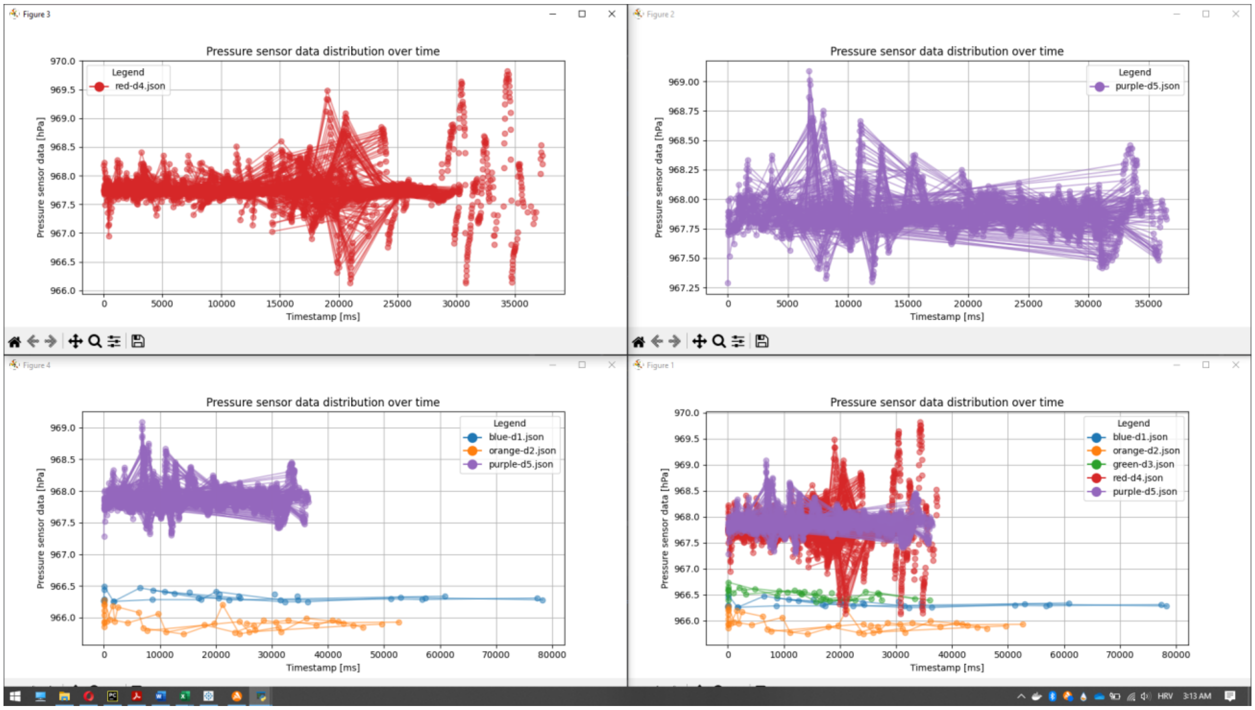
Conclusion:

Small and constant adjustments in the gyroscope data give info on the non-human activity for the following datasets:

- "Blue-d1.json"
- "Orange-d2.json"
- "Red-d4.json"

This info confirms our conclusion from the analysis of the touch event data.

2.4.4. Analyze- pressure sensor data



Conclusion:

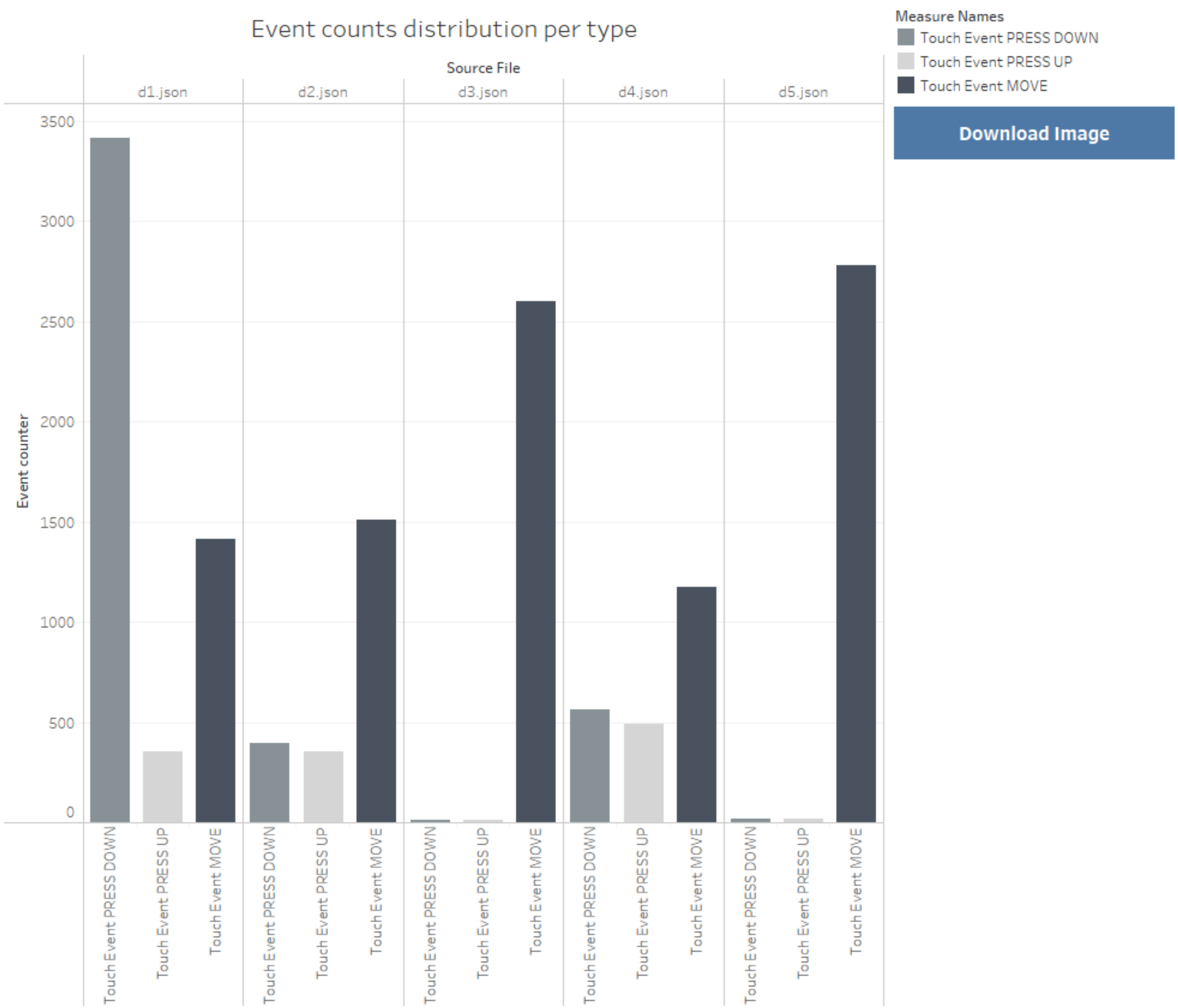
From the plotting pressure sensor data is possible to see a pretty regular-precise pattern in the red and blue dataset which can indicate non-human behavior.

Nothing can be concluded with a high probability from this data kind of plot. A further analysis of other events is done in the next chapters.

2.4.5 Analyze- touch events count distribution per dataset

For provided datasets, fewer “Press” events and more “Move” events are characteristic of the human-only behavior, as is the characteristic of provided datasets:

- "Green-d3.json"
- "Purple-d5.json"



3 Conclusion

Conclusions per performed analysis steps:

Dataset	Categorization of: 2.4.1. Analyse - touch events	Categorization of: 2.4.2 Analyse - accelerometer	Categorization of: 2.4.3 Analyse - gyroscope data	Categorization of: 2.4.4 Analyze - pressure sensor data	Categorization of: 2.4.5 Analyze - touch events count distribution per dataset
Blue - d1.json	Auto-clicker	Not possible to categorize	Auto-clicker	Auto-clicker	Auto-clicker
Orange - d2.json	Auto-clicker with anti-detection enabled	Not possible to categorize	Auto-clicker	Not possible to categorize	Auto-clicker
Green - d3.json	Human-only behavior	Human-only behavior	Human-only behavior	Not possible to categorize	Human-only behavior
Red - d4.json	Auto-clicker	Not possible to categorize	Auto-clicker	Auto-clicker	Auto-clicker
Purple - d5.json	Human-only behavior	Human-only behavior	Human-only behavior	Not possible to categorize	Human-only behavior

In final conclusion the provided datasets are categorized as:

Dataset	Final dataset category
Blue - d1.json	Auto-clicker
Orange - d2.json	Auto-clicker with anti-detection enabled
Green - d3.json	Human-only behavior
Red - d4.json	Auto-clicker
Purple - d5.json	Human-only behavior