SENSORS DATA FORMAT SPECIFICATION

All the smartphone associated data has been acquired through a self-developed application running on an Asus ZenFone AR (ZS571KL) with Android™ 7.0. This application relies on the standard APIs provided by the operating system for collecting the measurements from the available device sensors: magnetometer, accelerometer, gyroscope, ambient light, ambient temperature, air pressure, ambient noise level, and rotation vector. This last measurement represents the orientation of the device as a combination of an angle and an axis.

Additionally, we also gather all data related to nearby phone cell towers, phone signal intensity, device location, nearby WiFi networks and visible GPS satellites. Nearby WiFi netorks and phone cell towers information is acquired every 5 seconds due to the fact that the system limits the frequency at which an application is able to query the network status. Ambient noise level is also measured every 5 seconds, whereas the rest of measurements are acquired at the fastest rate provided by the system, which may vary for other devices.

GPS and cellphone data may not be present in some sequences depending on the availability of the connection.

Note that all sensors data files in the dataset contain extra fields for the position (x, y, z), rotation (x, y, z in radians) and time at the end of each line.

STANDARD SENSOR COORDINATE SYSTEM

When the device is held in its default orientation (vertical for the ASUS Zenphone AR), the axis are defined as shown in the figure. The X axis is horizontal and points to the right, the Y axis is vertical and points up, and the Z axis points toward the outside of the screen.

The axes are not swapped when the device's screen orientation changes, thus the sensor's coordinate system never changes as the device moves.

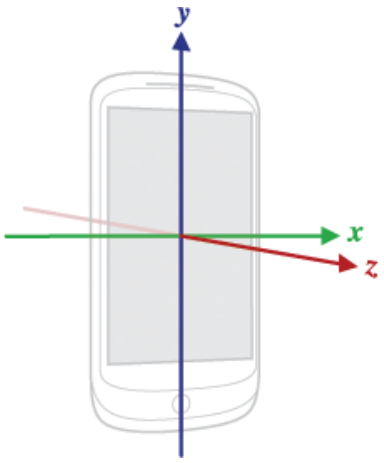

Figure from Android Documentation

ACCURACY

The accuracy field in the end of the data blocks described for all sensors contains a flag defined by the Android's API SensorManager class and can take one of the following values:

- SENSOR_STATUS_ACCURACY_HIGH = 3
  This sensor is reporting data with maximum accuracy
- SENSOR_STATUS_ACCURACY_MEDIUM = 2
  Calibration with the environment may improve the readings
- SENSOR_STATUS_ACCURACY_LOW = 1
  Calibration with the environment may be required.

MOTION SENSORS

MAGNETOMETER

The magnetic field sensor reports the ambient magnetic field measured in micro-Tesla (μT) along the three sensor axes (x, y and z). Factory calibration, temperature compensation and hard iron calibration is applied to the magnetic field measurements.

Data format (separated with spaces):

MAG	Timestamp	Magnetic Field @ X (μT)	Magnetic Field @ Y (μT)	Magnetic Field @ Z (μT)	0	0	Accuracy

ACCELEROMETER

The accelerometer measures the acceleration that is applied to a device by measuring the forces (including gravity) that are applied to the sensor itself.

Data format (separated with spaces):

ACC	Timestamp	Acceleration @ X (m/s^2)	Acceleration @ Y (m/s^2)	Acceleration @ Z (m/s^2)	0	0	Accuracy

GYROSCOPE

The gyroscope measures the rate of rotation around the device's x, y, and z axis. Measured rotation is positive in the counter-clockwise direction. We do not apply any filtering or correction for noise and drift to the rotational data.

Data format (separated with spaces):

GYR	Timestamp	Angular speed @ X (rad/s)	Angular speed @ Y (rad/s)	Angular speed @ Z (rad/s)	0	0	Accuracy

ATTITUDE (ROTATION VECTOR)

The rotation vector represents the orientation of the device. It is defined by an angle θ, in which the device has rotated around an axis (x, y, or z). This sensor uses the same coordinate system as the accelerometer.

The three elements of the rotation vector are equivalent to the last three components of a unit quaternion defined as (cos(θ/2), x*sin(θ/2), y*sin(θ/2), z*sin(θ/2)).

Data format (separated with spaces):

ATT	Timestamp	x * sin(θ/2)	y * sin(θ/2)	z * sin(θ/2)	cos(θ/2)	Heading Accuracy	Accuracy

In contrast to other sensors, the rotation vector is related to a specific coordinate system. This system has the following characteristics:

- X is defined as the vector product Y x Z. It is tangential to the ground at the device's current location and points approximately East.
- Y is tangential to the ground at the device's current location and points toward the geomagnetic North Pole.
- Z points toward the sky and is perpendicular to the ground plane.

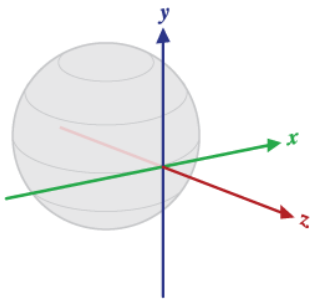

Figure from Android documentation

ENVIRONMENT SENSORS

AMBIENT LIGHT

Ambient illuminance measured in lux.

Data format (separated with spaces):

LUX	Timestamp	Illuminance (lux)	0	0	0	0	Accuracy

AMBIENT TEMPERATURE

Ambient temperature measured in °C.

Data format (separated with spaces):

TMP	Timestamp	Temperature (°C)	0	0	0	0	Accuracy

AMBIENT AIR PRESSURE

Ambient air pressure measured in hPa.

Data format (separated with spaces):

PRS	Timestamp	Ambient Air Pressure (hPa)	0	0	0	0	Accuracy

AMBIENT NOISE LEVEL

Ambient noise level measured by averaging the sound level registered by the device's microphone every 5 seconds.

Data format (separated with spaces):

MIC	Timestamp	Noise Level (dB)

NETWORKS

NEARBY WIFI NETWORKS

Nearby WiFi networks information is acquired every 5 seconds due to the fact that the system limits the frequency at which an application is able to query the network status.

Data format (separated with spaces):

WIFI	Timestamp	Number of detected networks	NET_1	NET_2	…	NET_N

For every detected network (NET_i):

SSID	BSSID	Level	Capabilities	CenterFreq0	CenterFreq1	Channel Width	Frequency	Operator Friendly Name	Venue Name

- **SSID**: network name.
- **BSSID**: address of the access point.
- **Level**: detected signal level in dBm.
- **Capabilities**: authentication, key management, and encryption schemes supported by the access point.
- **CenterFreq0**:
  - If the AP uses 40, 80 or 160 MHz, this is the center frequency (in MHz).
  - If the AP uses 80 + 80 MHz, this is the center frequency of the first segment (in MHz).
- **CenterFreq1**:
  - If the AP uses 80 + 80 MHz, this is the center frequency of the second segment (in MHz).
- **Channel Width**:
  - CHANNEL_WIDTH_20MHZ = 0
  - CHANNEL_WIDTH_40MHZ = 1
  - CHANNEL_WIDTH_80MHZ = 2
  - CHANNEL_WIDTH_80MHZ_PLUS_MHZ = 4 (Channel width is 80MHZ + 80MHZ)
- **Frequency**: the primary 20 MHz frequency (in MHz) of the channel over which the client is communicating with the access point.
- **Operator Friendly Name**: Passpoint operator name published by access point.
- **Venue Name**: Passpoint venue published by access point.

NEARBY CELL TOWERS

Nearby Phone cell towers information is acquired every 5 seconds due to the fact that the system limits the frequency at which an application is able to query the network status.

Data format (separated with spaces):

CELL	Timestamp	Number of detected towers	NET_1	NET_2	…	NET_N

For every detected GSM tower (NET_i):

gsm	Timestamp	Registered	Level	ASU level	dBm	CID	LAC	MNC	MCC	PSC	ARFCN	BSIC

- **gsm**: “gsm” text string.
- **Registered**: True if the phone is registered to a mobile network that provides service on this cell.
- **Level**: abstract level value for the overall signal quality 0..4.
- **ASU level**: RSSI in ASU 0..31, 99, or UNAVAILABLE.
- **dBm**: the RSSI of the measured cell.
- **CID**: 16-bit GSM Cell Identity described in TS 27.007, 0..65535, UNAVAILABLE if unavailable.
- **LAC**: 16-bit Location Area Code, 0..65535, UNAVAILABLE if unavailable.
- **MNC**: 2 or 3-digit Mobile Network Code, 0..999, UNAVAILABLE if unavailable.
- **MCC**: 3-digit Mobile Country Code, 0..999, UNAVAILABLE if unavailable.
- **PSC**: UNAVAILABLE - undefined for GSM
- **ARFCN**: 16-bit GSM Absolute RF Channel Number, UNAVAILABLE if unavailable.
- **BSIC**: 6-bit Base Station Identity Code, UNAVAILABLE if unavailable.

NOTE: UNAVAILABLE = 2147483647

For every detected WCDMA tower (NET_i):

wcdma	Timestamp	Registered	Level	ASU level	dB	CID	LAC	MNC	MCC	PSC	UARFCN

- **wcdma**: “wcdma” text string.
- **Registered**: True if the phone is registered to a mobile network that provides service on this cell.
- **Level**: abstract level value for the overall signal quality 0..4.
- **ASU level**: RSSI in ASU 0..31, 99, or UNAVAILABLE.
- **dBm**: the RSSI of the measured cell.
- **CID**: 16-bit GSM Cell Identity described in TS 27.007, 0..65535, UNAVAILABLE if unavailable.
- **LAC**: 16-bit Location Area Code, 0..65535, UNAVAILABLE if unavailable.
- **MNC**: 2 or 3-digit Mobile Network Code, 0..999, UNAVAILABLE if unavailable.
- **MCC**: 3-digit Mobile Country Code, 0..999, UNAVAILABLE if unavailable.
- **PSC**: 9-bit UMTS Primary Scrambling Code described in TS 25.331, 0..511, UNAVAILABLE if unavailable.
- **UARFCN**: 16-bit UMTS Absolute RF Channel Number, UNAVAILABLE if unavailable.

NOTE: UNAVAILABLE = 2147483647

For every detected LTE tower (NET_i):

lte	Timestamp	Registered	Level	ASU level	dB	CI	PCI	MNC	MCC	TAC	EARFCN

- **lte**: “lte” text string.
- **Registered**: True if the phone is registered to a mobile network that provides service on this cell.
- **Level**: abstract level value for the overall signal quality 0..4.
- **ASU level**: RSSI in ASU 0..31, 99, or UNAVAILABLE.
- **dBm**: the RSSI of the measured cell.
- **CI**: 28-bit Cell Identity, UNAVAILABLE if unavailable.
- **PCI**: Physical Cell Id 0..503, UNAVAILABLE if unavailable.
- **MNC**: 2 or 3-digit Mobile Network Code, 0..999, UNAVAILABLE if unavailable.
- **MCC**: 3-digit Mobile Country Code, 0..999, UNAVAILABLE if unavailable.
- **TAC**: 16-bit Tracking Area Code, UNAVAILABLE if unavailable.
- **EARFCN:** 18-bit Absolute RF Channel Number, UNAVAILABLE if unavailable.

NOTE: UNAVAILABLE = 2147483647

PHONE SIGNAL STRENGTH

Phone signal strength is represented as a single integer from 0 to 4, where 0 refers to very poor signal and 4 represents a very strong signal strength.

Data format (separated with spaces):

PHONE	Timestamp	Level	0	0	0	0	0

DEVICE LOCATION

A precise device location obtained through Android API LocationManager class. All location data is guaranteed to contain valid timestamp, longitude and latitude fields, while all other fields are optional.

Data format (separated with spaces):

LOC	Timestamp	Longitude	Latitude	Altitude	Speed	Provider	Number of satellites	Accuracy

GPS / GNSS STATUS

Data format (separated with spaces):

GPS	Timestamp	Number of visible satellites