CATEGORIZING AND ADJUSTING DATA QUALITY INDICATORS FOR ANDROID MOTION AND ENVIRONMENT SENSORS

- Ankan Mookherjee, M.S. Computer Science, axm3244@g.rit.edu, 312-508-8693

-

R·I·T
B. Thomas Golisano College of Computing and Information Sciences

PROJECT GOALS

- Capture motion and environment sensor data of Android Devices
- Carry out Data Quality analysis on sensor data to verify accuracy and trustworthiness
- Extract Data Quality (DQ) Indicators from the data and perform continuous integration of the data

MOTIVATION FOR WORK

- Smartphones produce immensely valuable environment sensor data like Barometer, Relative Humidity and Temperature data which helps in critical scientific pursuits across a wide variety of disciplines.
- Motion sensor data produced in Android phones can have calibration error due to damaged device.

DATA QUALITY INDICATOR

Attributes to extract & calculate DQ Metric

- ***Pressure*** - Ambient Real Time Air pressure vs Air Pressure from Open Weather API
- ***Gravity*** - Gravity Calculated From Phones Vs Actual Gravity at the latitude
- ***Gravity vs Accelerometer***
- ***Magnetic Orientation using Sensor Fusion***
  - Comparison with Gyroscope data
  - Comparison with Accelerometer data
- ***Battery Température*** vs Ambient Air Température

WORKFLOW

- Creating ***SensorTest*** app to record Sensor Data.
- Device for Cloud integration – Storing device data in Cloud Data Base.
- Calculating calibration error for motion sensor data.
- Calculating percentile of the *average difference* of a device and *current difference* with overall data collected from various devices.
- Calculating DQ indicators from percentiles.
- Weighing the data from all indicators to create a complete DQ score for a device.
- Creating results and continuous integration of DQ Indicator.
- Readjustment of DQ indicators with new incoming data by scheduling a time based cron job as worker role.

ARCHITECTURE AND DESIGN IMPLEMENTATION

Figure 1. High-Level Architecture

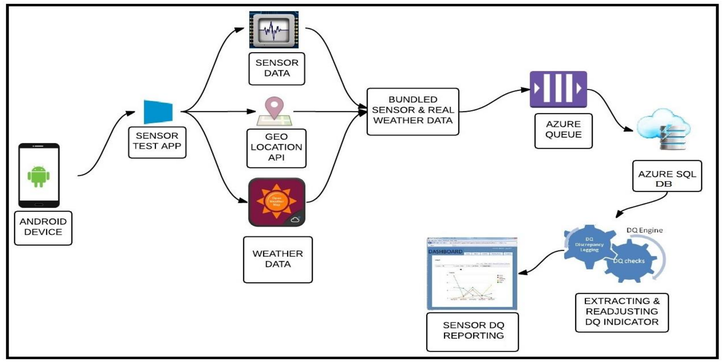


Figure 2. Sensor Fusion

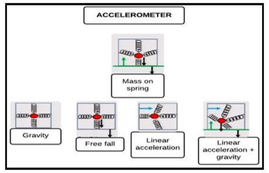


Figure 3. Technology

Figure 4. Gravity & Accelerometer Sensor

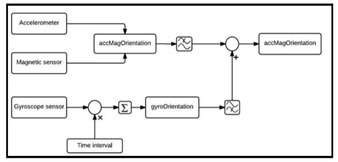
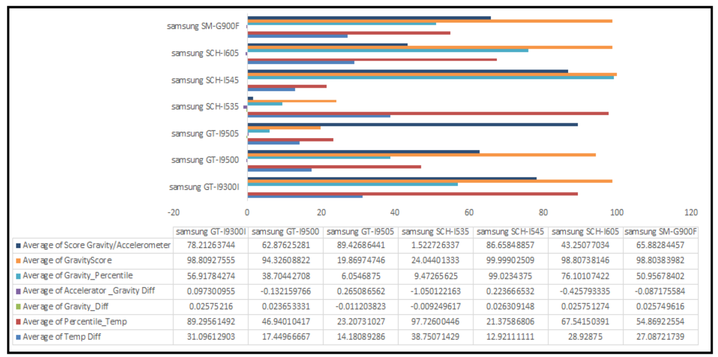
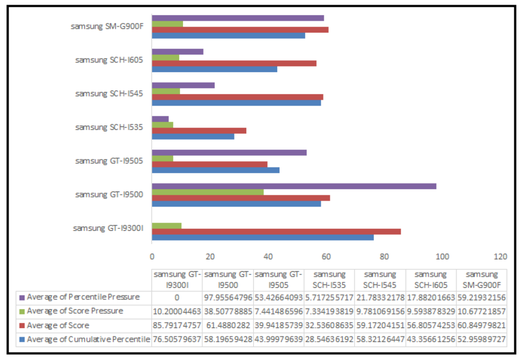


Figure 6. Result 1: Motion Sensor & Temperature Sensor Device DQ Indicator



	samsung GT-I9300I	samsung GT-I9500	samsung GT-I9505	samsung SCH-I535	samsung SCH-I545	samsung SCH-I605	samsung SM-G900F
Average of Score Gravity/Accelerometer	78.21263744	62.87625281	89.42686441	1.522726337	86.65848857	43.25077034	65.88284457
Average of GravityScore	98.80927555	94.32608822	19.86974746	24.04401333	99.99902509	98.80738146	98.80383982
Average of Gravity_Percentile	56.91784274	38.70442708	6.0546875	9.47265625	99.0234375	76.10107422	50.95678402
Average of Accelerator _Gravity Diff	0.097300955	-0.132159766	0.265086562	-1.050122163	0.223666532	-0.425793335	-0.087175584
Average of Gravity_Diff	0.02575216	0.023653331	-0.011203823	-0.009249617	0.026309148	0.025751274	0.025749616
Average of Percentile_Temp	89.29561492	46.94010417	23.20731027	97.72600446	21.37586806	67.54150391	54.86922554
Average of Temp Diff	31.09612903	17.44966667	14.18089286	38.75071429	12.92111111	28.92875	27.08721739

Figure 7. Result 2: Pressure Sensor & Cumulative Device DQ Indicator



	samsung GT-I9300I	samsung GT-I9500	samsung GT-I9505	samsung SCH-I535	samsung SCH-I545	samsung SCH-I605	samsung SM-G900F
Average of Percentile Pressure	0	97.85564796	53.42664093	5.717255717	21.78332178	17.88201663	59.21932156
Average of Score Pressure	10.20004463	38.50778885	7.441486596	7.334193819	9.781069156	9.593878329	10.67721857
Average of Score	85.79174757	61.4880282	39.94185739	32.53608635	59.17204151	56.80574253	60.84979821
Average of Cumulative Percentile	76.50579637	58.19659428	43.99979639	28.54636192	58.32126447	43.35661256	52.95989727

LESSONS LEARNED

- The Gravity Sensor was unable to supply a lot of useful information. The filters in the phone ensured that Raw data was acquired while the sensors were stationary on a stable desk.
- Ambient air temperature comparison using battery temperature was not extremely accurate strategy to generate sensor data quality as room temperature is different from external temperature (due to data collected indoor). However if trained with larger datasets and accurate ambient temperature information we can device machine learning system which has the potential to determine ambient air temperature and battery diagnosis tool in future.
- No correlation was found between phone security and sensor data calibration error due to lack of data from real untrustworthy source. However in theory intrusion detection and smartphone Trojan systems can acquire sensor data use it for unscrupulous purposes. Privacy information can be compromised with unauthorized use, such as by secretly collecting user geolocation and motion sensor information using an app.
- In theory, over a period of time sensors should degrade and relay less accurate information, this tool can be used to keep a track on degradation, do a diagnostic device sensor check and recalibration of the sensor devices.
- Data from pressure sensor has higher precession than geolocation based weather prediction API's. I believe if pressure sensor data is securely collected then it can lead to better meteorological data models and micro weather prediction systems as the data collect from android smartphone sensors has more precession.

RESULTS

- Device-to-cloud integration.
- Filtering and cleansing data and notification system within the app as well as at the DQ Engine.
- More than 1000 instances of collected sensor data from over 11 smartphones and 7 devices within 4 weeks.
- Data correlation for different devices and average DQ metric for a smartphone and a particular device.
- DQ Metric for an instance of data sent in the overall data collected for the device and that collected from various sources
- Samsung Galaxy S4 (SG-I9500) was the most reliable device out of the devices data collected so far.
- Sensor fusion helped ease out noise from gyroscope sensor and accelerometer data.

FUTURE WORK

- To release iPhone, Windows Phone SensorTest App. I have designed the android app using Xamarin with high code reusability factor.
- Diagnostic tool within app to get sensor calibration and accuracy information.
- Data broker to securely share high quality trustworthy sensor information.
- Want to carry out a study on Battery Temperature to predict ambient air temperature.
- Carry out further data mining algorithms on collected data and add sensor data for other devices.

ACKNOWLEDGEMENT

- Professor Leon Reznik, Project Advisor and Project , Coordinator, RIT Computer Science, lr@cs.rit.edu
- Professor Justin Cappos, Department of Computer Security, NYU Polytechnic, cappos@nyu.edu

REFERENCES

[1]J. Cappos, L. Wang, R. Weiss, Y. Yang, and Y. Zhuang. BlurSense: Dynamic fine-grained access control for smartphone privacy. Sensors Applications Symposium (SAS), 2014.

[2]] L. Reznik, "Integral Instrumentation Data Quality Evaluation: the Way to Enhance Safety, Security, and Environment Impact," presented at the 2012 IEEE International Instrumentation and Measurement Technology Conference, Graz, Austria, May 13-16, 2012, 2012.

[3] Sachs, David "Sensor fusion on android devices: A revolution in motion processing", GoogleTech Talks, 2010

[4] Milette Greg, Stroud Adam, "Professional Android Sensor programming", ISBN: 978-1-118-18348-9, John Wiley and Sons, 2012