Using a Reputation System to Produce Trustworthy Rainfall Estimates from Crowdsourced Data:

A Case Study in Durham, North Carolina

Alexander B. Chen
Jonathan L. Goodall
Flooding is becoming commonplace in cities worldwide.
~1000 square km
<10 agency rain gauges
Radar-derived rainfall is usually not accurate and has to be adjusted with rain gauges.
Personal Weather Station (PWS)

Temperature  Humidity  Wind speed  Rainfall
The number of PWS is 10 times higher than agency rain gauges.
<table>
<thead>
<tr>
<th></th>
<th>Agency Rainfall Networks</th>
<th>PWS Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner</td>
<td>Trained Personnel</td>
<td>Hobbyists</td>
</tr>
<tr>
<td>Siting</td>
<td>Adhere to specific standards</td>
<td>Based on owner’s effort</td>
</tr>
<tr>
<td>Sensor Type</td>
<td>Certified sensors</td>
<td>Based on affordability</td>
</tr>
<tr>
<td>Calibration</td>
<td>Rigorous</td>
<td>Unknown</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Routine</td>
<td>Unknown</td>
</tr>
</tbody>
</table>
How can we trust personal weather stations?
Forecast for Charlottesville, VA:

- Temperature: 68.0°F
- Feels Like: 63.0°F
- Dew Point: 45°F
- Humidity: 43%
- Precipitation Rate: 0.0 in
- Precipitation Accumulation: 0.0 in
- Pressure: 30.05 in

Current Conditions:

- Gusting Winds:
- Wind Speed:

Map showing Charlottesvile, VA with weather radar and a map legend.

Weather History for Charlottesville, VA [KVACHAR114]

Summary:
Oct 23, 2018
Trust:
Collective opinion of neighboring PWSs about the behavior of a subject PWS
Crowdsourced Personal Weather Station

Areal Rainfall Estimation Method

Trust score: Poor, Good

Trustworthy Rainfall Estimate
Reputation System for Crowdsourced Rainfall Networks

- Cluster
- Consensus
- Score
Cluster

Partition PWSs into groups that report similar data

Method: $k$-Means

Input:
- Latitude
- Longitude
- Elevation
Cluster

Partition PWSs into groups that report similar data

Method: *k*-Means

Input:

- Latitude
- Longitude
- Elevation
Consensus

Find the consensus and the deviation from the consensus within the cluster

Method: *Robust Averaging*

Simple Average \( t = 35.9 \)

Robust Average \( t = 40.4 \)

**Cooperative metric** \( C_i(t) = \frac{W_i(t) - \bar{W}}{\sigma(W)} \)

*\( W_i(t) \): robust weight  \* \( \bar{W} \): average weight  \* \( \sigma(W) \): std weight*

**Cooperative behaviors** \( C_i(t) > 0 \)

**Non-cooperative behaviors** \( C_i(t) < 0 \)
Score:

Manages and represents reputation information

Method: Beta Reputation System

\[ f(p|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} \]

TS (Trust Score) \sim E(p|\alpha, \beta) = \frac{\alpha}{\alpha + \beta}

Initial neutral trust score

\[ TS_{t=0} = 10 \times E(p|1,1) = 5.0 \]

Input: Cooperative metric

If \( C_t > 0 \)

\[ \alpha_{t+1} = \alpha_t \times ff + C_t, \quad \beta_{t+1} = \beta_t \times ff \]

If \( C_t < 0 \)

\[ \alpha_{t+1} = \alpha_t \times ff, \quad \beta_{t+1} = \beta_t \times ff + |C_t| \]

Forgetting Factor (\( ff \)): \( 0 \leq ff \leq 1 \)

Output: Trust score

\[ TS_t = 10 \times E(p) = \frac{\alpha_t}{\alpha_t + \beta_t} \]

\[ TS_{t+1} = 10 \times E(p) = \frac{\alpha_{t+1}}{\alpha_{t+1} + \beta_{t+1}} \]
4 trustworthy vs. 1 untrustworthy PWS

Trust score evolution
3 trustworthy vs. 2 untrustworthy PWSs

Trust scores evolution
Reputation System Results
Case study: Durham, North Carolina, USA

107 PWSs
149 days of rainfall data
K-Means cluster using longitude, latitude, and elevation
Trustworthy Rainfall Estimates
2018/07/29 Rainfall Observation (inch)

PWS Trust Score < 5.0
Replace the observation with the robust average
USGS Rain Gauge: 2.39 in.
Simple IDW: 2.09 in.
### All 149 Days

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE(in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple IDW</td>
<td>0.313</td>
</tr>
<tr>
<td>Trust Score IDW</td>
<td>0.225</td>
</tr>
</tbody>
</table>

### Days with rainfall > 1 in

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE(in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple IDW</td>
<td>0.611</td>
</tr>
<tr>
<td>Trust Score IDW</td>
<td>0.361</td>
</tr>
</tbody>
</table>
Conclusion
• Crowdsourced PWSs are filling in data gaps of agencies data but introduce trust gap for utilizing them.

• A reputation system method can effectively bridge this trust gap by evaluating the trustworthiness of the crowdsourced PWSs.

• Using trustworthy rainfall estimate method can improve the knowledge of rainfall patterns in areas with dense PWSs.