



# Techniques for Fault Detection and Visualization of Telemetry Dependence Relationships for Root Cause Fault Analysis in Complex Systems

---

**Nathaniel Guy**

**(With support of Dr. Mehran Mesbahi and Dr. Jeffrey Heer)**

**Robotics, Aerospace, and Information Networks Lab**

**University of Washington - William E. Boeing Aeronautics and Astronautics**

**MSAA Thesis Presentation – March 16<sup>th</sup>, 2016**

- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

- **Motivation and Objectives**
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

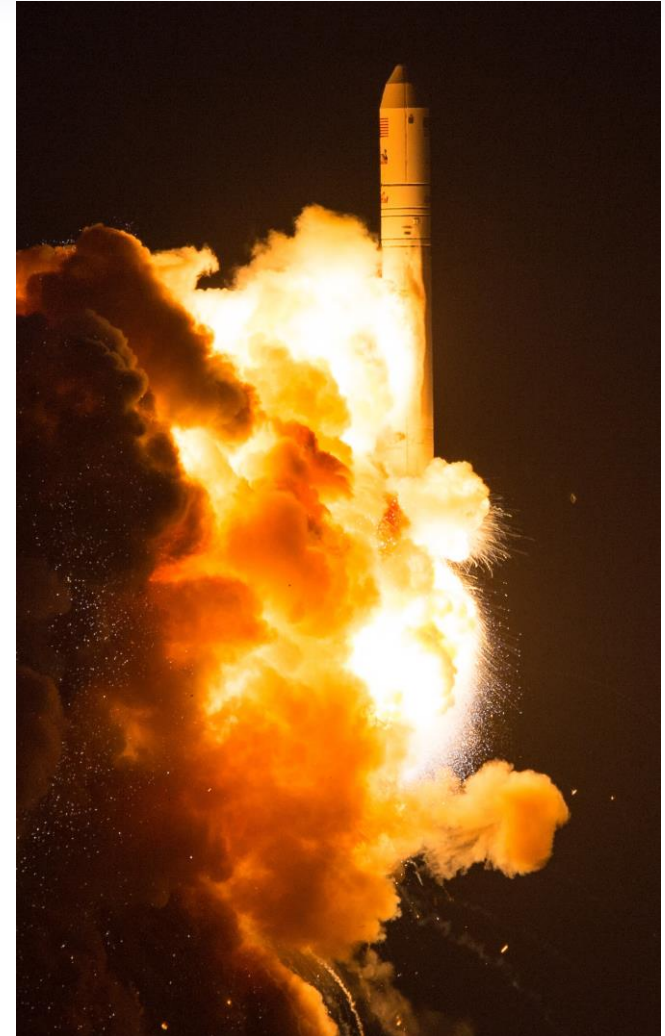
# Complex, Remote-Operated Systems

- Many types of systems must be teleoperated
  - Space systems
  - Underwater ROVs
  - Robots for radioactive environments
- Difficult to troubleshoot
- Large amounts of data
- Data analysis and visualization becomes difficult
- Vehicles such as the Dragon spacecraft have thousands of data channels!



# Antares CRS Orb-3

- Orbital Sciences October 2014, ISS cargo delivery
- Large explosion 6 seconds after launch
- Catastrophic failure; cargo completely destroyed
- What caused this?
- Investigation immediately launched
- Still no conclusive root case given; only vaguely linked to engine



- Determining root cause for a problem very difficult
- Problem detection often doesn't capture cause
- Very human-driven, but tools too simple
  - Many tools simply consist of plotting data channel time series
- Need better tools for improving data review and linking detected problems to uncaptured causes

- Find analytical techniques that can improve root cause analysis
- Make user-facing tools to improve human insight into system dynamics
- Test above techniques on actual aerospace system data

- Motivation and Objectives
- **Model Definition**
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References



- Data streams in discretely from a remote system
- This data is “telemetry”
  - Combination of sensor readings and software metrics
  - Velocity, temperature, rotations, attitude, dropped packets, etc.
  - Higher-level telemetry: camera images, spectrograms, etc.
- Often thousands of telemetry channels with frequent updates



- State modeled discretely:

$$x(t+1) = A(t)x(t) + B(t)u(t) + E_1n_1(t)$$

$$y(t) = C(t)x(t) + D(t)u(t) + E_2n_2(t)$$

- $x(t)$ : system state at time  $t$
  - $y(t)$ : system measurement at time  $t$
  - $u(t)$ : system input at time  $t$
  - $n_1(t)$  and  $n_2(t)$ : system disturbance vectors at time  $t$
  - $A(t), B(t), C(t), D(t), E_1(t), E_2(t)$ : linear dynamics
- 
- Telemetry data streams in, giving us values for  $y(t)$
  - Individual elements of  $y(t)$  are called **telemetry channels**

- Motivation and Objectives
- Model Definition
- **Fault Detection Background**
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

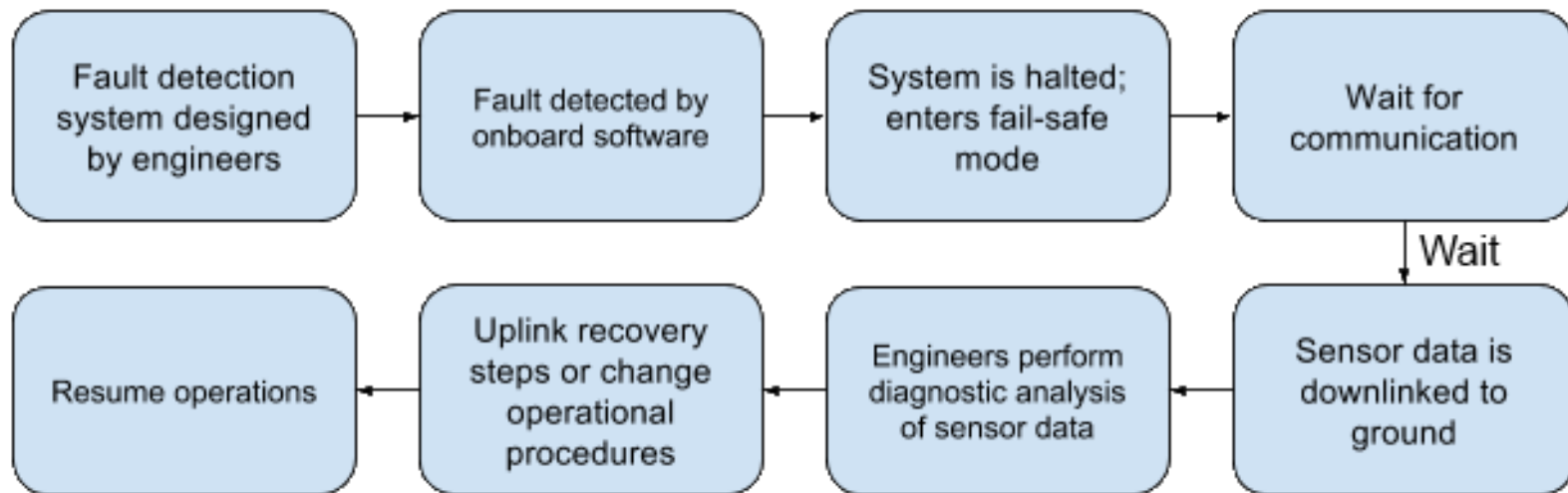
# Fault Detection, Isolation, and Recovery (Definitions)

- A **fault** is any system state that is undesirable
- **FDIR** is Fault...
  - **Detection**: deciding that a fault occurred
  - **Isolation**: determining where the fault occurred
  - **Recovery**: restoring the system to a non-faulted state
- Normal/expected data is **nominal**
- Otherwise, it's **anomalous**
- Faults can be divided into **fault levels** (by severity, or subsystem hierarchy)
- Faults can trigger notifications to human operators, called **alarms**

# Alarms



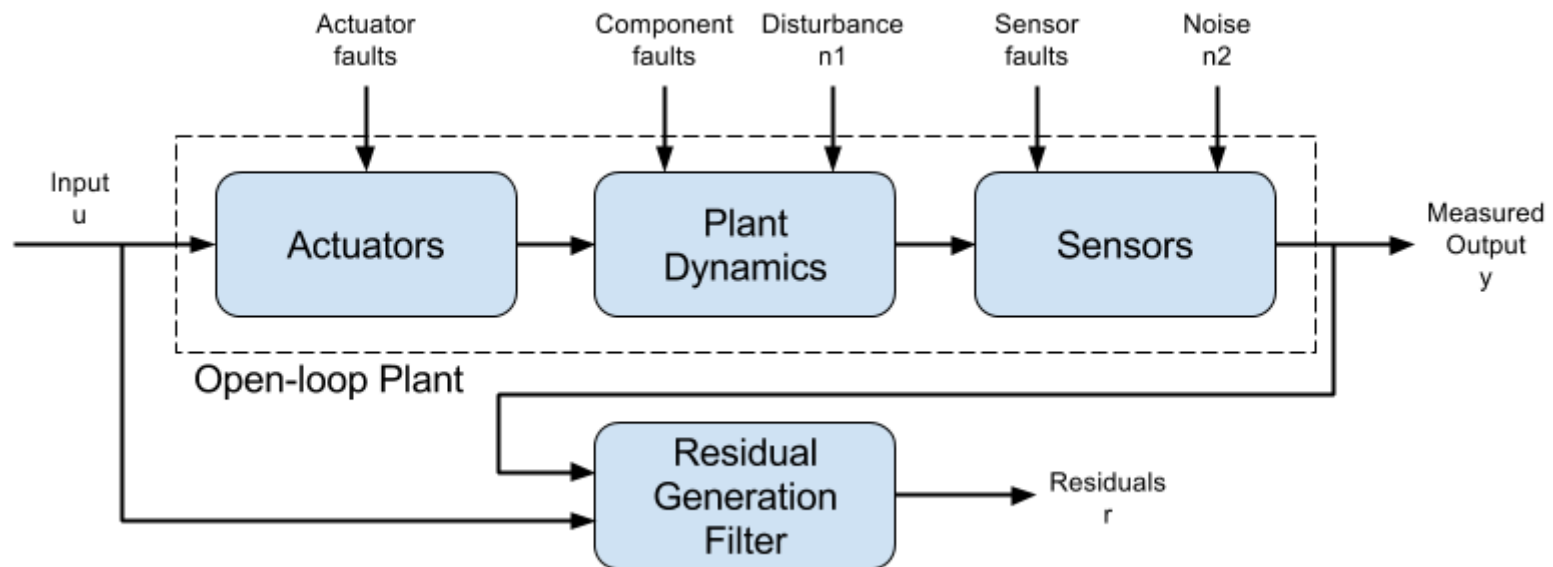
# Fault Detection, Isolation, and Recovery (Process)



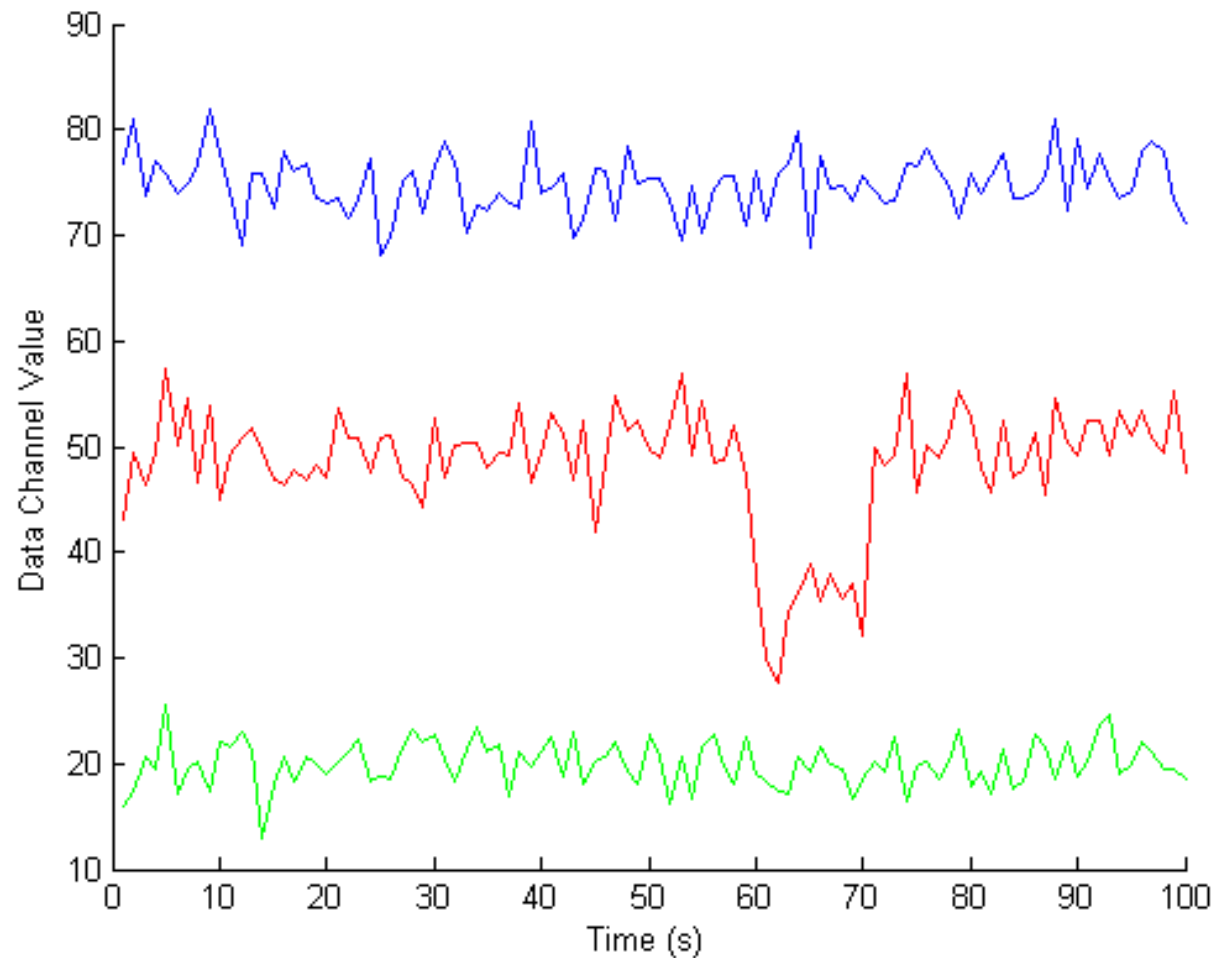
- Many faults cannot be recovered from automatically
- Complex issues require human intervention

# Residual Generation and Fault Detection

- Measured output compared to expected system model to find difference
- This “residual” used to detect faults
- Often, residual simply is a deviation from the expected mean
  - If residual falls outside a given threshold, a fault may be triggered (e.g., “high acceleration”)

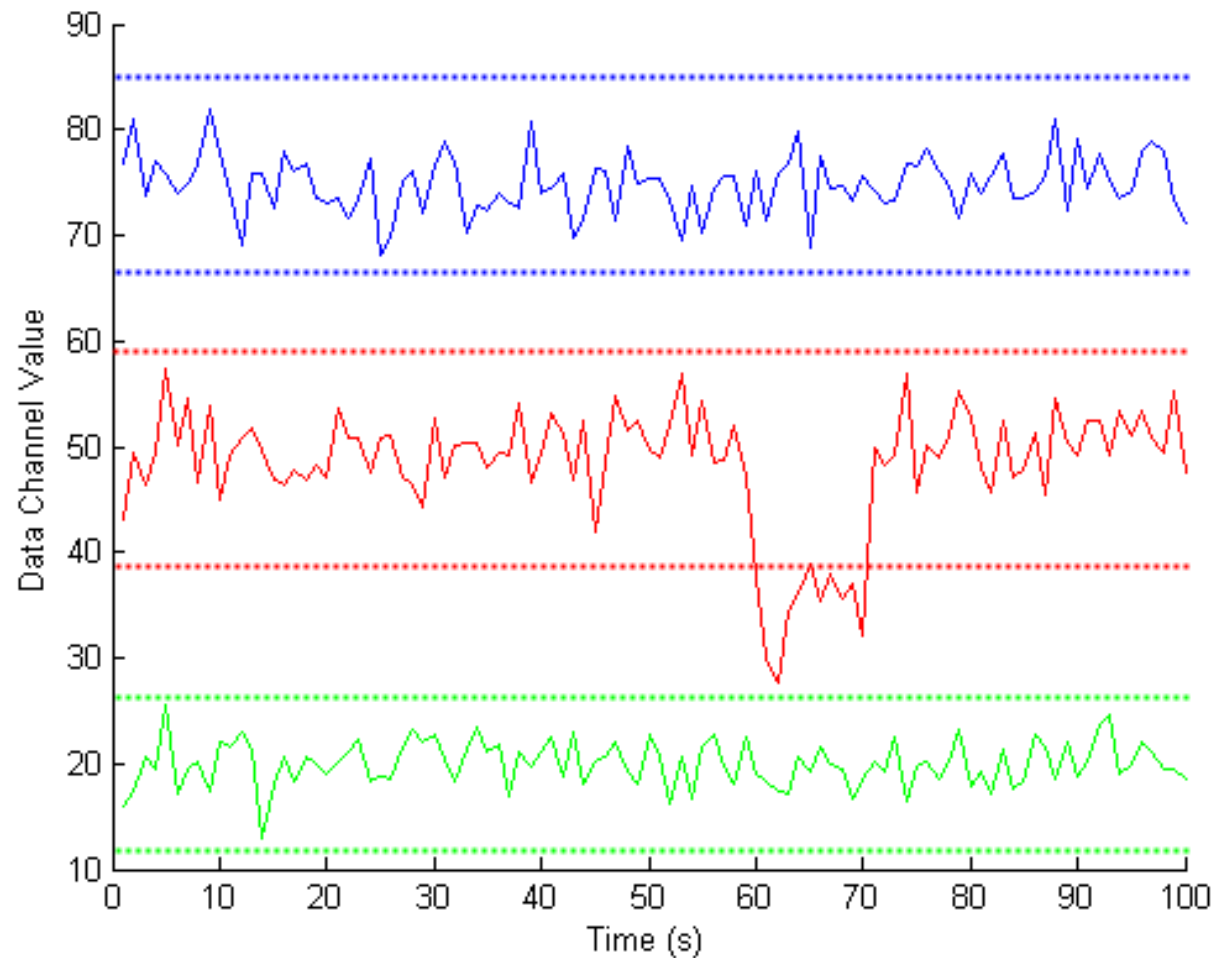


# Threshold Fault Detection Example

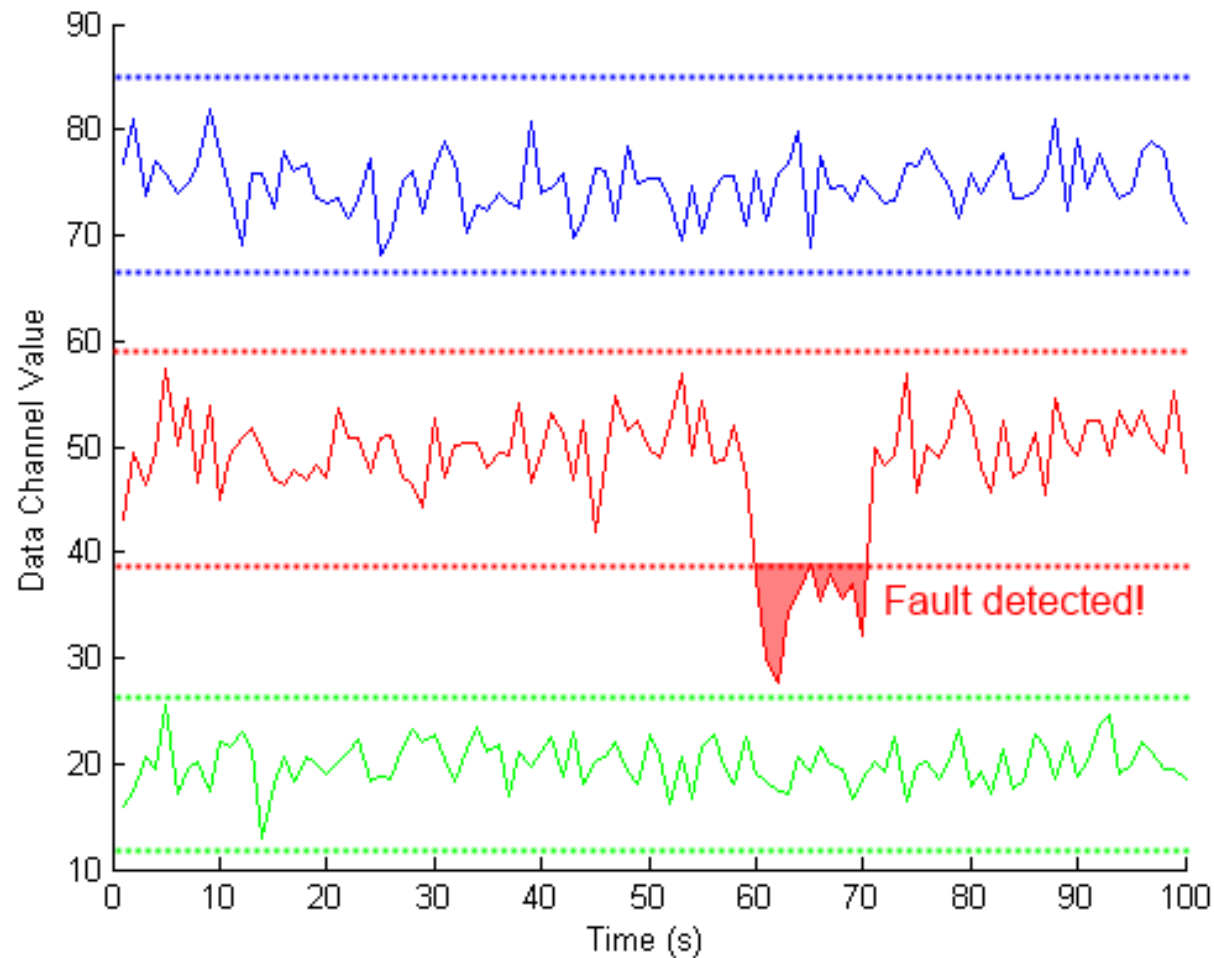




# Threshold Fault Detection Example



# Threshold Fault Detection Example



# Other Fault Detection Schemes

- Many other types of fault detection
  - Combinations of complex rules
  - Probabilistic models
  - Machine learning from nominal and anomalous data
  - Redundant hardware systems
- We'll be focusing on models to find deeper root causes which can be agnostic to underlying FD system

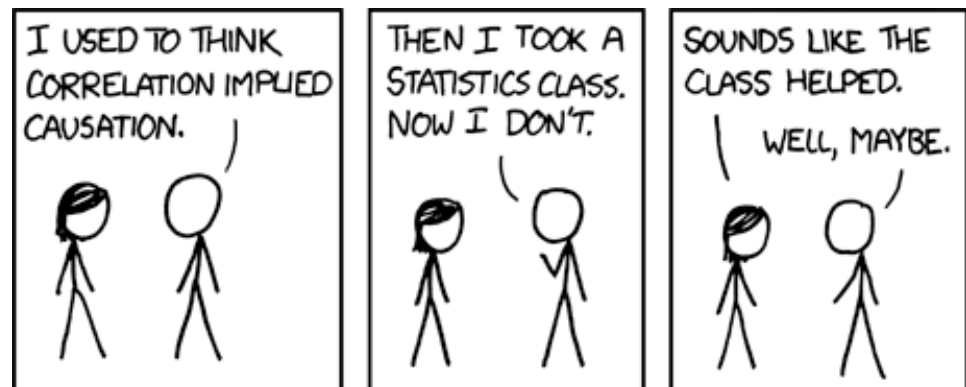
- Motivation and Objectives
- Model Definition
- Fault Detection Background
- **Correlative Analysis and Visualization Background**
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

# Correlative Analysis: Motivation

- Traditional FDIR only captures modeled faults
- Channels contributing a fault residual may be “linked” to other channels
- Or relationships at the time of a fault may be different from an earlier state, though raw telemetry is mostly nominal
- Example: temperature and wheel rotation
- Looking for correlations between channels may be able to find these links
  - Important to remember that correlation does not necessary imply causation, but can be suggestive

# Correlative Analysis: Motivation

- Traditional FDIR only captures modeled faults
- Channels contributing a fault residual may be “linked” to other channels
- Or relationships at the time of a fault may be different from an earlier state, though raw telemetry is mostly nominal
- Example: temperature and wheel rotation
- Looking for correlations between channels may be able to find these links
  - Important to remember that correlation does not necessary imply causation, but can be suggestive



- **Covariance** is strength of correlation between sets of random variables
  - Shows how they “change together”

$$\text{cov}(X, Y) = \mathbb{E}[(x - \mu_x)(y - \mu_y)]$$

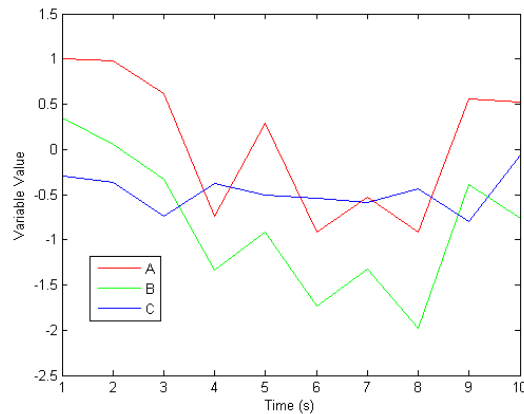
- **Covariance matrices** show all of the covariances between variables
  - For a vector of variables  $x \in \mathbb{R}^n$ , there is a symmetric covariance matrix  $\Sigma \in \mathbb{R}^{n \times n}$  such that  $\Sigma_{ij} = \text{cov}(x_i, x_j)$ .
  - Diagonal shows the variance of each variable

# Correlation Metrics: PCC

- We can scale a covariance matrix to get a normalized correlation between variables
- Most common method is Pearson Correlation Coefficient (PCC):

$$PCC_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

- PCC correlation “score” ranges from 1 (perfect positive correlation) to -1 (perfect negative correlation)
- 0 indicates no correlation

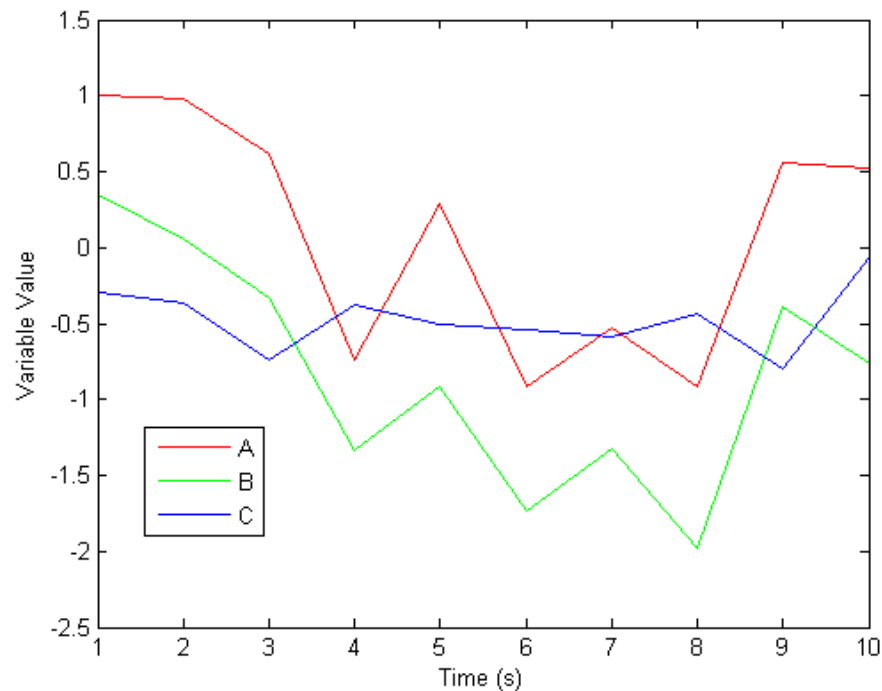


## PCC Correlations

	A	B	C
A	1.000	0.959	0.105
B	0.959	1.000	0.057
C	0.105	0.057	1.000



# Correlation Metrics: PCC



## PCC Correlations

	A	B	C
A	1.000	0.959	0.105
B	0.959	1.000	0.057
C	0.105	0.057	1.000

# Correlation Metrics: Rho and Tau

- PCC is limited as a correlation metric
  - Assumes linear relationship between variables
  - Assumes interval or ratio variables (not ordinal)
  - Assumes bivariate normal distribution
- Two common alternatives that don't have the above assumptions
- Rank correlations (depend on how values sort)—only need monotonicity
- Spearman's Rho

$$\rho_{X,Y} = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

$d_i$ : difference in ranking  
 $n$ : number of observations

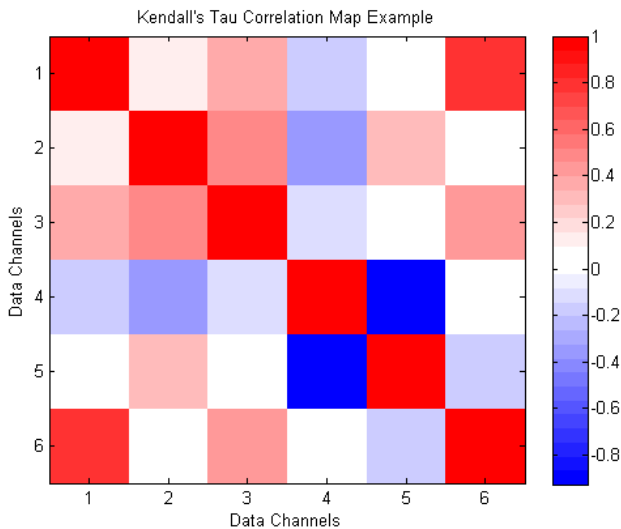
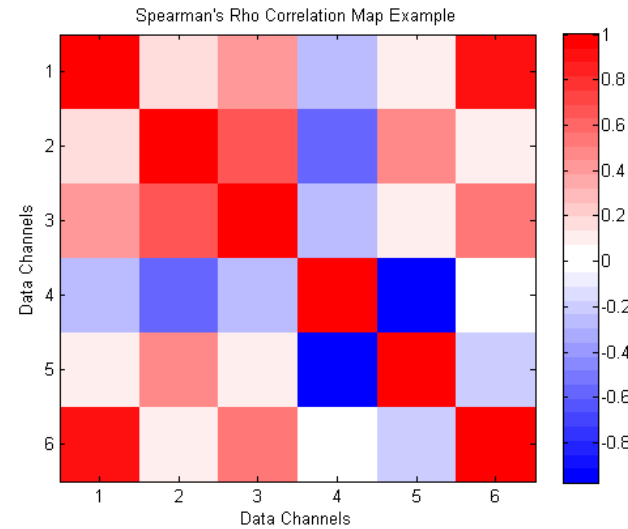
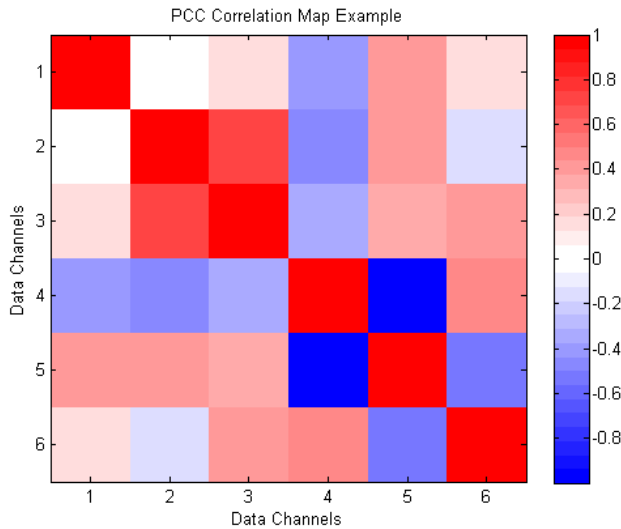
- Kendall's Tau

$$\tau_{X,Y} = \frac{n_c - n_d}{n(n^2 - 1)/2}$$

$n_c$ : # of pairs with consistent ranking  
 $n_d$ : # of pairs with inconsistent ranking  
 $n$ : number of observations

- Correlation matrices often shown as shaded grid (“corrgram”)
- Each row is a variable; each column is a variable
- Color of cell indicates correlation (usually dark for strong; light for weak)

# Correlation Metrics: Corrgram Comparison

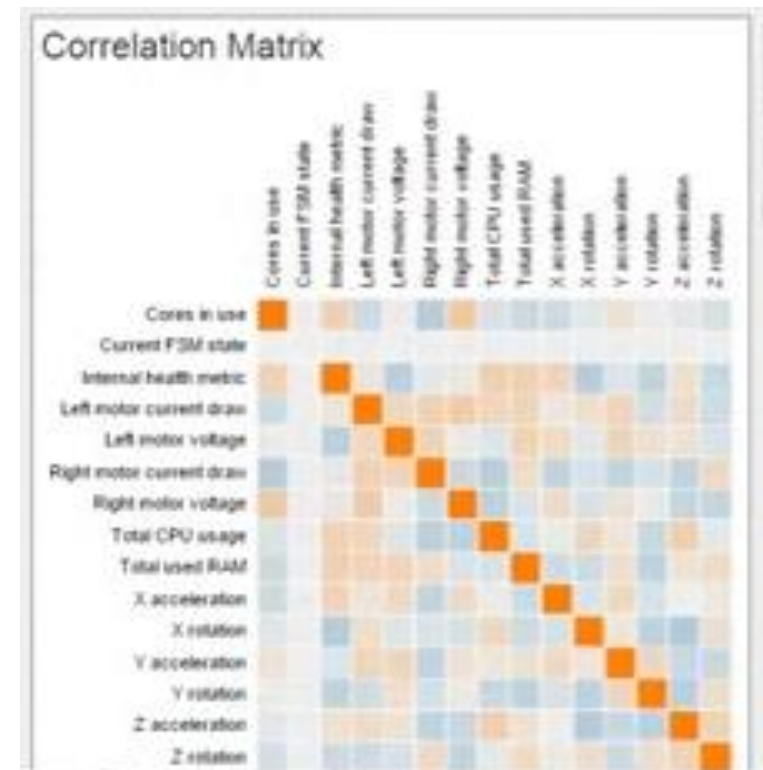


Correlated Data Values (Rows are Data Channels)

0	10	101	102	0	10	101	400
4.2333	4.2491	4.2437	4.1147	4.0428	4.0337	4.4442	4.1166
6.2585	6.2135	6.2618	6.2814	6.0419	6.1182	6.2942	6.1934
1.2	2.3	3.4	4.4	4.5	3.5	2.2	1.1
-10	-20	-30	-40	-40	-30	-20	-10
1	2	500	2000	1	100	200	201

# Animated Corrgrams

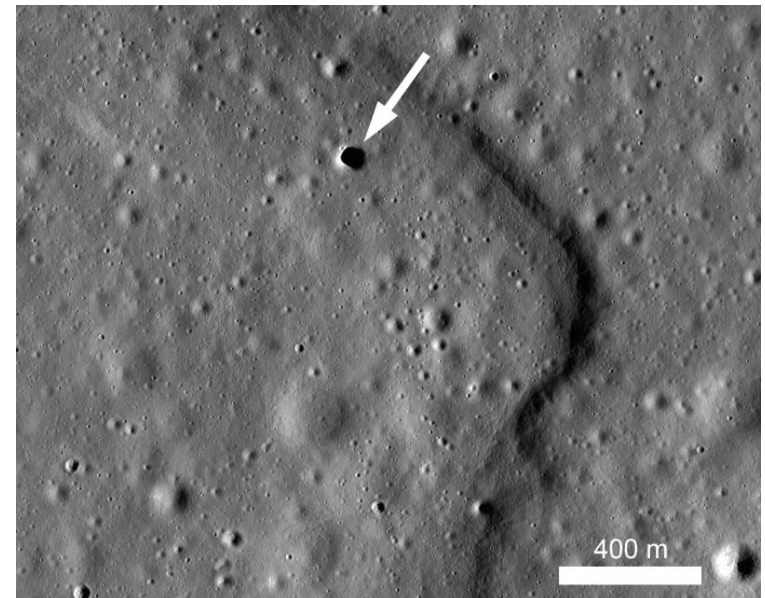
- Corrgrams can only capture data at one point in time
- If instantaneous correlation changes over time, we can capture this by sampling a time window
- We can then update a corrgram with the latest correlations
- Uncommon, but useful



- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- **Lunar Rover Ground Station Development**
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

# Hakuto and the Google Lunar XPRIZE

- Google Lunar XPRIZE competition
- Japanese team “Hakuto”
- Based at Tohoku University, Sendai, Japan
- Mission to the Moon in 2017
  - Complete GLXP requirements
  - Explore lunar caves



# Moonraker Rover

- Four-wheeled, 4kg micro-rover
- In development since 2010
- Four visible-spectrum cameras, 1 depth camera
- Many other sensors
- ~112 telemetry channels





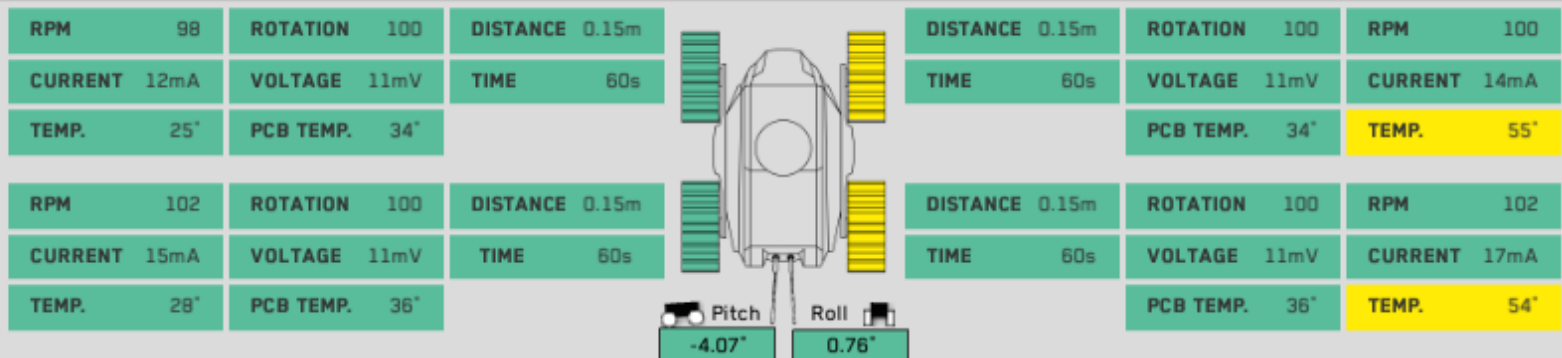
- Interface for communicating with Moonraker
- New system from scratch
- Priorities
  - Easy to understand
  - Easy commanding
  - Robust to human error
  - Maintainable



# GSN Feature: Telemetry Display

- Telemetry displayed on the screen when received
- Telemetry grouped by subsystem

## MOBILITY



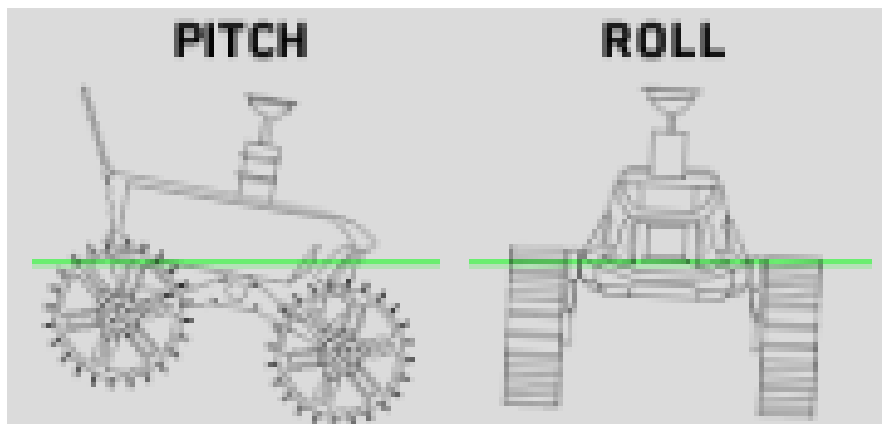
# GSN Feature: Telemetry Change Indicators

- Telemetry trends shown with arrow-based change indicator system

DISTANCE 16.70m ↑	ROTATIONS 10993 ↑	RPM 3001
TIME 219s ↑	VOLTAGE 12mV	CURRENT 16mA
	PCB TEMP. 44° ↓	TEMP. 44° ↓

# GSN Feature: Tachometer/Attitude Display

- Pitch and roll
  - Calculated from accelerometer telemetry
- RPM
  - Averaged from four-wheel motor telemetry



# GSN Feature: Saving/Loading/Playback

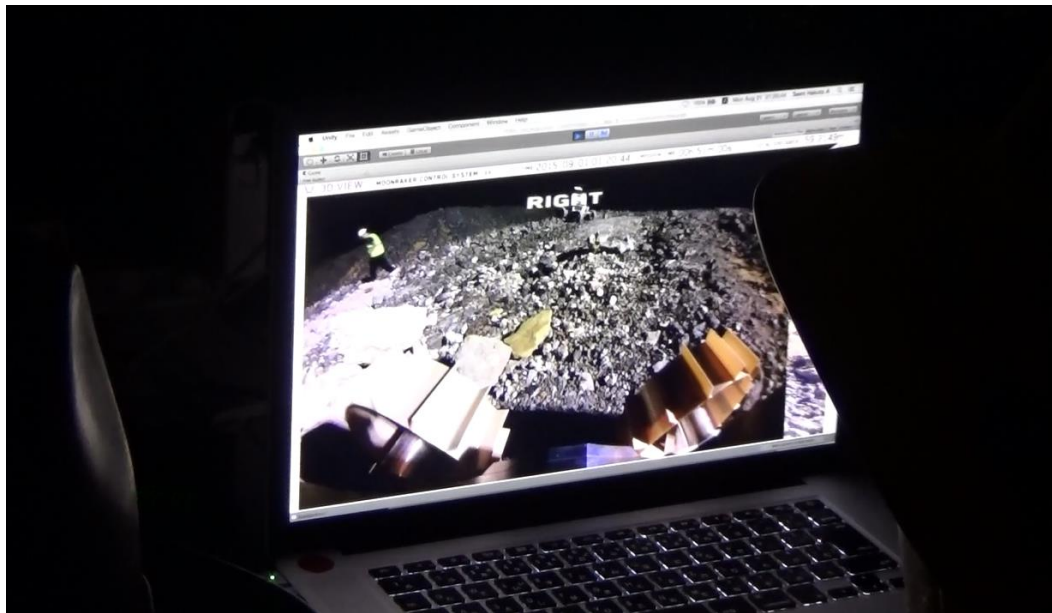
- Incoming telemetry can be paused
- Can scrub back and forth through received data
- Data can be saved/loaded, or exported to .csv format

Keyboard shortcuts:

A	S	P	D	F	
←←	←	PAUSE	→	→→	SAVE FOR LOADING
					SAVE TO CSV
					LOAD FROM FILE

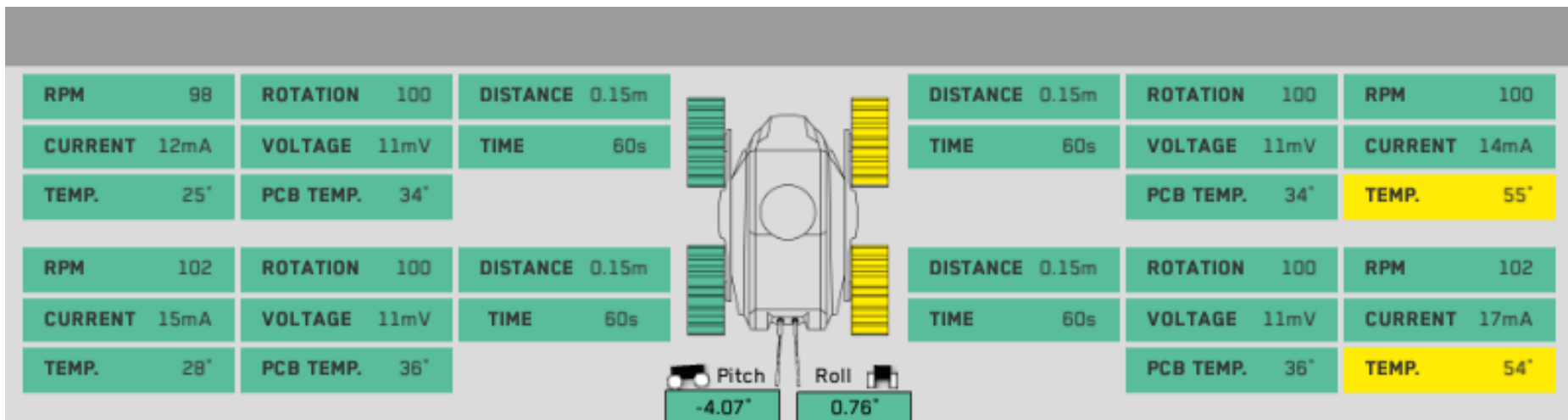
# Immersive 3D View

- User can rotate a camera inside a virtual 3D room
- Each of the four “walls” shows a camera image
- Gives intuitive understanding of environment



# Threshold-Basic Fault Detection

- Telemetry channels are grouped at the top of the screen
- Telemetry values exceeding warning/danger thresholds are clearly shown on the screen



# Detailed Fault Information

- Fault-specified data shown for selected channels
- Detailed descriptions can be provided
- Allows for knowledge transfer from experienced operators

## FAULT INFORMATION

**PDU\_vboost0:** Input voltage to charge the battery is extremely low. Solar cells may not be receiving any sunlight.

**PDU\_vboost1:** Input voltage to charge the battery is extremely low. Solar cells may not be receiving any sunlight.

**PDU\_vboost2:** Input voltage to charge the battery is extremely low. Solar cells may not be receiving any sunlight.

**PDU\_vbatt:** Battery voltage is low. The battery may not be charging.

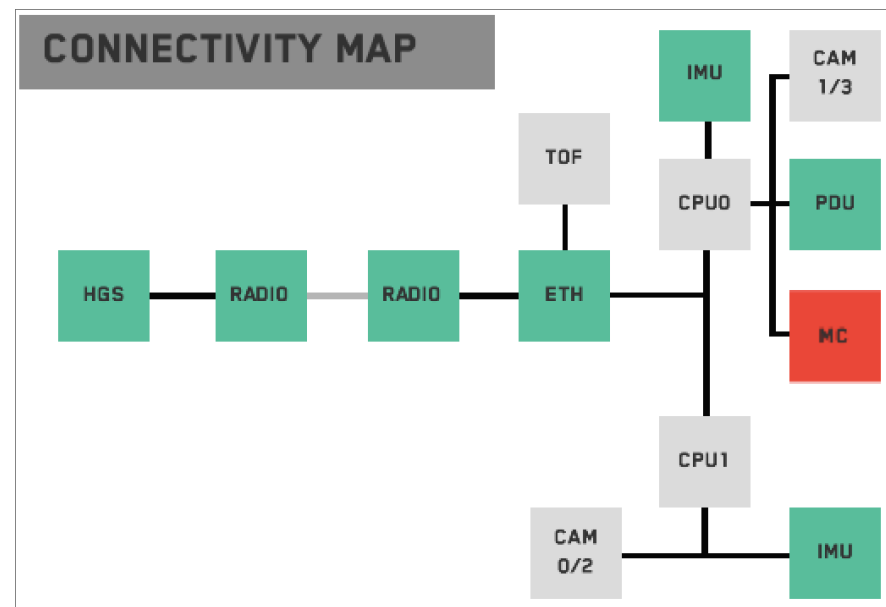
**MC\_voltage\_cpu0:** Nominal.

**MC\_voltage\_cpu1:** Nominal.



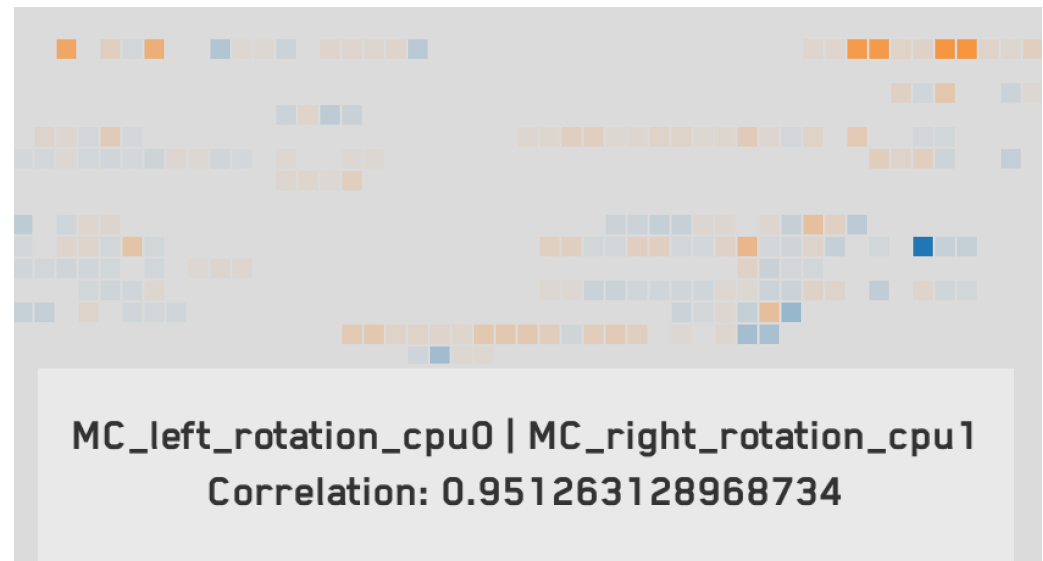
# Connectivity Map

- Visual display to show current status of modules
- Modules change color to reflect how recently they've been heard from

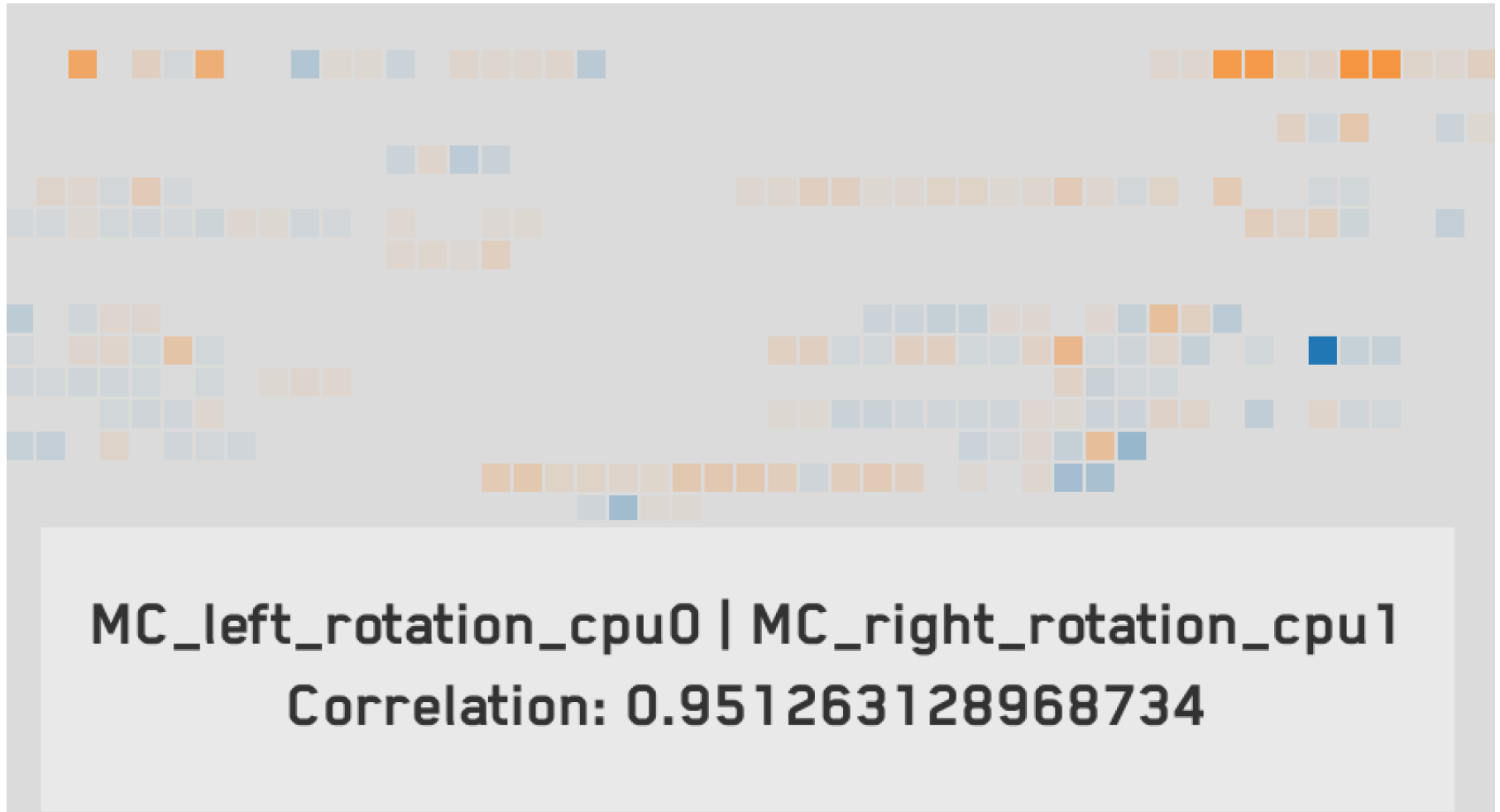


# Animated Corrgram

- 2D grid of blocks
- Each block shows the correlation between two channels
- PCC correlation score shown as hue:
  - High positive correlation: orange
  - High negative correlation: blue
- Grid updated with latest correlations



# Animated Corrgram Screenshot



---

**Algorithm 1** Animated Corrgram Generation Algorithm

---

```
1: procedure CORRGRAMGENERATOR( $M$ )                                ▷ Takes data matrix  $M \in \mathbb{R}^{n \times t}$ .
2:    $M_r \leftarrow \{U_{:,j}\}_{j=t-s}^n$                                 ▷ Reduce data to most recent  $s$  points.
3:    $C \leftarrow \mathbf{PCC}_s(M_r)$                                        ▷ Calculate a symmetric PCC matrix using last  $s$  points.
4:   for row = 1 to  $n$  do
5:     for col = 1 to  $n$  do
6:       color = GetColor( $C_{row,col}$ )                                ▷ Convert PCC score to a color.
7:        $\mathbf{Corr}_{row,col}.\text{color} \leftarrow \text{color}$                 ▷ Assign color to cell in corrgram.
8:     end for
9:   end for
10: end procedure                                                  ▷ Algorithm is re-run on every graphical update.
```

---

# First Integration



- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- **Intermediate Testing**
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- References

- Six Hakuto engineers went to Pittsburgh in August 2015
- Tested rover systems on moon-like terrain (rock quarry)
- Did tests in lunar lighting-like conditions
- Controlled Moonraker from my ground station for all tests





# Pittsburgh Field Testing





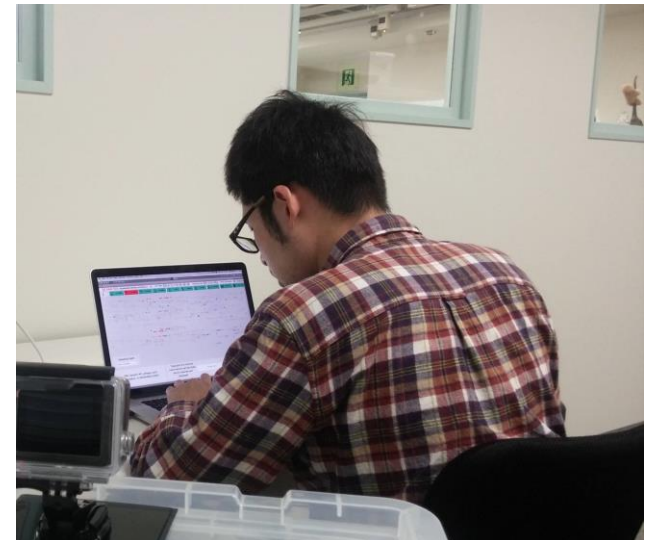
# Pittsburgh Field Testing Results

- Promising results
- All major goals successful
- Found many software bugs to fix
- “Immersive viewing” very useful
- “Connectivity map” excellent for troubleshooting
- User attention issues



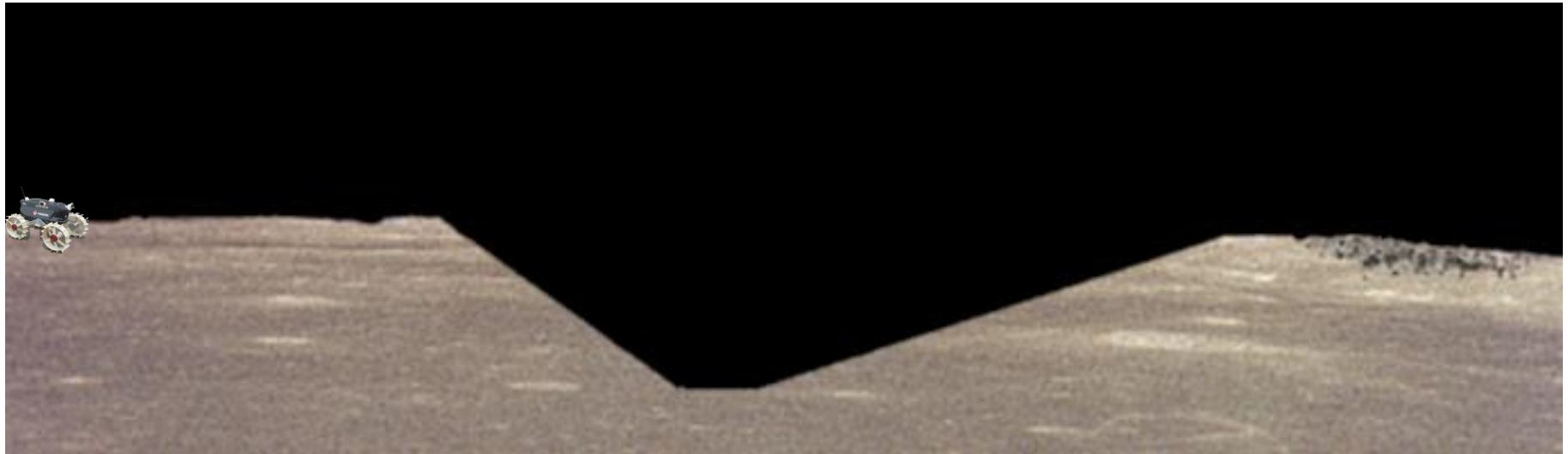
# Ground Station Usability Testing Overview

- November 2015 testing to find usability issues
- Emphasize on data monitoring, analysis
- Seven participants
  - All without extensive GSN experience
- 10-minute simulated Moonraker lunar run
- Each user watched this telemetry and analyzed it
- Users told story of what happened



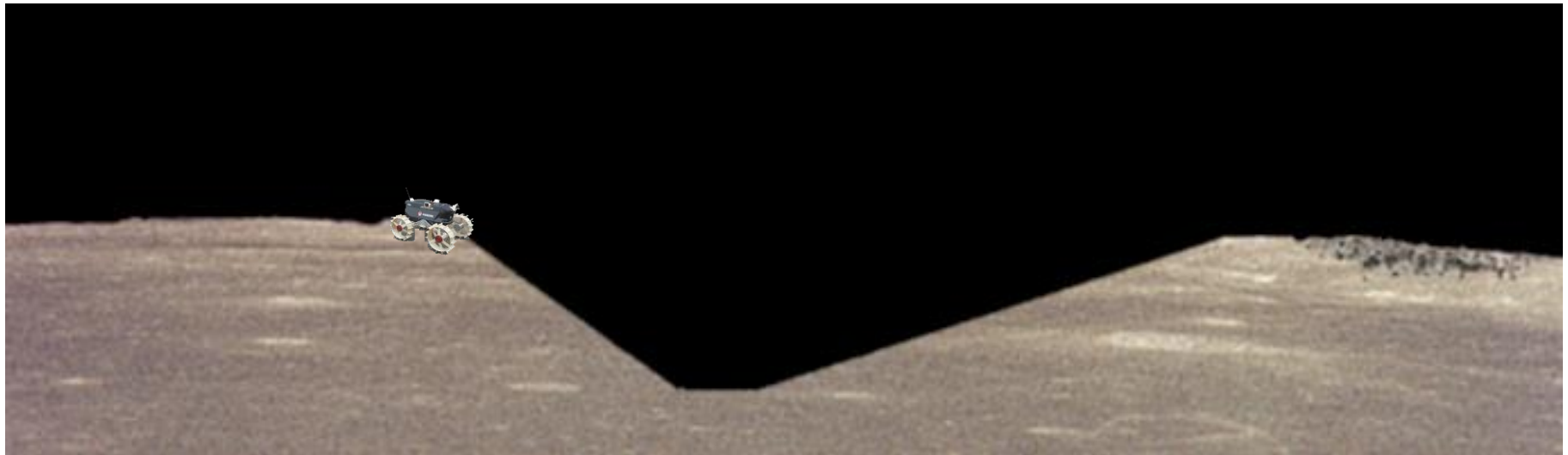
# Usability Testing Events

0s: Rover begins traveling forward along smooth terrain.

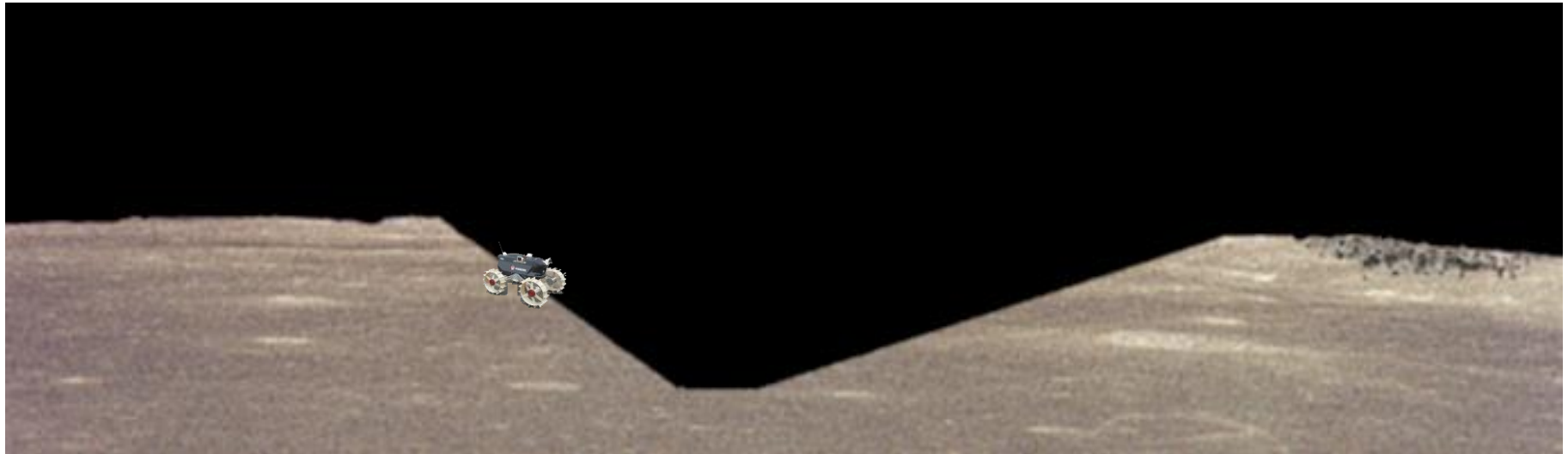


# Usability Testing Events

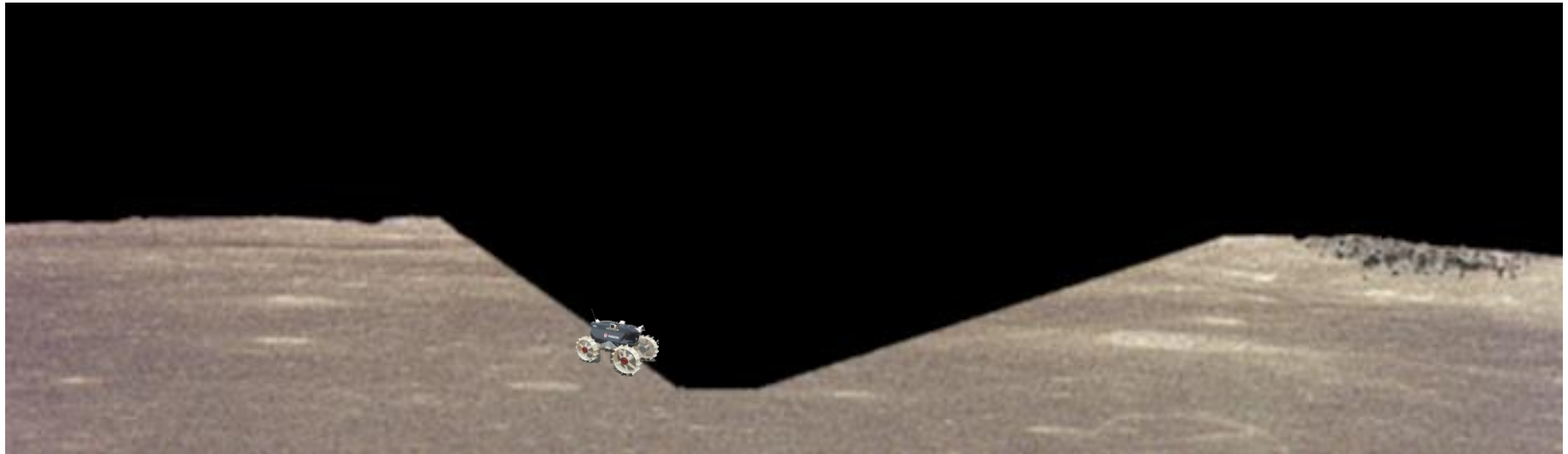
188s: Rover begins descending into crater.



223s: Rover loses line of sight with lander and packet drops begin.

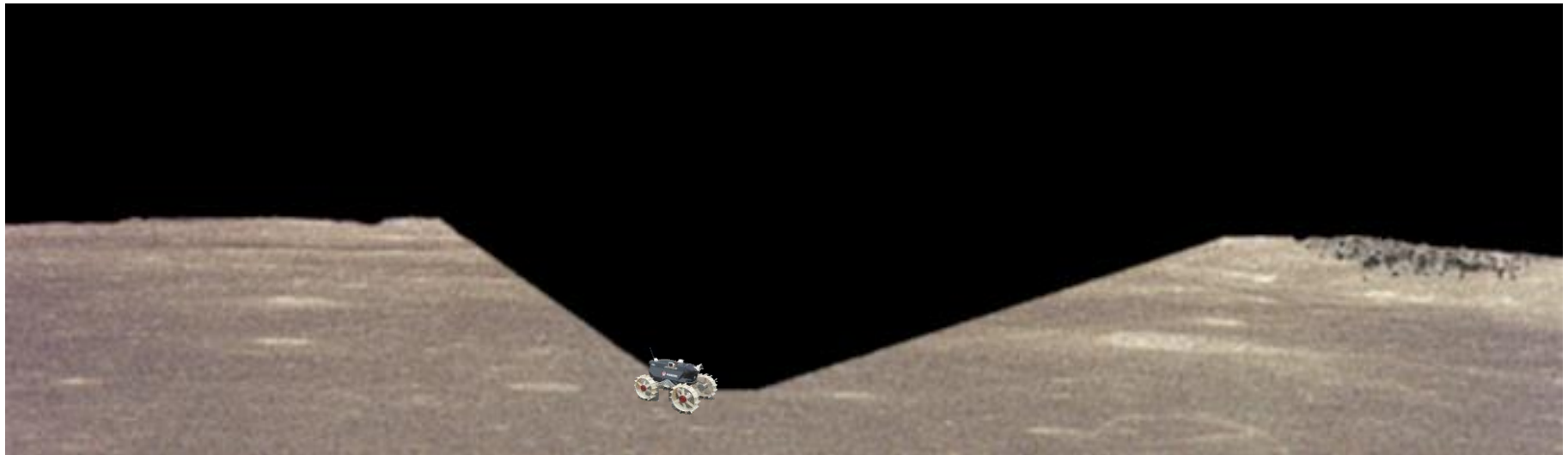


287s: Rover enters shade, causing temp, comms, and power drops.



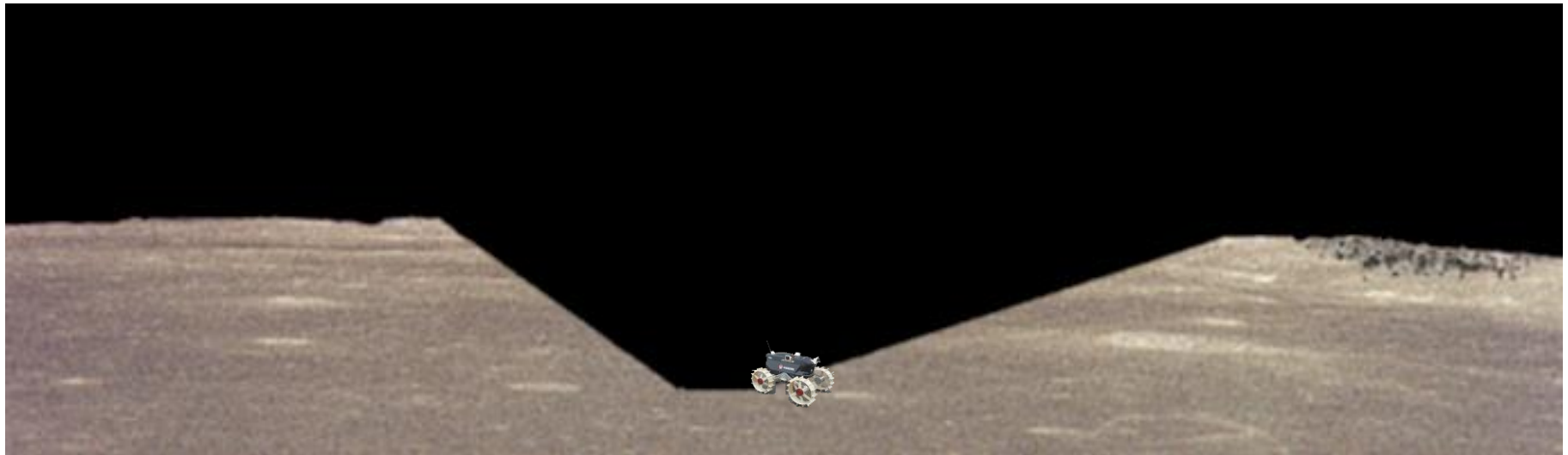
# Usability Testing Events

300s: Rover begins traversing smooth bottom of crater.



# Usability Testing Events

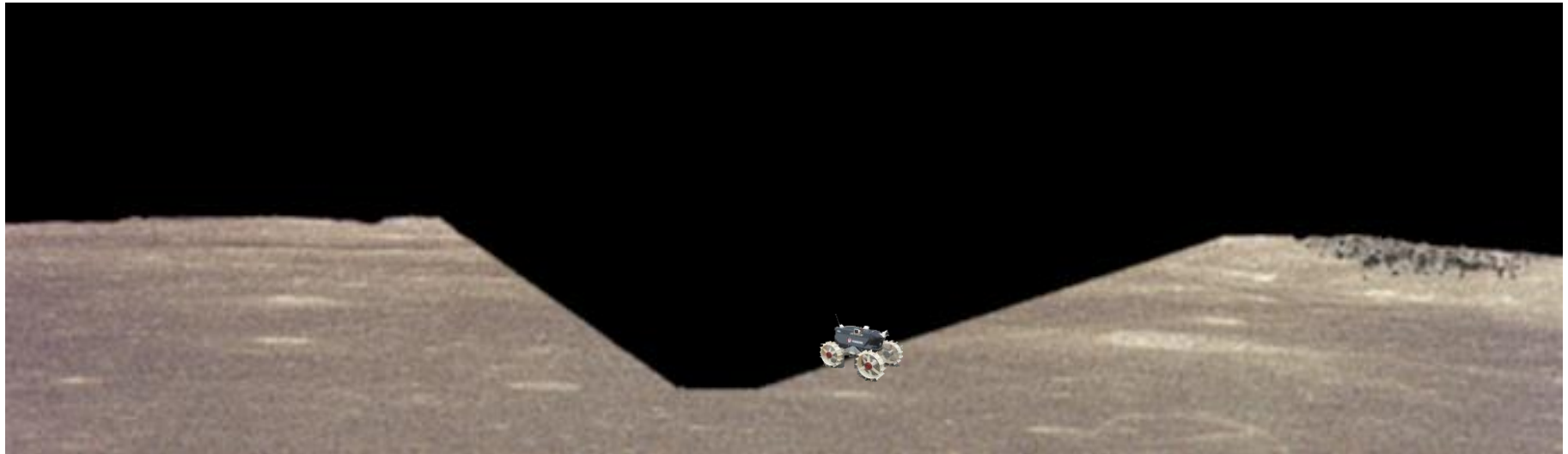
330s: Rover begins climbing out of crater.



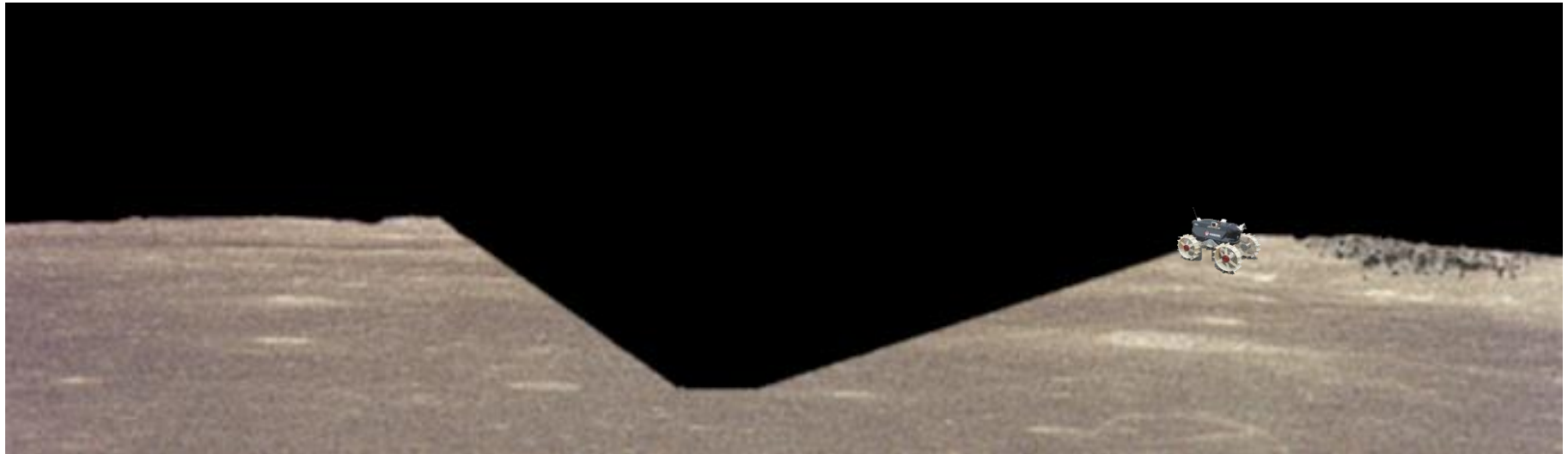


# Usability Testing Events

343s: Rover exits shade; rover continues uphill.

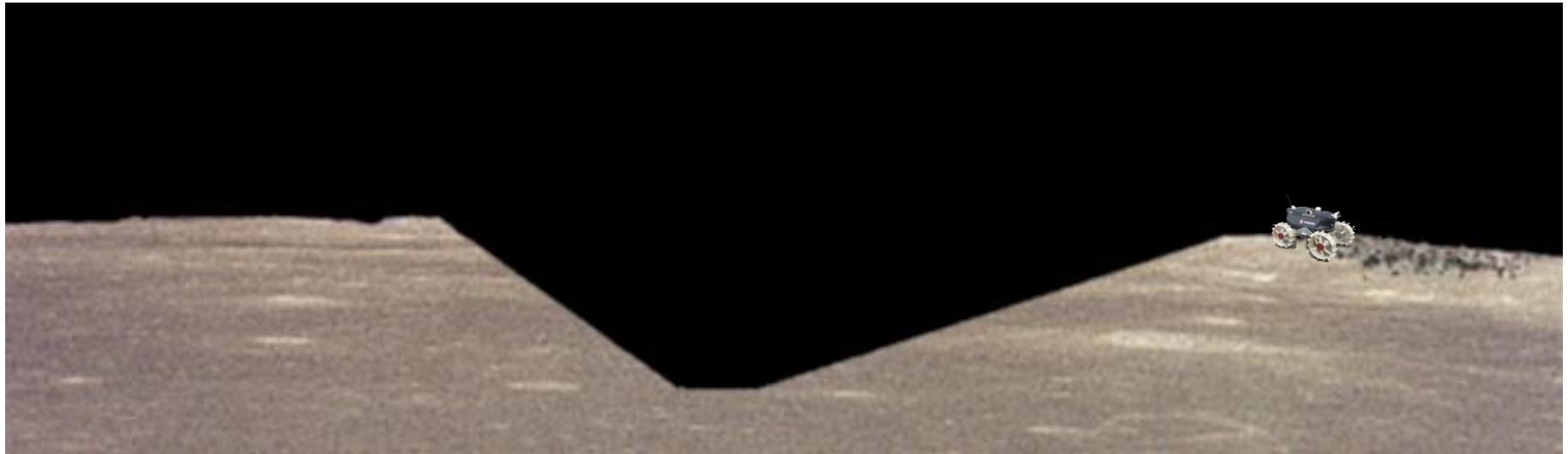


534s: Rover emerges from crater and enters smooth terrain.



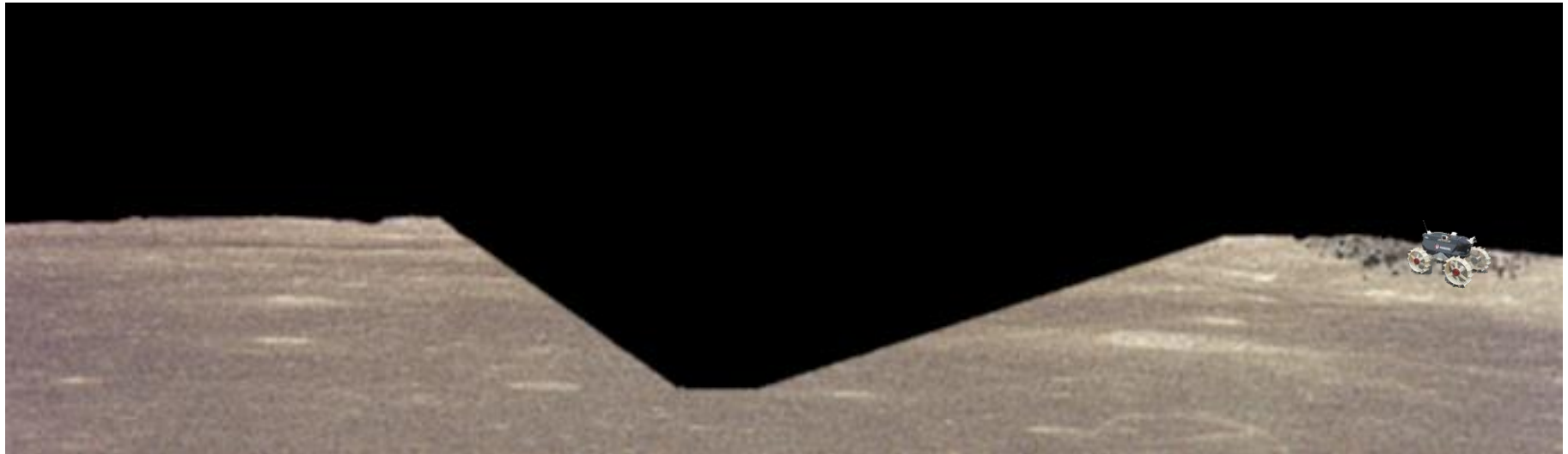
# Usability Testing Events

594s: Rover enters choppy terrain.



# Usability Testing Events

643s: Rover wheel has fault; rover stops moving.



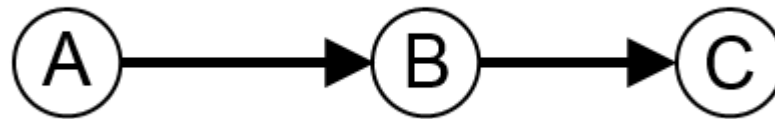
- User intuition matched planned simulation events
- Telemetry and fault detection-specific displays very useful
- Tunnel vision
- More fault descriptions requested
- Animated corrgram was hard to use, understand

- Ground station extremely useful for standard operations
- Needs improvement for root cause analysis to be possible
- Need to better capture expert knowledge about faults
- Corrgram has too much information to process quickly
- Corrgram only shows one point in time at a time
  - Need to show everything at once

- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- **2D Embedding Visualizations**
- Future Work and Conclusion
- Acknowledgements
- References

# Undirected Correlation Graphs

- Start with dependency graph
- 2D embedding where dependency represented with edge directions
- A depends on B; B depends on C:



- Variation: undirected graph where edges indicate correlation (each node is a data channel)



# Applying Undirected Correlation Graphs

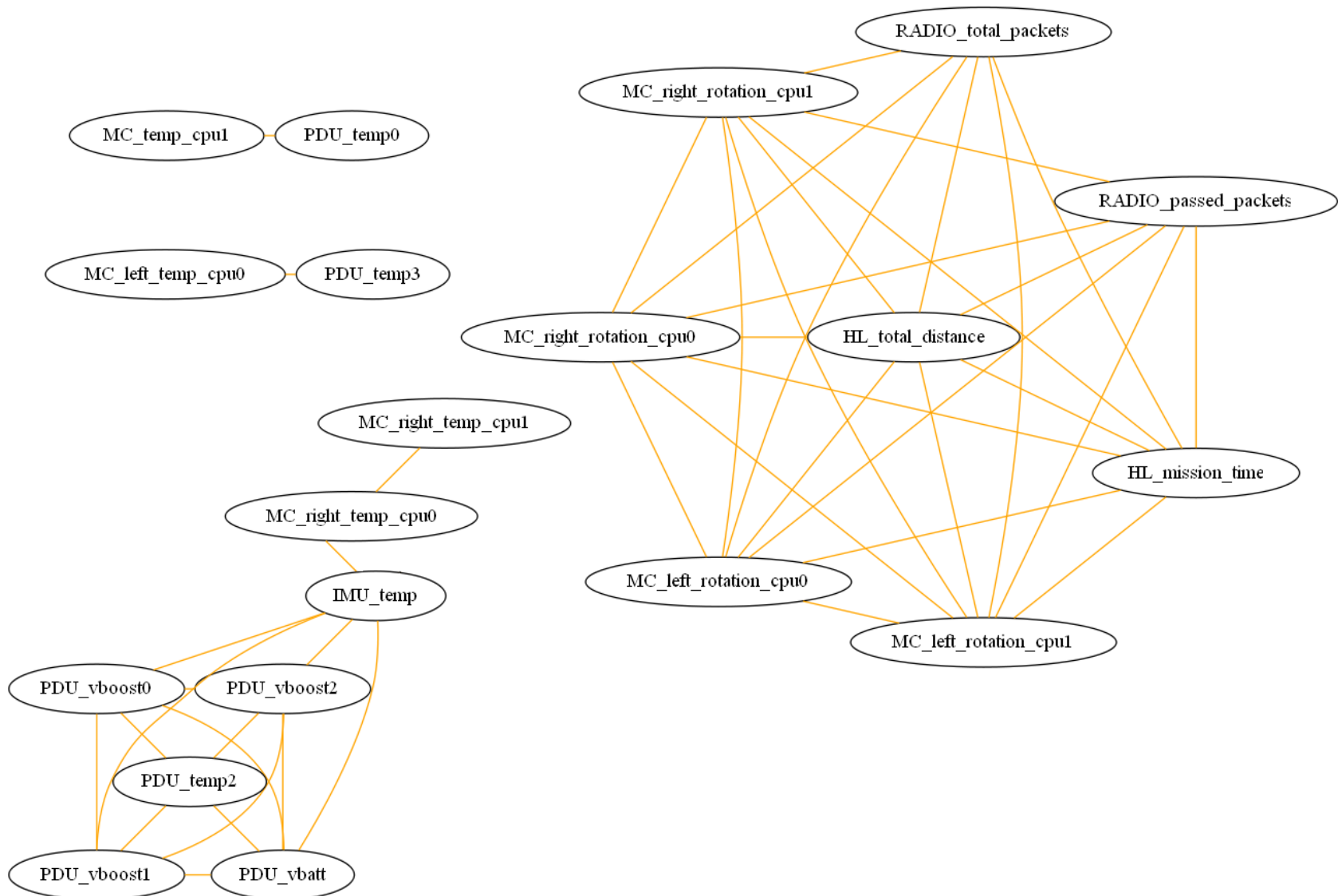


1. At a given point in time, generate a correlation matrix for all of the possible data channel pairs.
2. Generate an unconnected graph in which there exists a degree-0 node for each data channel.
3. For each node-node pair, add a connecting edge if they have a mutual correlation score above a reasonably high correlation threshold (e.g.,  $r_{PCC}^2 > 0.8$ ). This edge can be colored to show the sign of the correlation (e.g., blue for  $r_{PCC} < 0$  and red for  $r_{PCC} > 0$ ).
4. Finally, cull all nodes that are still of degree 0.

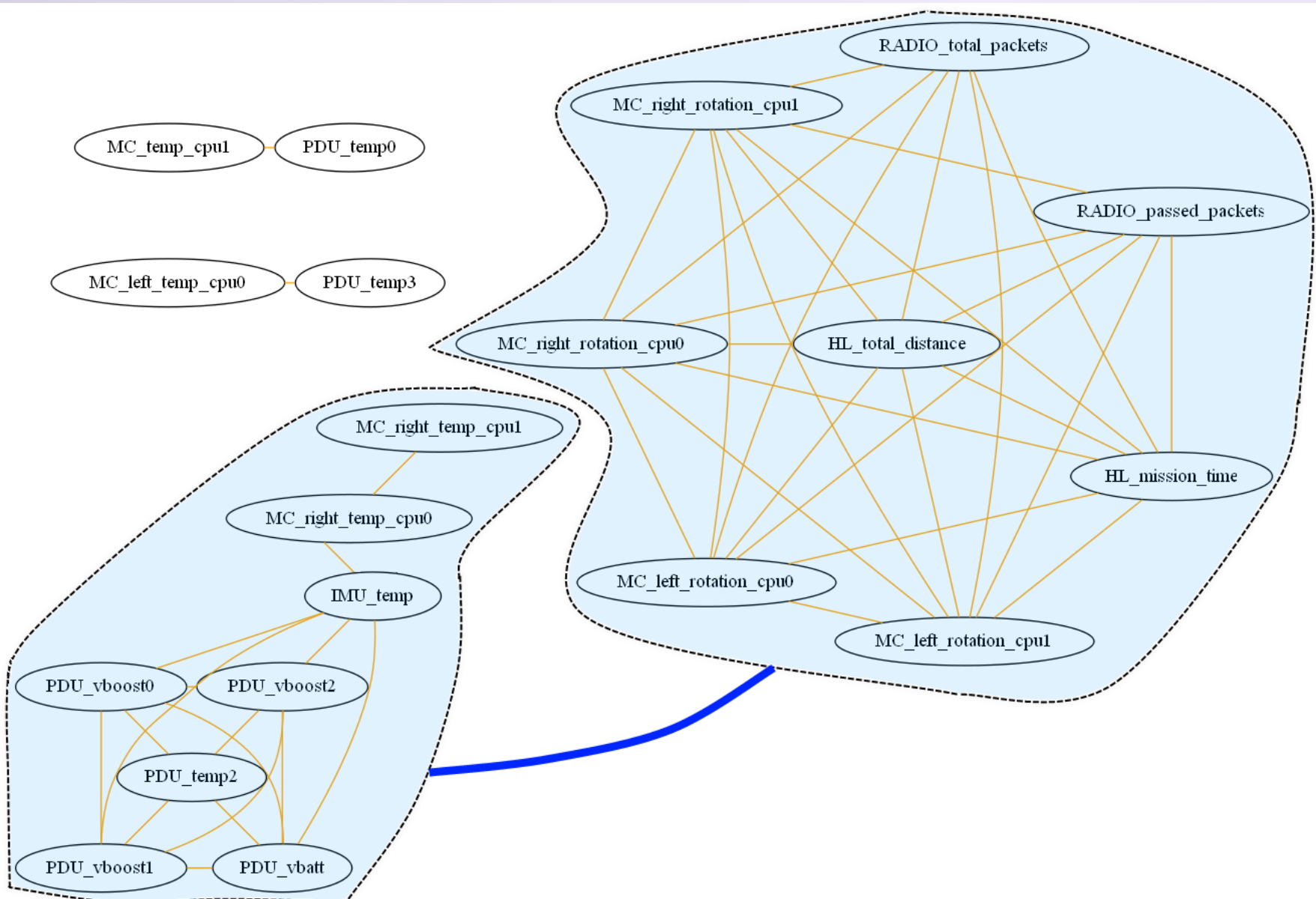
# UCGs: Full Correlation



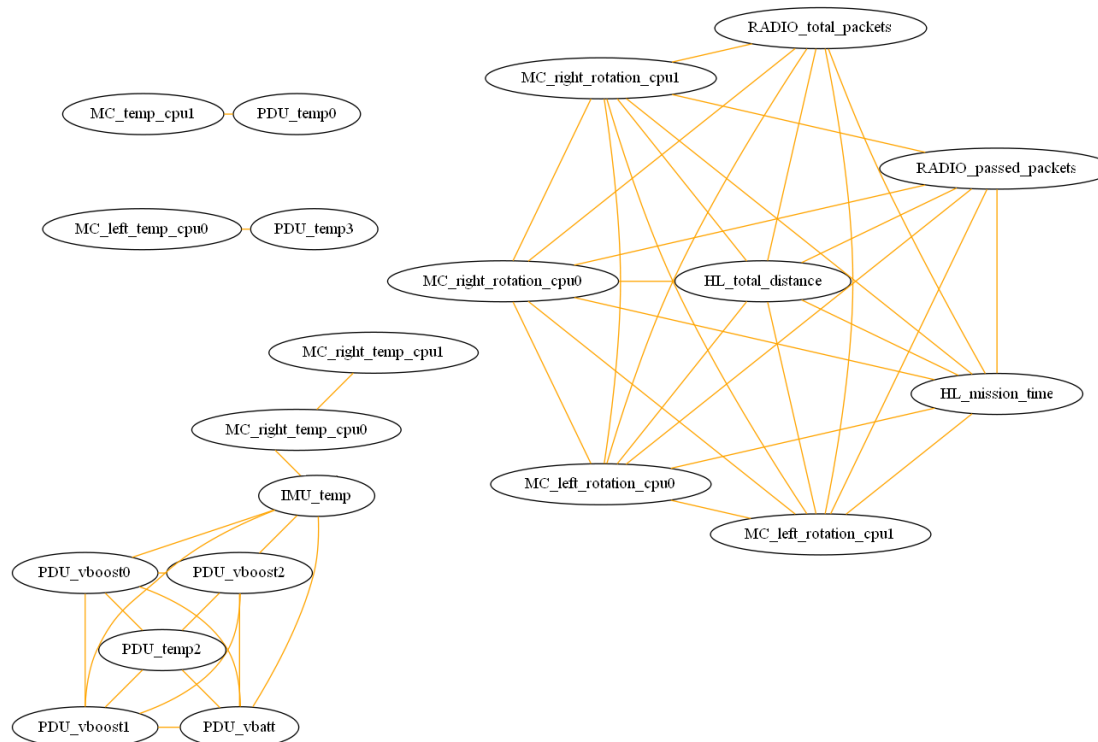
# UCGs: Positive Correlations Only



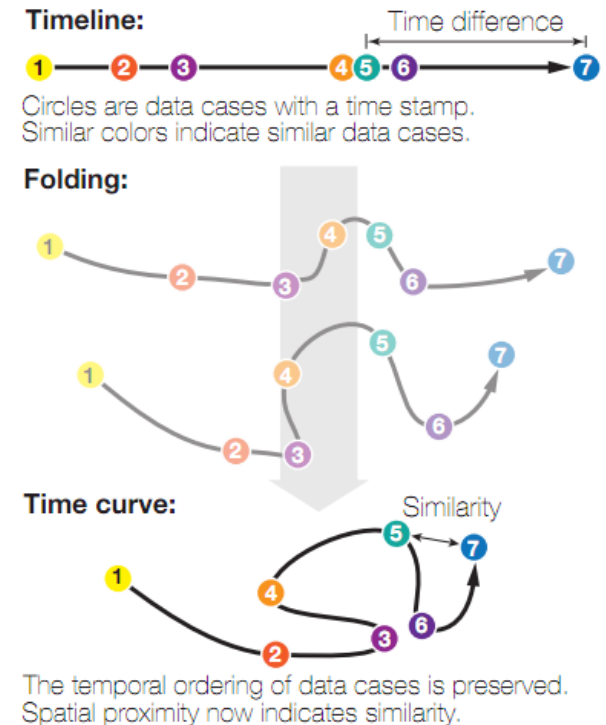
# UCGs: Positive Correlations with Cluster Correlations



- Correlations and clusters can be understood at a glance
- Positive correlations most intuitive
- Low-degree subgraphs worthy of second look
- Doesn't capture correlative state over time (needs to be animated)



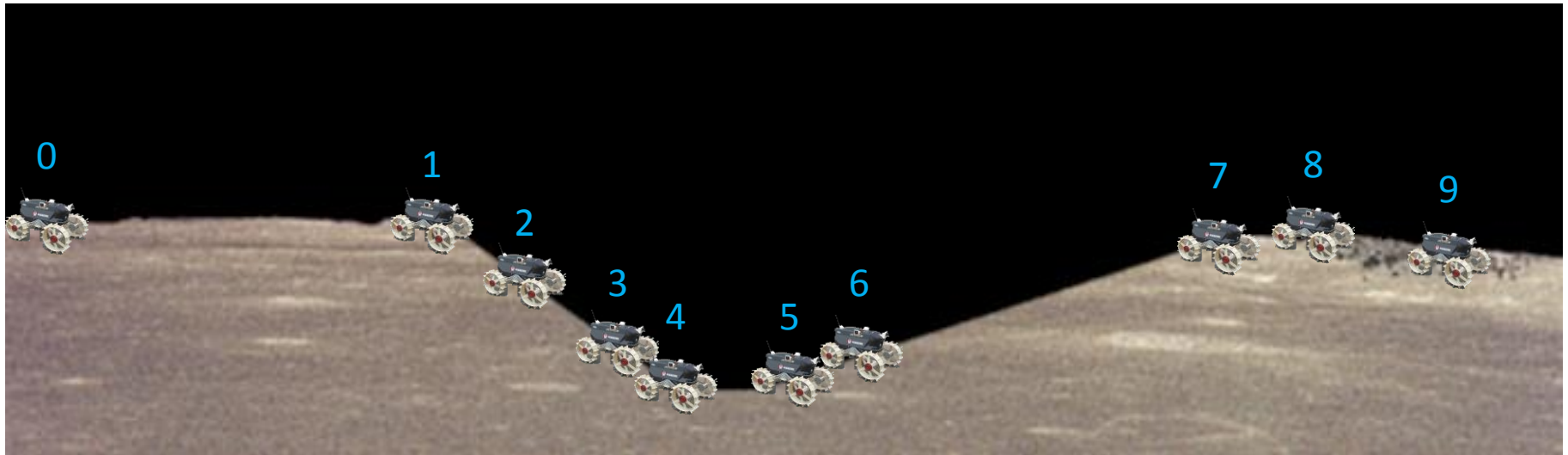
- Bach et al. algorithm from 2015
- Constructs 2D embedding showing changing system state over time
- Uses a set of “time points” and data about their distances to make a graph embedding that shows distance between states while maintaining temporal ordering



1. At regular intervals (e.g., once per second), record the telemetered system state into a vector  $\bar{x} \in \mathbb{R}^n$ .
2. Once the mission is complete, horizontally concatenate all state vectors to produce a state history matrix  $M \in \mathbb{R}^{n \times t}$ .
3. For a given window size  $w$ , generate  $t - w + 1$  windowed “sample matrices”  $W \in \mathbb{R}^{n \times w}$ , such that  $W_i = M_{:,i:i+w-1}$ .
4. For each sample matrix  $W$ , generate a correlative matrix  $C \in \mathbb{R}^{n \times n}$  (via PCC, Spearman’s Rho, or Kendall’s Tau), such that  $C_{ij} = C_{ji}$  gives the correlation score of the two data channels  $W_{i,:}$  and  $W_{j,:}$ . (Note that incalculable correlations, such as those with data channels of zero standard deviation, are set to be 0.)
5. Flatten each correlative matrix  $C$ , and concatenate the resulting column vectors horizontally into a matrix  $C_{all} \in \mathbb{R}^{n^2 \times (t-w+1)}$ , such that each column of  $C_{all}$  represents the total correlative state of the system at a time point.
6. Generate a symmetric distance matrix  $D \in \mathbb{R}^{(t-w+1) \times (t-w+1)}$ , such that  $D_{ij} = D_{ji}$  provides a pairwise distance score representing the difference between the states  $C_{:,i}$  and  $C_{:,j}$ .
7. Use the time curves algorithm to generate a 2D embedding of the  $(t - w + 1)$  states described in  $D$ .

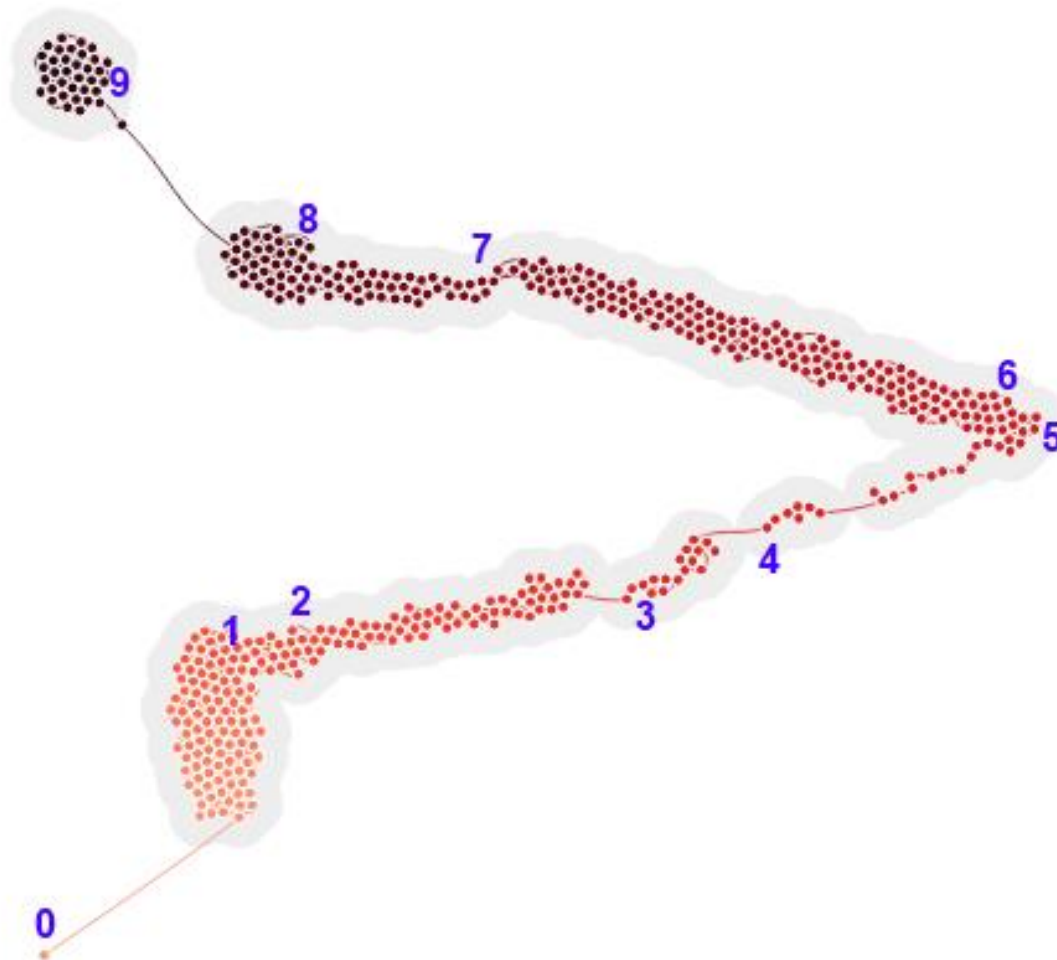
# Time Curve Data to Visualize

- Used simulated rover data (based closely on real telemetry)

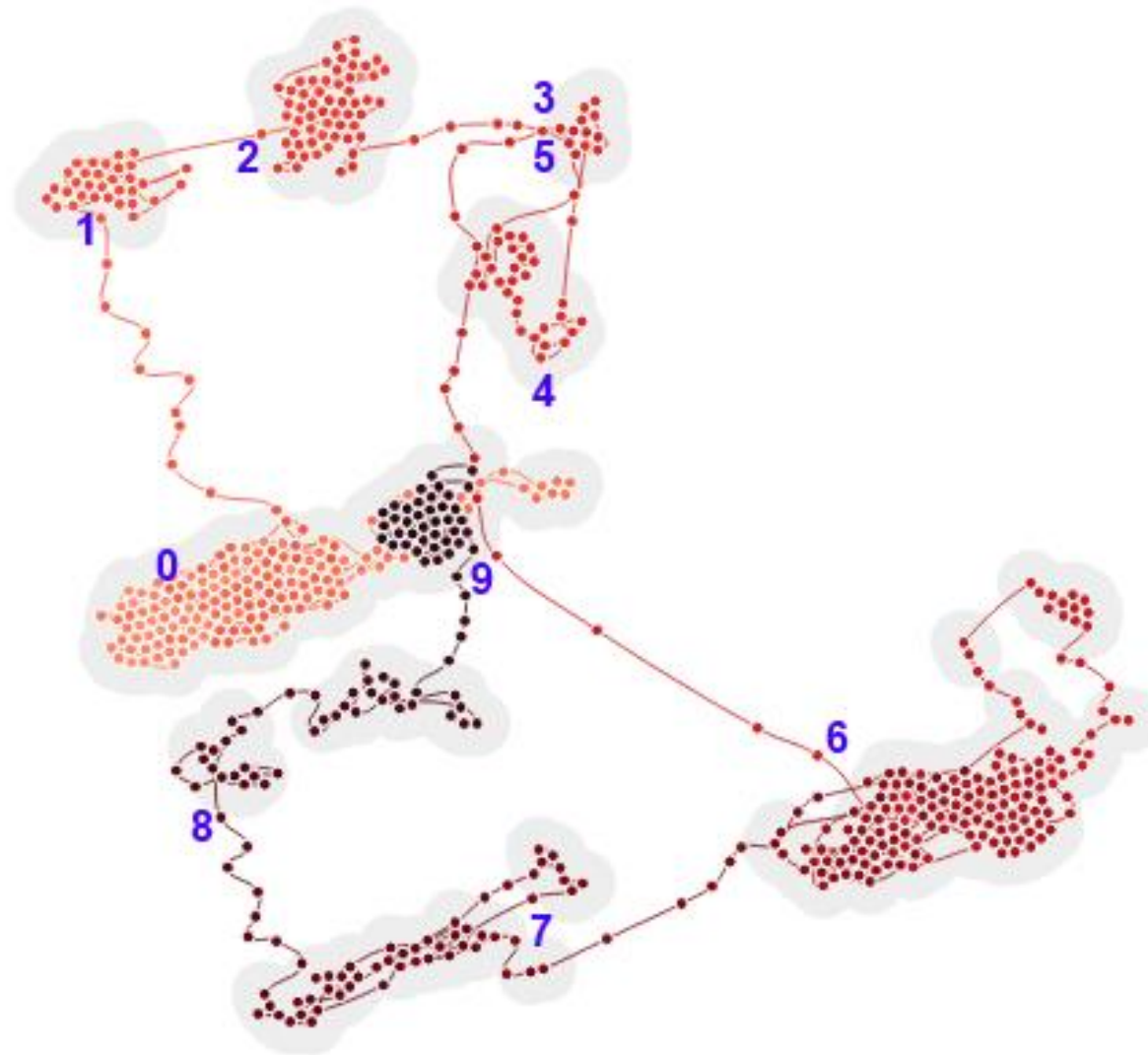




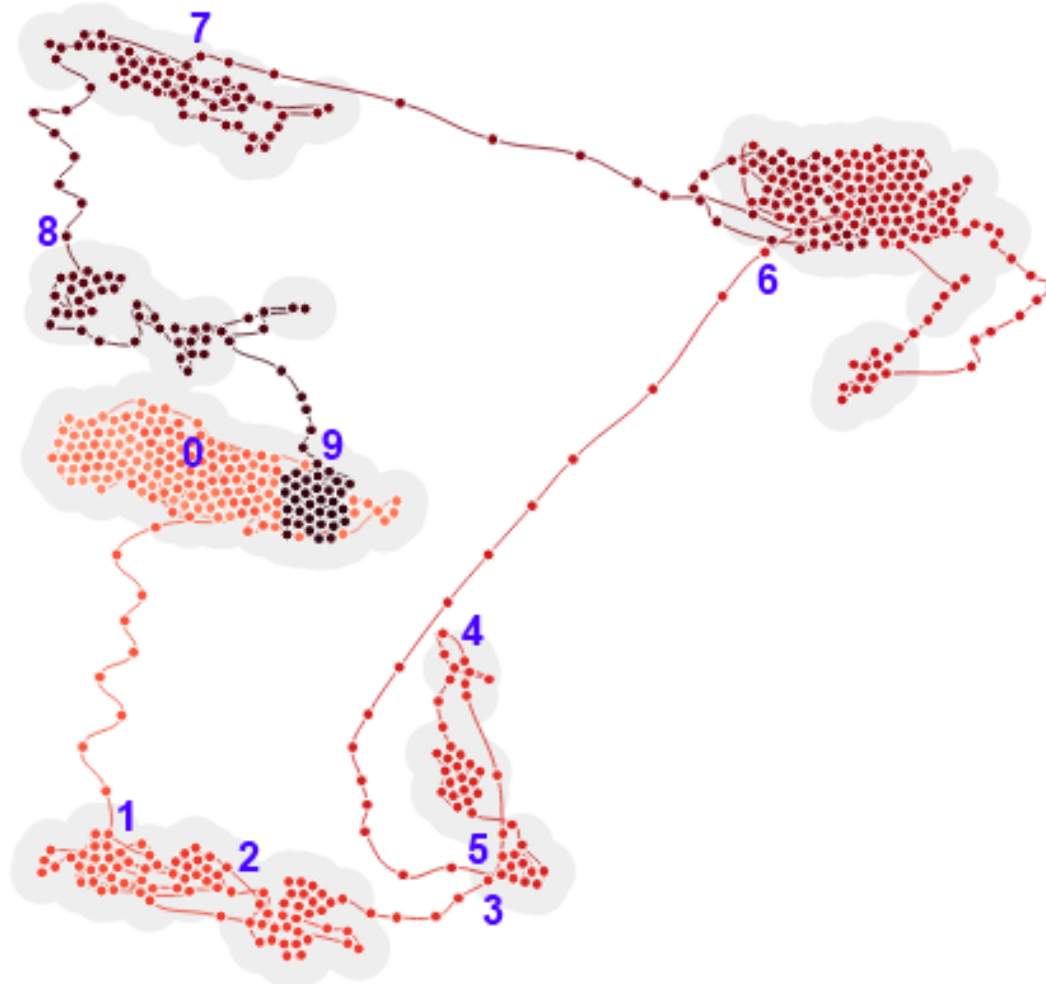
# Time Curves on Raw Telemetry State Data



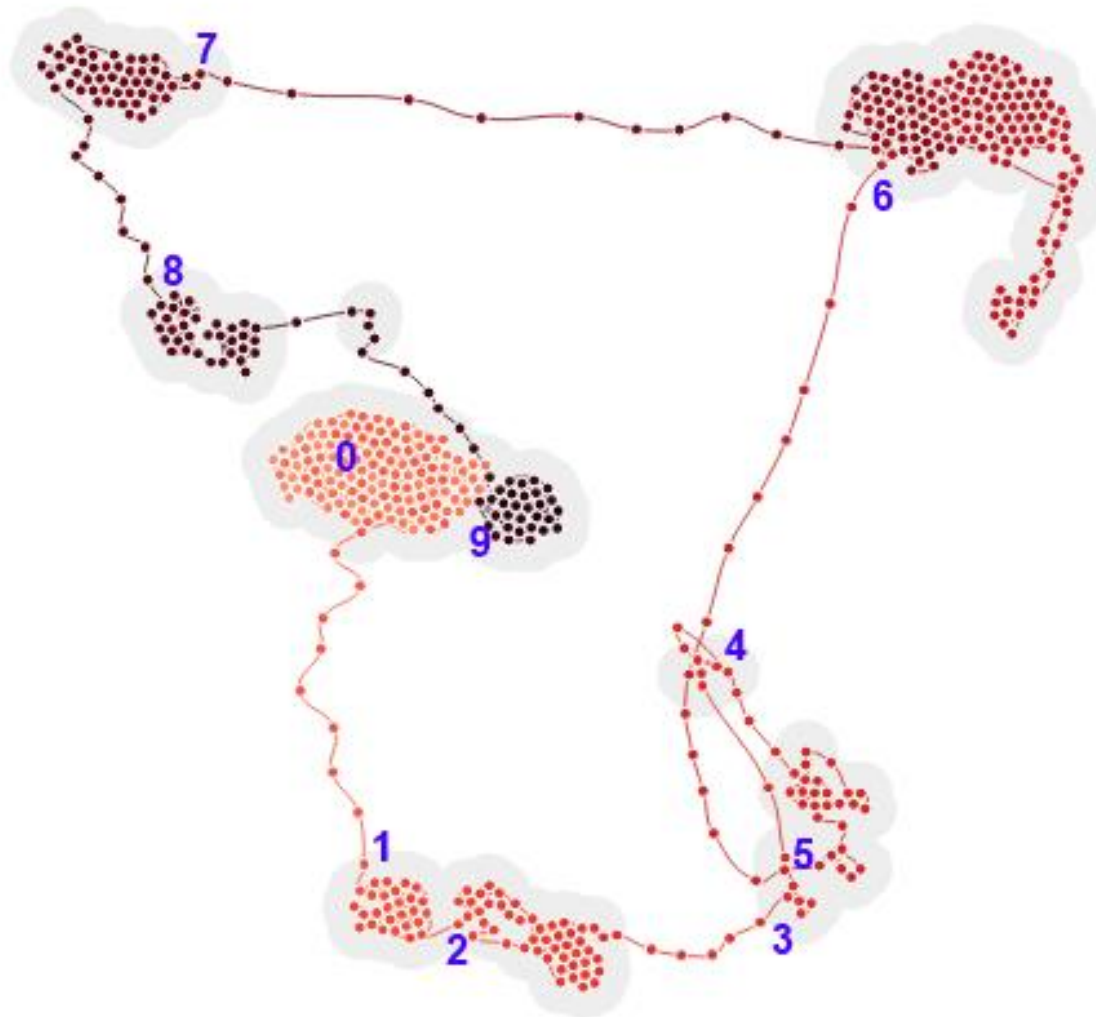
# Time Curves on PCC Correlation State



# Time Curves on Rho Correlation State



# Time Curves on Tau Correlation State



- Viewer can easily follow state progression and see operational modes
- Major moves correspond cleanly to actual timestamps
- Correlative analysis progression easier to understand
  - Internal relationships (rather than environmental ones) more easily captured by correlation
- Rank correlative measures better than PCC
- Able to “discover” unique states that aren’t expected
- Very promising and capable of suggesting correlative changes

- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- **Future Work and Conclusion**
- Acknowledgements
- References

- Integrating visualizations into a cohesive flow
  - Use data channels at portals between visualizations
- Make entire visualization flow automated, real-time
- Test on more data sets
  - Currently talking to University of Michigan, University of Tokyo
- Re-test usability with new analytical and visualization features

- Built ground station for real aerospace system
  - Incorporated high-level FDIR troubleshooting
  - V&V'ed in 100+ hours of field testing and usability testing
- Built analytical system to find correlative state
- Created three new types of correlative visualizations
- Showed that these visualizations give insights that hold up to scrutiny
- Tools show promise for future troubleshooting



- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- **Acknowledgements**
- References

# Acknowledgements



- Motivation and Objectives
- Model Definition
- Fault Detection Background
- Correlative Analysis and Visualization Background
- Lunar Rover Ground Station Development
- Intermediate Testing
- 2D Embedding Visualizations
- Future Work and Conclusion
- Acknowledgements
- **References**

1. Spaceflight 101. NASA Report on Antares Launch Failure places Blame on AJ26 Engines. <http://spaceflight101.com/nasa-report-on-antares-launch-failure-places-blame-on-aj26-engines/>, 2015.
2. Benjamin Bach, Conglei Shi, Nicolas Heulot, Tara Madhyastha, Tom Grabowski, and Pierre Dragicevic. Time curves: Folding time to visualize patterns of temporal evolution in data. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):559–568, 2016.
3. George Cancro, Russell Turner, Lilian Nguyen, Angela Li, Deane Sibol, John Gersh, Christine Piatko, Jaime Montemayor, and Priscilla McKerracher. An Interactive Visualization System for Analyzing Spacecraft Telemetry. In *Aerospace Conference, 2007 IEEE*, pages 1–9. IEEE, 2007.
4. Dauna Coulter. Down the Lunar Rabbit-hole. [http://science.nasa.gov/science-news/science-at-nasa/2010/12jul\\_rabbithole/](http://science.nasa.gov/science-news/science-at-nasa/2010/12jul_rabbithole/), 2010.

5. X PRIZE Foundation. Google Sponsors Lunar X PRIZE to Create a Space Race for a New Generation. <http://lunar.xprize.org/press-release/google-sponsors-lunar-xprize-create-space-race-new-generation>, 2007.
6. Michael Friendly. Corrgrams: Exploratory displays for correlation matrices. *The American Statistician*, 56(4):316–324, 2002.
7. Inseok Hwang, Sungwan Kim, Youdan Kim, and Chze Eng Seah. A survey of fault detection, isolation, and reconfiguration methods. *Control Systems Technology, IEEE Transactions on*, 18(3):636–653, 2010.
8. James Kurien and María DR Moreno. Intrinsic hurdles in applying automated diagnosis and recovery to spacecraft. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 40(5):945–958, 2010.

9. Randall Munroe. xkcd: Correlation. <https://xkcd.com/552/>, 2016.
10. DJ Murdoch and ED Chow. A graphical display of large correlation matrices. *The American Statistician*, 50(2):178–180, 1996.
11. RJ Rummel. Understanding Correlation. 1976.
12. Laird Statistics. Kendall's Tau-b using SPSS Statistics. <https://statistics.laerd.com/spss-tutorials/kendalls-tau-b-using-spss-statistics.php>, 2016.
13. Laird Statistics. Spearman's Rank-Order Correlation. <https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php>, 2016.

14. Laird Statistics. Pearson Product-Moment Correlation.  
<https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>, 2016.
15. NASA Independent Review Team. Orb3 Accident Investigation Report. Technical report, National Aeronautics and Space Administration, October 2015.
16. John Walker, Nathan Britton, Kazuya Yoshida, Toshiro Shimizu, Louis Burtz, and Alperen Pala. Update on the Qualification of the Hakuto Micro-Rover for the Google Lunar X-Prize. In *Proceedings of The 10th Conference on Field and Service Robotics (2015)*, 2015.

Thank you!

Questions?

