

Topic Modeling of NASA Space System Problem Reports

Practice Paper



Lucas Layman

Fraunhofer CESE



Allen P. Nikora

Jet Propulsion Laboratory
California Inst. of Technology



Joshua Meek

Fraunhofer CESE



Tim Menzies

N.C. State University

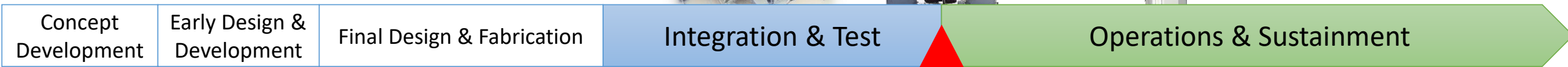
Context & Motivation

Systems are **expensive**

Spaceflight systems must be **reliable and fault tolerant**

NASA goal: Learn from problem trends from I&T and operations to **improve future design**

This is black



NASA Problem Reports

A rich source on how systems fail due to **hardware, software, human error, or environment**

- Date/time of event
- Criticality
- **Description**
- Investigation
- **Cause**
- Corrective Action

The good stuff is in “**natural**” language, the rest is unreliable

System Date: 08/16/2013 System Time: 11:55:51

SOAR # S-██████████0650 (E-12055)

Originated : 12/11/2011 23:03:47 (██████████)

Assigned To : ██████████

Last Updated : 04/18/2012 16:46:54 (██████████)

[View the STATUS LOG](#)

O B S E R V A T I O N	* Spacecraft : ██████████	Project								
	Anomaly Time : 2011-342-16:52:27	* Date								
	* Anomaly Category : FLIGHT SEGMENT	Sun/Da								
	<table border="1"><thead><tr><th>Polar Coordinates</th><th>ECI Coordinates</th></tr></thead><tbody><tr><td>Latitude :</td><td>X :</td></tr><tr><td>Longitude :</td><td>Y :</td></tr><tr><td>Altitude :</td><td>Z :</td></tr></tbody></table>	Polar Coordinates	ECI Coordinates	Latitude :	X :	Longitude :	Y :	Altitude :	Z :	Revolu
	Polar Coordinates	ECI Coordinates								
Latitude :	X :									
Longitude :	Y :									
Altitude :	Z :									
		Asc/De								

* Anomaly Title : RTS18 Inactive Buffer Validation Completed Too Late for SC_COPY_L

* Anomaly Description :
During execution of OAR-3504 ("Activate High-Beta Load File to Solar Array Gimbal (SC_SA_GBLCMD_BETA_ANGLE's call to STOL procedure SC_COPY_LOADBINARYTA RTS18 generated the following spacecraft event messages:
SC_EVENT: 11-342-16:52:27.544 CFE_TBL : ID:0081 TYPE:ERROR :Cannot activate SC_EVENT: 11-342-16:52:27.862 CFE_TBL : ID:0036 TYPE:INFO :SC_APP validation

SOAR State : CLOSED	SOAR Current Status :
** Subsystem/Instrument : Flight Software	
Assign To ██████████	
Is this a recurring anomaly? : No	
Assembly Name :	Part Name
Assembly ID :	Part Num
Assembly Serial Number :	Drawing F
Assembly Manufacturer :	

Investigation Log :
[12/09/2011: STS] Performed a "grep" of all event logs over the history of the miss

*** Cause of the Anomaly :
The SC_COPY_LOADBINARYTABLE STOL procedure loads, validates, and activates t similar STOL procedure SC_LOADBINARYTABLE, no telemetric verification is perform commands are separated by "WAIT 3" statements.

This duration is probably sufficient for validating many tables, but not always so for a other RTS table load file activations shows validations requiring only a second or two

In this case, the procedure commanded activation of the table load file before the

*** Corrective Action :

I N V E S T I G A T I O N		

R
E
S
O
L
U
T
I
O
N

Topic Modeling – Latent Dirichlet Allocation (LDA)

Collections of words, i.e., topics

Topic ID	Text
28	malindi fucino jsc line network mal comm router asi transfer houston asinet equipment est
12	circuit safeti trip bright star wider beta sirius cen grism mag april cathod great pcthold trend
3	usn socket carrier cds router ptp establish nmc noop record los power combin equipment n
52	xrt dpu chart xend xrtpos utc window build ack geni centroid offset mcp submodenon cathc

Manually create meaningful topic labels

Topic: spacecraft communication



Distribution of topics across documents

Problem Report	Topic_2	Topic_3	Topic_4	Topic_7	Topic_8	Topic_9	Topic_10	Topic_18
1						0.763037		0.029689
2				0.100154	0.597908			
3						0.384182		
4		0.910828						

*** Anomaly Title :** RTS18 Inactive Buffer Validation Completed Too Late for SC_COPY_

*** Anomaly Description :**
 During execution of OAR-3504 ("Activate High-Beta Load File to Solar Array Gimba
 SC_SA_GBLCMD_BETA_ANGLE's call to STOL procedure SC_COPY_LOADBINARYT
 RTS18 generated the following spacecraft event messages:
 SC_EVENT: 11-342-16:52:27.544 CFE_TBL : ID:0081 TYPE:ERROR :Cannot activa
 SC_EVENT: 11-342-16:52:27.862 CFE_TBL : ID:0036 TYPE:INFO :SC_APP validat

SOAR State : CLOSED **SOAR Current Status :**

**** Subsystem/Instrument :** Flight Software
Assign To ██████████
Is this a recurring anomaly? : No

Assembly Name : **Part Nar**
Assembly ID : **Part Num**
Assembly Serial Number : **Drawing**
Assembly Manufacturer :

Investigation Log :
 [12/09/2011: STS] Performed a "grep" of all event logs over the history of the m

**** Cause of the Anomaly :**
 The SC_COPY_LOADBINARYTABLE STOL procedure loads, validates, and activate
 similar STOL procedure SC_LOADBINARYTABLE, no telemetric verification is perfor
 commands are separated by "WAIT 3" statements.

 This duration is probably sufficient for validating many tables, but not always so for
 other RTS table load file activations shows validations requiring only a second or tv

 In this case, the procedure commanded activation of the table load file before the

**** Corrective Action :**
 After the SC_COPY_LOADBINARYTABLE proc ends, send /SWTSACTIVAT: TABL

 In March 2012, ██████MOT changed the logic in the SC_COPY_LOADBINARYTABLE
 rather than the previous 3 seconds.

Follow-on Actions/Recommendations :
 Recommending that the SC_COPY_LOADBINARYTABLE procedure be modified to
 SWTS_ACTIVATETABLE procedures in terms of the command verification checks
 longer to complete, but with the added benefit of command verification rather th
 to the next command.

Corpora

16,669 problem reports from 6 ongoing spaceflight missions

Mission type	Reports	Date span	Lifecycle phase
Space telescope	312	83 months	Operations
Planetary orbiter	715	40 months	Operations
Space telescope	908	91 months	Operations
Robotic explorer	3683	124 months	Integration & Test
Planetary orbiter	3885	227 months	Integration & Test
Robotic explorer	7166	93 months	Integration & Test

Findings from Individual Missions



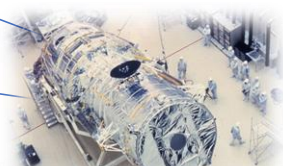
Does a problem report contain a topic? **Yes/no**

Tallied the Top 10 most frequent topics for **each of the 6 missions**

1) **One-third** of the Top 10s are **software-related**

2) #1 topic during I&T is **cabling and connectors**

3) **Flight software** is a common source of problem reports during I&T, but not during Operations



4) **Ground problems** are most common during operations

- station equipment
- network communications
- data-processing software



Concept Development

Early Design & Development

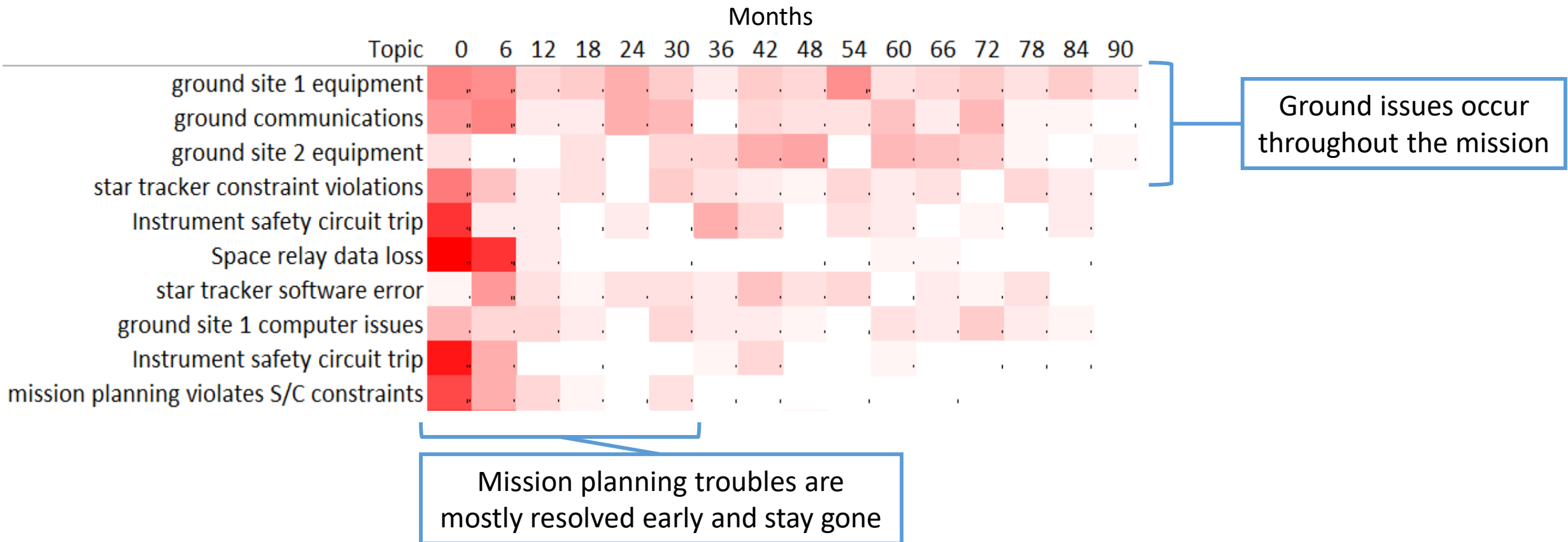
Final Design & Fabrication

Integration & Test

Operations & Sustainment

Changes in Topic Frequency over Time

Topic density for a Space Telescope over a 7.5 year Operations phase.



Mining Common Topics from All Projects

Topic	Reports
Cabling and connectors	1690
Avionics electrical performance	1507
Flight software – data mgmt	1149
Power generation and distribution	958
Actuators and motors	901
Flight software - telemetry	866
Mechanical or structural integrity	865
Flight software – power mgmt	861
Flight software – S/C fault detection	829
Flight software – power mgmt	828

Topic modeling of the combined corpus of 16,669 problem reports

- I&T and Operations
- Space telescopes, explorers, orbiters
- Developed at two different NASA centers

Flight software topics are **5 of the Top 10**

Mining Common Topics from All Projects

Topic	Reports	ST1	ST2	PS1	PS2	RE1	RE2
		312	908	715	3885	3683	7166
Cabling and connectors	1690						
Avionics electrical performance	1507						
Flight software – data mgmt	1149						
Power generation and distribution	958						
Actuators and motors	901						
Flight software - telemetry	866						
Mechanical or structural integrity	865						
Flight software – power mgmt	861						
Flight software – S/C fault detection	829						
Flight software – power mgmt	828						

But some of the missions had many more reports than others...

Mining Common Topics from All Projects

Topic	Reports	ST1	ST2	PS1	PS2	RE1	RE2
		312	908	715	3885	3683	7166
Cabling and connectors	1690				9.4	14.3	10.9
Avionics electrical performance	1507				11.1	8.3	10.6
Flight software – data mgmt	1149				25.2		
Power generation and distribution	958				13.2		
Actuators and motors	901					9.3	6.5
Flight software - telemetry	866					13.1	
Mechanical or structural integrity	865				8.8		
Flight software – power mgmt	861	5.4				14.4	
Flight software – S/C fault detection	829						10.6
Flight software – power mgmt	828						10.6

Topics extracted from the missions with the greatest #s of reports dominate the Top 10 list

Thus...

Cells show where >5% of problem reports for the mission had the topic

Topic Modeling Pitfalls and Challenges



1. **Please, know your data** – mixing seemingly similar data generated from different processes is **dangerous!**
2. Selecting the number of topics to model is an art and not a science – we need **tools and empirical guidance**
3. When creating meaningful semantic labels for topics, **documents must be read** to ensure label accuracy
4. Useful topics are both **cohesive and distinctive**, but there are no measures for semantic distinction?

Thanks and Credits

Supported by:

- NASA OSMA Software Assurance Research Program grants NNX11AP93G and NNX15AC81G
- NSF grant 1302169
- JPL's Office of the Management System and JPL's Assurance Technology Program Office

Image credits:

- Hubble Space Telescope – NASA, <http://goo.gl/55zZr4>
- Space shuttle – Matt Stroshane/Getty Images North America, <http://goo.gl/MbzafV>
- Hubble Space Telescope integration – NASA, <http://goo.gl/jQ256l>
- Ground Terminal – European Space Agency, <http://goo.gl/pGxONc>
- Waveform – Josh King/Commotion Wireless, <https://goo.gl/StWBqJ>