

A study of hierarchical glitch categorization

Eiichi Hirose
Peter R. Saulson
Syracuse University

LSC/Virgo data analysis meeting
Dec 11, 2007

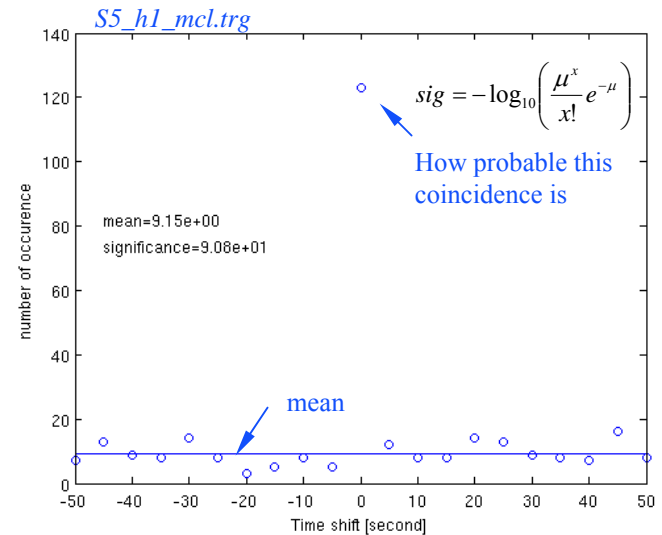
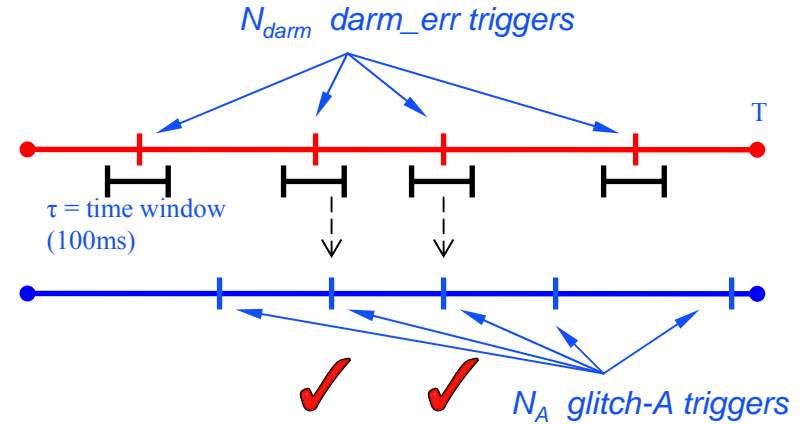
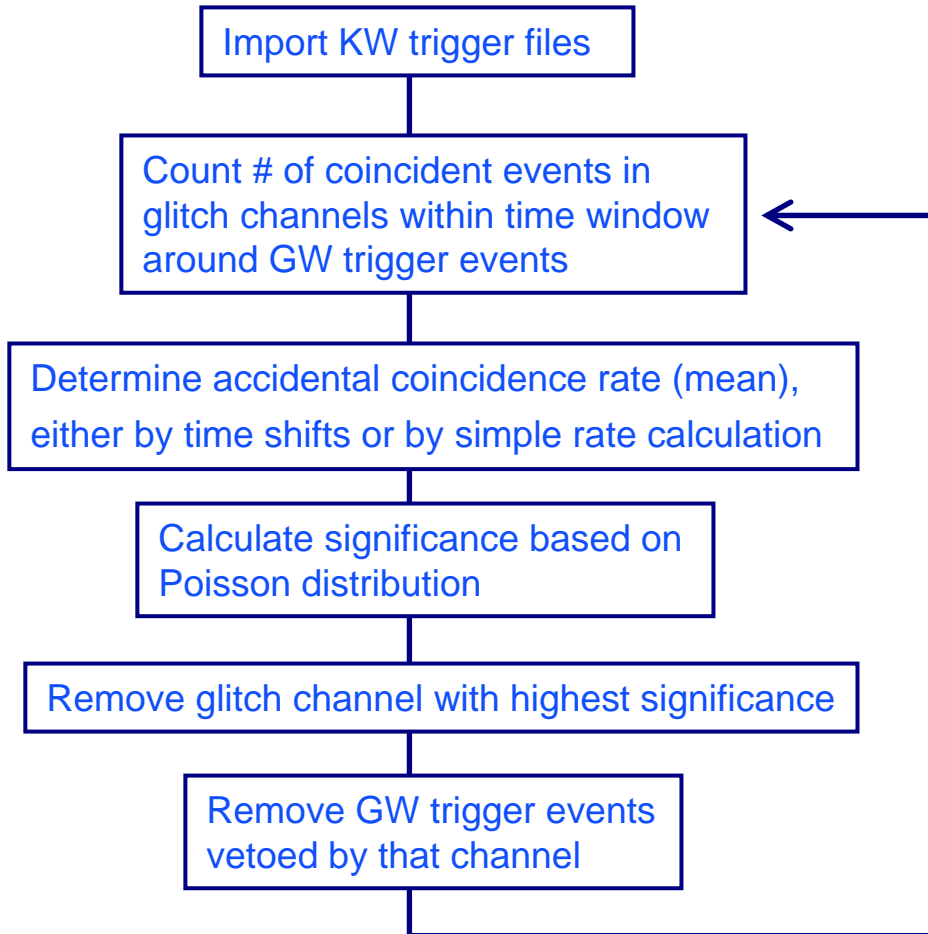
- First result of glitch categorization method proposed by Peter Shawhan



- » Find most significant veto channel, and remove all DARM_ERR events associated with it
- » Find most significant channel on cleaned version of DARM_ERR list, then repeat

- Code developed (Matlab) and applied to first five month KW trigger data on H1, using IFO channels (not PEM)
- Visualization discloses correlation between channels

- Many channels have significant correlation with DARM_ERR, but do they each veto events that are different, or mostly the same events?
- Is there a rational way to find a finite set of veto channels that is effective?
- Can veto studies help us understand the origin of glitches?



FLOW	JOBS	LINK
<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;">Raw KW trigger files of first five months (815155200 - 828288000)</div> <div style="text-align: center;">↓</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px; width: fit-content; margin: auto;">DQ-cut</div> <div style="text-align: center;">↓</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px; width: fit-content; margin: auto;">Matlab codes</div> <div style="text-align: center;">↓</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px; width: fit-content; margin: auto;">Output file</div>	<p>Apply data quality cut (category 1 and category 2)</p> <p>Count # of coincident events, calculate significance, remove glitch file with the highest significance, take out darm_err events vetoed by the channel, then repeat</p> <p>Store results in text format</p>	<p>http://apex.ligo-wa.caltech.edu/~ehirose/glitch/dq_cut3.m</p> <p>http://apex.ligo-wa.caltech.edu/~ehirose/glitch/glitch_project.m</p> <p>http://apex.ligo-wa.caltech.edu/~ehirose/glitch/significance.m</p> <p>http://apex.ligo-wa.caltech.edu/~ehirose/glitch/make_new_set.m</p> <p>http://apex.ligo-wa.caltech.edu/~ehirose/glitch/glitch_out.txt</p>

Note: simple rate calculation used this time

(Timeshift sometimes makes zero coincidence -> zero mean -> infinite significance)

$$Expected \# \text{ of chance coincidences} = N_A \frac{N_{DARM} \tau}{T}$$

Round	Chan. w/ highest significance	significance	Channels whose significance hugely reduced	Efficiency (cum.)	Deadtime (cum.)
0	prcctrl	639	pobi, reflq, pobq, michctrl, ..	5.89%	0.00349%
1	wfs2iy	85.7	wfs2qy, wfs2ip, wfs2qp, itmyy, wfs1qy, wfs1qp, itmxy, ..	6.68%	0.00396%
2	wfs1qp	36.8	itmxy, itmyy, wfs1qy, ..	6.99%	0.00409%

- First result of hierarchical glitch categorization
- Trend of significance groups channels
- Potentially useful in control rooms applying to much shorter time duration (~24 hrs)
- Future works
 - » Try different time window and different KW threshold
 - » Use all trigger files (PEM etc)
 - » Apply to full S5 data
 - » Study sets of DARM_ERR events that fall in different veto families, to look for clues about their origin

Acknowledgment

Peter Shawhan

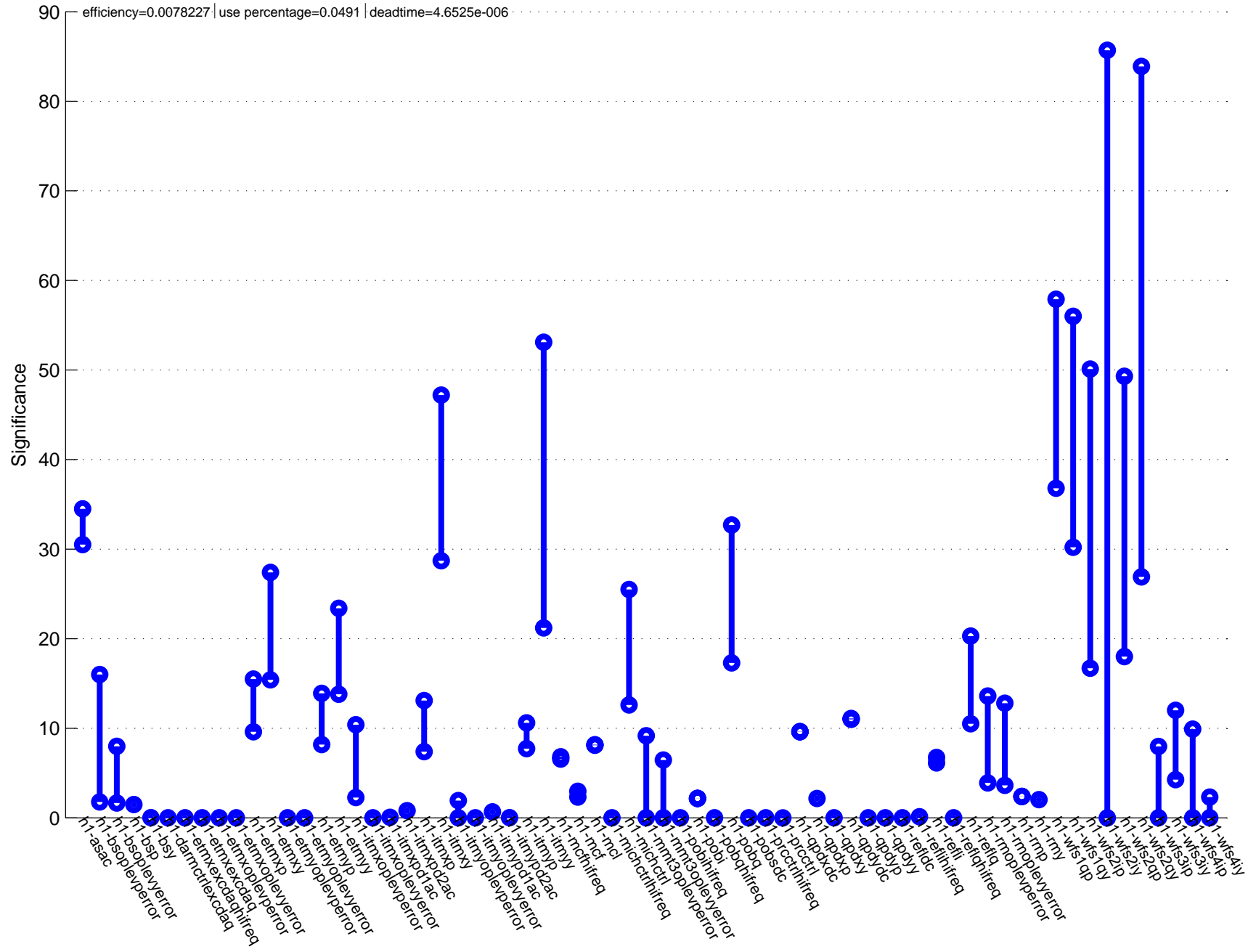


Appendix



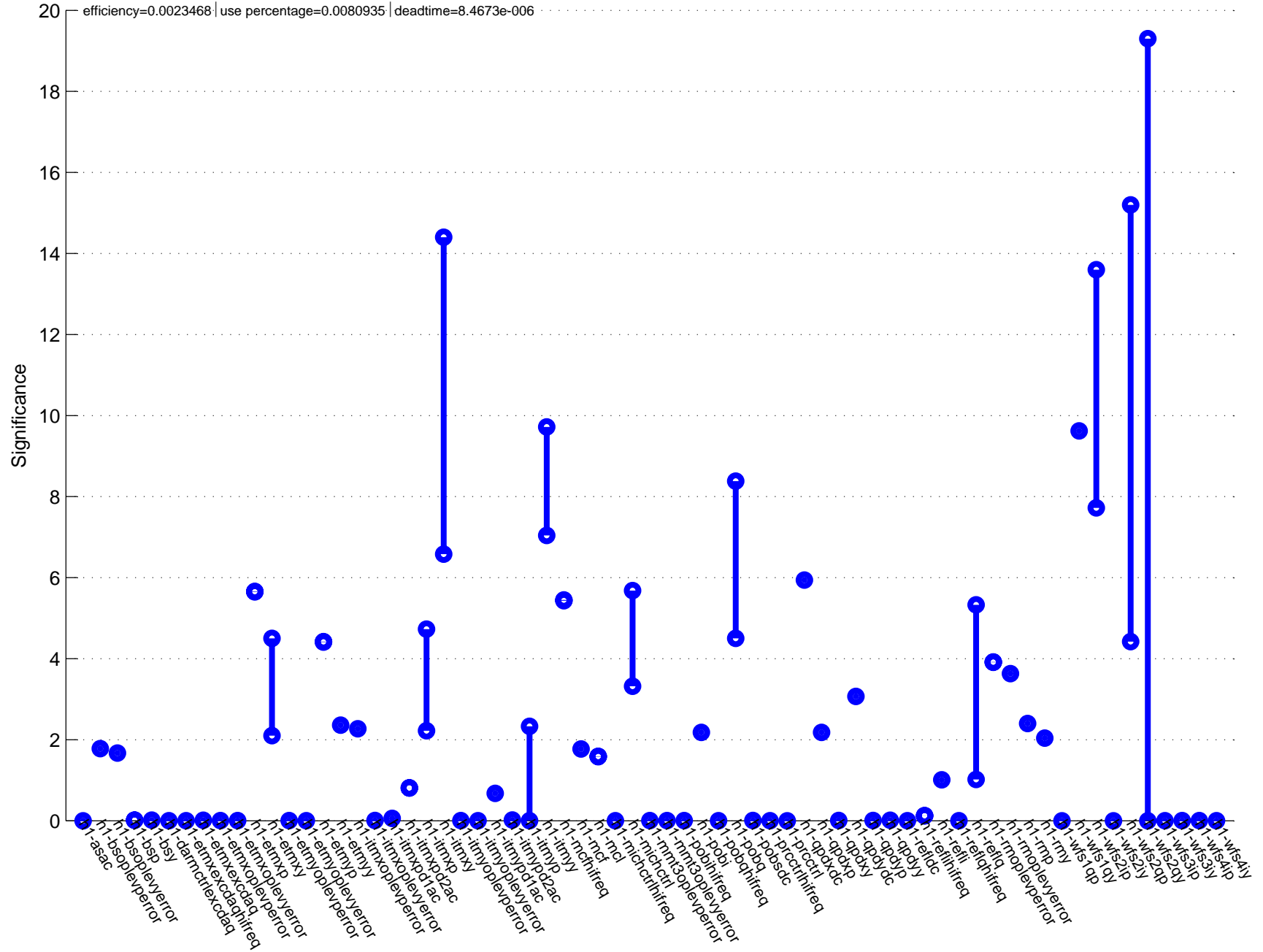
http://apex.ligo-wa.caltech.edu/~ehirose/glitch/plots_binded.pdf

Round 1 -> Round 2



Round 4 -> Round 5

efficiency=0.0023468 | use percentage=0.0080935 | deadtime=8.4673e-006



Round 6 -> Round 7

