

Practical Software and Systems Measurement

Objective Information for Decision Makers



***Causal Search in
Observational Data***

September 11, 2018

***Mike Konrad & Robert Stoddard,
SEI***

Causal Search in Observational Data - Intro

Objectives of the Workshop

The workshop will produce the following:

- ***Group statement to the PSM community on next steps to enlighten the full community on causal learning and encourage adoption***
- ***Accomplish the above by engaging in working discussions in small groups followed by a final large group summary of:***
 - a. ***Research questions and hypotheses to be investigated through causal learning (confirm/debunk conventional wisdom)***
 - b. ***Data sources helpful in causal learning research***
 - c. ***Next Steps and discussion of participants' datasets***

Workshop Format: Alternating Presentation & Brainstorming

What is Causal Learning?

- ***Activity 1: Identify a research question/topic of interest (Slide 22)***

What Are Causal Discovery Algorithms?

- ***Activity 2: Analyze a dataset (Slide 31)***

What Example Results has SEI obtained? (Case Study 1)

- ***Activity 3: Identify promising sources for datasets (Slide 39)***

What Example Results has SEI obtained? (Case Study 2)

- ***Activity 4: Define a causal learning adoption roadmap (Slide 46)***
- ***Activity 5: Establish a causal learning adoption user group (Slide 47)***

Conclusion

Workshop Background -1

PSM history in this area

- ***Causal learning is a novel research approach supported by many algorithms and tools that draws cause-effect structure from observation data – no prior PSM history***

Where we're heading

- ***We'd like to work with early adopters and collaborate on research that produces a causal model covering various topics in SW engineering***

Issues, questions, and topics

- ***Beginning with this workshop, we're eliciting research topics of significance and relevance to SW engineering***
- ***Build self-sufficiency among early adopters***

Workshop Background -2

The SEI is leading a three-year research project (SCOPE) that seeks to:

- ***Apply modern advances in causal learning (search and estimation)***
- ***Go beyond traditional correlation and regression analyses and accurately identify the causal relations among software process factors and product outcomes***

With this workshop, we intend to:

- ***Enlighten the practical measurement community***
- ***Encourage joint collaboration in the early adoption of causal learning to improve the quality of systems engineering and software engineering research.***

Intended Output

- ***List of candidate research topics and questions to be pursued with causal learning***
- ***Potential sources for datasets to give insight into these research topics and questions***
- ***Identification of potential early adopters***
- ***Roadmap of activities to enlighten SE and SW engineering researchers about causal learning***
- ***Identification of an early user group to meet in future PSM workshops and hold teleconferences***

Causal Search in Observational Data - Outbrief

Workshop Participants

Name	Email	Company
Robert Stoddard	rws@sei.cmu.edu	SEI
Emile Ettedgui	Emile.e.attedgui.civ@mail.mil	AFCAA
Allen Bouchard	Allen.bouchard1@us.af.mil	ALCMC/HNJJ
Lauren Lindros	Lauren.h.lindros@aero.org	Aero Corp.
Mike Konrad	mdk@sei.cmu.edu	SEI

Summary

Candidate Research Topics (example)

- ***What practices makes Agile successful vs. not so successful?***
 - ***How to define Agile and Success?***
 - ***Iterations, earlier user interaction, shorter feedback loops,...***

Sources for datasets (examples)

- ***CADE for ACAT1 programs (masked data?)***
- ***AFIT and NPS for students pursuing research***
- ***JIRA***
- ***Aerospace for Space***
- ***For the above, access to results from causal search—but maybe not the data itself***

Conclusions, Recommendations, and Results

- ***Continue eliciting candidate topics, research questions from practical software measurement community***
- ***Give priority to including software defect data, technical debt***
- ***General research approach on topic: literature search => characteristic practices, priorities => collect from data sources and by survey => causal search what practices etc. cause different definitions of project success***
- ***Roadmap for engaging practical SW measurement community (next slide)***

Next Steps/Action Items

- 1. Send out email to workshop attendees with follow-up info (Konrad, 21 Sep)***
- 2. Establish a wiki site to facilitate future PSM Causal Learning sharing and communications (Stoddard, 30 Sep)***
- 3. Send out email invitation to PSM community to join the wiki site communications (Stoddard, 15 Oct)***
- 4. Begin posting semi-monthly updates on progress using causal learning within software measurement projects (Konrad, Stoddard, 1 Nov)***
- 5. Identify several PSM causal learning projects with targeted results to be shared at PSM 2019 (Konrad, Stoddard, 1 Dec)***
- 6. Post a suggested training path on Wiki for PSM early adopters of causal learning (Konrad, Stoddard, 15 Dec)***
- 7. Schedule 1-2 free, public F2F Causal Learning tutorials/classes (Konrad, Stoddard, 15 Jan 2019)***