Generic and Robust Localization of Multi-Dimensional Root Causes

Zeyan Li, Chengyang Luo, Yiwei Zhao, Yongqian Sun,
Kaixin Sui, Xiping Wang, Dapeng Liu, Xing Jin, Qi Wang , Dan Pei

ISSRE 2019

Outline

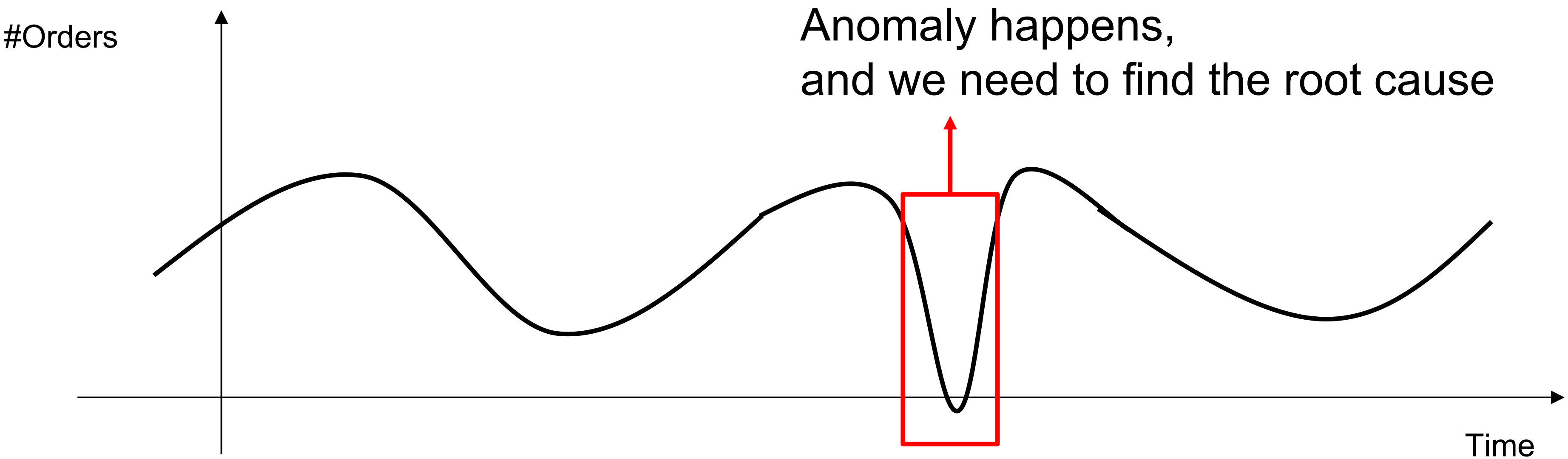
Background → Methodology → Experiment → Summary

Outline

Background → Methodology → Experiment → Summary

Background

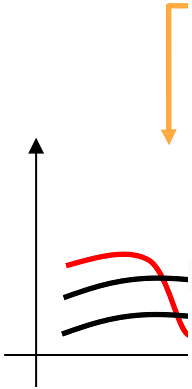
- KPI: key performance indicator

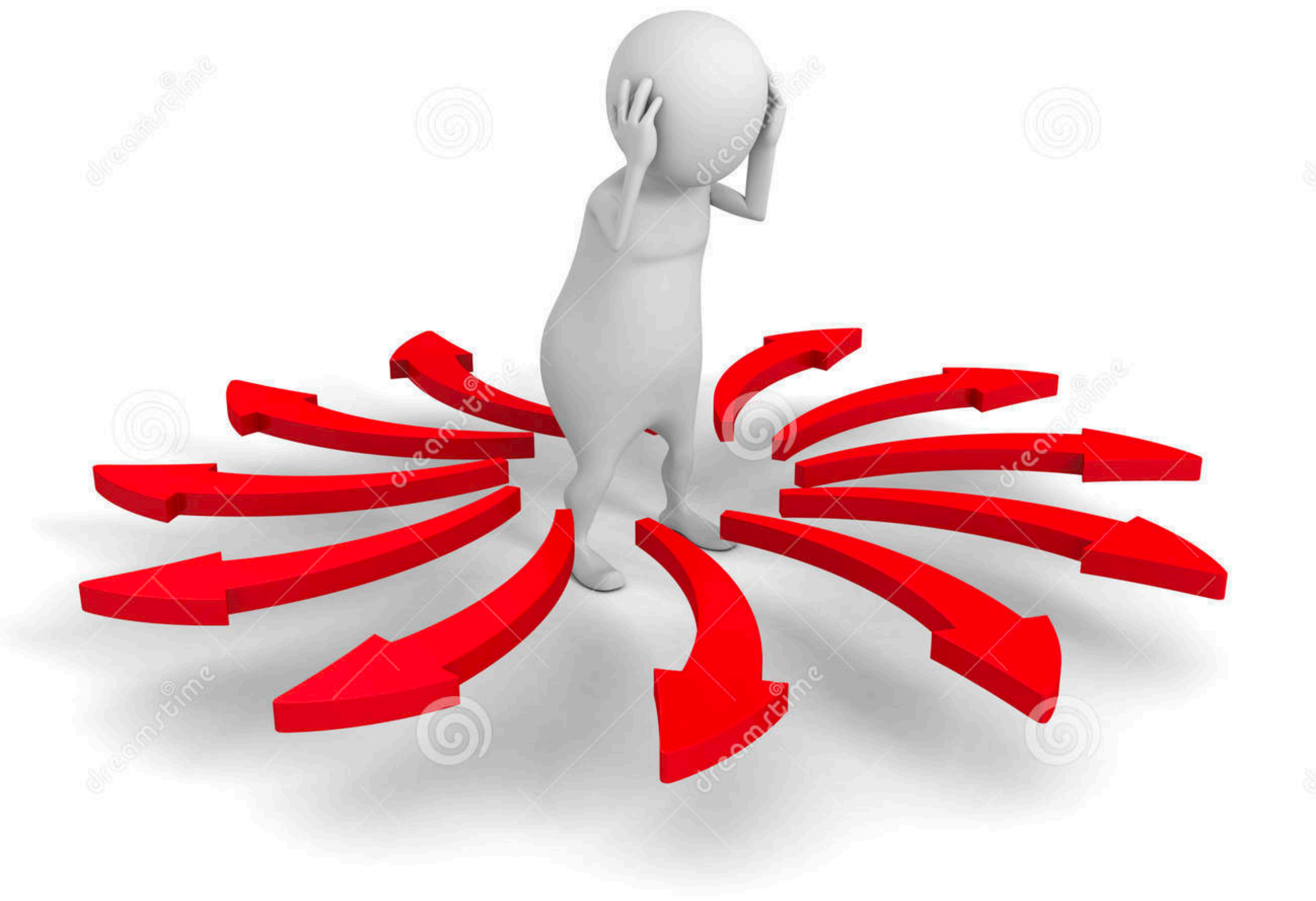


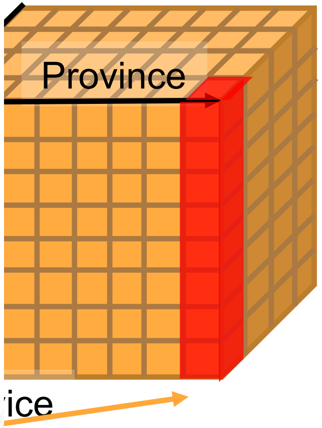
Motiv

Raw log for a

	....
	....

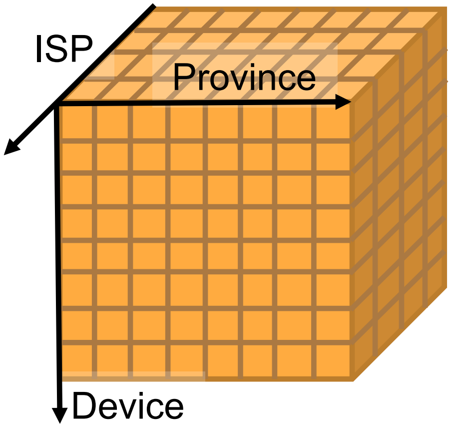






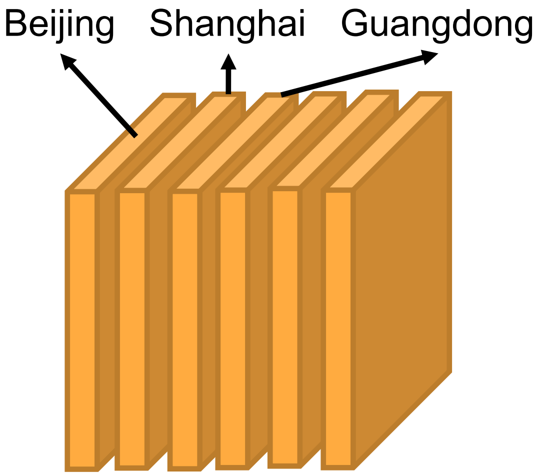
Beijing & China Unicom

Multi-dimensional Data

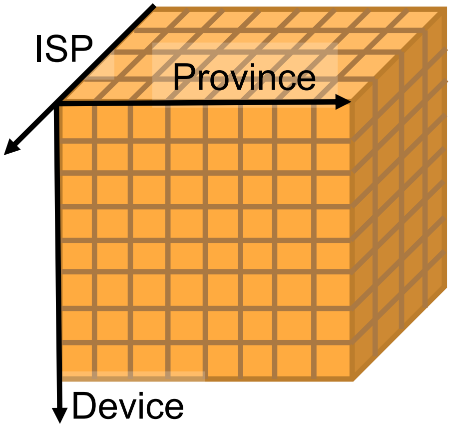


- Cuboid: a way to slice the multi-dimensional data
- Attribute combination: elements in a cuboid

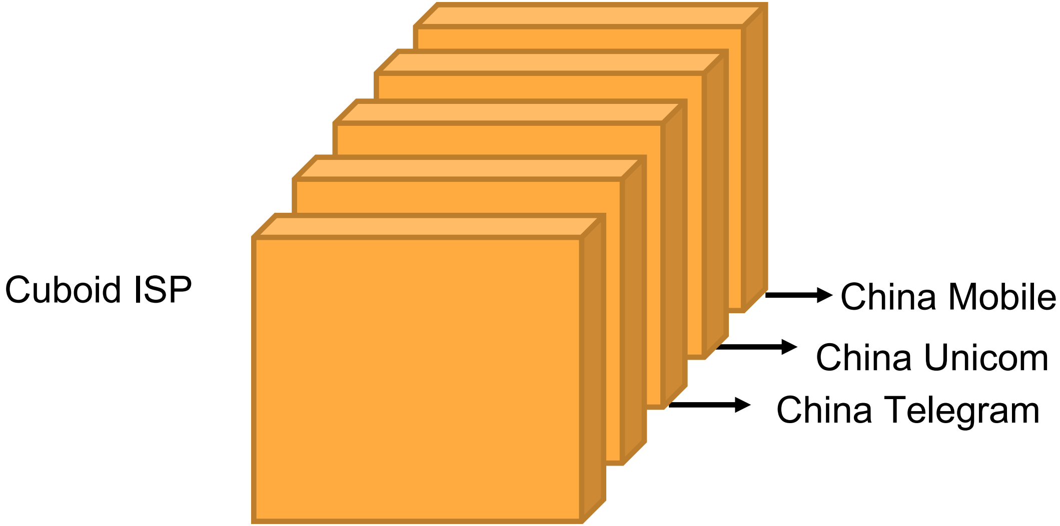
Cuboid Province



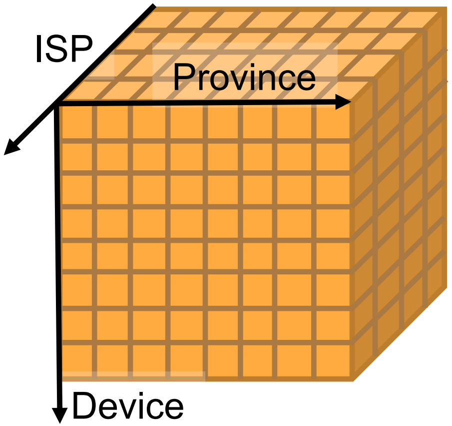
Multi-dimensional Data



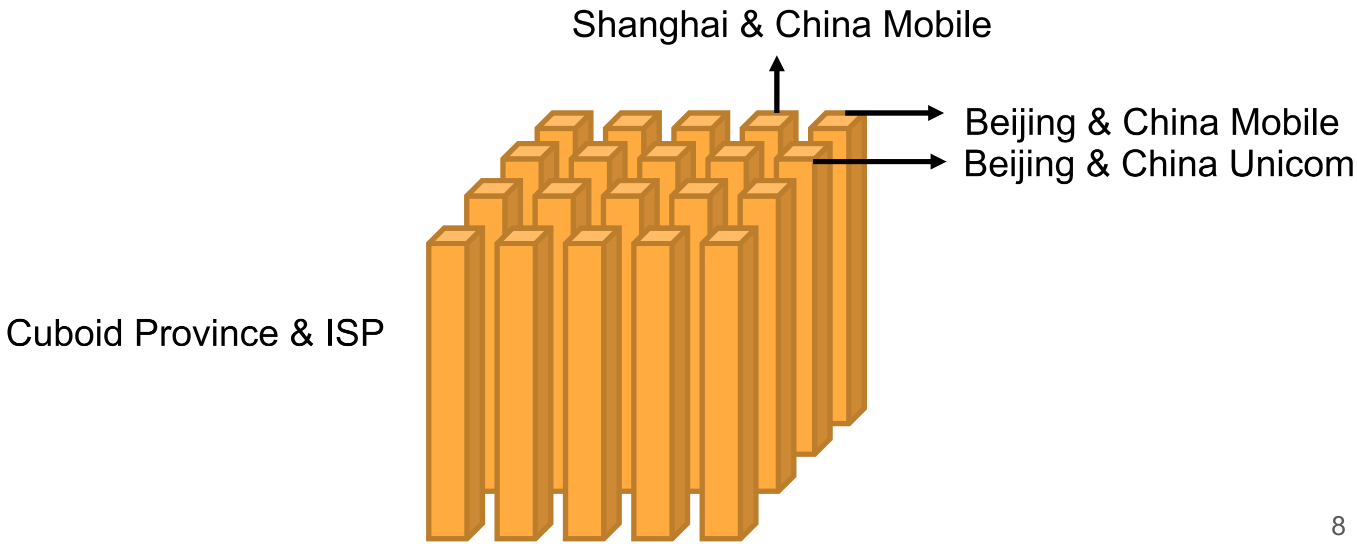
- Cuboid: a way to slice the multi-dimensional data
- Attribute combination: elements in a cuboid



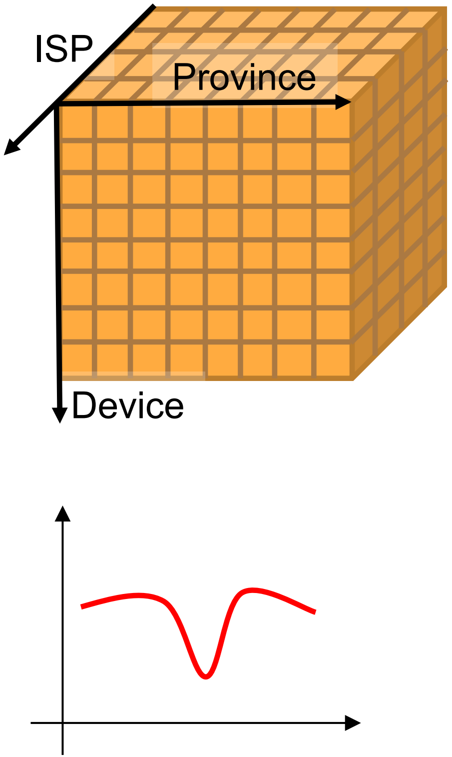
Multi-dimensional Data



- Cuboid: a way to slice the multi-dimensional data
- Attribute combination: elements in a cuboid

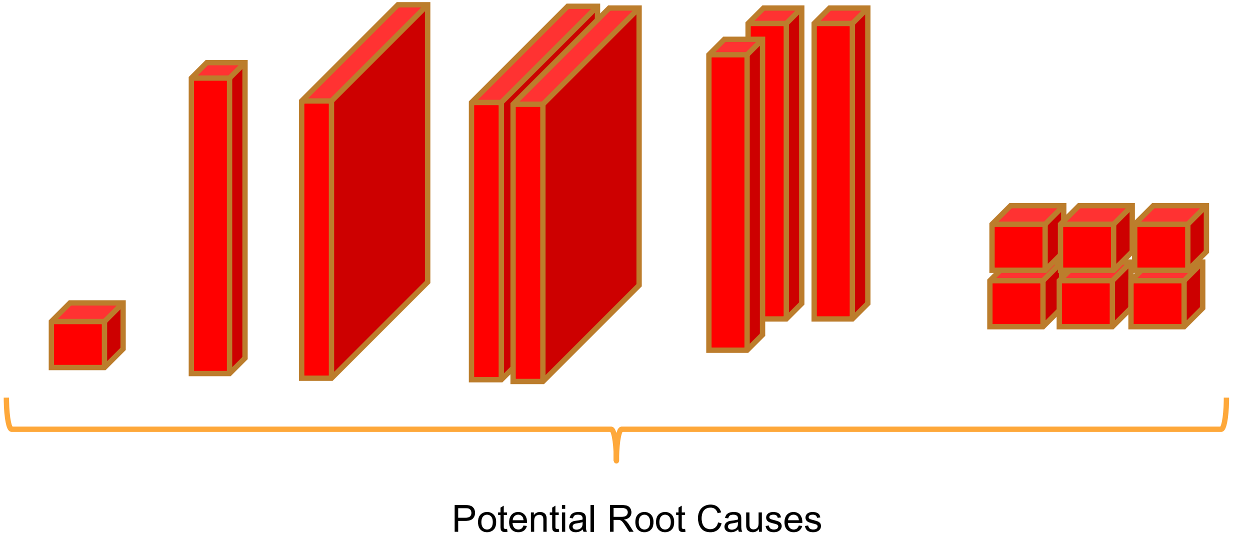


Problem Statement



The KPI of the whole cube is abnormal,
but where is the root cause?

Root cause is a set of attribute combinations



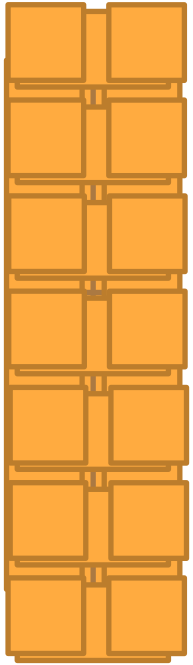
Challenge: Huge Search Space

Root Cause: a set of attribute combinations

How many potential root cause for a simple 2-d data?

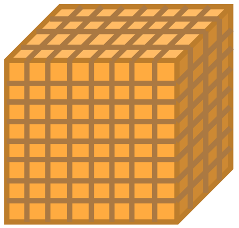
$$2^2+7+14-1$$

$$2^{2^2+7+14-1}$$

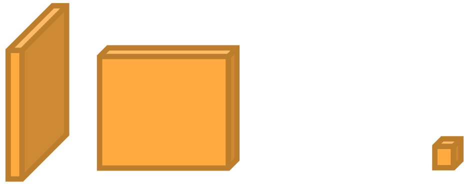


Previous Approaches

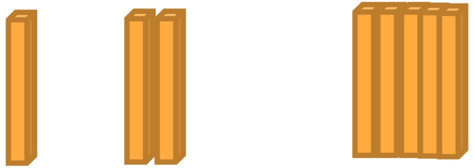
Algorithm	**Root Cause Assumption**
Adtributor (NSDI, 2014)	single attribute
Recursive Adtributor (Master Thesis, 2018)	none
iDice (ICSE, 2016)	one or two attribute combinations
Apriori (TON, 2017)	none
HotSpot (IEEE Access, 2018)	all attribute combinations of the root cause in one cuboid
Squeeze (ISSRE, 2019)	those which cause the same changes are in one cuboid



Adtributor

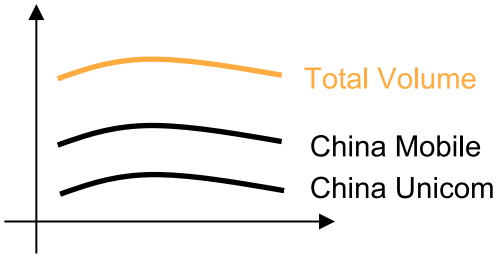


iDice

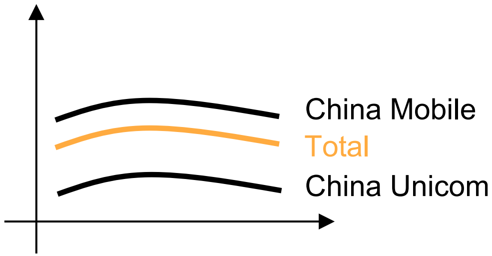


Previous Approaches

Algorithm	Measure
Adtributor (NSDI, 2014)	fundamental & derived (quotient)
Recursive Adtributor (Master Thesis, 2018)	fundamental & derived (quotient)
iDice (ICSE, 2016)	fundamental only
Apriori (TON, 2017)	fundamental & derived
HotSpot (IEEE Access, 2018)	fundamental only
Squeeze (ISSRE, 2019)	fundamental & derived (quotient, product)



Orders
fundamental, additive

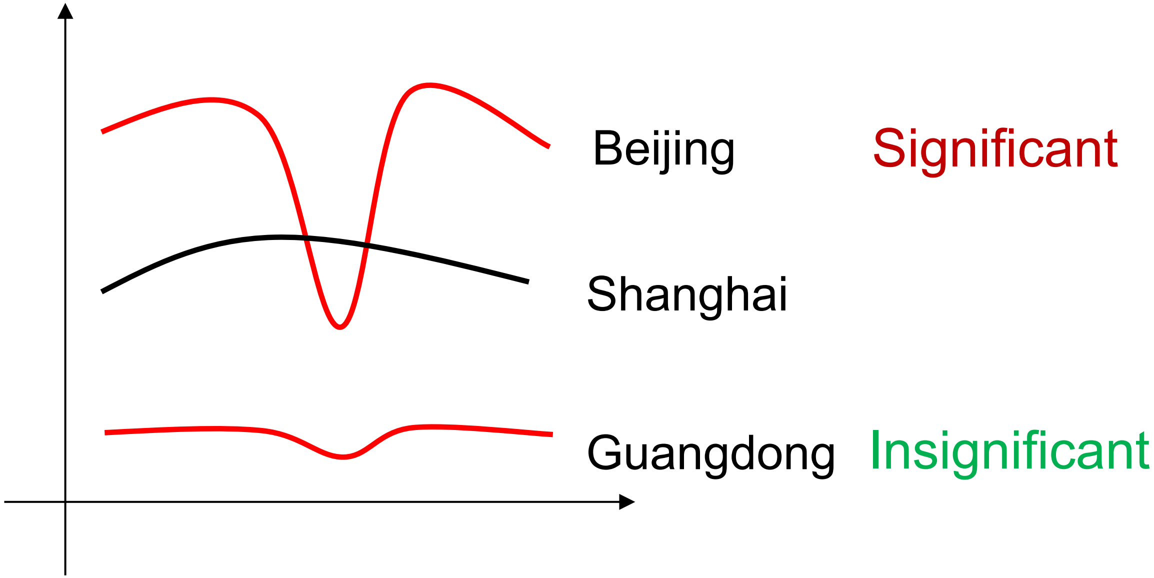


% Success Rate
derived, not additive

iDice and HotSpot rely on addition,
thus cannot handle derived measures

Previous Approaches

Algorithm	Change Magnitude
Adtributor (NSDI, 2014)	significant
Recursive Adtributor (Master Thesis, 2018)	significant
iDice (ICSE, 2016)	significant
Apriori (TON, 2017)	any
HotSpot (IEEE Access, 2018)	significant
Squeeze (ISSRE, 2019)	any



Previous Approaches

Algorithm	Parameter Fine Tuning
Adtributor (NSDI, 2014)	no
Recursive Adtributor (Master Thesis, 2018)	yes
iDice (ICSE, 2016)	no
Apriori (TON, 2017)	yes
HotSpot (IEEE Access, 2018)	no
Squeeze (ISSRE, 2019)	no

Some approaches perform badly without parameter fine tuning

Previous Approaches

Algorithm	**Time Cost**
Adtributor (NSDI, 2014)	very short
Recursive Adtributor (Master Thesis, 2018)	short
iDice (ICSE, 2016)	very short
Apriori (TON, 2017)	always too long
HotSpot (IEEE Access, 2018)	sometimes long
Squeeze (ISSRE, 2019)	short

Some approaches cost too much time

Previous Approach

Algorithm	Root Cause Assumption	Measure	Change Magnitude	Parameter Fine Tuning	Time Cost
Adtributor (NSDI, 2014)	single attribute	fundamental & derived (quotient)	significant	no	very short
Recursive Adtributor (Master Thesis, 2018)	none	fundamental & derived (quotient)	significant	yes	short
iDice (ICSE, 2016)	one or two attribute combinations	fundamental only	significant	no	very short
Apriori (TON, 2017)	none	fundamental & derived	any	yes	always too long
HotSpot (IEEE Access, 2018)	all attribute combinations of the root cause in one cuboid	fundamental only	significant	no	sometimes long
Squeeze (ISSRE, 2019)	**those which cause the same changes are in one cuboid**	**fundamental & derived (quotient, product)**	**any**	**no**	**short**

Design Goals

Root Cause Assumption	Measure	Change Magnitude	Parameter Fine Tuning	Time Cost

Squeeze

has no impractical assumptions

handles both fundamental and derived measures

handles anomalies with any change magnitude

does not need parameter fine tuning

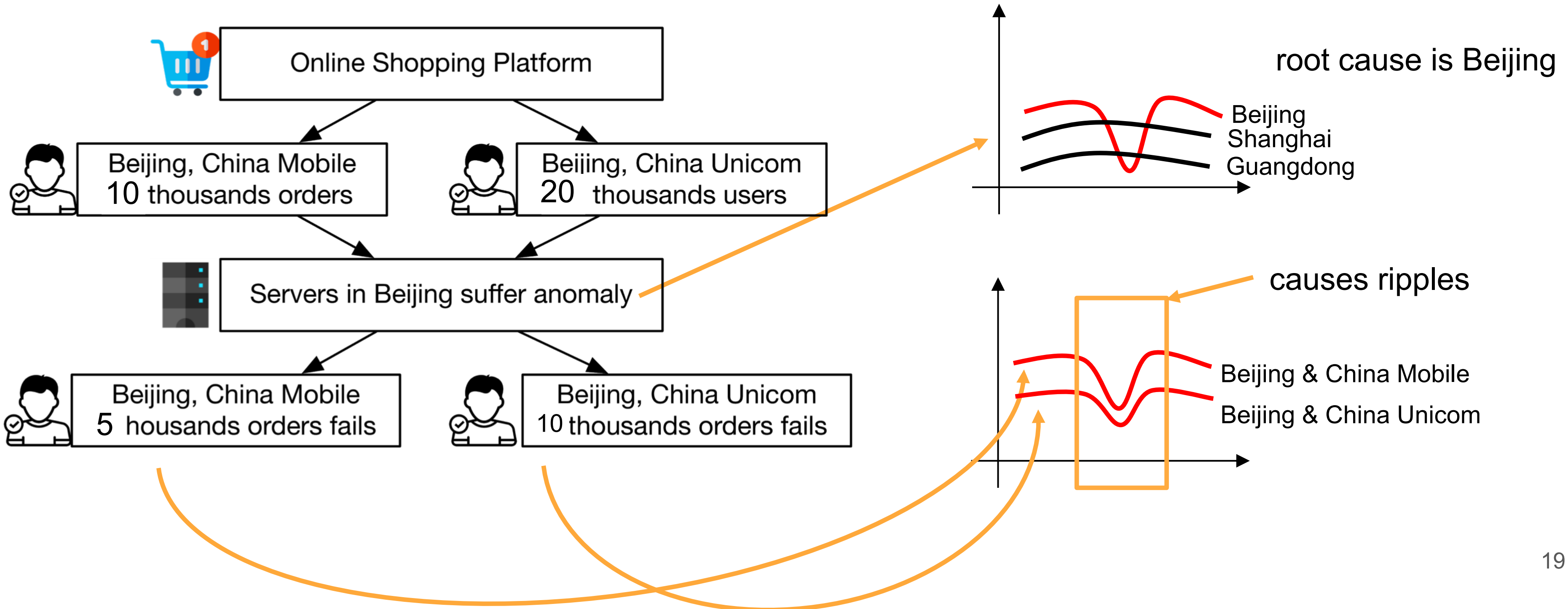
is consistently fast in all cases

Outline

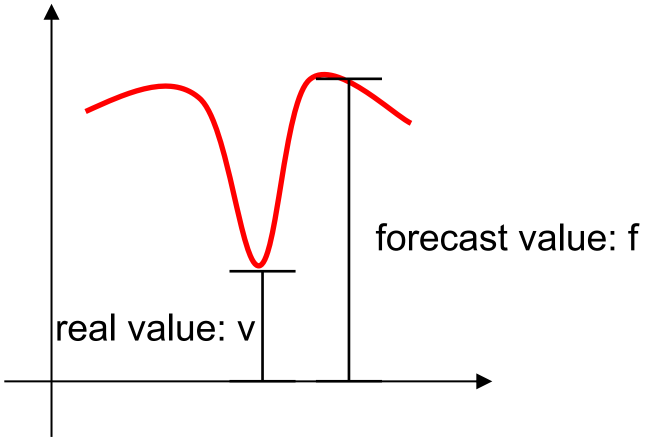
Background → Methodology → Experiment → Summary

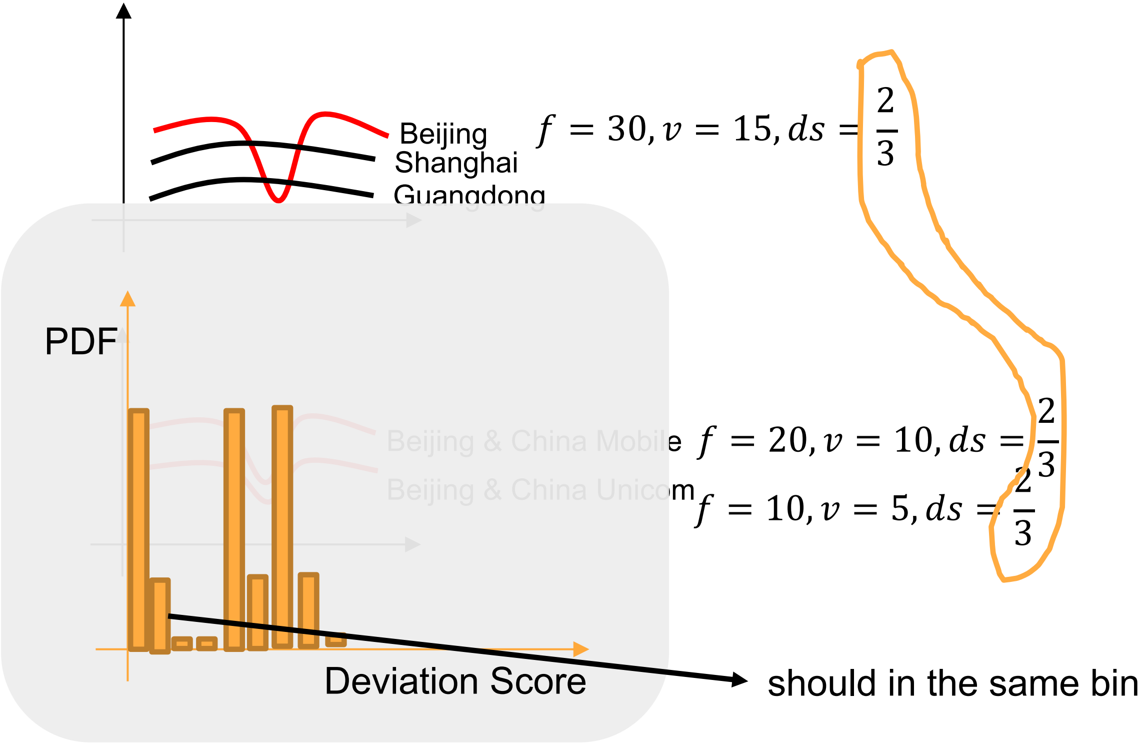
Core Idea: Generalized Ripple Effect (GRE)

With idea from HotSpot[IEEE Access 2018], we propose generalized ripple Effect



Core Idea: GRE & Deviation Score





$f = 30, v = 15, ds = \frac{2}{3}$

$f = 20, v = 10, ds = \frac{2}{3}$

$f = 10, v = 5, ds = \frac{2}{3}$

$$deviation\ score = 2\frac{f - v}{f + v}$$

Core Idea: GRE in Real World Cases

successful orders drops down after an update

By manually analysis, root cause is *ServiceType*=020020

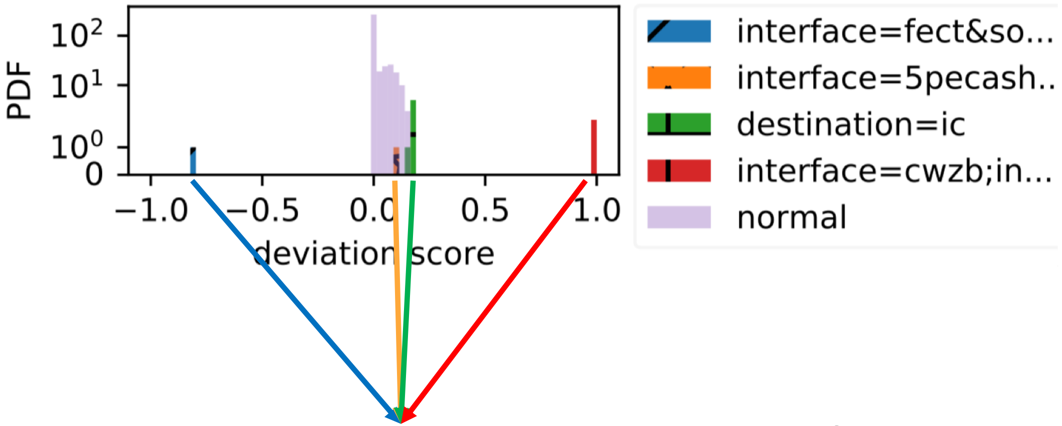
Their deviation scores are in the same bin, which supports GRE

Core Idea: GRE in Real World Cases

successful orders drops down

4 root cause attribute combinations

Case 2



The data shows that deviation scores of the same root cause are in the same bin

Generalized Ripple Effect

Does *GRE* holds for both fundamental and derived measures?

Yes. Please see the details in the paper.

$$\Delta_{M_3}(e)=\frac{f_{M_1}(e)}{f_{M_2}(e)}-\frac{v_{M_1}(e)}{v_{M_2}(e)}=\frac{\Delta_{M_1}(e)v_{M_2}(e)-\Delta_{M_2}(e)v_{M_1}(e)}{v_{M_2}(e)f_{M_2}(e)}$$

similarly, $\Delta_{M_3}(S)=\frac{\Delta_{M_1}(S)v_{M_2}(S)-\Delta_{M_2}(S)v_{M_1}(S)}{v_{M_2}(S)f_{M_2}(S)}$

$\because$ ripple effect, $\therefore \Delta_{M_i}(e)=\frac{f_{M_i}(e)}{f_{M_i}(S)}\Delta_{M_i}(S), i=1,2$

$$\therefore \Delta_{M_3}\frac{f_{M_3}(e)}{f_{M_3}(S)}=\frac{\Delta_{M_1}(S)v_{M_2}(S)-\Delta_{M_2}(S)v_{M_1}(S)}{v_{M_2}(S)f_{M_2}(S)}\frac{f_{M_2}(S)}{f_{M_1}(S)}\frac{f_{M_1}(e)}{f_{M_2}(e)}$$

$$=\frac{\Delta_{M_1}(e)-\frac{\Delta_{M_2}(S)}{v_{M_2}(S)}v_{M_1}(e)}{f_{M_2}(e)}=\frac{\Delta_{M_1}(e)-\frac{\Delta_{M_2}(e)}{v_{M_2}(e)}v_{M_1}(e)}{f_{M_2}(e)}=\Delta_{M_3}(e) \quad \square$$

Core Idea: Generalized Potential Score

Evaluate how likely a set of attribute combination is the root cause

$$GPS = 1 - \frac{\mathrm{avg}(|\boldsymbol{v}(S_1) - \boldsymbol{a}(S_1)|) + \mathrm{avg}(|\boldsymbol{v}(S_2) - \boldsymbol{f}(S_2)|)}{\mathrm{avg}(|\boldsymbol{v}(S_1) - \boldsymbol{f}(S_1)|) + \mathrm{avg}(|\boldsymbol{v}(S_2) - \boldsymbol{f}(S_2)|)}$$

Core Idea: Generalized Potential Score

partition

Sorted candidates list: $\underbrace{\text{(Beijing,*),}}_{\text{S1: abnormal}}$ | $\underbrace{\text{(Shanghai, *), (Guangdong, *), (Zhejiang, *, ...)}}_{\text{S2: normal}}$

→ KPI value should be expected by GRE

→ $\frac{v(\text{Beijing})}{f(Beijing)} = 0.5$, half fails

→ $a(Beijing, China\ Mobile) = f(Beijing, China\ Mobile) * 0.5 = 5$

→ $a(Beijing, China\ Unicom) = f(Beijing, China\ Unicom) * 0.5 = 10$

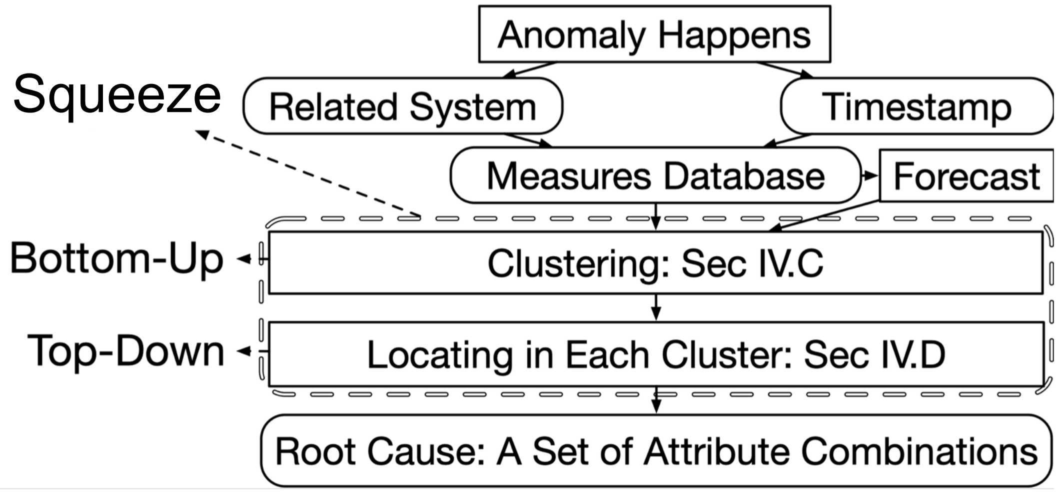
→ forecast value and real value should be close

→ f(S2) – v(S2) ~ 0

$$GPS = 1 - \frac{\overset{0}{\operatorname{avg}(|\boldsymbol{v}(S_1) - \boldsymbol{a}(S_1)|)} + \overset{0}{\operatorname{avg}(|\boldsymbol{v}(S_2) - \boldsymbol{f}(S_2)|)}}{\operatorname{avg}(|\boldsymbol{v}(S_1) - \boldsymbol{f}(S_1)|) + \operatorname{avg}(|\boldsymbol{v}(S_2) - \boldsymbol{f}(S_2)|)}$$

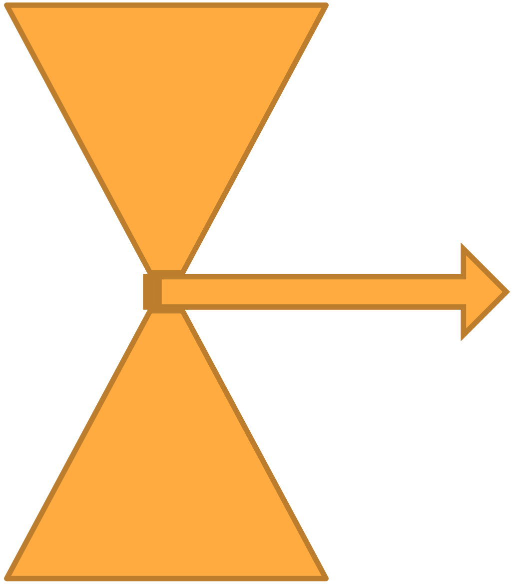
normalization

Overall Architecture



Squeeze

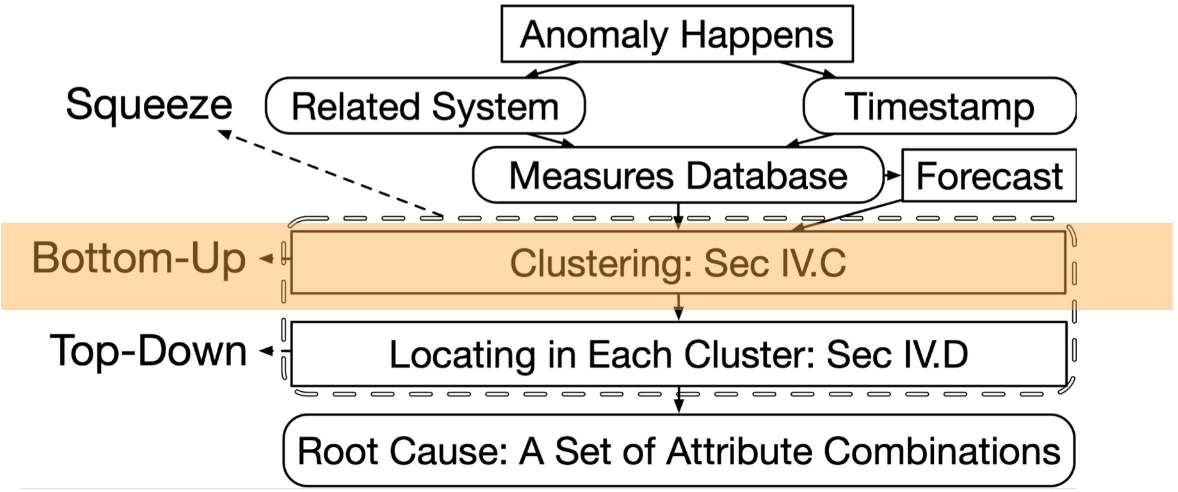
Top to Bottom: Search in each cluster



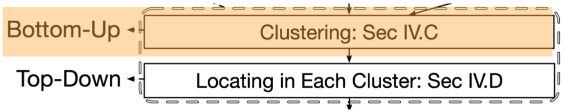
Root Causes

Bottom to Top: clustering for leaf attribute combinations

Clustering



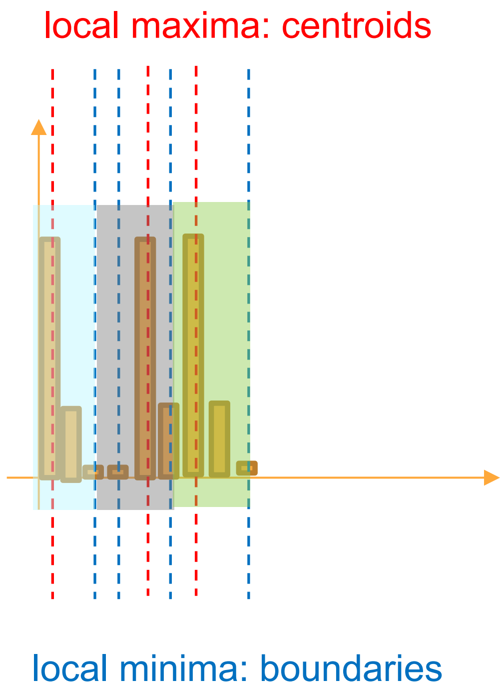
Anomaly Happens
Squeeze
Related System
Timestamp
Measures Database
Forecast
Bottom-Up
Clustering: Sec IV.C
Top-Down
Locating in Each Cluster: Sec IV.D
Root Cause: A Set of Attribute Combinations



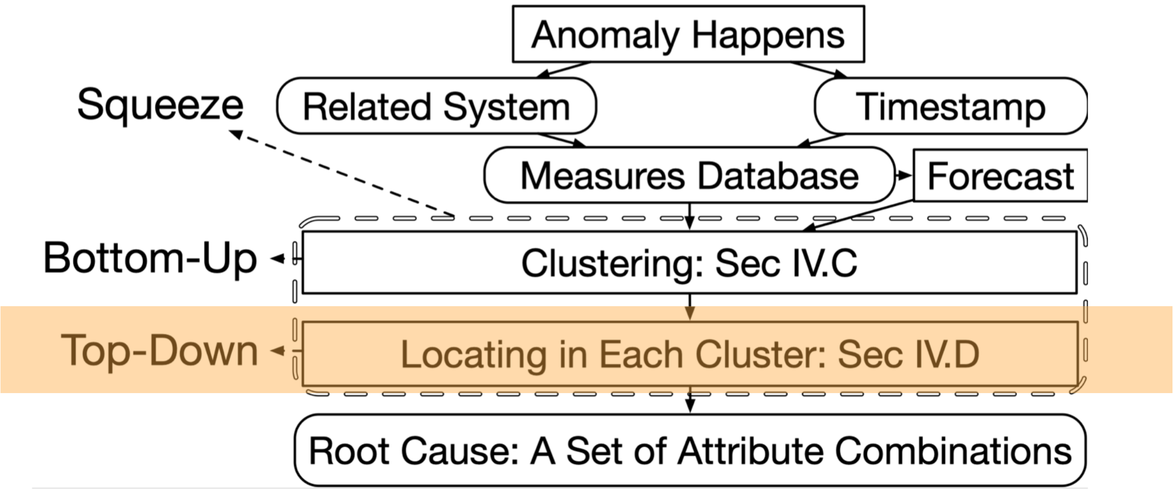
Clustering

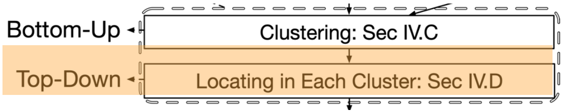
Find attribute combinations affected by the same root cause

Find attribute combinations have similar deviation scores

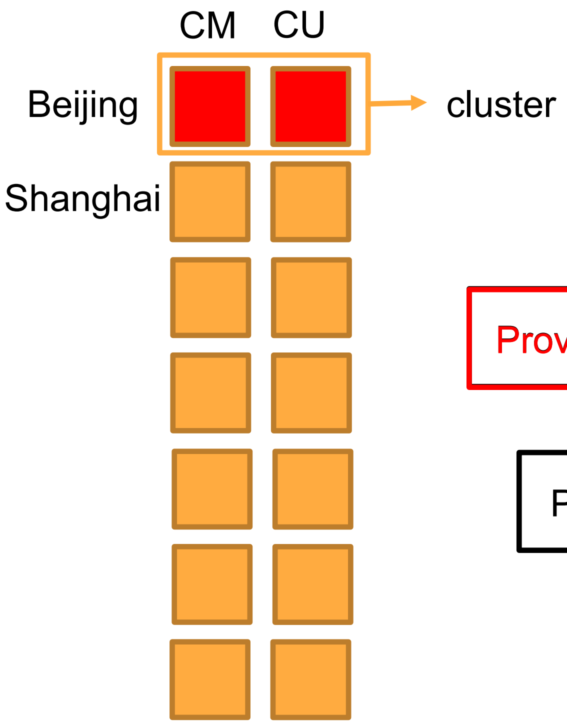


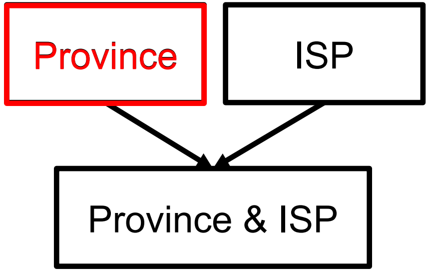
Localize in Each Cluster

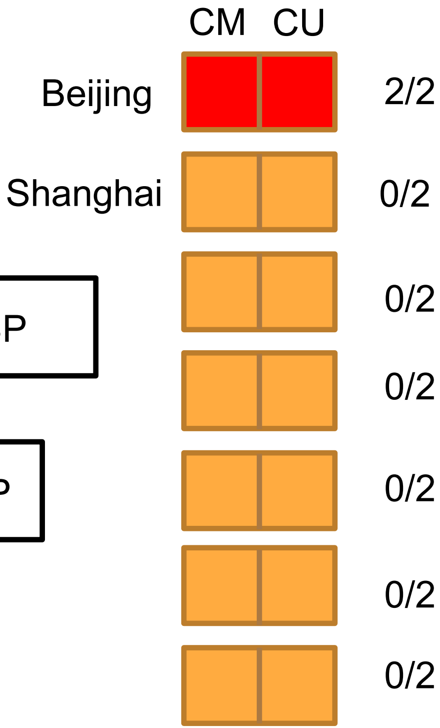




Localize in Cluster







Sorted List:
Beijing, Shanghai, ...

Top-K items in this list
with highest GPS

Beijing, GPS = 1, Root Cause

Outline

Background → Methodology → Experiment → Summary

Experiment Setup

We use

- real KPI datasets from 2 companies;
- synthetic anomalies => 7 semi-synthetic datasets
- Moving average as the forecasting algorithm.

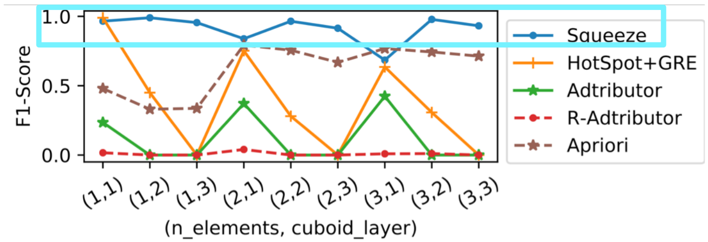
Effectiveness

Squeeze achieves relatively good F1-score on both fundamental & derived measures.

Two of Fundamental Measure Datasets

F1-score		(n_elements, cuboid_ayer)								
Dataset	Algorithm	(1, 1)	(1, 2)	(1, 3)	(2, 1)	(2, 2)	(2, 3)	(3,1)	(3,2)	(3, 3)
$\mathcal{A}$	*Squeeze*	**0.8632**	**0.7827**	**0.4932**	**0.7584**	**0.6361**	**0.4097**	**0.6441**	**0.5145**	**0.3618**
	HotSpot	0.6856	0.4389	0.2158	0.5085	0.3433	0.2043	0.3988	0.2916	0.1768
	Adtributor	0.3892	0.0000	0.0000	0.4010	0.0000	0.0000	0.3857	0.0000	0.0000
	R-Adtributor	0.0180	0.0020	0.0016	0.0075	0.0049	0.0294	0.0081	0.0067	0.0410
	iDice	0.0000	0.0036	0.0425	0.0000	0.0065	0.0437	0.0000	0.0007	0.0172
	Apriori	0.1036	0.0580	0.0001	0.1427	0.0926	0.0019	0.1537	0.0882	0.0062
$\mathcal{B}_0$	*Squeeze*	0.9041	**0.9327**	**0.9231**	**0.9604**	**0.9799**	**0.9333**	**0.9631**	**0.9371**	**0.9228**
	HotSpot	**0.9950**	0.4928	0.1215	0.7588	0.3934	0.0577	0.5961	0.3043	0.0775
	Adtributor	0.3044	0.0000	0.0000	0.4226	0.0000	0.0000	0.4654	0.0000	0.0000
	R-Adtributor	0.0639	0.0000	0.0000	0.0114	0.0000	0.0000	0.0177	0.0000	0.0000
	iDice	0.0000	0.0517	0.0488	0.0000	0.0409	0.0618	0.0000	0.0228	0.0959
	Apriori	0.4430	0.5116	0.7523	0.8490	0.6853	0.7351	0.8743	0.8087	0.7368

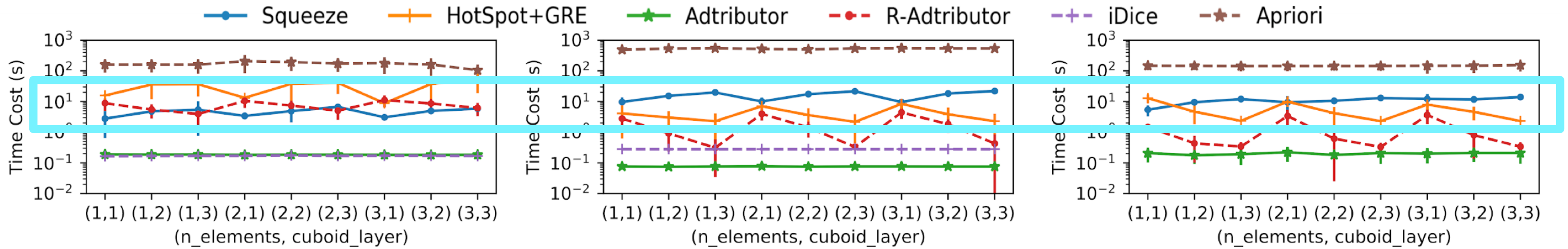
Derived Measure Dataset



Efficiency

Squeeze is fast enough consistently in all cases.

Squeeze costs only ten to twenty seconds consistently in all cases.

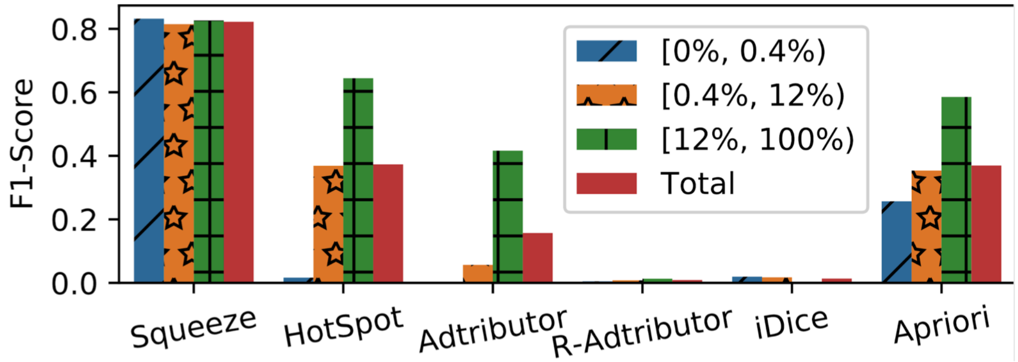


(a) $\mathcal{A}$ (b) $\mathcal{B}_0$ (c) $\mathcal{D}$

Various Anomaly Change Magnitude

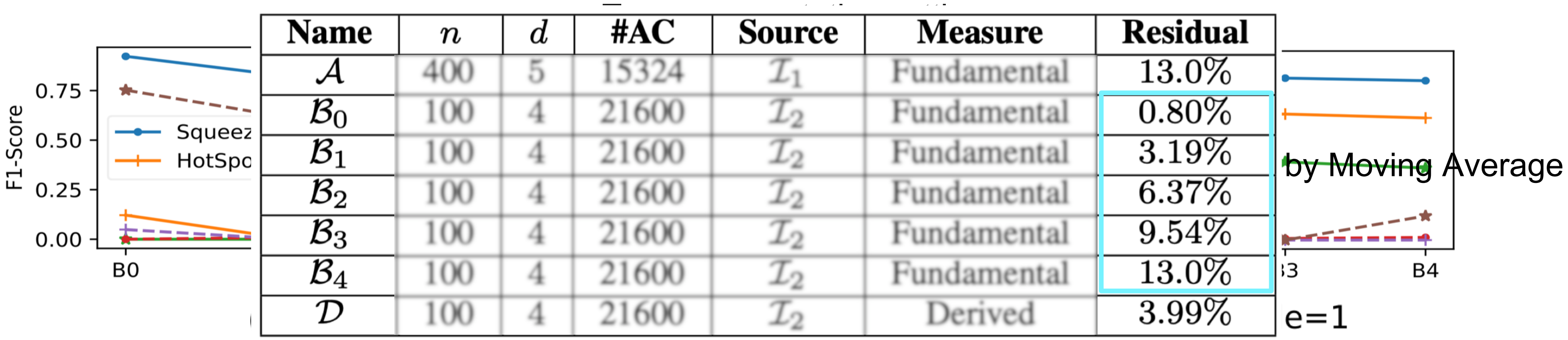
Squeeze performs well regardless of anomaly change magnitudes

0.4% and 12% are 25 and 75 percentile of change magnitudes



Various Forecasting Residual

Squeeze performs well under various residuals, and always outperforms others.



Name	n	d	#AC	Source	Measure	Residual
$\mathcal{A}$	400	5	15324	$\mathcal{I}_1$	Fundamental	13.0%
$\mathcal{B}_0$	100	4	21600	$\mathcal{I}_2$	Fundamental	0.80%
$\mathcal{B}_1$	100	4	21600	$\mathcal{I}_2$	Fundamental	3.19%
$\mathcal{B}_2$	100	4	21600	$\mathcal{I}_2$	Fundamental	6.37%
$\mathcal{B}_3$	100	4	21600	$\mathcal{I}_2$	Fundamental	9.54%
$\mathcal{B}_4$	100	4	21600	$\mathcal{I}_2$	Fundamental	13.0%
$\mathcal{D}$	100	4	21600	$\mathcal{I}_2$	Derived	3.99%

Outline

Background → Methodology → Experiment → Summary

Summary

- Bottom-up & Top-down => Squeeze
- Contributions:
  - Generalized ripple effect
  - Squeeze algorithm.
  - Experimental study on real world data and semi-synthetic data show Squeeze is both **effective** and **efficient.**
- Future Works
  - focus on numerical attributes
  - show GRE for more types of derived measures

References

- *Ahmed, F., Erman, J., Ge, Z., Liu, A., Wang, J., Yan, H.* (2017). **Detecting and Localizing End-to-End Performance Degradation for Cellular Data Services Based on TCP Loss Ratio and Round Trip Time** IEEE/ACM Transactions on Networking (TON) 25()
- *Bhagwan, R., Kumar, R., Ramjee, R., NSDI, V.* (2014). **Adtributor: Revenue Debugging in Advertising Systems.**
- *Lin, Q., Lou, J., Zhang, H., Zhang, D.* (2016). **iDice: Problem Identification for Emerging Issues** 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) https://dx.doi.org/10.1145/2884781.2884795
- *Rudenius, L., Persson, M.***Anomaly Detection and Fault Localization.** Master's thesis, 2018, goteborg : Chalmers University of Technology.
- *Sun, Y., Zhao, Y., Su, Y., Liu, D., Nie, X., Meng, Y., Cheng, S., Pei, D., Zhang, S., Qu, X., Guo, X.* (2018). **HotSpot: Anomaly Localization for Additive KPIs With Multi-Dimensional Attributes** IEEE Access 6(), 10909-10923. https://dx.doi.org/10.1109/ACCESS.2018.2804764

Thank you. Q&A