Event Detection using Customer Care Calls

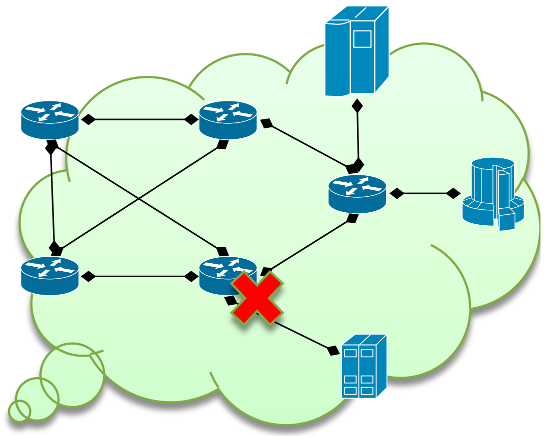
Yi-Chao Chen[1], Gene Moo Lee[1], Nick Duffield[2], Lili Qiu[1], Jia Wang[2]
The University of Texas at Austin[1], AT&T Labs – Research[2]

04/17/2013

IEEE INFOCOM 2013

Motivation

- **Customer care call** is a direct channel between **service provider** and **customers**
  - Reveal **problems** observed by customers
  - Understand **impact** of network events on customer perceived performance
    - Regions, group of users, services, …



Motivation

- **Service providers** have strong motivation to understand customer care calls
  - **Reduce cost:** ~$10 per call
  - **Prioritize and handle anomalies** by the impact
  - Improve **customers' impression** to the service provider

Problem Formulation

- **Goal**
  - Automatically detect **anomalies** using **customer care calls.**
- **Input**
  - **Customer care calls**
  - Call agents label calls using **predefined categories**
    - ~10,000 categroies
  - A call is labeled with **multiple categories**
    - Issue, customer need, call type, problem resolution

Problem Formulation

- **Output**
  - **Anomalies**
  - **Performance problems** observed by **customers**
  - e.g. ***Service outage*** due to DOS attack, ***power outage***, ***low bandwidth*** due to maintenance

Example: Categories of Customer Care Calls

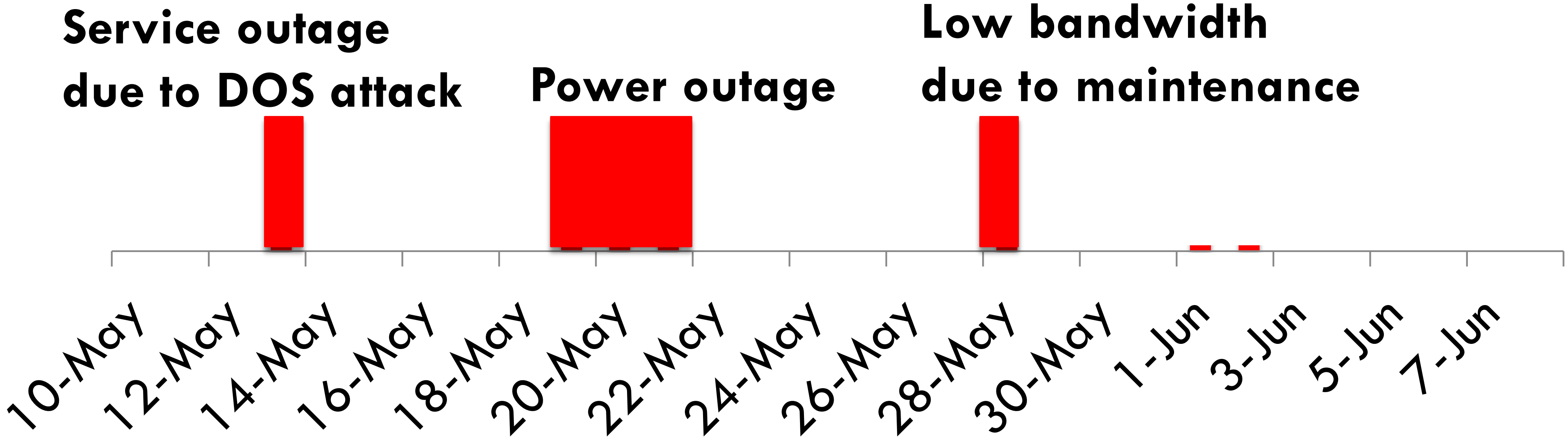
- Equipment
- Call Disconnected
- Cannot make calls
- Educated - How to use
- Equipment inquiry
- Feature
- Service Plan

Input: Customer Care Calls

$$
\begin{array}{c}
 \\
Categories_1 \\
Categories_2 \\
Categories_3 \\
\ldots \\
Categories_N
\end{array}
\begin{array}{c}
\begin{array}{ccccc} \mathrm{T}_1 & \mathrm{T}_2 & \mathrm{T}_3 & \ldots & \mathrm{T}_t \end{array} \\
\begin{bmatrix}
82534 & 76462 & 71950 & \ldots & 25290 \\
126913 & 112925 & 96535 & \ldots & 92442 \\
64997 & 58767 & 52807 & \ldots & 48270 \\
\ldots & & & & \\
84909 & 79901 & 79671 & \ldots & 76100
\end{bmatrix}
\end{array}
$$

# of calls of **category n** in **time bin t**

Example: Anomaly

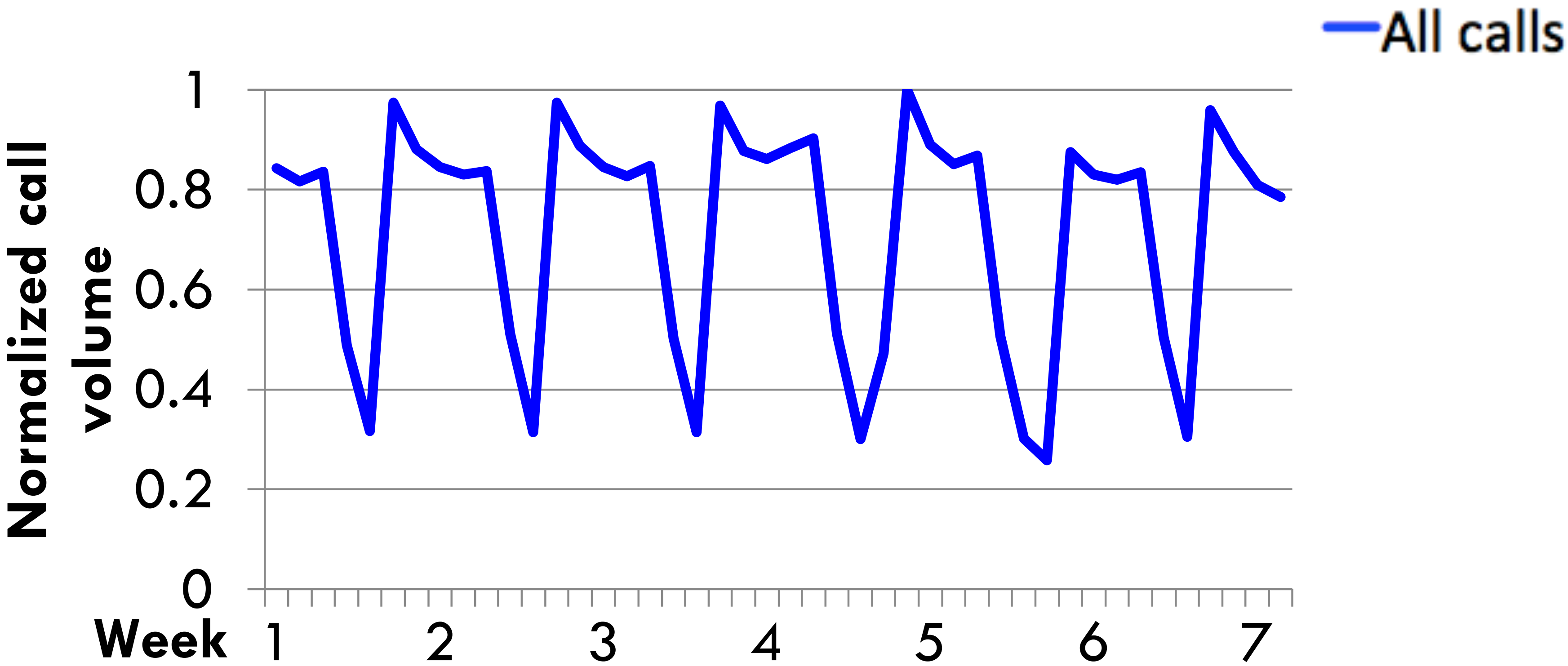


Output: anomaly indicator =

[0 0 0 1 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0]

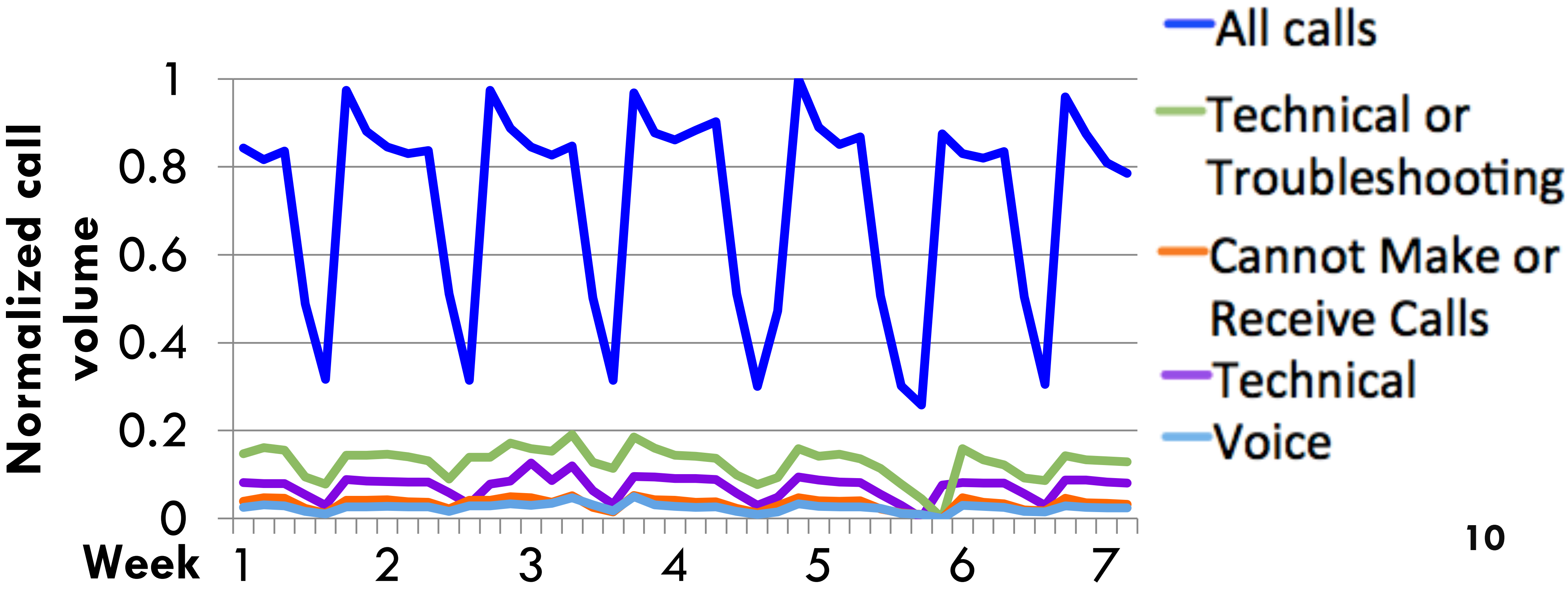
Challenges

- Customers respond to an anomaly in different ways.
- Events may not be detectable by a **single category**.



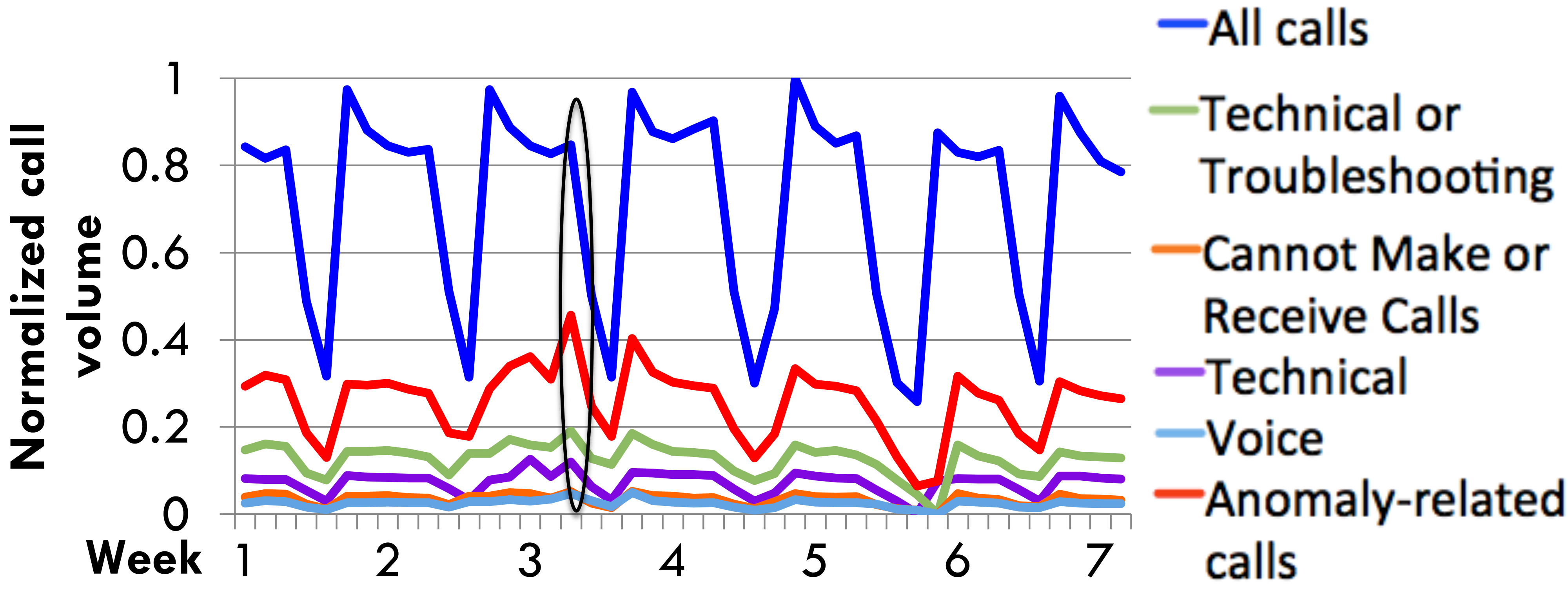
Challenges

- Customers respond to an anomaly in different ways.
- Events may not be detectable by a **single category**.



Challenges

- Customers respond to an anomaly in different ways.
- Events may not be detectable by a **single category**.
- There are **thousands of categories**.



Our Approach

- We use **regression** to approach the problem by casting it as an inference problem:

$$Ax = b$$

A(*t*,*n*): **# calls** of category *n* at time *t*

x(*n*): **weight** of the *n-th* category

b(*t*): **anomaly** indicator at time *t*

Our Approach

- **Training set:** the history data
  - **A**: Input timeseries of customer care calls
  - **b**: Ground-truth of when anomalies take place
  - **x**: The weight to learn

$$Ax = b$$

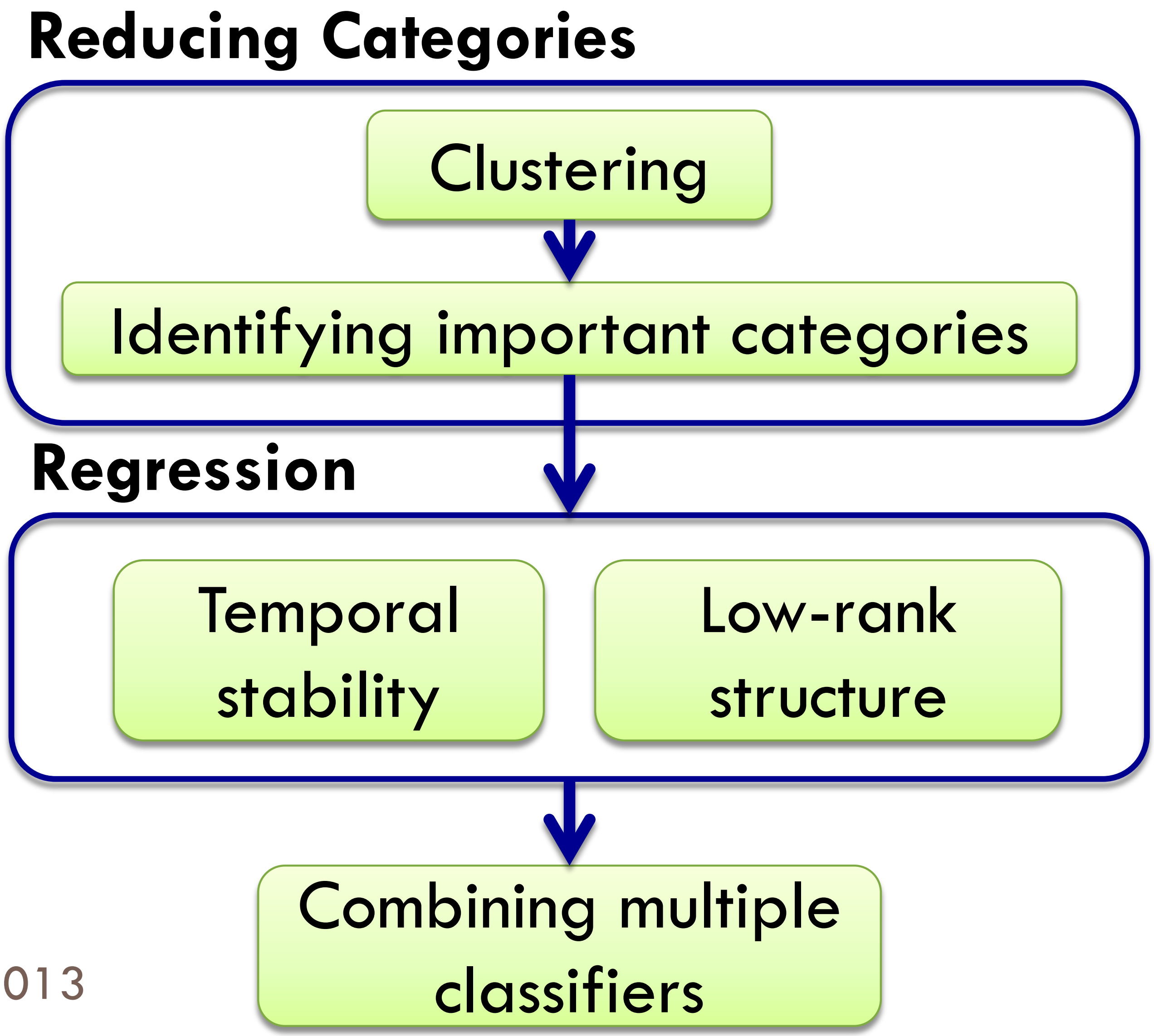
- **Testing set:**
  - **A'**: The latest customer care call timeseries
  - **x**: The weight learning from training set
  - **b'**: The anomaly to detect

$$A'x = b'$$

Issues

- **Dynamic x**
  - The relationship between customer care calls and anomalies may change.
- **Under-constraints**
  - # categories can be larger than # training traces
- **Over-fitting**
  - The weights begin to memorize training data rather than learning the generic trend
- **Scalability**
  - There are thousands of categories and thousands of time intervals.
- **Varying customer response time**

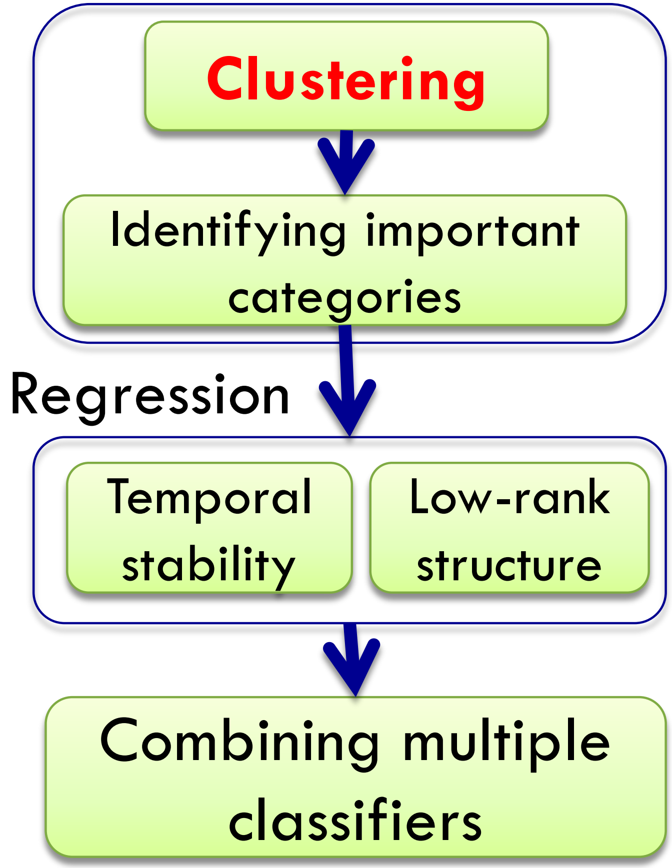
System Overview



Clustering

- Agents usually classify calls based on the **textual names** of categories.
  - e.*g*. **"Equipment"**, **"Equipment problem"**, and **"Troubleshooting- Equipment"**
- Cluster categories based on the similarity of their textual names
  - **Dice's coefficient**

Reducing Categories

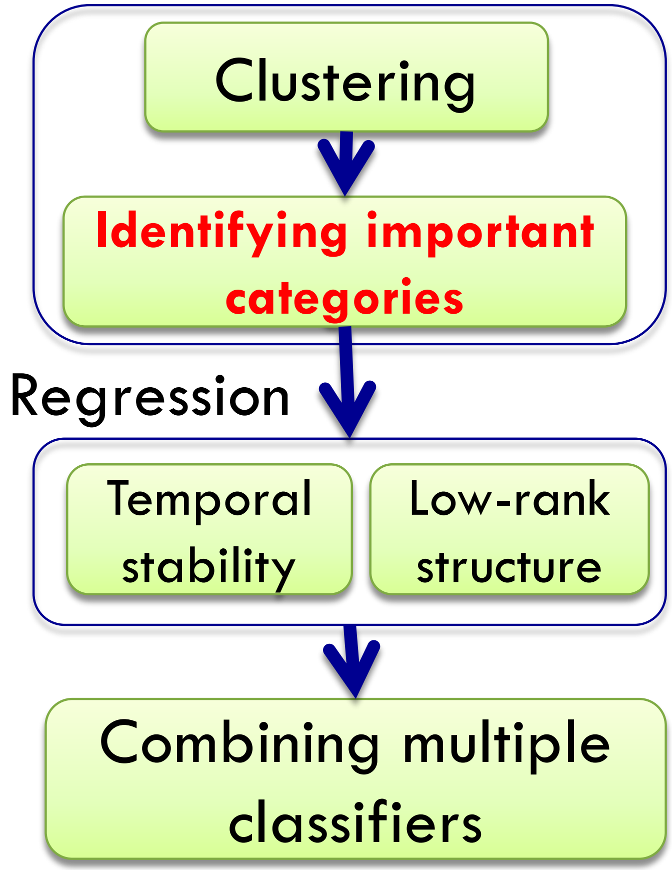


Identify Important Categories

- **L_1-norm regularization**
  - Penalize all factors in x equally and make x **sparse**
  - Select categories with corresponding value in x is not 0

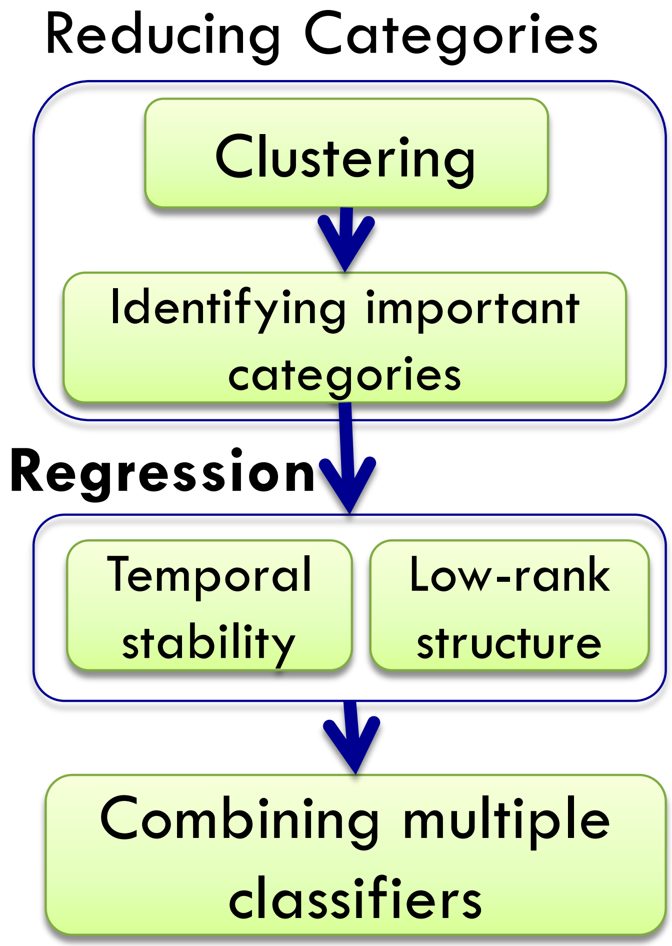
$$\min_{x} \|Ax - b\|_2^2 + \lambda \|x\|_1$$

Reducing Categories



Regression

- Impose additional structures for **under-constraints** and **over-fitting**
  - The weight values are stable
  - Small number of factors of dominate anomalies



- **Fitting Error**

$$f(X) = \sum_{d} \|A_d x_d - b_d\|_2^2$$

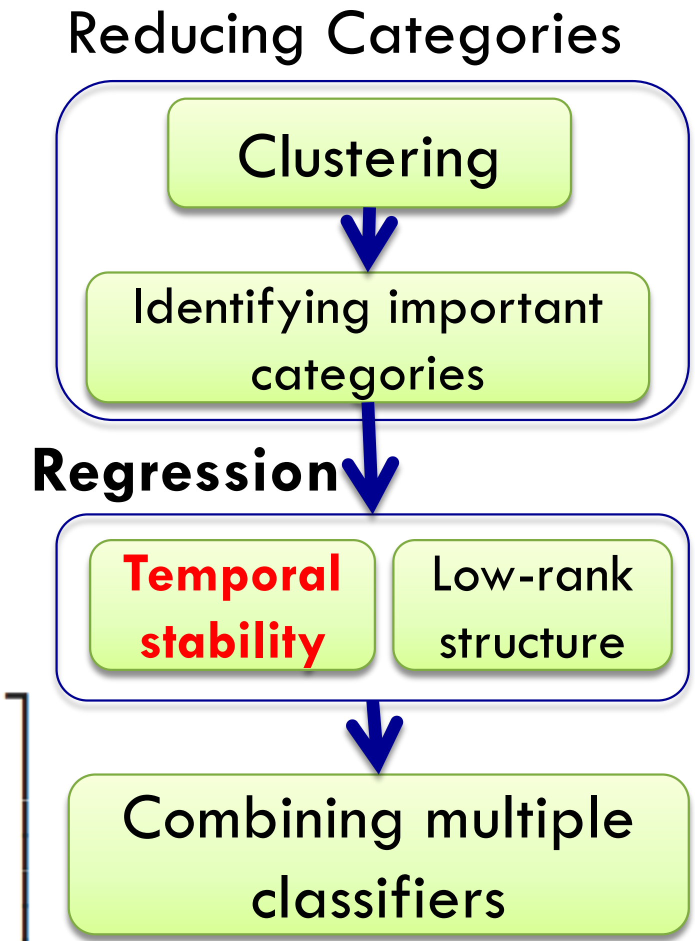
Regression

- **Temporal stability**
  - The weight values are stable across consecutive days.

$$g(X) = \left\| X \times T^T \right\|_2^2$$

$$X = [x_1 x_2 .. x_d]$$

$$T = \begin{bmatrix} 1 & -1 & 0 & \dots & \\ 0 & 1 & -1 & \ddots & \\ 0 & 0 & 1 & -1 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \end{bmatrix}$$

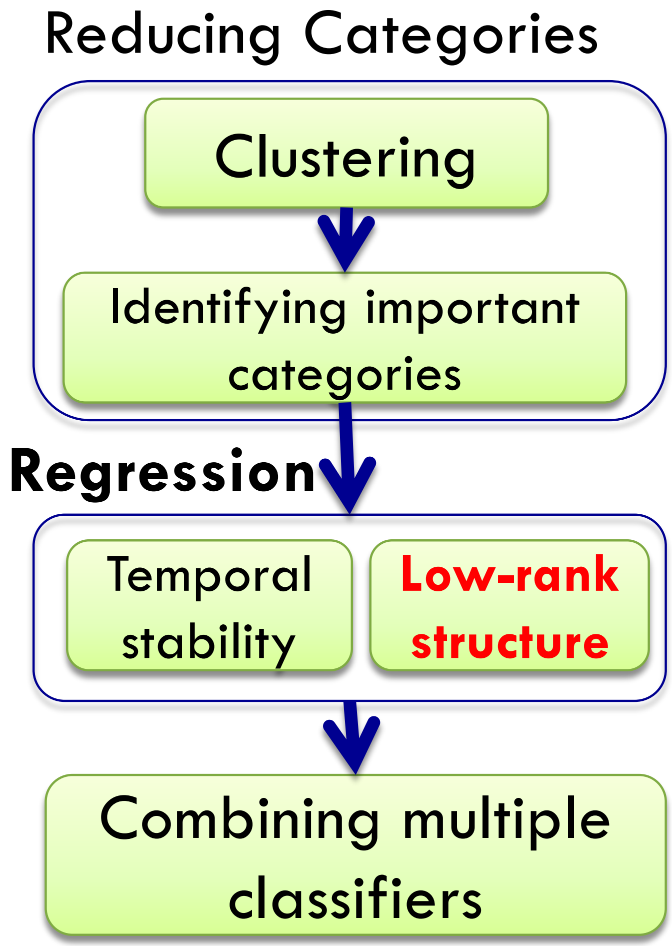


Regression

- **Low-rank structure**
  - The weight values exhibit low-rank structure due to the *temporal stability* and the small number of dominant factors that cause the anomalies.

$$h(X, U, V) = \left\| X - U \times V^T \right\|_2^2$$

$$X \approx U \times V^T$$



Regression

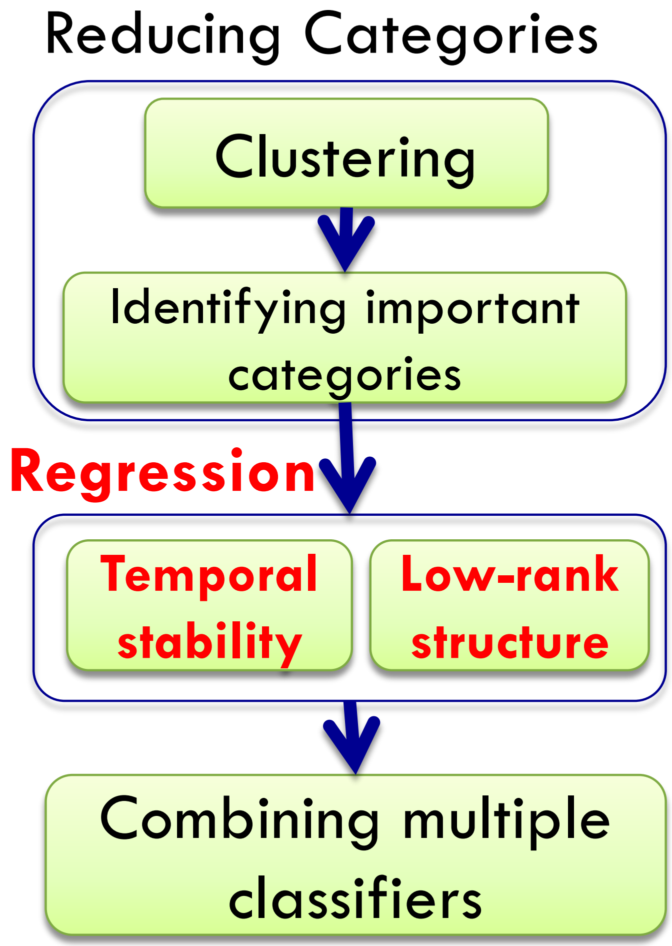
- **Find the weight X that minimize:**

$$o(X,U,V) = f(X) + \alpha \cdot g(X) + \beta \cdot h(X,U,V)$$

f(X): **Fitting error**

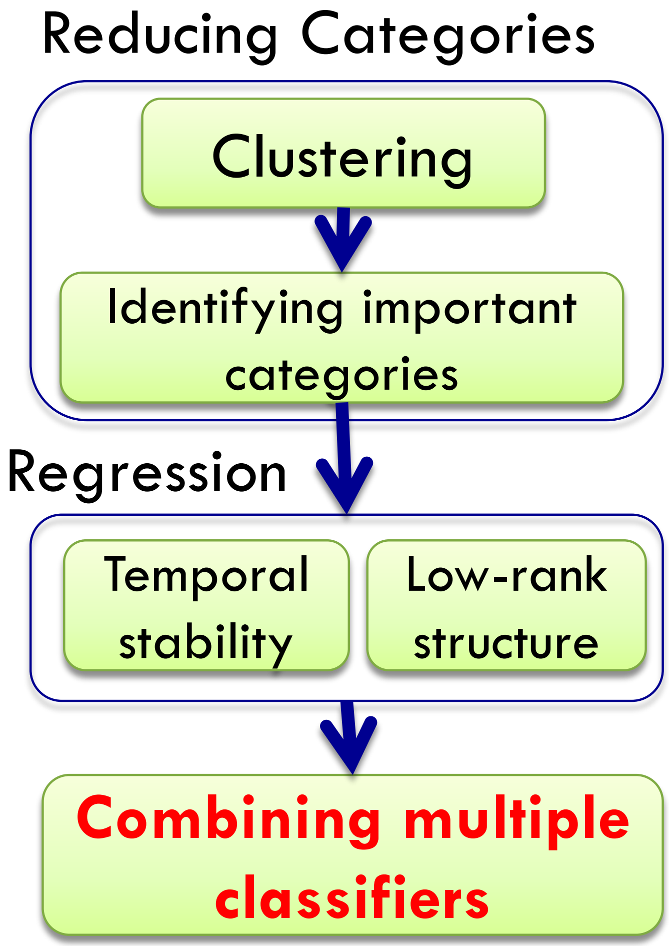
g(X): **Temporal stability**

h(X,U,V): **Low-rank**



Combining Multiple Classifiers

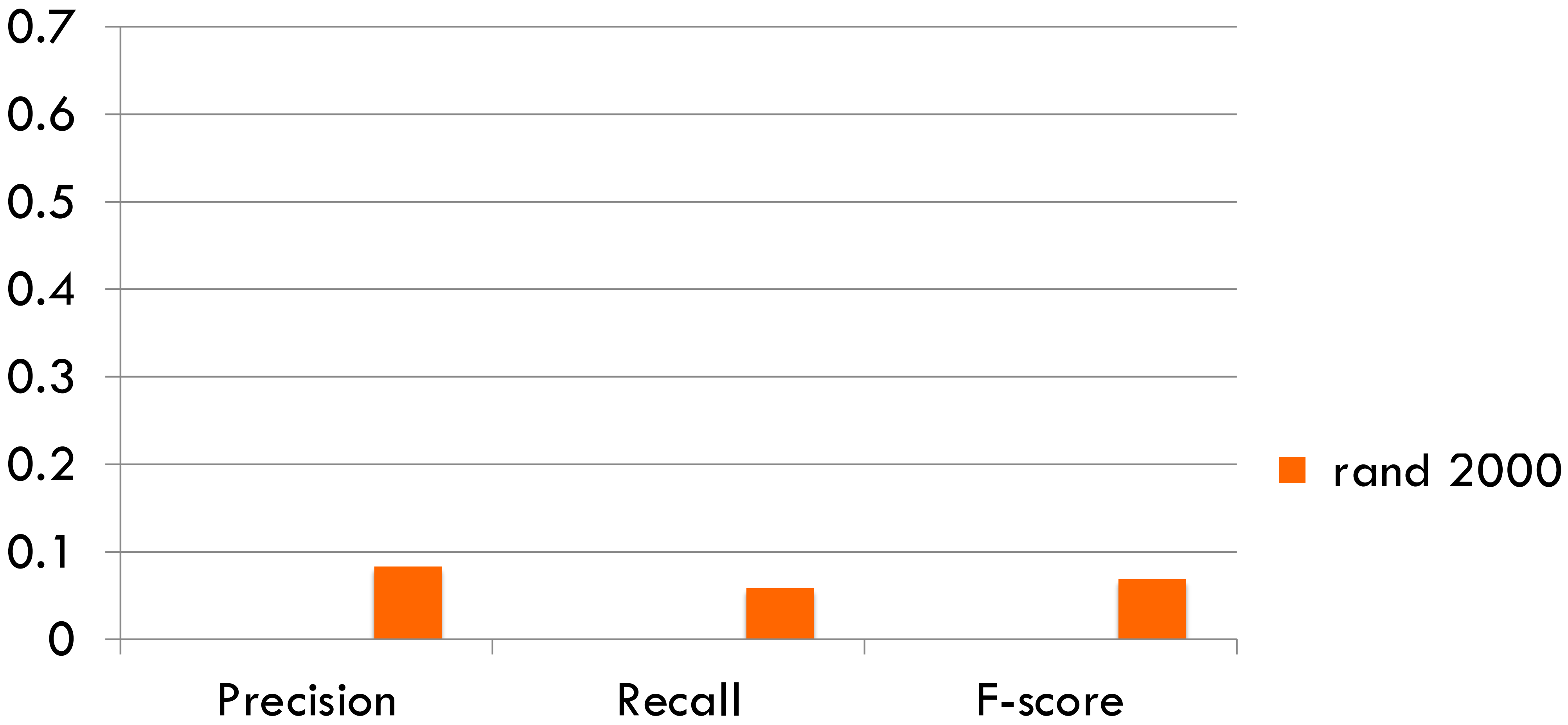
- What **time scale** should be used?
  - Customers do not respond to an anomaly immediately
  - The response time may differ by hours
- Include calls made in previous *n* (*1~5*) and next *m* (*0~6*) hours as additional features.



Evaluation

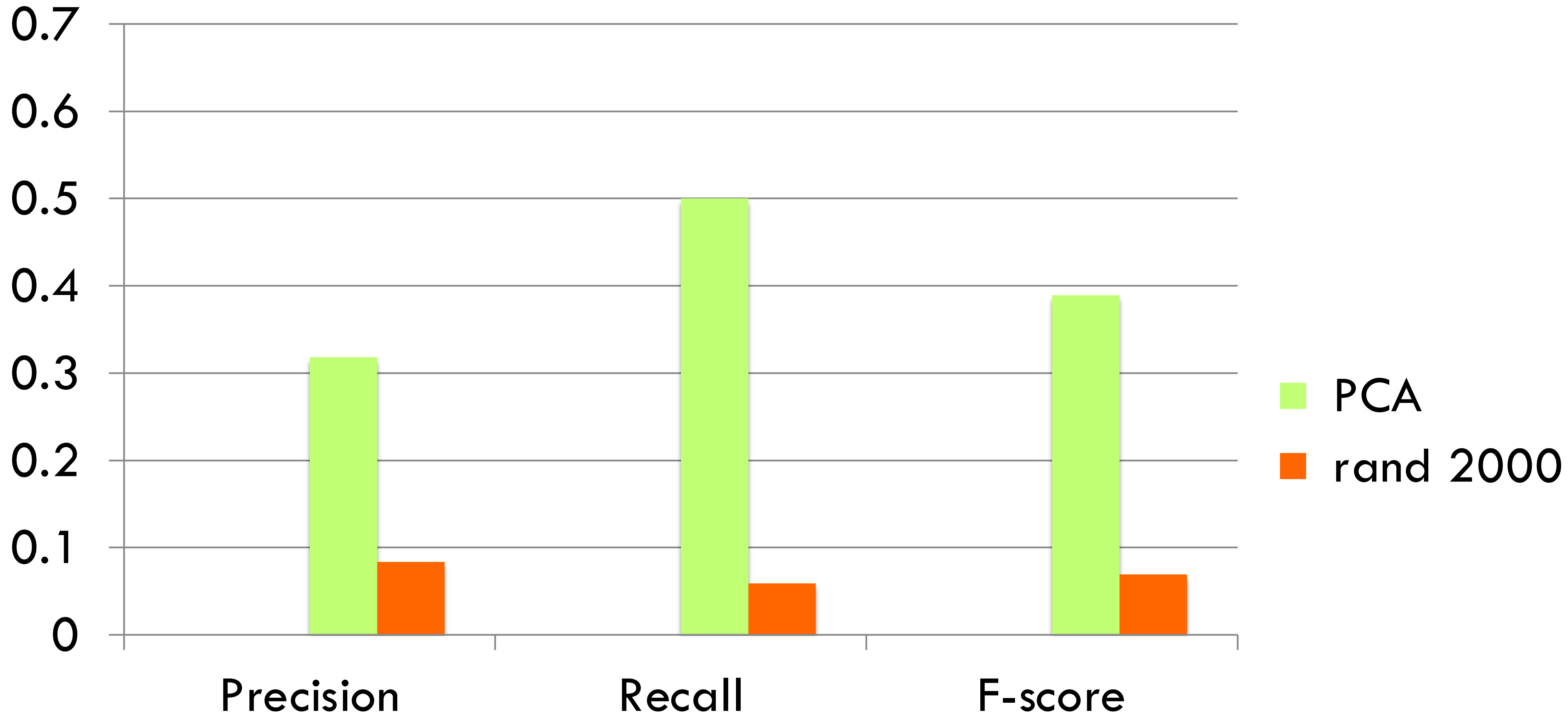
- **Dataset**
  - **Customer care calls**: data from a large network service provider in the US during **Aug. 2010~ July 2011**
  - **Ground-truth anomalies**: all anomalies reported by Call Centers, Network Operation Centers, and etc.
- **Metrics**
  - **Precision**: the fraction of claimed anomalies which are real anomalies $\frac{tp}{tp+fp}$
  - **Recall**: the fraction of real anomalies are claimed $\frac{tp}{tp+fn}$
  - **F-score**: the harmonic mean of precision and recall $\frac{2}{\frac{1}{precision}+\frac{1}{recall}}$

Evaluation –
Identifying Important Features



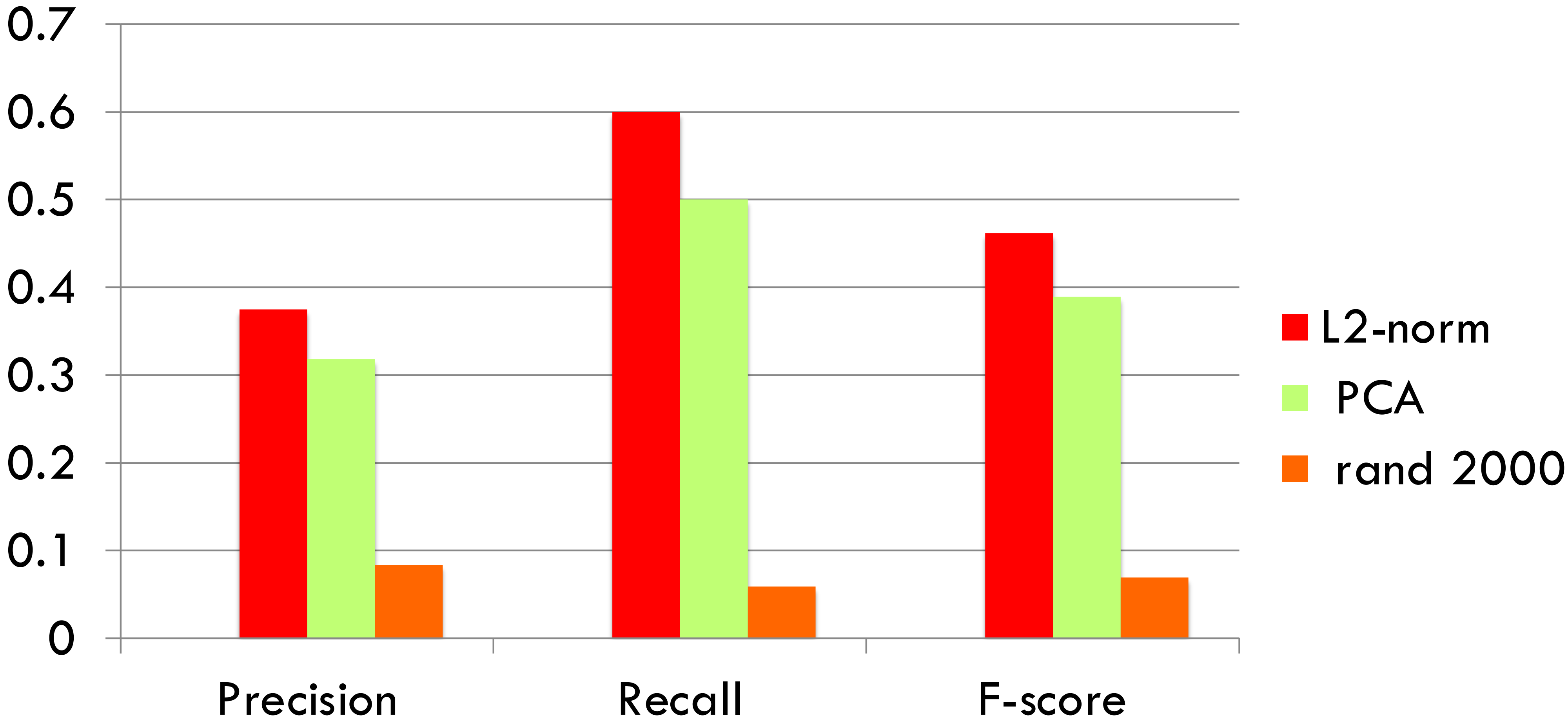
$$\frac{tp}{tp+fp} \qquad \frac{tp}{tp+fn} \qquad \frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

Evaluation – Identifying Important Features



$$\frac{tp}{tp+fp} \qquad \frac{tp}{tp+fn} \qquad \frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

Evaluation – Identifying Important Features

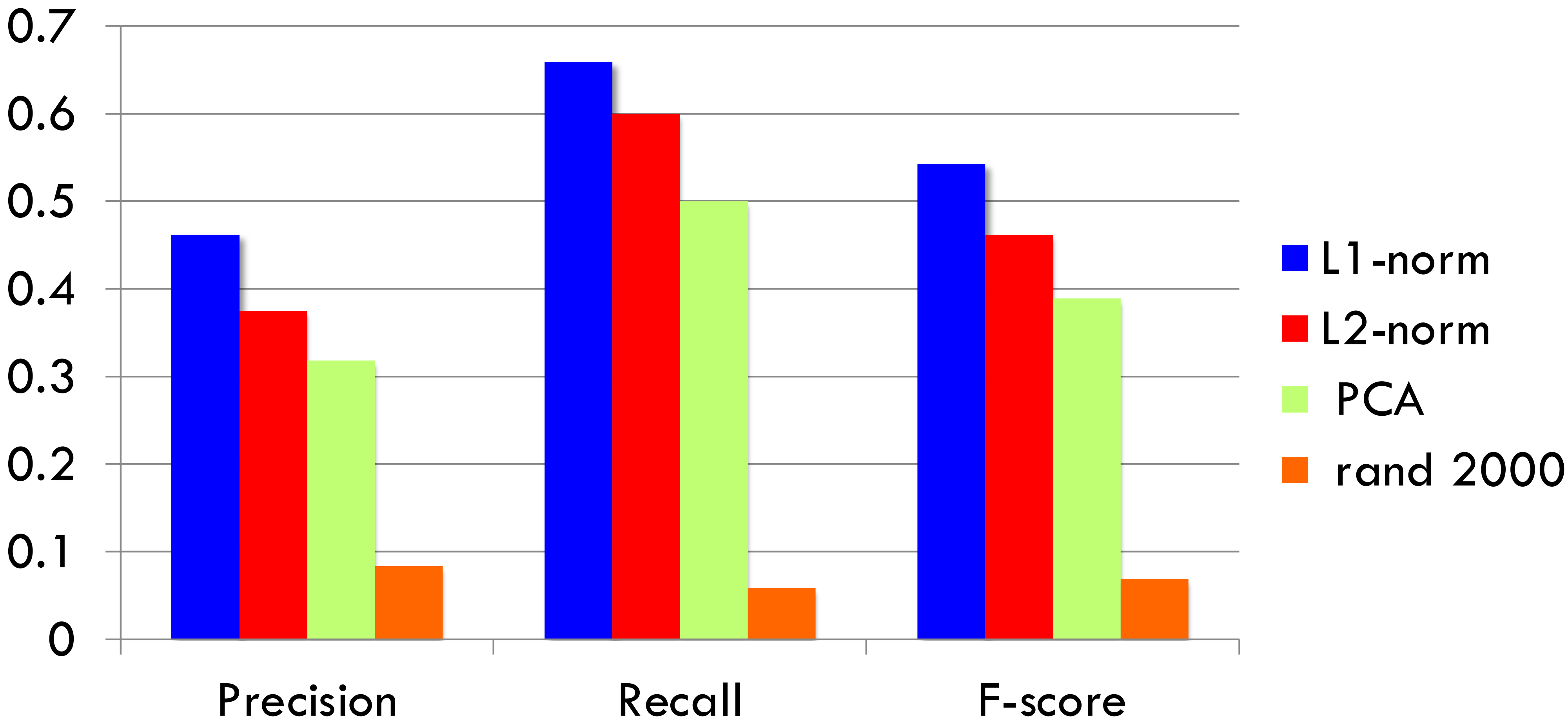


$$\frac{tp}{tp+fp}$$

$$\frac{tp}{tp+fn}$$

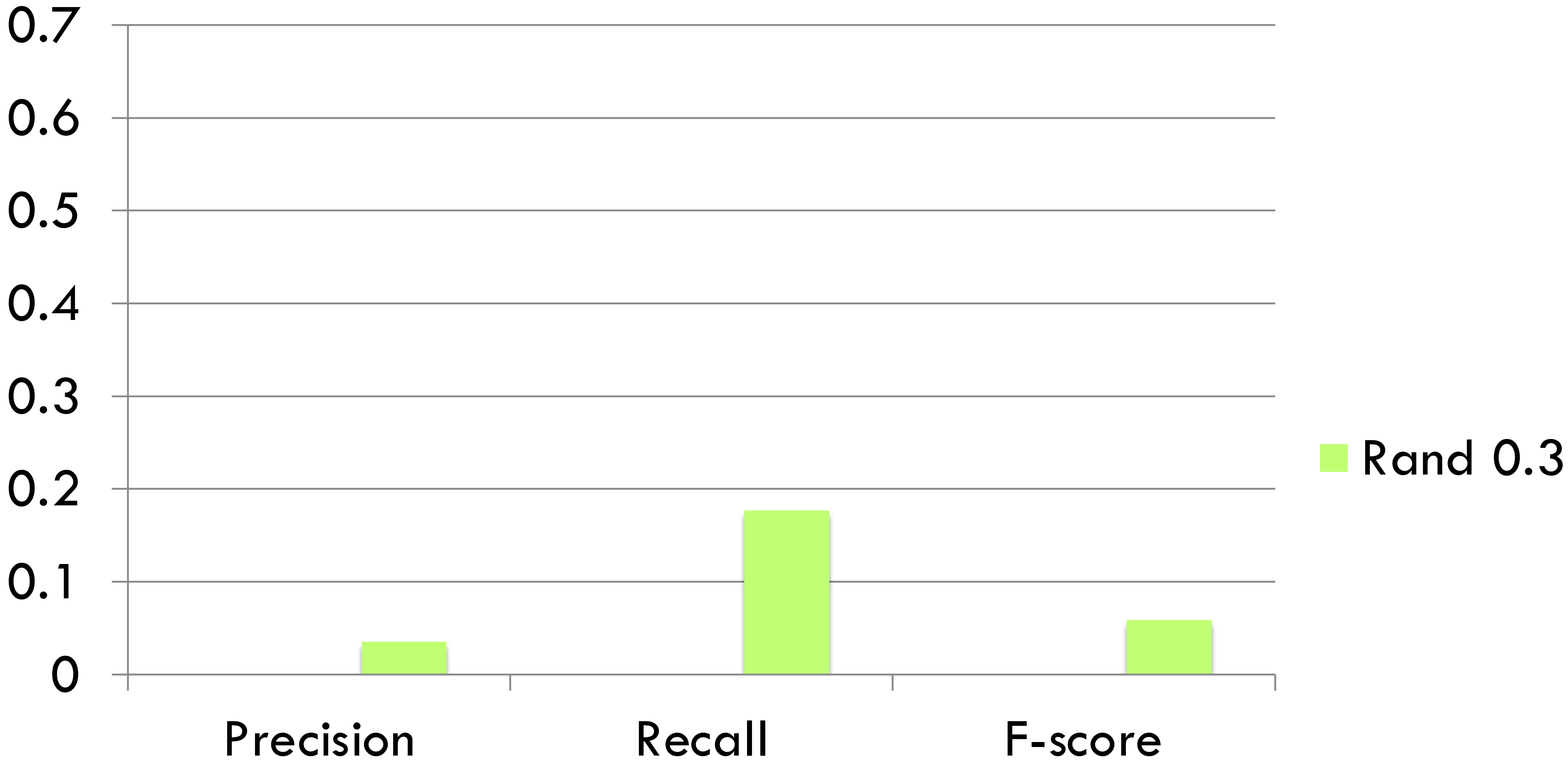
$$\frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

Evaluation – Identifying Important Features



L1-norm out-performs **rand 2000, PCA,** and **L2-norm** by **454%, 32%,** and **10%**

Evaluation – Regression

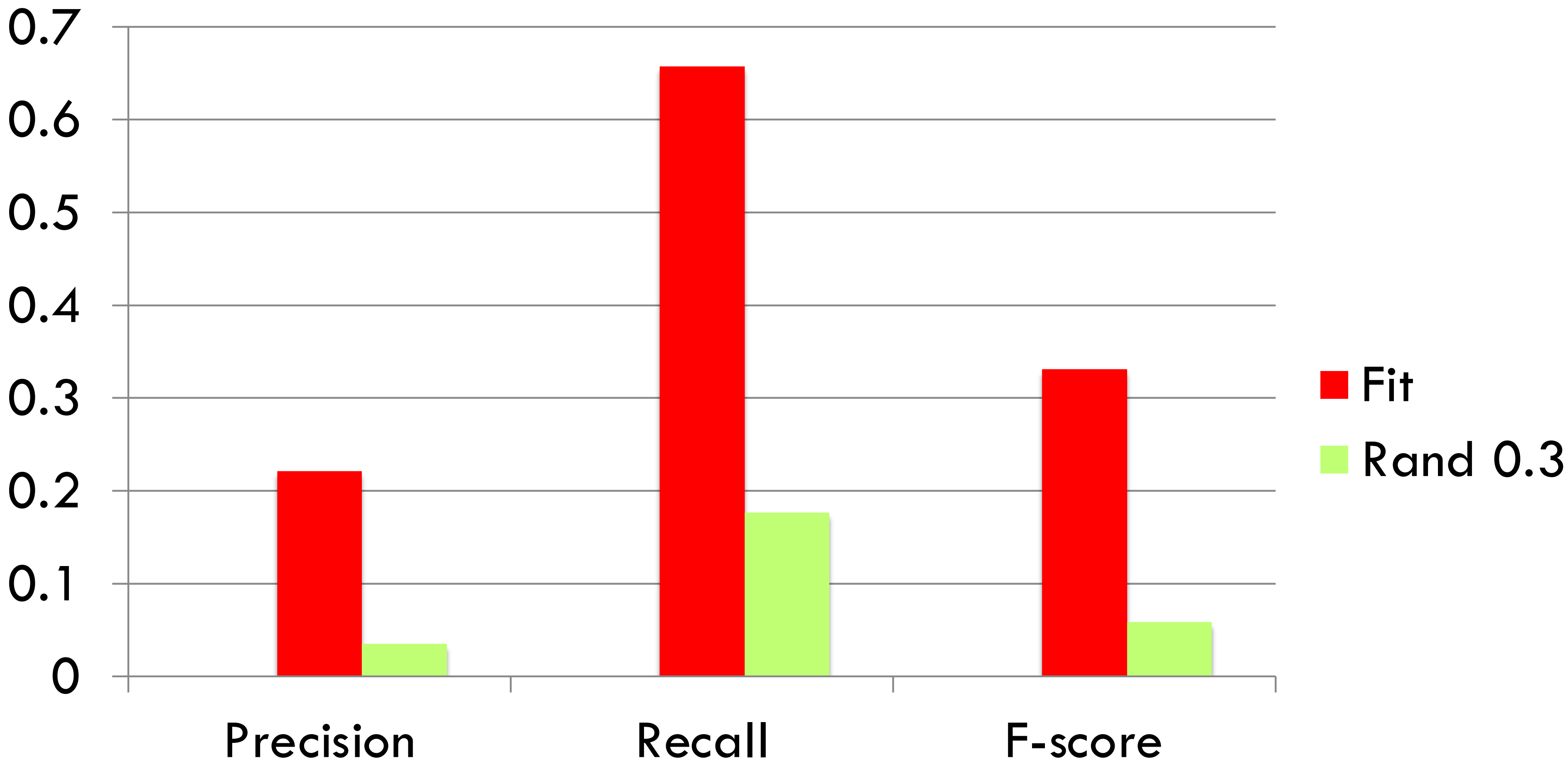


$$\frac{tp}{tp+fp}$$

$$\frac{tp}{tp+fn}$$

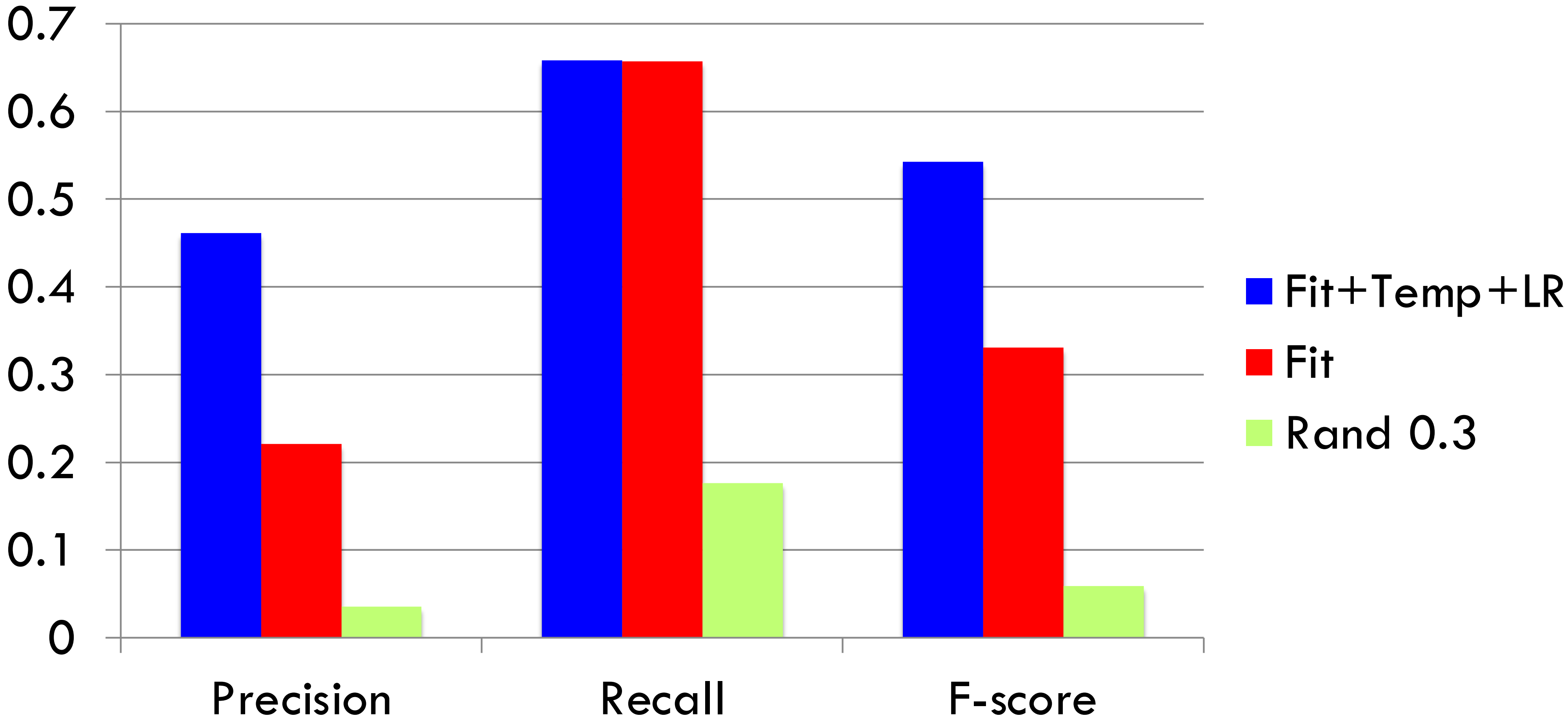
$$\frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

Evaluation – Regression



$$\frac{tp}{tp+fp} \qquad \frac{tp}{tp+fn} \qquad \frac{2}{\frac{1}{precision}+\frac{1}{recall}}$$

Evaluation – Regression



Fit+Temp+LR out-performs **Random** and **Fit** by **823%** and **64%**

Contributions

- Propose to use **customer care calls** as a complementary source to network metrics.
  - A direct measurement of QoE perceived by customers
- Develop a systematic method to automatically detect **events** using **customer care calls.**
  - Scale to a large number of features
  - Robust to the noise

Thank You!

yichao@cs.utexas.edu

Backup Slides


GLOBAL NETWORK
at&t
OPERATIONS
at&t
Console Master

Twitter

- Leverage **Twitter** as external data
  - Additional features
  - Interpreting detected anomalies
- Information from a tweet
  - Timestamp
  - Text
    - Term Frequency - Inverse Document Frequency (TF-IDF)
  - Hashtags: keyword of the topics
    - Used as features
    - e.g. #ATTFAIL
  - Location

Interpreting Anomaly - Location

Describing the Anomaly

- Examples
  - 3G network outage
    - **Location:** New York, NY
    - **Event Summary:** service, outage, nyc, calls, ny, morning, service
  - Outage due to an earthquake
    - **Location:** East Coast
    - **Event Summary:** #earthquake, working, wireless, service, nyc, apparently, new, york
  - Internet service outage
    - **Location:** Bay Area
    - **Event Summary:** serviceU, bay, outage, service, Internet, area, support, #fail

How to Select Parameters

- K-fold cross-validation
  - Partition the **training data** into **K equal size parts.**
  - In *round i*, use the *partition i* for **training** and the remaining *k-1 partitions* for **testing.**
    - The process is repeated k times.
  - Average k results as the evaluation of the selected value