“Is my Internet down?”:
Sifting through User–Affecting Outages with Google Trends

Ege Cem Kirci    Martin Vahlensieck    Laurent Vanbever

ETHzürich
What do you do when you have an Internet problem?
What do **you** do when you have an Internet problem? It turns out many people ask Google!
Can users’ aggregated search activity act as a “user indicator” for Internet outages?
Can users’ aggregated search activity act as a “user indicator” for Internet outages?
This is how the `<Internet outage>` search topic looks like in Texas.
This is how the <Internet outage> search topic looks like in Texas

Semantic clustering of related search queries, e.g., #InternetOutage
This is how the <Internet outage> search topic looks like in Texas

Semantic clustering of related search queries, e.g., #InternetOutage

Normalized popularity index:
The search term’s proportion among all the searches
This is how the <Internet outage> search topic looks like in Texas.
This is how the *Internet outage* search topic looks like in Texas.

![Search Interest Graph](image-url)
This is how the <Internet outage> search topic looks like in Texas.
This is how the <Internet outage> search topic looks like in **Texas**.
This is how the <Internet outage> search topic looks like in Texas.
This is how the <Internet outage> search topic looks like in Texas

Internet recovers after massive outage; Verizon acknowledges ‘issue’ as source

Networks are struggling in Texas amid historic winter storms
T-Mobile, Spectrum, and others are working to restore service

By Jay Peters | @jaypeters | Feb 15, 2021, 3:33pm EST
To analyze this data, we need to extract and evaluate it in a **systematic and quantifiable** way.
To analyze this data, we need to extract and evaluate it in a **systematic and quantifiable** way.

Construct time series...
To analyze this data, we need to extract and evaluate it in a **systematic** and **quantifiable** way.
To analyze this data, we need to extract and evaluate it in a **systematic and quantifiable** way.
Introducing…
SIFT
Construct time series…

Identify spikes…

Interpret the metadata…

Processing

Detection

Analysis
Google Trends provides **limited and partial data**

**Input**

- **Time range:** 01 Jan 2020 - 31 Dec 2021
- **Area:** California
- **Search term:** Internet outage
SIFT fetches \textit{partial} time frames covering the request.

Input

- Time range: 01 Jan 2020 - 31 Dec 2021
- Area: California
- Search term: Internet outage

Time frames

Time range

Location

Search term

Partial time frames
Google Trends returns **weekly** time frames

**Input**
- Time range: 01 Jan 2020 - 31 Dec 2021
- Area: California
- Search term: Internet outage

**Google Trends**
- 168 points
- Rising terms: spectrum internet outage, internet down, metro pcs outage, san jose power outage

**Pieceswise normalized**

**Individual time frames**
- One week of data

**Rising suggestions**
SIFT appends and calibrates time frames into continuous time series.

**Input**
- **Time range:** 01 Jan 2020 - 31 Dec 2021
- **Area:** California
- **Search term:** Internet outage

**Google Trends**
- **Time frames**
  - 0h, 6h, 12h, 18h, 0h
- **168 points**

**Rising terms**
- spectrum internet outage: 100
- internet down: 76
- metro pcs outage: 242
- san jose power outage: 90

**Processing Pipeline**
- **Reconstruction**
  - x1.9
- **Stitching**

**Rescaling**
- Weight:
  - spectrum internet outage: 100
  - internet down: 76
  - metro pcs outage: 242
  - san jose power outage: 90

---

![Diagram showing the processing pipeline with SIFT appends and calibrates time frames into continuous time series.](image-url)
Construct time series…

Identify spikes…

Interpret the metadata…

Processing

Detection

Analysis
We get inspired by **topographic prominence** in spike detection

### Input
- **Time range:** 01 Jan 2020 - 31 Dec 2021
- **Area:** California
- **Search term:** Internet outage

### Google Trends
- **Time frames:**
  - 0h
  - 6h
  - 12h
  - 18h
  - 0h
- **Rising terms**
  - spectrum internet outage: 100
  - internet down: 76
  - metro pcs outage: 242
  - san jose power outage: 90
- **168 points**

### Processing Pipeline
- **Detection:**
  - peak time
  - start time
- **Reconstruction:**
  - x1.9
  - reconstructed time series
Construct time series...

Processing

Identify spikes...

Detection

Interpret the metadata...

Analysis
Rising suggestions help us annotate spikes with context.
Construct time series…

Identify spikes…

Interpret the metadata…

Processing

Detection

Analysis
What are the **takeaways** and **insights**?
Spikes exhibit a skewed distribution in locations

Decreasing order: California, Texas, Florida, New York...
Half of the outages originate from 10 out of 50 states.
Spikes exhibit a skewed distribution in durations.
10% of the outages attract user interest for at least 3 hours.

Texas, 15 Feb. 2021-10h, 45h duration Winter storm
Spikes exhibit a skewed distribution in the area.
11% of the outages simultaneously affect at least 10 distinct states.
Internet outages may originate from physical or logical causes
Although spikes lasting at least five hours comprise the top 3.5% of all the spikes…

California heat wave, wildfires
... 73% of these spikes are annotated by

<Power outage> rising suggestion
Users’ **perceptions** can help with **understanding** the outages
Users’ perceptions can help with understanding the outages.

SIFT: a systematic and quantifiable approach
Detect, evaluate, and interpret user-affecting outages
Users’ perceptions can help with understanding the outages

SIFT: a systematic and quantifiable approach
Detect, evaluate, and interpret user-affecting outages

SIFT: first step in a long journey
Inter-disciplinary studies: security, politics, socio-economics
Decoupling the Internet from the power infrastructure
Users’ **perceptions** can help with understanding the outages

**SIFT: a systematic and quantifiable approach**
Detect, evaluate, and interpret user-affecting outages

**SIFT: first step in a long journey**
Inter-disciplinary studies: security, politics, socio-economics
Decoupling the Internet from the power infrastructure

**SIFT: an open-source community tool**
[https://github.com/nsg-ethz/SIFT](https://github.com/nsg-ethz/SIFT)
“Is my Internet down?”:
Sifting through User–Affecting Outages with Google Trends

Ege Cem Kirci    Martin Vahlensieck    Laurent Vanbever

ETHzürich