Why are blackouts in power grids heavy-tailed?

Bert Zwart

CWI and TU/e

August 31, 2020

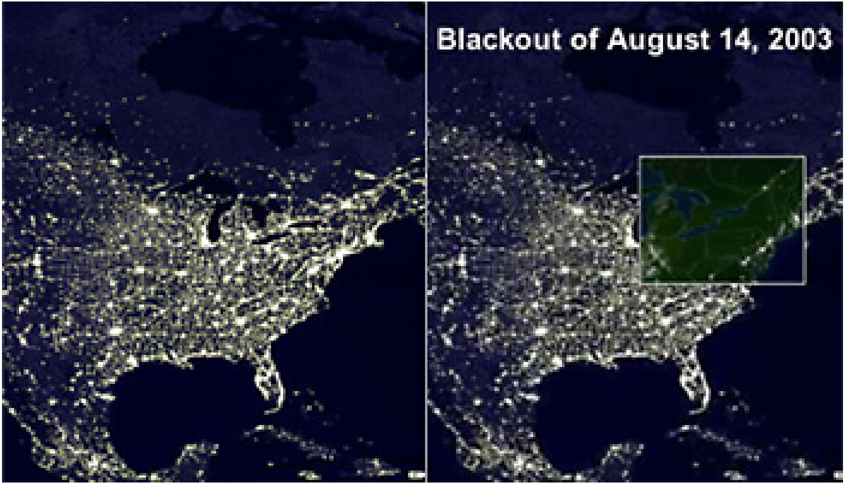
Joint work with Tommaso Nesti and Fiona Sloothaak

Centrum Wiskunde & Informatica

Overview

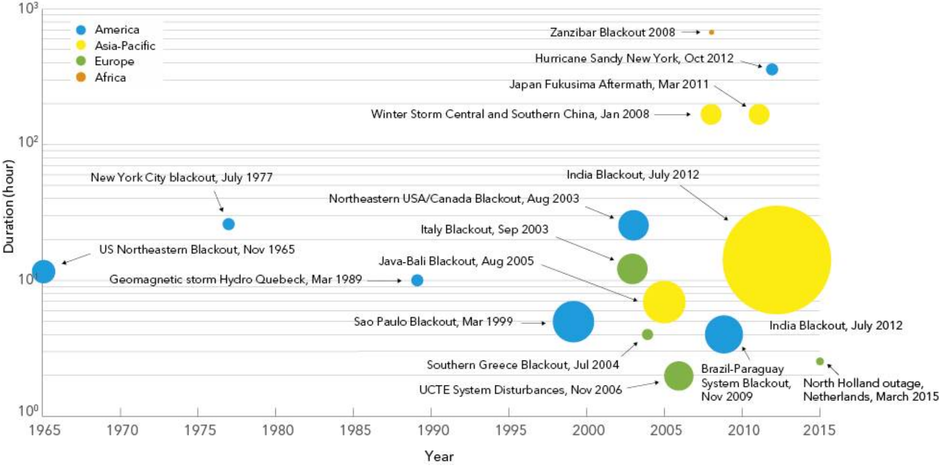
- Background: energy networks, blackouts, heavy tails
- Main insight of this talk
- Mathematical model and analysis
- Case study
- Conclusion and outlook

Blackouts in power grids



$80B$ of annual economic damage to US economy from blackouts

Blackouts in the past fifty years



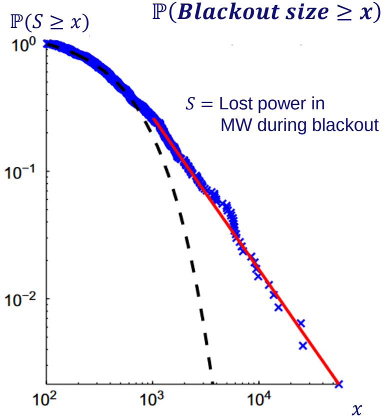
(source: dnv-gl)

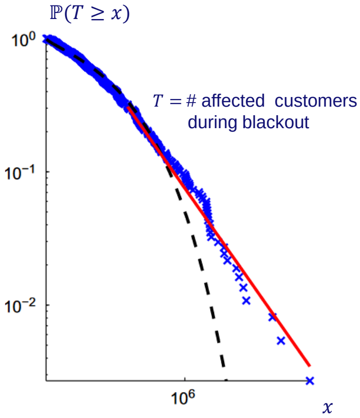
Can we mathematically understand blackouts?

- It took TNO 5 months to figure out the cause of the 2018 Schiphol power outage (the TenneT investigation still has not been concluded)
- "It is not complex, but complicated"
- "It is not possible to come up with a both interesting and useful result"
- To predict and detect anomalies, should we use simple black box methods from machine learning or sophisticated high-dimensional nonlinear models?
- At least one feature of blackouts is not complicated

Pareto laws in power grids (Hines 09)

$\mathbb{P}(\boldsymbol{Blackout\ size} \geq \boldsymbol{x}) = \boldsymbol{c}\, \boldsymbol{x}^{-\boldsymbol{\alpha}}$





WHY?

Rare Events depend on "Tail Behaviors"

Light-Tailed Distributions

- Extreme Values are Very Rare
- Normal, Exponential, etc

Heavy-Tailed Distributions

- Extreme Values are Frequent
- **Pareto Law**, **Weibull**, etc

Rare Events depend on "Tail Behaviors"

Light-Tailed Distributions

- Extreme Values are Very Rare
- Normal, Exponential, etc

Heavy-Tailed Distributions

- Extreme Values are Frequent
- **Pareto Law**, **Weibull**, etc

Heavy tails are not as well understood as light tails.

Rare Events depend on “Tail Behaviors”

Light-Tailed Distributions

- Extreme Values are Very Rare
- Normal, Exponential, etc

Heavy-Tailed Distributions

- Extreme Values are Frequent
- **Pareto Law**, **Weibull**, etc

Systemwide rare events

arise because

EVERYTHING goes wrong.

(Conspiracy Principle)

Heavy tails are not as well understood as light tails.

Rare Events depend on “Tail Behaviors”

Light-Tailed Distributions

- Extreme Values are Very Rare
- Normal, Exponential, etc

Systemwide rare events

arise because

EVERYTHING goes wrong.

(Conspiracy Principle)

Heavy-Tailed Distributions

- Extreme Values are Frequent
- **Pareto Law**, **Weibull**, etc

Systemwide rare events

arise because of

A FEW Catastrophes.

(Catastrophe Principle)

Heavy tails are not as well understood as light tails.

Examples and properties of heavy tails

As $x \to \infty$:

- Pareto tails (or power tails): $P(X > x) \approx x^{-\alpha} = e^{-\alpha \log x}$

- Lognormal tails: $P(X > x) \approx e^{-\alpha(\log x)^2}$

- Weibull tails: $P(X > x) \approx e^{-\alpha x^{\beta}}$, $\beta \in (0,1)$.

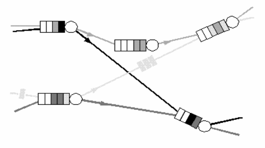
Key properties:

$$E[e^{\varepsilon X}] = \infty, \qquad \varepsilon > 0.$$

$$P(X_1 + \ldots + X_n > x) \sim P(\max_{i=1,\ldots,n} X_i > x).$$

Heavy Tails are Everywhere:

Computer systems

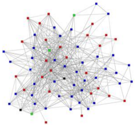


delays, files, …

Finance

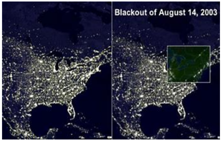
losses

Social networks



popularity, contagion

Energy Systems



blackouts

How do heavy tails occur?

Heavy tails can occur in many ways

- Exogenous factors (e.g. job sizes in queueing)
- Multiplication (e.g. gains or losses in finance)
- Preferential attachment (social, and other networks)
- (Self-organized) criticality
- ...

Existing work on blackouts, based on model simulation output data, show long-range correlations in outages, and attributes this to criticality. Earlier work suggests the usage of critical branching processes.

Our contribution: a different explanation

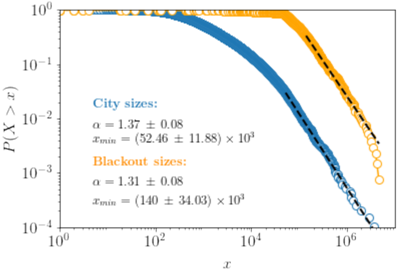
Let C be the size of a city, in terms of number of people, and let T be the size of a blackout, in terms of number of customers affected
Both have statistically significant, almost identical power law for US:

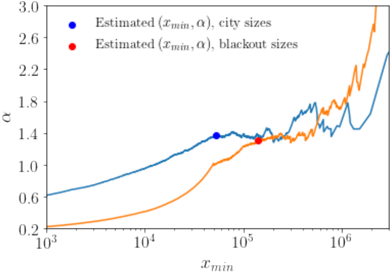
$$P(C > x) \approx x^{-1.37} \qquad\qquad P(T > x) \approx x^{-1.31}.$$

German city sizes: power law with index 1.28

2015 rank	City	State	2015 Estimate	2011 Census	Change	2015 land area	2015 po
1	**Berlin**	Berlin	3,520,031	3,292,365	+6.91%	891.68 km² 344.28 sq mi	
2	**Hamburg**	Hamburg	1,787,408	1,706,696	+4.73%	755.3 km² 291.6 sq mi	
3	**Munich** (München)	Bavaria	1,450,381	1,348,335	+7.57%	310.7 km² 120.0 sq mi	
4	*Cologne* (Köln)	North Rhine-Westphalia	1,060,582	1,005,775	+5.45%	405.02 km² 156.38 sq mi	
5	*Frankfurt am Main*	Hesse	732,688	667,925	+9.70%	248.31 km² 95.87 sq mi	

log-log plots and Hill plots





US city size data (2000 census) and US outage data (NERC, 2002-2018). Cutoff chosen according to the PLFIT method of Clauset et. al (2009).

Intuition

1. Initially the network is fully connected and supply equals demand
2. A failure occurs. Energy is rerouted according to laws of physics, resulting in additional overloads/failures. A cascade occurs.
3. At some point the network stops being connected. At least one of the network components has a shortage.
4. Demand in each network component is proportional to sum of the cities in that component
5. Since city sizes are heavy-tailed, the sum roughly equals the maximum. A heavy-tailed number of consumers (in a big city) will have a shortage.

To make this rigorous, we need to show that the **mismatch** between supply and demand after a network cut is heavy-tailed.

Mathematical model

- Graph with heavy-tailed sinks:
  Demand at node i: X_i, with $P(X_i > x) \sim cx^{-\alpha}$.
  $\mathbf{X} = (X_1, \ldots, X_n)$.

- To model electricity, we use the DC load flow model. Network topology and reactances are all encoded in the load-flow matrix $\mathbf{V}$, supply vector equals $\mathbf{g}$, leading to power flows $\mathbf{V}(\mathbf{g} - \mathbf{X})$.

- We consider three stages in our model:
  - Planning
  - Operation
  - Emergency

Model: operational stage (DC-OPF)

Generation g_i in each node i is computed by solving

$$\min \frac{1}{2} \sum_{i=1}^{n} g_i^2$$

$$\sum_i g_i = \sum_i X_i$$

$$-\bar{\mathbf{f}} \le \mathbf{V}(\mathbf{g} - \mathbf{X}) \le \bar{\mathbf{f}}.$$

This determines the network flows $\mathbf{F} = \mathbf{V}(\mathbf{g} - \mathbf{X})$ which play a role in the cascade.

We determine the line limits $\bar{\mathbf{f}}$ in a planning problem.

Model: planning stage

Given n cities with random sizes $X_1, \ldots, X_n$ and given a network topology, we determine line limits $\bar{\mathbf{f}}$ by solving an unconstrained OPF problem:

$$\min \frac{1}{2} \sum_{i=1}^{n} g_i^2$$

subject to the balance constraint

$$\sum_i g_i = \sum_i X_i.$$

The planning problem has solution $g_i = \bar{X}_n$ for $i = 1, \ldots, n$, with $\bar{X}_n = (1/n) \sum_{i=1}^{n} X_i$ the average city size. We now let $\lambda \in (0,1)$ and set

$$\bar{\mathbf{f}} = \lambda \mathbf{V}(\bar{X}_n \mathbf{e} - \mathbf{X}) = -\lambda \mathbf{V}\mathbf{X},$$

This vector will be used in the operational stage

Model: operational stage (DC-OPF)

$$\min \frac{1}{2}\sum_{i=1}^{n} g_i^2$$

$$\sum_i g_i = \sum_i X_i$$

$$-\bar{\mathbf{f}} \leq \mathbf{V}(\mathbf{g}-\mathbf{X}) \leq \bar{\mathbf{f}}$$

with

$$\bar{\mathbf{f}} = \lambda \mathbf{V}(\bar{X}_n \mathbf{e} - \mathbf{X}) = -\lambda \mathbf{V}\mathbf{X}.$$

This leads to actual line flows $\mathbf{F} = \mathbf{V}(\mathbf{g}-\mathbf{X})$.

Model: emergency stage

Given: line flows $\mathbf{F} = \mathbf{V}(\mathbf{g} - \mathbf{X})$

- Start with one random line outage.
- Recompute power flows.
- Additional lines fail if new line flow exceeds $\bar{f}_i/\lambda$.
- If islands occur, load or generation is shed proportionally.
- T: size of total load shed once cascade is over.

Main result

Let X be a generic city size, with $P(X > x) \sim C_X x^{-\alpha}$.
Note that T is the blackout size [in terms of number of customers affected]

$$P(T > x) \sim C_T x^{-\alpha}, \qquad x \to \infty, \tag{1}$$

$$C_T = C_X n \sum_{j=1}^{n} P(|A_1| = j)(1 - j\lambda/n)^{\alpha}. \tag{2}$$

A_1 denotes the (random) set of nodes making up the island with the largest city in the network.

Proof idea: heavy-tailed large deviations theory allows us to consider the case of a single big city, and many small cities, reducing the analysis of the cascade to a single-sink network.

Rigorous proof for almost all λ [when C_T is continous in λ].

Numerical studies

Our result holds up against several simulation studies

- Generation constraints
- Extending DC to AC
- No heavy tailed blackout size if city sizes are uniformly distributed
- IEEE test networks
- Synthetic scalable networks, tailored to power grids [Wang, Scaglione, and co-authors]

Critical assumption: frozen vs random city sizes [quenched vs annealed]

SciGRID case study - Impact of λ

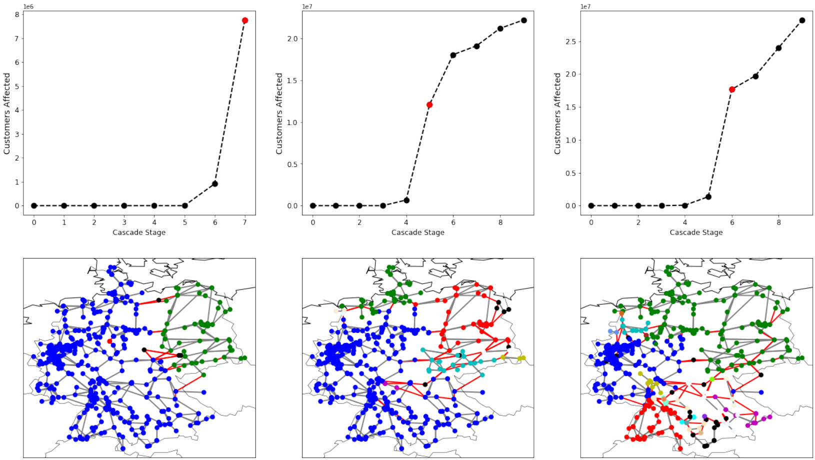


Figure: Dissection of biggest blackout for loading factors $\lambda = 0.7$ (left panels), $\lambda = 0.8$ (middle) and $\lambda = 0.9$ (right) in terms of the cumulative number of affected customers at each consecutive stage as displayed in the top charts with the biggest jump colored red.

Concluding remarks

- Main insight
  - Network upgrades only make the pre-factor in front of power law smaller. This provides limited effect in preventing large blackouts.
  - Duration of blackout is light-tailed, and seems independent of size.
  - It therefore seems to make more sense to invest in making cities more resilient, rather than to invest in network upgrades.
- City sizes are not (very) random. Simulation studies suggests Pareto laws occur in frozen networks of $\geq 10^4$ cities [this includes North America]. For smaller networks (like Germany), the principle of one big jump still seems to hold.
- New class of network models, of which topologies may not be scale free, but scale-free phenomena occur due to city sizes. Currently looking at other transportation networks.