Package 'bursts'

July 22, 2025

Type Package

Title Markov Model for Bursty Behavior in Streams

Version 1.0-2

Date 2022-07-14

Author Jeff Binder [aut, cre]

Maintainer Jeff Binder <extruded@gmail.com>

Imports graphics

Description An implementation of Jon Kleinberg's burst detection algorithm (Kleinberg (2003) <doi:10.1023/A:1024940629314>). Uses an infinite Markov model to detect periods of increased activity in a series of discrete events with known times, and provides a simple visualization of the results.

License MIT + file LICENSE

Copyright (c) 2012 NYU School of Medicine; (c) 2012-22 Jeff Binder

NeedsCompilation no

Repository CRAN

Date/Publication 2022-07-19 11:30:02 UTC

Contents

---

bursts-package   *Markov Model for Bursty Behavior in Streams*

Description

An implementation of Jon Kleinberg's burst detection algorithm. Uses an infinite Markov model to detect periods of increased activity in a series of discrete events with known times, and provides a simple visualization of the results.

Details

Package:	bursts
Type:	Package
Version:	1.0-2
Date:	2022-07-14
License:	MIT

The function `kleinberg` performs the analysis, returing a data frame containing a list of all of the 'bursts' and their intensities. The function `plot.bursts` can be used to display a simple visualization of the hierarchical burst structure.

Author(s)

Jeff Binder

Maintainer: Jeff Binder <extruded@gmail.com>

References

Kleinberg, J. (2003). "Bursty and Hierarchical Structure in Streams." Data Mining and Knowledge Discovery **7**: 373-397. <doi:10.1023/A:1024940629314>

http://www.cs.cornell.edu/home/kleinber/kdd02.html

---

kleinberg   *Detects periods of increased activity in a stream of events*

Description

This function attempts to identify 'bursts' of activity in a series of discrete events that take place at known times, based on an infinite hidden Markov model. An optimal state sequence is computed that balances the total trasition cost against the probability of the observed event timing.

Usage

```
kleinberg(offsets, s = 2, gamma = 1)
```

Arguments

`offsets`	a vector containing the time offsets (numeric or Date) of the relevant events.
`s`	the base of the exponent used to determine event frequencies in a given state.
`gamma`	a coefficient that modifies the cost of a transition to a higher state.

Details

The model is a hidden Markov process in which, after each event, the state of the system probablistically determines how much time will pass until the next event occurs. While the system is in state `i`, the gaps between events are assumed to be drawn from an exponential distribution with expected value proportional to `s ** -i`. The value of `s` may be modified; higher values increase the strictness of the algorithm's criterion for how dramatic an increase of activity has to be to be considered a 'burst.'

The cost of a state change is proportional to the increase in state number; this proportion can be modified by setting the parameter `gamma`. Higher values mean roughly that bursts must be sustained over longer periods of time in order for the algorithm to recognize them.

Note that the algorithm will not work if there are two events that occur at the same time.

Value

Returns a data frame of class 'bursts.' Each row represents a (maximal) interval of time in which the system was at or above a given level of activity. The first row always indicates a period at level 1+ lasting from the time of the first event to the time of the last; subsequent rows always indicate levels greater than 1, and thus represent 'bursts.'

The 'start' time of a burst is defined as the time of the event that precedes the state change. This is so that the time interval of a burst contains the first unusually short gap between events, rather than beginning immediately after it.

As per Kleinberg, if the system goes through a burst that increases the level by more than 1, a separate row is created for each level that it goes through.

Author(s)

Jeff Binder

References

Kleinberg, J. (2003). "Bursty and Hierarchical Structure in Streams." Data Mining and Knowledge Discovery **7**: 373-397. <doi:10.1023/A:1024940629314>

http://www.cs.cornell.edu/home/kleinber/kdd02.html

See Also

plot.bursts

Examples

```
offsets <- c(seq(0, 400, 100), seq(410, 450, 5), seq(451, 470, 2),
             seq(480, 600, 5), 700, seq(710, 800, 5), 900, 1000)
bursts <- kleinberg(offsets)
plot.bursts(bursts)
```

---

plot.bursts	*Plots a burst model*

Description

This method produces a simple plot of the burst structure produced by the `kleinberg` function.

Usage

```
## S3 method for class 'bursts'
plot(x, ...)
```

Arguments

x   a data frame produced by the `kleinberg` function.

...   other parameters are passed along to `plot`.

Details

Horizontal bars represent the 'bursts,' with their vertical position indicating the level (i.e. intensity) and the x axis representing time. Bars appear above others to represent the hierarchical structure that emerges when long bursts contain sub-bursts of even higher activity.

Value

No return value; called to produce a plot.

Author(s)

Jeff Binder

References

J. Kleinberg. Bursty and Hierarchical Structure in Streams. Proc. 8th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2002.

http://www.cs.cornell.edu/home/kleinber/kdd02.html

See Also

kleinberg

Examples

```
offsets <- c(seq(0, 400, 100), seq(410, 450, 5), seq(451, 470, 2),
             seq(480, 600, 5), 700, seq(710, 800, 5), 900, 1000)
bursts <- kleinberg(offsets)
plot(bursts)
```

Index