I wrote some notes hopefully the idea is right I'm still not done reading the paper as I want to understand the algorithms used more in depth but here is my sense of the paper so far:

ok so my understanding is that in order to find censorship patterns we identify similar patterns on groups of networks that block the same website in the same way

(page 6 formula)
we can use the formula for a censorship event (Cθ) is:

Cθ=(Rθ,{dθ1,dθ2,...,dθn})
where
Rθ = A blocking rule where, when and how of the censorship event.
{dθ1,dθ2,...,dθn} = A list of websites that were affected by the blocking rule
here we can say
censorship event Cθ =
(Rθ,{dθ1,dθ2,...,dθn})
A blocking rule that affects a list of websites

(page 6 equation on "Discover simultaneous blocking through decision trees")

collects censorship data per website
website (di) country (c) over time period (t1,t2,... tm)
each column represent a vantage point or a network that we monitor censorship from
each row represents a point in time (t1,t2,... tm)
each cell (res(j,k)) is the censorship status for website (di) at vantage point vp and time tk

(page 6 equation 2 on "Aggregation on IP organizations")

Decision trees is a model in which makes decision using "yes/no" questions at each step so by using sklearn gini impurity decision tree to build a decision tree for each website (d)
decision tree (DT) represents how website (d) was blocked over a certain period of time and location

how to check if websites are blocked in the same way
by using one websites decision tree to classify another website censorship behavior and finding if two websites are blocked in the similar way
using equation 4 (page 7) takes the best prediction accuracy from either decision tree using max function and use a distance measure and as the lower the distance measure becomes the more similar the censorship behavior becomes

the decision trees are used to help detect censorship patterns by splitting data based on location and time

DBSCan clustering used to group website that show similar censorship patterns and removes uncensored websites
in algorithm 1 (page 7) we can see how clustering of decision trees take place the goal is to group websites based on how they are censored
first identifies 'innocent websites' called 'innocent trees' and will be removed from the dataset so only blocked websites remain
second the clustering process groups websites that show similar censorship patterns if websites have the same blocking pattern they are put into a cluster
third extracts event information