System Evolution Analytics: Evolution and Change Pattern Mining of Inter-Connected Entities

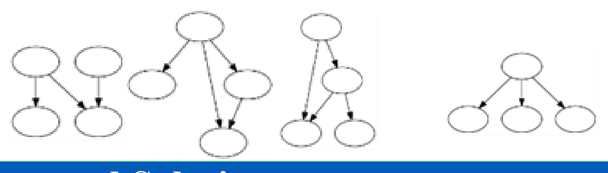
Animesh Chaturvedi, Aruna Tiwari
Indian Institute of Technology, Indore
animesh.chaturvedi88@gmail.com, artiwari@iiti.ac.in

Basic Keywords

Evolving System evolves over time

Data mining is a process of extracting information

- **Evolution and Change mining**
- **Network Rule Mining**, as A => B, where 'A' and B' are set of antecedent and consequent
- **Network Subgraph Mining**



Proposed Problem and Solution

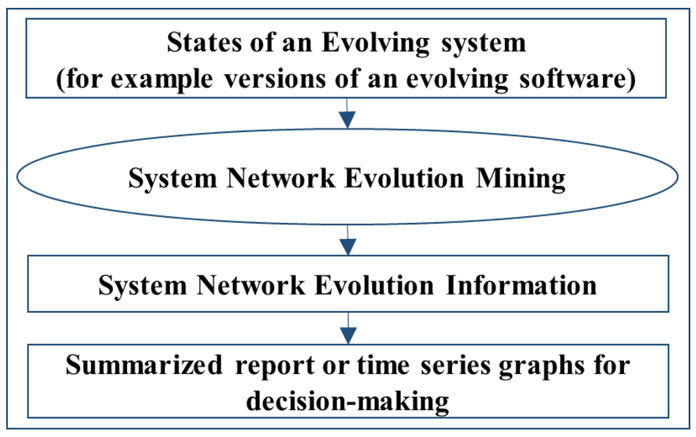
Problem There are following two challenges in evolving system mining:

A. Evolving systems are complex, thus requires pre-processing by domain expert.

B. Mining algorithm for multiple states of an evolving system, thus algorithm needs customization.

Solution

- Preprocess a state series of an evolving system to make a series of evolving networks EN = {EN_1, EN_2… EN_N}
- Retrieve mining information about each evolving network using well-known network rule mining and network subgraph (motif) mining.
- Merge the retrieved information to make an aggregated evolution information, which describe mining evolution information about the state series of an evolving system.
- As a summarized report, create time series graphs, charts, or tables. Then, use the report to analyze the evolution of an evolving



Algorithm ***System_Evolution_Mining***(*EvolvingSystem*)

Initialize *i* ∈ integer 1 to N

For each retrieved N states (S_1, S_2… S_N) of the *EvolvingSystem* and stored in a directory *DirRepository*

{ *evolvingNetworks* = ***Preprocess***(*DirRepository*)
mining_info = ***NEM***(*evolvingNetworks*)
network_evolution_infos = ***merge***(*mining_info*) } NEM
}

report_time_series = ***info_present***(*network_evolution_info*)

System Evolution Analytics: Pattern Mining of Inter-Connected Entities

State Series (SS) is a collection of states, such that SS = {S_1, S_2 … S_N} at various time points {t_1, t_2, t_3… t_N}.

- The SS can be pre-processed to make an evolving network series as EN = {EN_1, EN_2… EN_N} corresponding to the {S_1, S_2 … S_N}.
- Exist a relationship between (S_i, EN_i, t_i), system state S_i and evolving network EN_i at the i^{th} time point t_i, where 'i' varies from 1 to N.
- An example of an evolving state series is software version series.

Network Evolution Mining: "mining of an evolving system that retrieves network evolution and change(s) information about evolving entity-sets in a state series of an evolving system network"

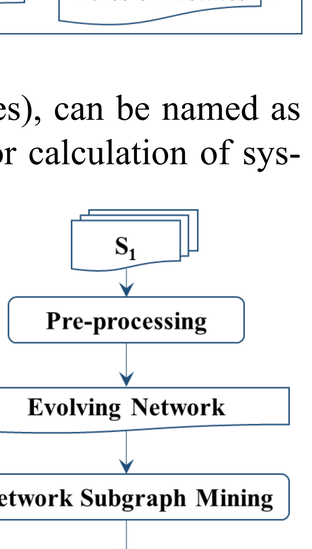
Network Evolution Rule (NER): "set of evolution rule that contains the evolution and change(s) information about if evolution and change(s) occur to a source entity-set, then evolution or change(s) is likely to occur in its target entity-sets for a system network over a state series" e.g. {e_1, e_3, e_4}=>{e_2, e_5}.

Network Evolution Subgraph (NES): "set of subgraph containing the evolution and change(s) information about evolving entity-sets in evolving system network" Subgraph information (H_j, $freq_j$), Subgraph key {H_1, H_2… H_M} Subgraph frequency {$freq_1$, $freq_2$… $freq_M$}

System Evolution Analytics: Pattern Mining of Inter-Connected Entities

A. Network Evolution Rule Mining

- Pre-process a series of evolving networks to make system network databases with two ordered sets of source and target entities.
- For each state, use sequential rule mining to generate sequential rules for source and target entities; according to thresholds: minimum support and minimum confidence . The sequence rules are the network rules, such that each rule contains set of source entities followed by set of target entities.
- Merge all the network rules of different states to make a collection of NERs.
- Retrieve the NERs that are occurring in most number of states. These frequent NERs (over the states), can be named as stable (or persistent) rules. These stable rules are helpful for calculation of system stability over states.



B. Network Evolution Subgraph Mining

- Pre-process to make a series of evolving networks compatible to the network-motif mining algorithm.
- For each state, use network subgraph mining and generate subgraph information that contains combination of subgraph key and subgraph frequency.
- Merge all the network subgraph information of different states to make a collection of NESs. Each NES has its aggregate frequency of occurrence in the series of evolving networks.

C. System Evolution Analytics Tool

We built two tools, which follow system evolution analytics model to render reports with NERs and NESs.

Results & Discussions

We applied our tools on six evolving systems: Hadoop HDFS-Core, Bible Translation, Multi-sport Events, Retail market system, Positive sentiment of movie genres, and Negative sentiment of movie genres.

We made evolving networks for each evolving system. Then, we evaluated Network Evolution Rules and Network Evolution Subgraphs for the six evolving systems. Such aggregated system information helps in decision-making.

Our Relevant Publications

A. Chaturvedi and A. Tiwari. "System Evolution Analytics: Evolution and Change Pattern Mining of Inter-Connected Entities". ***IEEE International Conference on Systems, Man, and Cybernetics (SMC)*** **2018.**

In addition to this paper, we also has two other accepted publications

- A. Chaturvedi and A. Tiwari. "System Evolution Analytics: Deep Evolution and Change Learning of Inter-Connected Entities". *IEEE International Conference on Systems, Man, and Cybernetics (SMC)* 2018.
- A. Chaturvedi and A. Tiwari. "System Network Complexity: Network Evolution Subgraphs of System State series". *IEEE Transactions on Emerging Topics in Computational Intelligence* (in press).