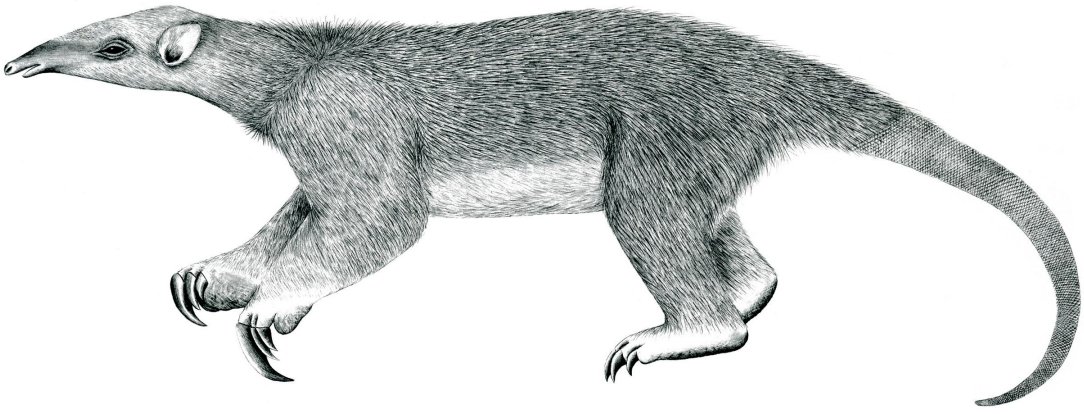
Tranalyzer2

centrality

Centrality



Tranalyzer Development Team

Contents

1 centrality

1.1 Description

This plugin produces a connection matrix from pcap files and calculates the centrality of each IP address. The centrality is defined by the corresponding entry in the eigenvector to the largest eigenvalue of the adjacency matrix (Section 1.3).

1.2 Dependencies

1.2.1 Other Plugins

This plugin requires the `basicStats` plugin if `CENTRALITY_MATRIXENTRIES` ≥ 2.

1.3 Eigenvector centrality

Let $G = (V,E)$ be a graph, where V is a set of vertices v_i and E a set of edges connecting those vertices. We define the *adjacency matrix* $A \in \mathbb{N}_0^{n \times n}$ matrix labelled by the vertices of G, where the entry $A_{i,j} = 1$, if there is a connecting edge between v_i and v_j.
Eigenvalues are defined as roots of the characteristic polynomial $p(\lambda) = det(A - \lambda * I)$, where I is the identity matrix and A the adjacency matrix.
Each eigenvalue $\lambda \in \mathbb{C}$ has a corresponding eigenvector $\mathbf{x} \in \mathbb{C}^{n \times 1}$ which satisfies the equation $A * \mathbf{x} = \lambda * \mathbf{x}$.
The centrality of vertex i is defined by: $centrality(v_i) := \mathbf{x}_i$, where $\mathbf{x}$ is the eigenvector to $\lambda_{max} := max(\lambda \in \mathbb{R})$. All values of $\mathbf{x}$ are strictly positive.
This Plugin uses a simple version of the power iteration to calculate this eigenvector:
$\mathbf{x}_{n+1} = \frac{A * \mathbf{x}_n}{\|A * \mathbf{x}_n\|}$. For $n \to \infty$, $\mathbf{x}_n$ converges to $\mathbf{x}$.

1.4 Configuration Flags

The following flags can be used to control the output of the plugin:

Variable	Default	Description
`CENTRALITY_MATRIXENTRIES`	1	0: $A_{i,j} \in \{0,1\}$ 1: $A_{i,j}$ = number of flows between v_i and v_j 2: $A_{i,j}$ = number of bytes sent between v_i and v_j 3: $A_{i,j}$ = bytes asymmetry between v_i and v_j 4: $A_{i,j}$ = number of packets sent between v_i and v_j 5: $A_{i,j}$ = packets asymmetry between v_i and v_j
`CENTRALITY_TIME_CALC`	1	0: centrality calculated at application termination, $\mathbb{N} \ni n \neq 0$: centrality calculated every n seconds (dump time)
`CENTRALITY_IP_FORMAT`	1	0: IPs as unsigned integer 1: IPs as hex 2: IPs in compressed format, e.g., 1.2.3.4
`CENTRALITY_MATRIXFILE`	0	1: Write a file with triplet matrix
`CENTRALITY_TRAVIZ`	0	1: Traviz output mode

In addition, `CENTRALITY_SUFFIX` and `MATRIX_SUFFIX` can be used to control the suffix for the centrality and matrix output files (`"_centrality.txt"` and `"_matrix.txt"` respectively).

1.5 Flow File Output

There is no output to the flow file.

1.6 Centrality File Output

If `CENTRALITY_TRAVIZ=1`, then the first row is as follows:

```
% number_of_rows time IP centrality
```

Where `number_of_rows` is the actual number of entries in the file and `time`, `IP` and `centrality` represent the name of the various columns.

Column	Type	Description	Flags
1	U32/U64	Time (seconds)	`CENTRALITY_TRAVIZ=1`
2	IP4/U32	IPv4 address	
3	D	Centrality	

1.7 Matrix File Output

This File shows the adjacency matrix of your network in triplet matrix format. You can restore the original matrix A by assigning $A_{row.column} = value$ and $A_{i.j} = 0$ for all unassigned indices.
The first row and column will be full of ones as this is the maximum centrality host used for normalization.

Column	Type	Description
1	I	Matrix row
2	I	Matrix column
3	I	value of $A_{row,column}$

1.8 TODO

- Add support for IPv6
- Analyzing centralities in known networks
- Classification of networks by their centrality