

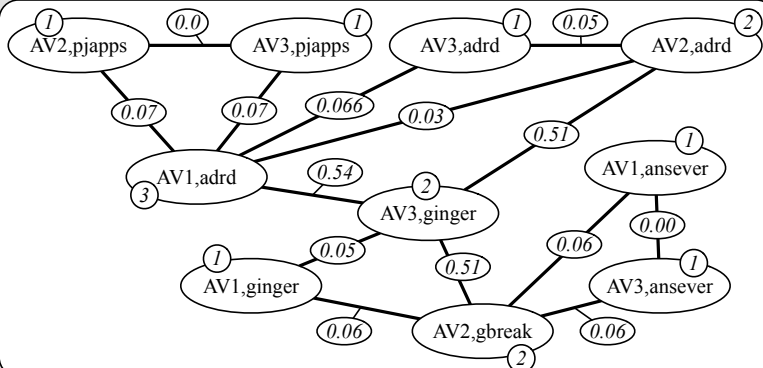
Multiple family names per samples

	AV1	AV2	AV3
app1	adrd	pjapps	pjapps
app2	adrd	adrd	adrd
app3	ginger	gbreak	ginger
app4	anserver	gbreak	anserver
app5	adrd	adrd	ginger

Clustering of Family Names

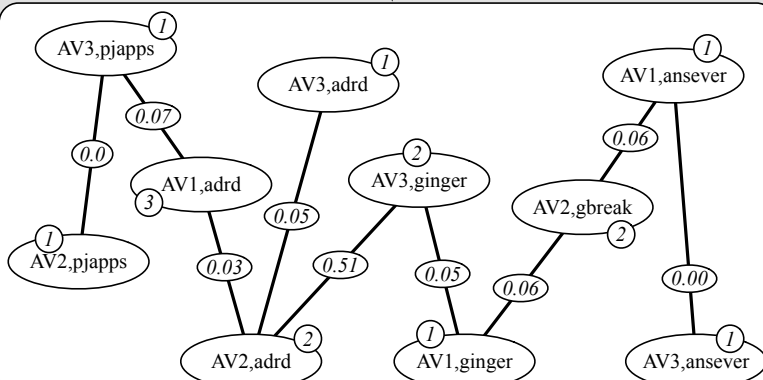
**Transform
to
Weighted
Graph**

Sample Set
Similarity Metric

Graph Structure
(associations)

Transform to Tree

TRIM Step

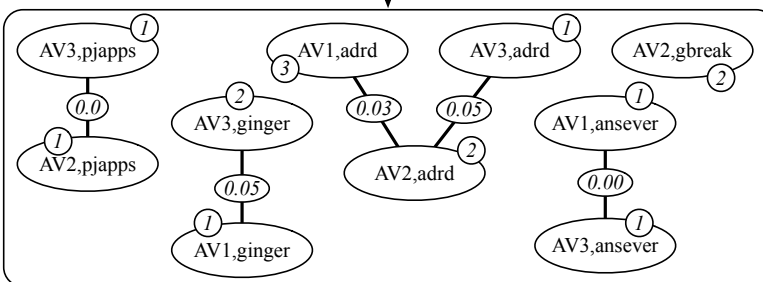


Tree Structure

Threshold

Cut

CUT Step



Family Name Selection

	Inferred Name
app1	pjapps
app2	adrd
app3	ginger
app4	anserver
app5	adrd

Single family
name per sample