Combing Density and Overlap (CoDO)

Assumptions

An overlap is significant if ...

1. It is at least as dense as the containing graphs
2. It is "large enough"

Definition

$$p_{CoDO} = \Pr[|\hat{A} \cap \hat{B}| \geq |Z| \cap \delta(\hat{A} \cap \hat{B}) \geq \delta(Z)]$$

where Z is the set of vertices in the overlap subgraph and $\delta()$ measures the density of a graph, i.e. $\frac{|E|}{C(|V|,2)}$.

Combing Density and Overlap (CoDO)

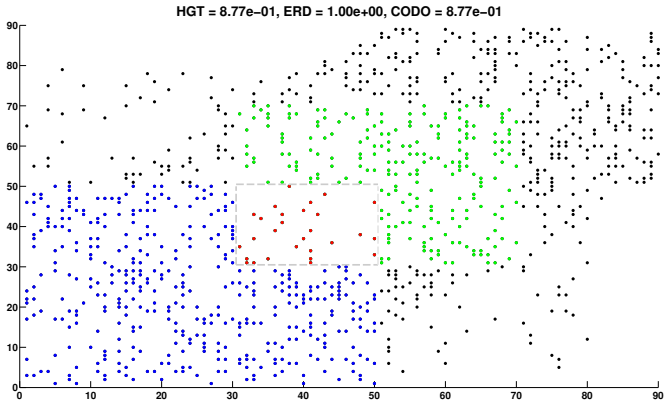
Expansion

By conditioning on the size of the overlap, we can get an explicit formula for this p-value in terms of hypergeometric tails:

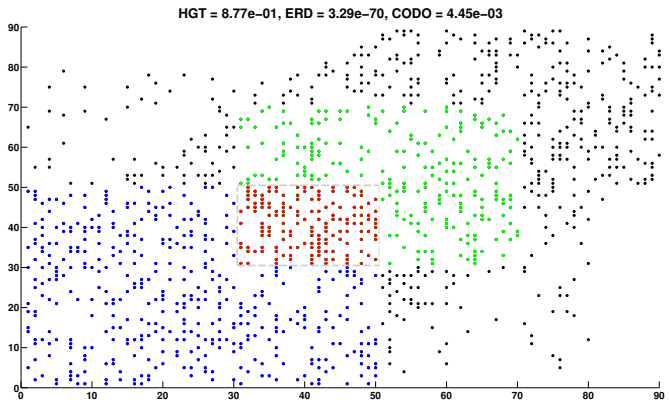
$$p_{CoDO} = \sum_{j=|Z|}^{\min\{|\hat{A}|,|\hat{B}|\}} \Pr\left[|\hat{A}\cap\hat{B}| = j\right] \cdot \Pr\left[\delta(\hat{A}\cap\hat{B}) \geq \delta(Z) \big| |\hat{A}\cap\hat{B}| = j\right]$$

Combing Density and Overlap (CoDO)

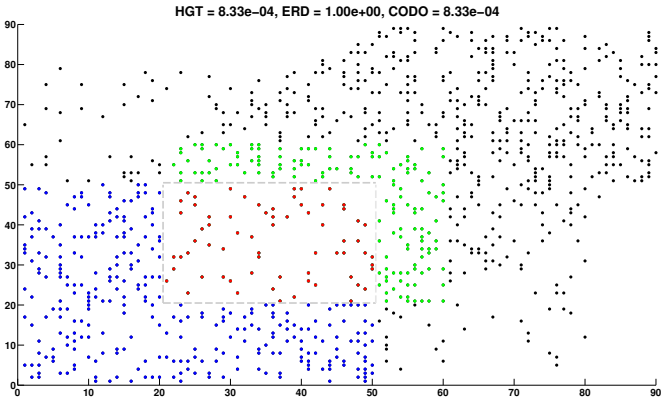
Example



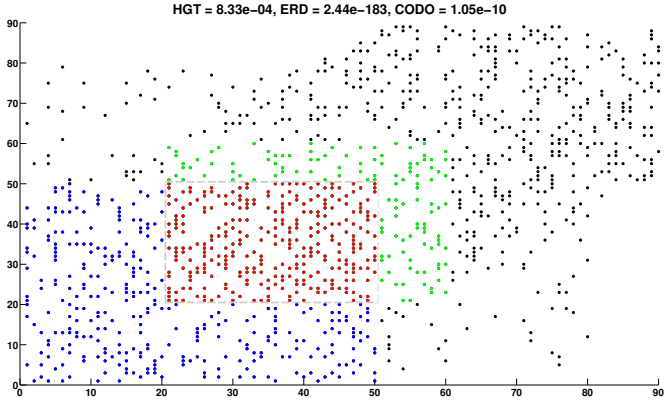
(a) Insignificant Overlap/Insignificant Density



(b) Insignificant Overlap/Significant Density



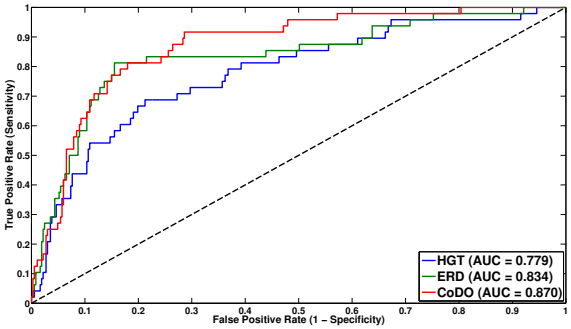
(c) Significant Overlap/Insignificant Density



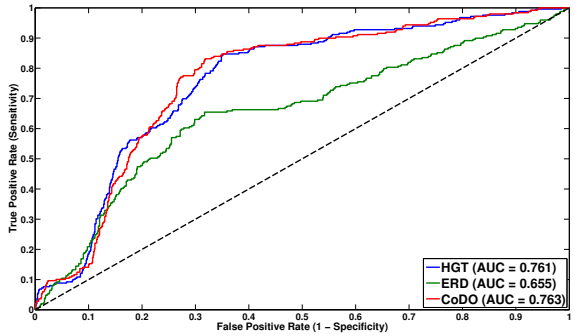
(d) Significant Overlap/Significant Density

Combing Density and Overlap (CoDO)

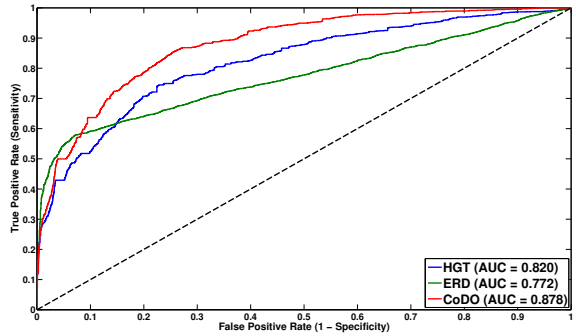
Application– Social networks



(e) Facebook



(f) Google+



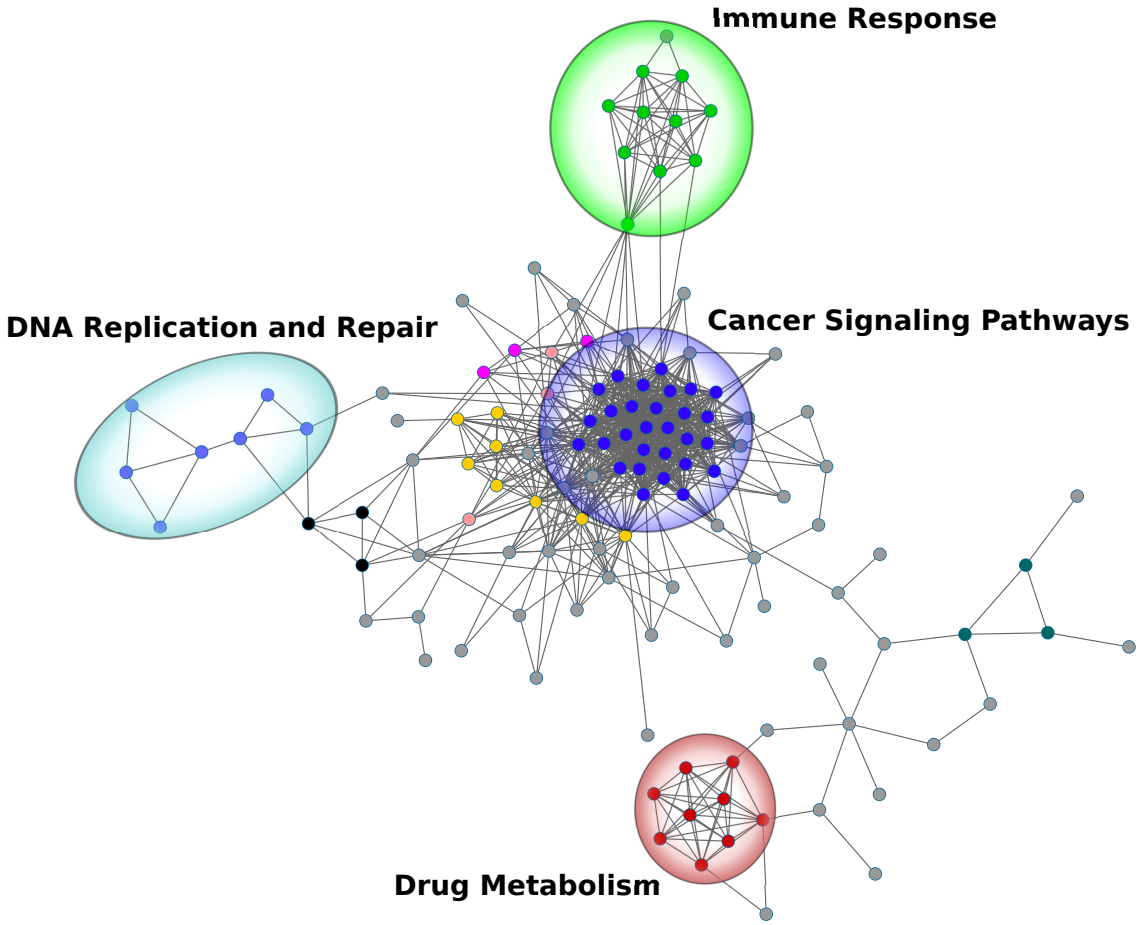
(g) Twitter

Definition

Ego Net is the induced subgraph among friends (alters) of a given user (ego)

Combing Density and Overlap (CoDO)

Application– Biological networks



Overlap among KEGG pathways is an indicator of pathway cross-talk