The proposed method exploits two sources (views) of information (primary data and user interactions) to identify which aspects in the two views are related and which are specific to only one of them.

Learning Algorithm:

\[ p(D, F, \Theta) \propto \prod_{i=1}^{N} \prod_{j>i} \prod_{i'j' \in \Theta} p_{i,j} p_{i,j}' q_{i,j} \]

Graphical Model

\[ D \] and \( F \) are two relational count data sets representing similarities between pairs of \( N \) samples, \( \{x_i\}_{i=1}^{N} \) from two different views (\( D \) for data view and \( F \) for user view). The two views, \( D \) and \( F \), are modeled with distributions \( p \) and \( q \), respectively.

\begin{align*}
    p_{i,j} & \propto \exp(-||z_i - z_j||^2 - ||z_i^{(D)} - z_j^{(D)}||^2) \\
    p_{i,j}' & \propto \exp(-||z_i - z_j||^2 - ||z_i^{(P)} - z_j^{(P)}||^2)
\end{align*}

By defining \( y_i = [z_i, z_i^{(D)}, z_i^{(P)}] \):

\[ p_{i,j} \propto \exp(- (y_i - y_j)^T \mathbf{W}^{(D)} \mathbf{W}^{(D)\top} (y_i - y_j)) \]

\[ p_{i,j}' \propto \exp(- (y_i - y_j)^T \mathbf{W}^{(P)} \mathbf{W}^{(P)\top} (y_i - y_j)) \]

Locations on the display, \( y_i \), can be estimated by maximizing the following log-likelihood function:

\[ \mathcal{L} = \lambda \sum_{i=1}^{N} \sum_{j>i} \tilde{d}_{i,j} \log p_{i,j} + (1 - \lambda) \sum_{i'j' \in \Theta} \tilde{f}_{i',j'} \log q_{i',j'} \]

CONCLUSION

- We have presented a statistical principle to identify and visualize aspects of data relevant to the user by exploiting statistical relations found between the primary data, and user-provided auxiliary data.
- A main future goal is to use similar technique in interactive visualization where user interaction data will be measured all the time and visualization needs to react faster.

REFERENCES