Subject: Similarity Analysis query
Posted by sansun on Tue, 20 Aug 2024 10:19:58 GMT
View Forum Message <> Reply to Message

Hi Thomas,

I have created a similarity chart using Similarity Analysis/Activity Cliff. With no activity data, I have created only the similarity chart (attached). I have two questions regarding the interpretation.

1) After reading the original paper I understand that molecules with >80% similarity (w.r.t. to the chosen descriptor) show connection lines. Is this a correct interpretation? I have manually checked some of the connected compounds and have not seen any that display similarity lower than 80% to each other.

2) Can it be said that molecules not connected to any other and occupy space towards the edge (e.g. red square) are unique since they do not have close neighbours? This is based on the following sentence in the original paper.

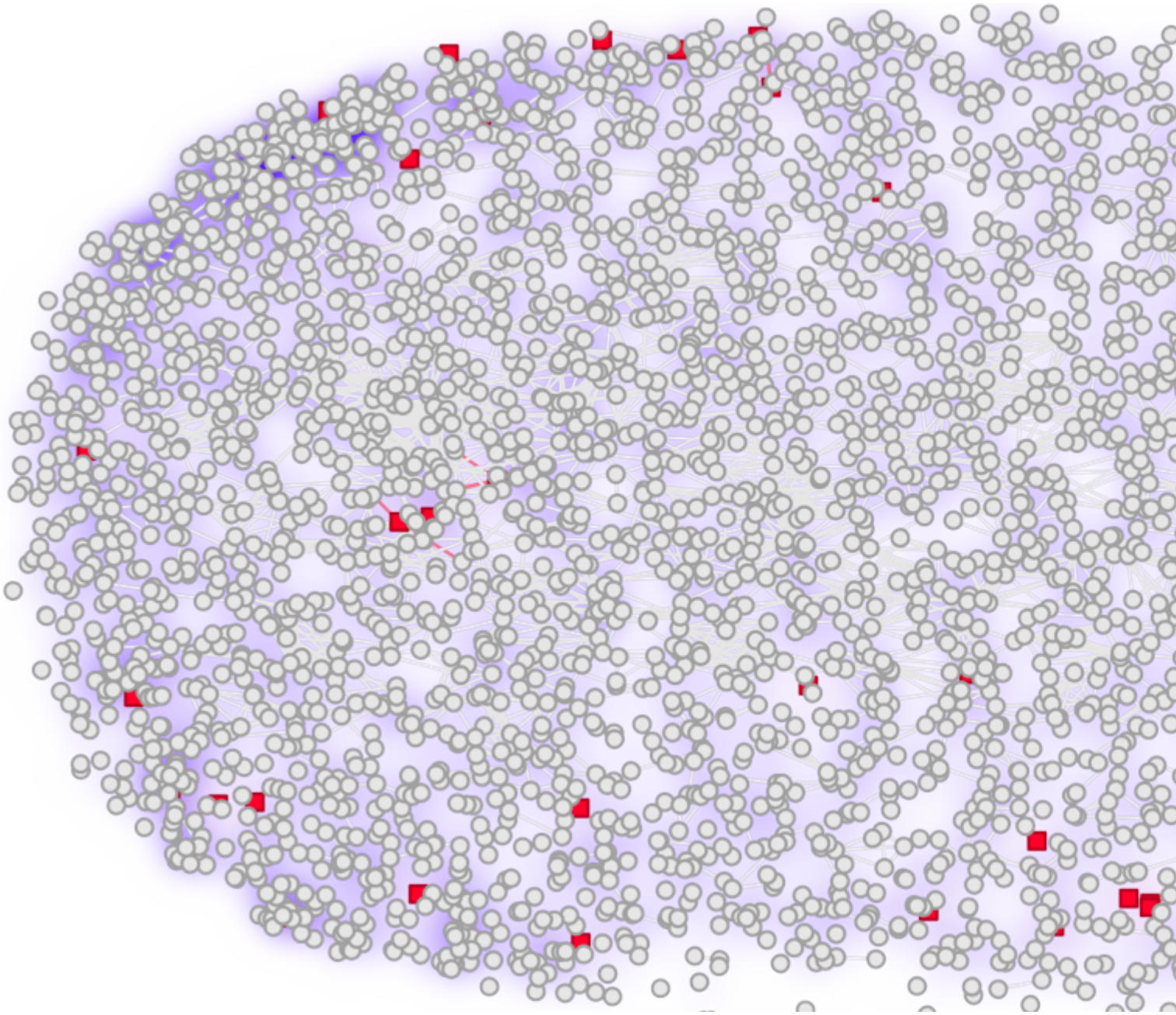
"Therefore, an individual compound, when being dragged toward the mean location of its neighbors, is often dragged toward the center"

However, the next sentence also says that a remedy is included in the algorithm to avoid crowding at the center. Therefore, I am little confused about the second interpretation.

Sorry, for the long message.
Thank you

File Attachments

1) similarity map.png, downloaded 1329 times



Subject: Re: Similarity Analysis query
Posted by thomas on Thu, 29 Aug 2024 18:37:55 GMT
View Forum Message <> Reply to Message

Hi Sansun,

1) The cutoff similarity limit for connection lines is at least 80%, but it is dynamically increased depending on the number of highly similar pairs. If most molecules are considered similar to most of the others, the algorithm cannot work.

2) It is true that molecules not connected to any other molecules tend to be more often close to the edge, because they are not drawn to the center. Since the center density is much higher at the end of the attraction phase, a correction is applied that moves all molecules towards the edge to achieve a similar density at all areas.

Hope this explains it...

Thomas
