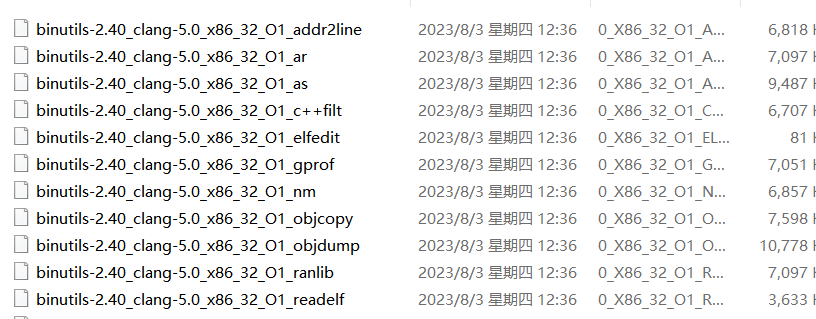
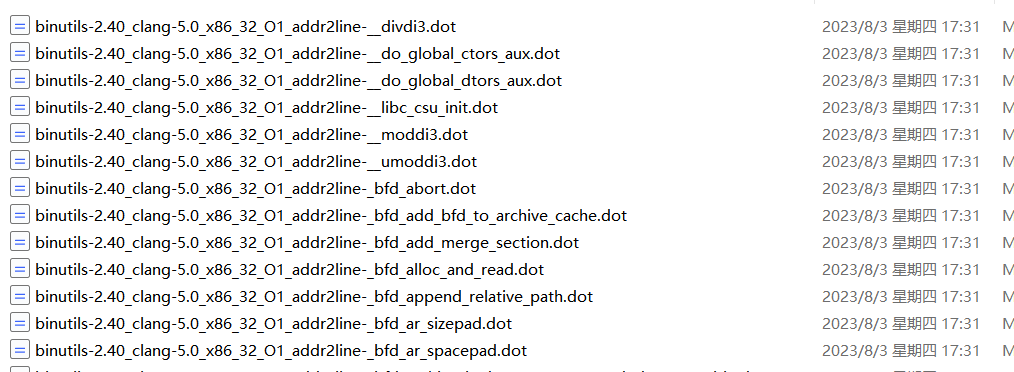
原始数据集：



经过工具生成图之后为：



其中 binutils-2.40\_clang-5.0\_x86\_32\_01\_addr2line为二进制文件名，divdi3为函数名。因为图是以函数为单位的。  
解析文件名为：

Project：binutils-2.40  
compiler：clang-5.0  
artitecture：x86\_32  
optimizer：01  
filename：addr2line  
function\_name:divdi3

如果需要跨编译器的数据集配对方式就是：

相似对：

寻找project、artitecture、optimizer、filename、function都一样一样的文件，但compiler不一样例如clang、gcc

同理跨优化选项也是。

不相似对：除了相似之外的，随机匹配，但是得去除相似的