**Constructing social networks for digital forensic investigations**

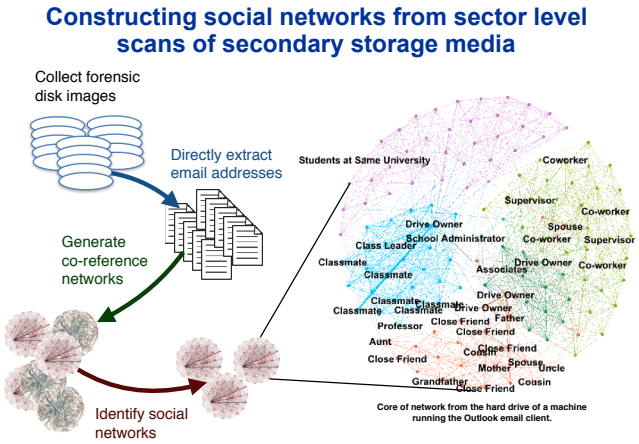
Gregory Allen[1], Ralucca Gera[2], Janina L. Green [1], and Michael McCarrin,[2]

1. US Army
2. Naval Postgraduate School

Digital forensic investigations can benefit greatly from an understanding of a suspect's social network. Secondary storage devices, such as hard drives, often contain a wealth of information that can be used to deduce the social network of their primary users. Typically, this information is gathered through painstaking manual analysis of user files in the logical file system, or extracted through a linear scan for email addresses and other relevant artifacts in the raw media. The latter technique, though more efficient, produces noisy results containing many artifacts that are unrelated to user behavior, which the analyst must manually examine and discard. Even after useful artifacts are identified, this method obscures the relationship between parties associated with the artifacts.

We improve this process with a novel technique for constructing social networks directly from artifacts produced by the linear scan. This technique is file system agnostic, easily parallelized, and robust against file fragmentation. The work proceeds in two stages. First, we generate co-reference networks based on byte-offset proximity between email addresses. Second, we use supervised learning techniques to separate social networks from networks of unrelated artifacts.

We find that correlating email addresses based solely on byte offset proximity is sufficient to distinguish between addresses that are artifacts of installed software and those which are connected to individual user logins and communication [1]. Further, on drives where the user employed email clients, this method can often successfully automatically generate a visualization of the user's social network [2].




The network construction process and an example network built from user email addresses. We observe that classic measures of centrality are effective for identifying important nodes, and community detection methods often meaningfully partition complex components.

# References

[1] Gregory Allen. Constructing and classifying email networks from raw forensic images. Master's thesis, Monterey, California: Naval Postgraduate School, 2016.

[2] Janina L Green. Constructing social networks from secondary storage with bulk analysis tools, 2016.