J3 Fragment Stack Analysis Techniques for Efficient Reconstruction of Ripped-Up Documents

Patrick De Smet, PhD, Ghent University, Department TELIN/TW07, Sint-Pietersnieuwstraat 41, Ghent, B-9000, Belgium*

This presentation will demonstrate several guidelines and (computerassisted) techniques that can enable a more efficient reassembly procedure for recomposing multiple stacks of recovered ripped-up document fragments.

This presentation will impact the forensic community and/or humanity by developing and discussing formal guidelines that can aid in faster reassemby of ripped-up documents.

When ripped-up documents are recovered at a crime scene, manual reassembly of the fragments can be a very time consuming task. Automatic and semi-automatic techniques have been and are being studied in order to reduce the complexity of this “jigsaw puzzle” problem, but currently such an automated reassembly process often still needs manual intervention and correction.

Hence, recent research has explored other possibilities for deducing more structured procedures and guidelines for reassembling recovered stacks of ripped-up document fragments. More specifically, the authors have investigated how a piece of paper is typically ripped-up and how this process can be reversed by discovering the subsequent ripping steps that were used. Obviously, it is assume that the stacks of fragments have been collected without disturbing their original and relative ordering or positioning relationships.

As it turns out, (sets of) fragments are often stacked on top of each other while a document is being ripped-up. Hence, by using a formal search procedure for matching fragments from the different substacks back together, ripped-up documents can be quickly reassembled. Obviously, the numbering indexes that separate the different substacks are unknown when the matching process is initiated, but after successfully matching some of the fragments together, the substack boundaries can easily be deduced and the subsequent matching process can be speeded-up further. The approach turns out to be quite efficient when compared to an often random search procedure as would typically be used by any inexperienced person reassembling a ripped-up document.

First, researchers studied the analysis and the reconstruction of several real but “ideal” stacks of fragments. As an example the following case can be considered. Suppose a person decides to rip-up a single page of a document. This person first rips the page in half. The two resulting fragments are stacked on top of each other and the stack is then torn in two again. This stacking and ripping process is repeated until either the fragments become too small or the physical thickness of the stack inhibits further iterations. Clearly, the final stack will contain 2^i fragments, for i ripping iterations. If a theoretical and small scale example is considered, e.g. for i=3, a binary ripping sequence would yield a stack sequence “1,2” for iteration 1, “1a,2a,1b,2b” for iteration 2, and “1aA,2aA,1bA,2bA,1aB,2aB,1bB,2bB” for the last iteration. The reassembly problem will need to rearrange the fragments “1,2,3,4,5,6,7,8” that are individually positioned on top of a reassembly working surface or desk; the fragments are labeled consecutively here since the ripping sequences, e.g. the a/b and A/B labeling, are still unknown at the start of the reassembly process. The first matching experiment in the proposed reassembly procedure is given by (1,5). Since this would yield a valid matching result, the two fragments can be recomposed into a new fragment “I” that is put on top of a second reassembly desk. Next, researchers match and recompose the subsequent pairs of fragments, i.e. (2,6), (3,7), and (4,8). This yields the new fragments II, III and IV that are also put on top of the second reassembly surface. The entire procedure can now be repeated for the new set of fragments, etc.

Next, more complex reassembly and matching processes were reviewed by considering and introducing real-life problems such as: (i) uncertainty of the ordering of *multiple* substacks that could possibly be recovered independently at a crime-scene: what if one stack of three fragments and one stack of five fragments were recovered at different locations, should the first stack be labeled “1,2,3” or rather “6,7,8,” (ii) compensating for missing or (temporarily) unmatchable fragments: what if only seven fragments from the scene were recovered or if a certain fragment would not display any distinctive color, text or shape features, (iii) non-binary ripping iterations: what if a ripping sequence “1,2,3,” “1a,2a,3a,1b,2b,3b,1c,2c,3c” would be used, etc. Some of these more specific problems are still being studied further, but researchers believe the most important complications have already been resolved successfully.

Finally, researchers have developed an interactive computer program that helps its user to follow the formal matching rules and guidelines studied and developed. This enables the user not having to memorize or study all the properties and guidelines related to the authors’ formal reassembly process.

 ** Presenting Author*

An important goal of this abstract is to discuss and further interact with other forensic investigators and scientist in order to have research and results evaluated and commented on.

Reconstruction, Ripped-Up Documents, Fragment Matching

 ** Presenting Author*