Para conseguir apresentar uma lista com os locais de saida do extrato em xml analisado, foi considerada a expressão regular “<SAIDA>”.

Iniciando a filtragem pelas linhas começadas por *<SAIDA>*, foi de seguida feita a limpeza destas frases, removendo o seu inicio e fim, de forma a ficar só com a string que tem o nome do local de saida. Esta limpeza foi feita recorrendo ao metodo gsub.

De seguida, para considerar nomes de locais com mais do que uma palavra, separados pelo caracter espaço, redefeniu-se o parametro Field Separator com o “\n”, apesar de puder ter sido escolhido outro simbolo.

Tendo resolvida toda a questão do parsing, o nome dos locais foram guardados num array. Pela utilização deste array, e das funcionalidade que esta estrutura de dados apresenta em gawk, sempre que um determinado local de saida for intrepertado, é incrementado o indice do array correspondente àquele local.

Deste modo, no padrão especial END é impreesso ao utilizador final não só a lista dos locais de saida, como tambem, o número de saídas registadas para esse local.