

入学年度 平成 6 年度

学籍番号 06117952

氏名 三村 恵介

論文題目 情報記憶媒体の安全な廃棄と再利用に関する研究

中野 研究室

1 はじめに

ハードディスクは、(高レベル)フォーマットをしてもファイル管理情報が消去されるだけで、ファイルそのもののデータは消去されずに残ってしまう。市販されている復元ツールなどを用いれば容易にファイルを復元することができ、情報の漏洩の危険性がある。全てのデータを消去してしまえば、情報の漏洩の危険性はなくなるが、ハードディスク全体にアクセスしなければならなくなり時間がかかってしまう。本研究では、ハードディスク上のファイルをセクタ単位で消去した際に消去にかかる時間を数式化する。また、一般に最も多く使用されている Windows のファイルシステムにおいて、その構造をもとにして、戦略的に消去する場合にどの程度安全かつ効率的であるかを検討する。また、それらをもとにして安全かつ効率的な消去方法を提案する。

2 ハードディスクでのファイル管理

ハードディスクを読み書きする際、ハードディスク内ではセクタ(一般に1セクタ=512バイト)と呼ばれる単位ごとにアクセスが行われる。また、Windowsではファイルを管理するときにセクタごとではなく、いくつかのセクタを一まとまりにした、クラスタという単位でファイルアロケーションを行っている。クラスタのサイズは、同じOS、同じファイル管理システムを用いても、ハードディスクの総容量に応じて異なっている。

3 ハードディスク上のファイルサイズ分布

ハードディスク上のファイル数と、それらのサイズの分布を調査した。ここでは、OSとしてWindows Me、ファイルシステムとしてFAT32を用いているハードディスクを対象とした。横軸を1つのファイルが使用しているセクタ数、縦軸をそのセクタ数を使用しているファイル数の頻度および積算値を示す(図1)。

4 ランダム消去によるコスト見積り

ランダムにセクタを消去することにより、どれくらいのファイルが破壊されるかを数式化する。ここで、ファイル中の1セクタでも消去された時、ファイルは破壊されたものとする。ファイルサイズが k セクタであるファイル数を N_k 、全セクタ数に占める消去するセクタの割合を α とした時、 k セクタの長さのファイルが破壊されずに残る確率 r_k は、

$$r_k = (1 - \alpha)^k \quad (1)$$

k セクタの長さのファイルのうち、破壊されるファイル数は $(1 - r_k) \cdot n_k$ となり、全てのファイルのうちで破壊されるファイルの割合は、つぎのようになる。

$$\left(\sum_{k=1} (1 - r_k) \cdot n_k \right) / \sum_{k=1} n_k \quad (2)$$

5 戦略的消去手法

以下に示す6個の戦略について、評価を行った。

1. 全てのセクタを消去する
2. ランダムに選んだセクタの内容を消す
3. 各クラスタの先頭セクタを消す
4. サブディレクトリエントリを消す
5. テキストファイルのクラスタを全て消す
6. FATを消す

これらの6個の戦略について、それぞれの評価を図2に示す。横軸は、総セクタ数に対して消去したセクタ数の割合を示し、縦軸は、総ファイル数に対して破壊されなかったファイルの数の割合を示す。2.と5.以外の戦略については消去するセクタ数が定まっているので、そのセクタ数が総セクタ数に対して占める割合と、その時破壊されずに残るファイルの割合を求め、プロットした。

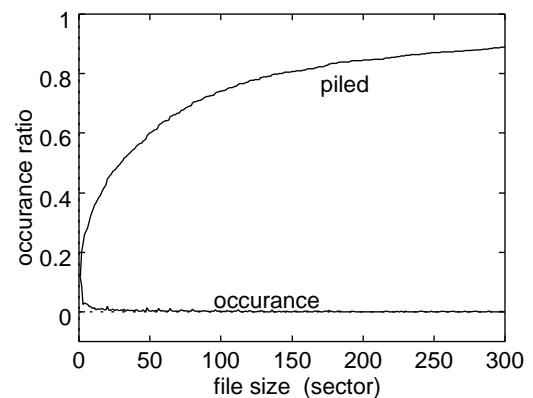


図 1: ハードディスク上のファイル分布

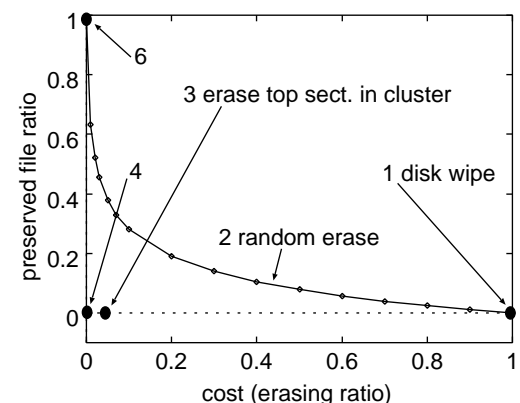


図 2: 戦略的にセクタを消去した時の各評価

6 おわりに

本研究では、セクタ単位でファイルを消去する方法について、ランダムにセクタを消去したとき、どの程度のファイルのデータ本体が破壊されるかを検討した。そして、それを基準にして他の戦略がどれくらいの安全性と効率性を持っているのか検討した。