In today’s lecture, we learned how to \textit{Compactly Represent Bags of Integers} using \textbf{Bloom filters}. The formal definition of the problem, and its solution, appear below, but let us first understand why this is necessary.

\section{1.1 A Sample Application}

Suppose we have a set of valid URLs, \(U\), and that \(|U| = n\). Suppose further that each URL is approximately 100 characters in length; then each URL requires 800 bits for proper representation. It should be clear that any set requires \(800n\) bits to represent the \(n\)-element set, \(U\).

Now consider the following basic caching structure from \textit{Fan et al.} where

\begin{table}[h]
\caption{A simple shared caching structure (picture omitted)}
\begin{itemize}
\item \(C_1, C_2, C_3\) are caches, each with a set of documents stored; documents are indexed by URL
\item edges represent the xfer of a cache’s set of URLs
\item a query for URL \(x\) can be made in one, some, or all of the caches.
\end{itemize}
\end{table}