https://www.cisa.gov/sites/default/files/FactSheets/NCCIC%20ICS_Factsheet_File_Hashing_S508C.pdf

The article “file hashing” explained how hashing works and why it's crucial in verifying the integrity of data, specifically in the cybersecurity world.

One of the key points I learned was how hashing creates a unique identifier, or signature, for data. This signature can then be used to ensure that the data hasn’t been altered. The idea that a small change in the data will completely change the hash value really emphasizes how precise and secure this method is.

The article also highlighted some of the practical applications of hashing in cybersecurity. For instance, when downloading files, hashes can be used to verify that the file you downloaded is exactly what was intended by the sender and hasn't been tampered with. This is something I've seen on download pages before but never really appreciated. Knowing that hashes can help protect against malicious alterations gives me a new perspective on their importance.

The concept of hash collisions was also touched on, which happens when two different inputs produce the same hash. It's a bit worrying to learn that some older algorithms like MD5 and SHA-1 are more prone to these collisions, which could be exploited by attackers.

Another cool thing I learned about was VirusTotal (VT). It’s a service where you can submit files or URLs to check if they’re malicious. I never knew that you could look up a file’s hash to see if it’s been flagged as dangerous before. However, the article also warned about being cautious when uploading files with sensitive information because once uploaded, it could be publicly accessible. This was a good reminder to think twice before sharing anything online, even if it’s for security checks.

Overall, this article was super informative and really expanded my understanding of how hashing works and its role in keeping data secure. It made me appreciate the behind-the-scenes work that goes into protecting our information online and highlighted the importance of staying informed about the tools and methods that can help us stay safe in the digital world.

Concept utilized: One new concept I utilized in this project was the use of a Hashtable to store and process baby names data. Additionally, I employed a hashing algorithm to ensure that the data is distributed evenly, minimizing collisions, and ensuring fast lookup times.

Challenges encountered:

Data Parsing: One of the significant challenges was reading and parsing the data from the provided files. The files contained baby names from 2016 to 2021, and each name had to be accurately mapped to its corresponding counts for each year. Ensuring the correct parsing of this data required careful attention to detail.

Observations gained: An interesting observation from processing the baby names data was the trend in popularity of certain names over the years. Some names consistently appeared at the top, indicating their popularity, while others fluctuated, reflecting changing naming trends.

Probing question: How do you think the concept of hashing and services like VirusTotal could be integrated into everyday cybersecurity practices for organizations to enhance data protection and threat detection?