

دانشکده مهندسی کامپیوتر

ارائه مطالب علمی و فنی

مستند پروژه

بررسی الگوریتم‌های فشرده‌سازی و کاربردهای آنها

نگارندگان:
محمد مهدی عرفانیان

۱۷ آذر ۱۳۹۸



چکیده

در مستندی که پیش روی خواننده عزیز قرار دارد تلاش شده تا مختصراً الگوریتم‌های فشرده‌سازی مختلف و کاربردهای آن‌ها در زمینه‌های مختلف مهندسی کامپیوتر در راستای انجام پروژه درس ارائه مطالب علمی و فنی بررسی شود. این درس در پاییز ۹۸ در دانشکده مهندسی کامپیوتر دانشگاه صنعتی شریف توسط دکتر همت‌یار ارائه شده است.

برای سهولت کار استاد محترم درس برای تحصیل اطمینان از درستی مستندسازی و همچنین استفاده دانش‌جویان علاقه‌مند، سیر پیشرفت مستند به همراه کدهای $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ در *Github* قرار گرفته است، لازم به ذکر است که این مستند به صورت متن‌باز ارائه شده و استفاده از آن بدون ذکر منبع برای همگان آزاد است. در انتها از استاد محترم درس، دستیار آموزشی ایشان و خوانندگان محترم تشکر می‌کنیم.

با آرزوی خوش‌وقتی برای تمامی خوانندگان این مستند
نگارنده

فهرست مطالب

فصل ۱

مقدمه

توضیحی اولیه مبنی بر تعریف کلی فشرده‌سازی، انواع الگوریتم‌ها و کاربردها

۱.۱ تعریف

الگوریتم‌های فشرده‌سازی، الگوریتم‌هایی هستند که می‌توان با استفاده از آن‌ها داده‌ها را طوری رمزنگاری کرد که در تعداد کمتری بیت نسبت به آرایش اولیه قابل ارائه باشند. برای مثال می‌دانیم که برای ذخیره هر بیت اسکی هشت بیت فضا لازم است، می‌توان با استفاده از نگاشتی متشکل از حروف استفاده‌شده در یک متن تعداد بیت‌های مورد نیاز برای نشان دادن هر حرف استفاده‌شده در متن را کاهش داد.

۲.۱ انواع الگوریتم‌های فشرده‌سازی

در یک دسته‌بندی الگوریتم‌های فشرده‌سازی را به دو نوع زیر افراز می‌کنند.

• Lossless

• Lossy

۱.۲.۱ الگوریتم‌های Lossless

در این سری الگوریتم‌ها داده ورودی بدون هیچ‌گونه هدررفتی از داده خروجی قابل بازیابی است، این الگوریتم‌ها در مواقعی که ثابت ماندن داده در طی فشرده‌سازی الزامی است استفاده می‌شوند، همچنین معمولاً برای بازیابی اطلاعات فشرده‌شده نیاز به داده‌هایی خارجی است که با کمک آن عمل بازیابی انجام می‌گیرد، از این رو می‌توان از این نوع الگوریتم‌ها در رمزنگاری نیز استفاده کرد. کاربرد اصلی این الگوریتم‌ها در فشرده‌سازی متون است که اشتباه شدن حتی یک حرف می‌تواند باعث بدخوانی و بدفهمی متن اصلی گردد. الگوریتم‌های مشهور کمپرس متن به شرح زیر اند.

- RLE
- Huffman Encoding
- Burrows Wheeler Transform

۲.۲.۱ الگوریتم‌های Lossy

در این الگوریتم‌ها پس از هر بار فشرده‌سازی مقداری از داده‌ها از دست می‌روند، معیار ارزیابی این الگوریتم‌ها مقدار فشرده‌سازی با توجه به میزان هدررفت داده می‌باشد، به علت هدررفت مقداری از داده این الگوریتم‌ها معمولاً در مواردی که هدررفت اندک داده توسط انسان یا ماشین قابل تشخیص نباشد استفاده می‌شوند، مثلاً تکنیک‌های ذخیره‌سازی تصاویر و ویدئوها در کامپیوترها مبتنی بر الگوریتم‌های Lossy است زیرا چشم انسان قادر به تشخیص عوض شدن تعدادی پیکسل در صفحه پس از بازیابی فایل فشرده‌شده نیست.

۳.۱ کاربردها

فشرده‌سازی داده‌ها در دنیای امروز مهندسی کامپیوتر در نقاط مختلفی نقش دارد، در ادامه مختصراً برخی از کاربردهای فشرده‌سازی توضیح داده می‌شود.

- مدیا

برای انتقال داده‌های صوتی-تصویری یا به عبارت دیگر مدیاهای مختلف نیاز به الگوریتم‌های فشرده‌سازی به شدت احساس می‌شود، حجم هر فیلم سینمایی بدون فشرده‌سازی می‌تواند تا صدها گیگابایت برسد و انتشار هر تصویر در فضای اینترنت بدون فشرده‌سازی می‌تواند هر بیننده را ساعت‌ها معطل کند. فرمت‌های مختلف فشرده‌سازی عکس مانند JPEG، PNG، EPS ... و همچنین فرمت‌های پخش ویدئو مانند MP4، MKV ... برای اکثر افراد فرمت‌های شناخته‌شده‌ای هستند که در بطن خود از الگوریتم‌های مختلف فشرده‌سازی استفاده می‌کنند.

- پردازش سیگنال

- ژنتیک

- امنیت