

دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

دانشکده مهندسی کامپیوتر

گزارش سمینار کارشناسی ارشد

مهندسی کامپیوتر-معماری سیستم‌های کامپیوتری

شتا بدھی سخت افزاری پیش‌بینی عمر باقی‌مانده مفید
دستگاه‌های دوار با استفاده از شبکه عصبی ترانسفورمر بر بستر
FPGA

نگارش

رضا آدینه پور

استاد راهنما

جناب آقای دکتر مرتضی صاحب‌الزمانی

۱۴۰۳ مرداد



اول

سپاس

از استاد بزرگوارم، جناب آقای دکتر مرتضی صاحب‌الزمانی، که به بنده اعتماد کردند و با کمک‌ها و راهنمایی‌های بی‌دریغ‌شان، مرا در به سرانجام رساندن این پایان‌نامه یاری دادند، سپاسگزاری و قدردانی می‌کنم. حمایت‌های مستمر و بی‌وقفه ایشان، نه تنها انگیزه و توان مضاعفی به من بخشید، بلکه مسیر پژوهشی این پایان‌نامه را با روشنایی دانش و تجربه‌شان هموار ساخت. بدون همراهی، مشورت‌ها و تشویق‌های استاد محترم، تحقق این هدف امکان‌پذیر نمی‌بود. از صمیم قلب برای ایشان آرزوی سلامتی، موفقیت و برکت دارم.

چنان مجنون شوم گویی که شهرآشوب دورانم
نه پندی و نه اندرزی به گوشم پنبه چسباندم
بماند یادگار این شعر
که من از دار این دنیا فقط این شعر می‌دانم.

چکیده

در محیط‌ها و کارخانه‌های صنعتی، همواره یکی از مهم‌ترین دغدغه‌ها، نگهداری و تعمیر ابزارآلات و دستگاه‌های صنعتی است. دستگاه‌هایی که هرکدام وظیفه‌ای مهم را در خط تولید هر شرکتی بر عهده دارند و کوچک‌ترین آسیب و خرابی می‌تواند خط تولید شرکت را مختل کند. بنابراین، نیاز است که همواره بتوانیم وضعیت فعلی سلامت دستگاه‌ها را مورد پایش قرار دهیم و بتوانیم زمان خرابی دستگاه و عمر مفید باقی‌مانده آن را پیش‌بینی نماییم.

پیش‌بینی عمر مفید باقی‌مانده مفید (Remaining Useful Life) یا به اختصار RUL وظیفه‌ای بسیار مهم در زمینه پیش‌بینی و مدیریت سلامت (PHM) ابزار و تجهیزات صنعتی است. پیش‌بینی دقیق RUL امری حیاتی و مهم است چرا که امکان تعمیر و نگهداری به موقع را فراهم می‌کند، مدت زمان خرابی دستگاه را کاهش می‌دهد و می‌توان بر اساس عمر مفید باقی‌مانده دستگاه، آن را تحت بار قرار داده و کارایی عملیاتی را بهبود بخشد.

کلیدواژه‌ها: عمر باقی‌مانده مفید، RUL، ترانسفرمر، FPGA

فهرست مطالب

۱	۱	مقدمه
۱	۱-۱	تعريف مسئله
۲	۲-۱	اهمیت موضوع
۲	۳-۱	اهداف پژوهش
۲	۴-۱	ساختار پایان نامه
۳	۲	مفاهیم اولیه
۳	۱-۲	عمر باقیمانده مفید
۴	۱-۱-۲	اولین زمان خرابی
۴	۲-۱-۲	عمر پایانی دستگاه
۵	۲-۲	داده‌ها
۶	۱-۲-۲	مجموعه داده XJTU-SY
۸	۲-۲-۲	مجموعه داده PRONOSTIA
۱۱	۳-۲-۲	مجموعه داده C-MAPSS
۱۲	۳	کارهای پیشین
۱۲	۱-۳	مسائل خوشبندی
۱۴	۲-۳	خوشبندی k -مرکز
۱۶	۳-۳	مدل جوییار داده

۱۷	۴-۳ تقریب‌پذیری
۱۸	۴ چالش‌ها و نوآوری‌ها
۱۹	۵ نتیجه‌گیری
۲۰	مراجع
۲۱	واژه‌نامه
۲۳	آ مطالب تکمیلی

فهرست جداول

۱-۲	پارامترهای بلبرینگ‌های آزمایش شده	۷
۲-۲	اطلاعات مجموعه‌داده XJTU-SY	۸
۳-۲	تقسیم بندی داده‌های جمع‌آوری شده در مجموعه داده PRONOSTIA	۱۱
۱-۳	نمونه‌هایی از کران پایین تقریب‌پذیری مسائل خوشبندی	۱۷

فهرست تصاویر

۱-۲	تابع تقریب زننده عمر باقی مانده یک دستگاه	۳
۲-۲	سیگنال ارتعاش بلبرینگ از ابتدای فعالیت تا تخریب کامل [۱]	۴
۳-۲	سیگنال RUL شکل ۲-۲ [۱]	۵
۴-۲	بستر تهیه مجموعه داده XJTU-SY [۲]	۶
۵-۲	تنظیمات نمونه برداری برای سیگنال های ارتعاشی	۷
۶-۲	سیگنال های ارتعاش افقی و عمودی و RUL بلبرینگ ۲-۱	۸
۷-۲	سیگنال های ارتعاش افقی و عمودی و RUL بلبرینگ ۲-۲	۹
۸-۲	بستر تهیه مجموعه داده PRONOSTIA [۳]	۹
۹-۲	تنظیمات نمونه برداری سیگنال ها در این آزمایش	۱۰
۱۰-۲	سیگنال خام تهیه شده در این آزمایش و بلبرینگ تخریب شده، [۳]	۱۰
۱-۳	نمونه ای از مسئله ۲-مرکز	۱۴
۲-۳	نمونه ای از مسئله ۲-مرکز با داده های پرت	۱۵

فصل ۱

مقدمه

پیش‌بینی عمر مفید باقی‌مانده^۱ (یا به اختصار RUL) یکی از مباحث کلیدی در حوزه مدیریت سلامت و پیش‌بینی^۲ (PHM) ابزارها و تجهیزات صنعتی است. RUL به مدت زمانی اشاره دارد که یک دستگاه یا ابزار قبل از رسیدن به نقطه خرابی و از کار افتادن نهایی، می‌تواند به طور مؤثر کار کند. این حوزه پژوهشی با استفاده از تکنیک‌های مختلف و پیشرفته در تلاش است تا به صنایع کمک کند تا بهره‌وری و کارایی خود را افزایش دهند و هزینه‌های ناشی از تعمیر و نگهداری غیرضروری را کاهش دهند.

۱-۱ تعریف مسئله

در صنایع مختلف، از جمله خودروسازی، هواپما، نفت و گاز و تولیدات صنعتی، تجهیزات و ماشین‌آلات به طور مداوم تحت شرایط کاری سخت و پیچیده قرار دارند. هرگونه خرابی ناگهانی این تجهیزات می‌تواند به وقفه‌های غیرمنتظره در تولید منجر شود که علاوه بر خسارات مالی، ممکن است اثرات زیان‌باری بر کیفیت محصول نهایی و رضایت مشتریان داشته باشد. بنابراین، نیاز است که وضعیت فعلی سلامت دستگاه‌ها به طور مستمر پایش شود و زمان خرابی دستگاه و عمر مفید باقی‌مانده آن با دقت بالایی پیش‌بینی شود. این پیش‌بینی نیازمند استفاده از تحلیل داده‌های حسگرها، مدل‌سازی ریاضی و الگوریتم‌های یادگیری ماشین^۳ و عمیق^۴ است.

¹Remaining Useful Life

²Prognostics and Health Management

³Machine Learning

⁴Deep Learning

۲-۱ اهمیت موضوع

اهمیت پیش‌بینی دقیق RUL در صنعت به دلیل تاثیر مستقیم آن بر بهره‌وری، کارایی و کاهش هزینه‌های تعمیر و نگهداری غیرضروری، به خوبی شناخته شده است. تکنیک‌های پیشرفته در حوزه هوش مصنوعی^۵ و تحلیل داده‌های بزرگ^۶ این امکان را فراهم کرده‌اند که داده‌های جمع‌آوری شده از تجهیزات به صورت بلادرنگ^۷ تحلیل شوند و مدل‌های پیش‌بینی RUL با دقت بالاتری ارائه شوند. این پیشرفتهای، به شرکت‌ها این امکان را می‌دهد که فرآیندهای نگهداری و تعمیرات خود را بهینه‌سازی کنند و تصمیم‌گیری‌های بهتری در زمینه مدیریت دارایی‌های خود انجام دهند.

۳-۱ اهداف پژوهش

این پژوهش بر توسعه یک راه حل شتابدهی سخت‌افزاری بر بستر FPGA پیش‌بینی RUL با استفاده از شبکه عصبی مصنوعی که وظیفه آن آموزش و یادگیری توالی و درنهایت پیش‌بینی آن است تمرکز دارد. با استفاده از قابلیت پردازش موازی FPGA و همچنین توان مصرفی بسیار پایین آن، هدف ما افزایش کارایی و مقیاس‌پذیری سیستم‌های پیش‌بینی RUL به ویژه برای دستگاه‌های دور است.

۴-۱ ساختار پایان‌نامه

این پایان‌نامه در پنج فصل به شرح زیر ارائه می‌شود. مفاهیم اولیه‌ی و نحوه‌ی کار شبکه عصبی ترانسفرمر در فصل «۲»^۸ اشاره شده است. فصل «۳»^۹ به مطالعه و بررسی کارهای پیشین مرتبط با پیش‌بینی عمر باقی‌مانده می‌پردازد. در فصل «۴»^{۱۰}، چالش‌های موجود در این پژوهش و راه حل‌های ما برای رفع این چالش‌ها ارائه شده است. فصل «۵»^{۱۱} به جمع‌بندی کارهای انجام شده در این پژوهش و ارائه‌ی پیشنهادهایی برای انجام کارهای آتی خواهد پرداخت.

⁵ Artificial Intelligence

⁶ Big Data Analytics

⁷ Real Time

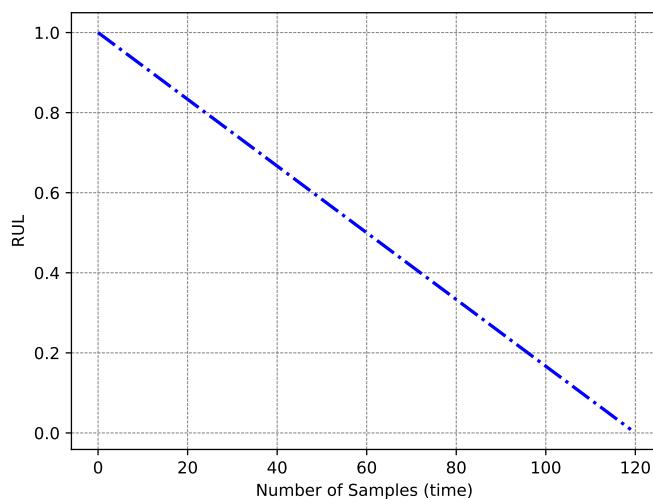
فصل ۲

مفاهیم اولیه

۱-۲ عمر باقی‌مانده مفید

مانند انسان‌ها، همه دستگاه‌ها و قطعات نیز عمری دارند و برای پایش سلامت دستگاه نیاز است که بتوانیم از عمر باقی‌مانده قطعه مطلع باشیم.

عمر یک قطعه را می‌توان به‌وسیله یک تابع خطی که آن را تابع RUL می‌نامیم «شکل ۳-۲» تقریب بزنیم.



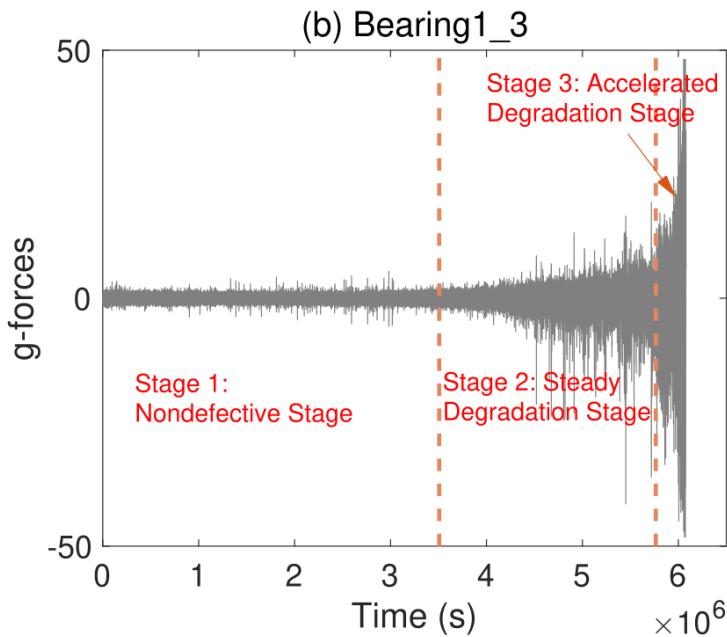
شکل ۱-۲: تابع تقریب زننده عمر باقی‌مانده یک دستگاه

محور عمودی در شکل «۳-۲» نشان‌دهنده میزان سلامت دستگاه و محور افقی نشان‌دهنده زمان است که معمولاً بر حسب دقیقه بیان می‌شود.

همه دستگاه‌ها زمانی که در آستانه بروز خطا و خرابی قرار می‌گیرند رفتار غیرعادی از خودشان نشان می‌دهند.

از جمله این رفتارها می‌توان به نوسانات غیرطبیعی، افزایش دمای دستگاه، افزایش سرعت^۱ در دستگاه اشاره نمود.

برای مثال در [۱] سیگنال ارتعاشات یک بلبرینگ به عنوان یکی از اصلی‌ترین قطعات صنعتی از ابتدای شروع به کار تا زمان بروز اولین تخریب^۲ و تخریب کامل^۳ در مدت زمان ۶۹ روز جمع‌آوری شده است. «شکل ۲-۲»



شکل ۲-۲: سیگنال ارتعاش بلبرینگ از ابتدای فعالیت تا تخریب کامل [۱]

۱-۱-۲ اولین زمان خرابی

اولین زمانی را که دستگاه دچار نوسانات شدید می‌شود را به عنوان اولین زمان شروع فرایند تخریب در نظر می‌گیریم و آن را «FPT» می‌نامیم.

۲-۱-۲ عمر پایانی دستگاه

با افزایش دامنه نوسانات ثبت شده از دستگاه، تخریب دستگاه بیشتر شده و دستگاه گرم‌تر می‌شود. از این فرایند به عنوان یک بازخورد^۴ مثبت یاد می‌شود که افزایش گرما، نوسانات را بیشتر کرده و نوسانات بیشتر نیز گرمای دستگاه را افزایش می‌دهد. با تشدید هرچه بیشتر نوسانات، دستگاه به پایان عمر خود نزدیک‌تر شده و درنهایت

¹Noise

²First Prediction Time

³End of Life

⁴Feedback

از کار می‌افتد. زمان از کار فتادن نهایی دستگاه را به عنوان زمان پایان زندگی «EOF» تعریف می‌کنیم.

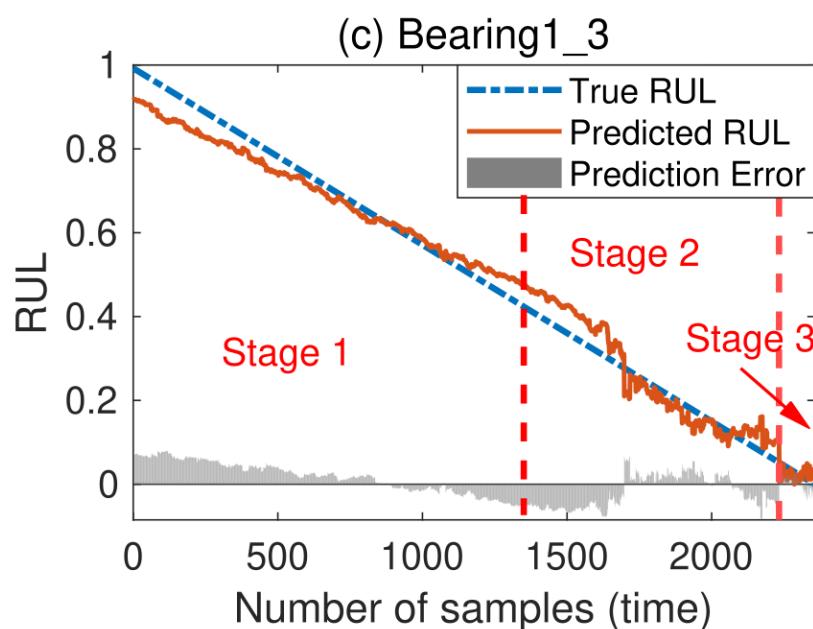
و درنهایت سیگнал RUL به صورت تفاضل این دوزمان تعریف می‌شود:

$$T_{RUL} = T_{EOF} - T_{FPT} \quad (1-2)$$

این سیگнал از جنس زمان است و مقدار RUL در این باز زمانی از رابطه زیر پیروی می‌کند:

$$RUL(t) = -t \quad (2-2)$$

برای مثال سیگнал RUL برای شکل «۲-۲» به صورت زیر محاسبه می‌شود:



شکل ۲-۳: سیگнал RUL شکل ۲-۲ [۱]

۲-۲ داده‌ها

برای پیش‌بینی عمر باقی‌مانده مفید، چندین مجموعه داده^۵ وجود دارد که در ادامه آنها را معرفی و بررسی می‌کنیم.

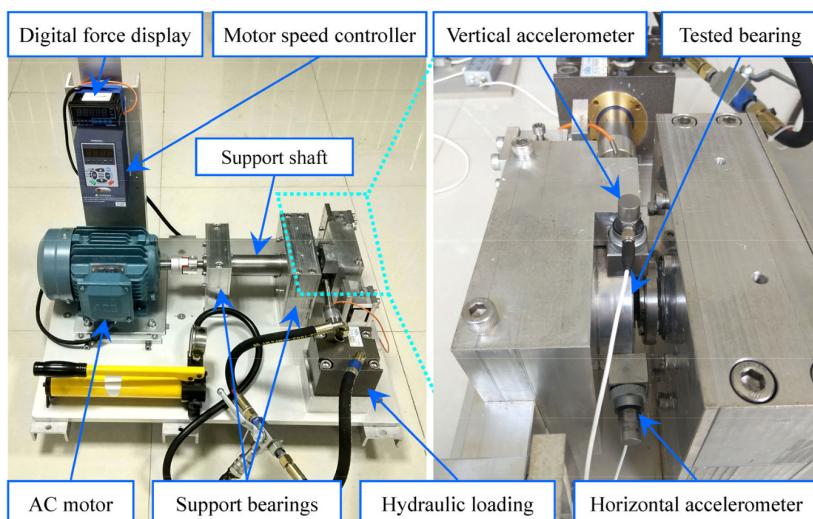
⁵Dataset

۱-۲-۲ مجموعه داده XJTU-SY

این مجموعه داده^۶ شامل داده های ثبت شده از ۱۵ بلبرینگ است که با انجام آزمایش های تخریب سریع، دچار تخریب شده اند. [۲]

این مجموعه داده توسط سیستمی که در شکل «۴-۲» نشان داده شده است، مشکل از یک موتور القایی جریان متناوب^۷، یک کنترل کننده سرعت موتور، یک محور^۸ پشتیبان، دو بلبرینگ پشتیبان (بلبرینگ های سنگین) و یک سیستم بارگذاری هیدرولیک تشکیل شده است.

این بستر آزمون برای انجام آزمایش های تخریب تسربی شده بلبرینگ، تحت شرایط مختلف عملیاتی (نیروی شعاعی و سرعت چرخشی مختلف) طراحی شده است. نیروی شعاعی توسط سیستم بارگذاری هیدرولیک تولید شده و به محفظه بلبرینگ های آزمایش شده اعمال می شود و سرعت چرخش نیز توسط کنترلر سرعت موتور القایی AC تنظیم و نگه داشته می شود.



شکل ۴-۲: بستر تهیه مجموعه داده XJTU-SY [۲]

بلبرینگ های مورد استفاده در این آزمایش از نوع LDK UER204 هستند که پارامترهای دقیق آنها در جدول «۱-۲» آورده شده است.

این آزمایش، تحت ۳ شرط عملیاتی مختلف انجام شده است و هر ۵ بلبرینگ موجود در این آزمایش تحت این سه شرط عملیاتی قرار گرفته اند. این شرایط عملیاتی شامل موارد زیر هستند:

• ۲۱۰۰ دور در دقیقه^۹ (۳۵ هرتز) و بار دینامیکی ۱۲ کیلونیوتون

^۶ می توانید این مجموعه داده را از اینجا دانلود کنید: biaowang.tech/xjtu-sy-bearing-datasets/

⁷ Alternating Current

⁸ Shaft

⁹ RPM

جدول ۱-۲: پارامترهای بلبرینگ‌های آزمایش شده

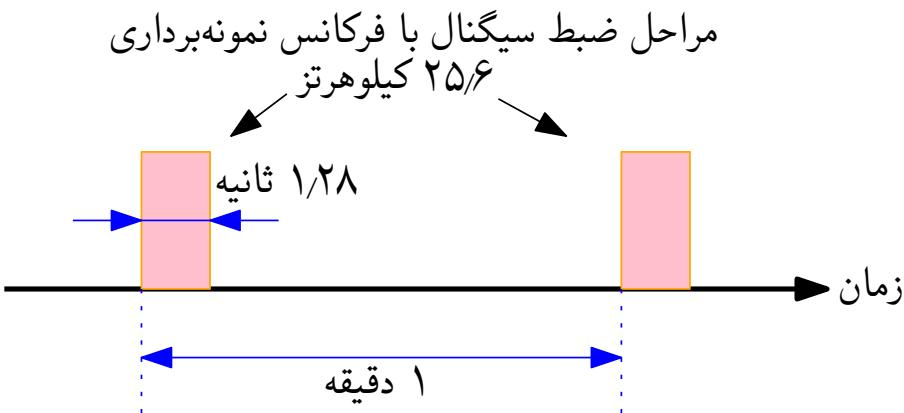
پارامتر	مقدار	پارامتر	مقدار
قطر مسیر بیرونی	۳۹/۸۰ mm	قطر مسیر داخلی	۲۹/۳۰ mm
قطر متوسط بلبرینگ	۳۴/۵۵ mm	قطر توب	۷/۹۲ mm
تعداد توب‌ها	۸	زاویه تماس	۰°
بار استاتیک	۶۶۵ kN	بار دینامیک	۱۲۸۲ kN

• ۲۲۵° دور در دقیقه (۳۷/۵ هرتز) و بار دینامیکی ۱۱ کیلونیوتون

• ۲۴۰° دور در دقیقه (۴۰ هرتز) و بار دینامیکی ۱۰ کیلونیوتون

برای جمع‌آوری سیگنال‌های ارتعاشی بلبرینگ‌های آزمایش شده، همان‌طور که در شکل «۴-۲» نشان داده شده است، دو شتاب‌سنج از نوع PCB۳۵۲۰۳۳ در زاویه ۹۰ درجه بر روی محفظه بلبرینگ‌های آزمایش شده قرار داده شده است، یعنی یکی بر روی محور افقی و دیگری بر روی محور عمودی نصب شده است.

همچنین فرکانس نمونه‌برداری بر روی ۲۵/۶ کیلوهرتز تنظیم شده است. همان‌طور که در شکل «۵-۲» نشان داده شده است، در مجموع ۳۲۷۶۸ نقطه داده (به مدت ۱/۲۸ ثانیه) برای هر نمونه‌برداری ثبت می‌شوند و دوره نمونه‌برداری برابر با ۱ دقیقه است.



شکل ۲-۵: تنظیمات نمونه‌برداری برای سیگنال‌های ارتعاشی

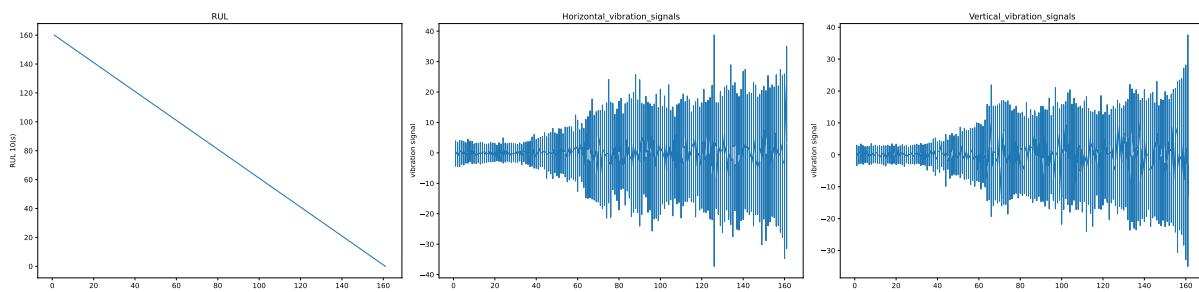
برای هر نمونه‌برداری، داده‌های به‌دست‌آمده در یک فایل csv ذخیره شده است که در آن ستون اول، سیگنال‌های ارتعاشی افقی و ستون دوم سیگنال‌های ارتعاشی عمودی را شامل می‌شود. جدول «۲-۲» اطلاعات دقیق هر بلبرینگ آزمایش شده، شامل تعداد فایل‌های csv، عمر بلبرینگ و عنصر خواری را فهرست می‌کند.

نقطه EOF در این مجموعه داده $A_H \times 10^\circ$ در نظر گرفته شده است که A_H بیشینه دامنه سیگنال ارتعاش عمودی و افقی در حالت کاری طبیعی است.

جدول ۲-۲: اطلاعات مجموعه داده XJTU-SY

عامل خطا	شرایط عملکرد	مجموعه داده های بلبرینگ	تعداد فایل ها	طول عمر بلبرینگ
بیرونی	شرایط ۱ ۳۵ هرتز، ۱۲ کیلونیوتون)	بلبرینگ ۱_۱	۱۲۳	۲ ساعت و ۳ دقیقه
بیرونی		بلبرینگ ۲_۱	۱۶۱	۲ ساعت و ۴۱ دقیقه
بیرونی		بلبرینگ ۳_۱	۱۵۸	۲ ساعت و ۳۸ دقیقه
قفسه		بلبرینگ ۴_۱	۱۲۲	۲ ساعت و ۲ دقیقه
داخلی و بیرونی		بلبرینگ ۵_۱	۵۲	۵ دقیقه
داخلی	شرایط ۲ ۳۷/۵ هرتز، ۱۱ کیلونیوتون)	بلبرینگ ۱_۲	۴۹۱	۸ ساعت و ۱۱ دقیقه
بیرونی		بلبرینگ ۲_۲	۱۶۱	۲ ساعت و ۴۱ دقیقه
قفسه		بلبرینگ ۳_۲	۵۳۳	۸ ساعت و ۵۳ دقیقه
بیرونی		بلبرینگ ۴_۲	۴۲	۴۲ دقیقه
بیرونی		بلبرینگ ۵_۲	۳۳۹	۵ ساعت و ۳۹ دقیقه
بیرونی	شرایط ۳ ۴۰ هرتز، ۱۰ کیلونیوتون)	بلبرینگ ۱_۳	۲۵۳۸	۴۲ ساعت و ۱۸ دقیقه
بلبرینگ ۲_۳		۲۴۹۶	۴۱ ساعت و ۳۶ دقیقه	داخلی، ساجمه، قفسه و بیرونی
داخلی		بلبرینگ ۳_۳	۳۷۱	۶ ساعت و ۱۱ دقیقه
داخلی		بلبرینگ ۴_۳	۱۵۱۵	۲۵ ساعت و ۱۵ دقیقه
بیرونی		بلبرینگ ۵_۳	۱۱۴	۱ ساعت و ۵۴ دقیقه

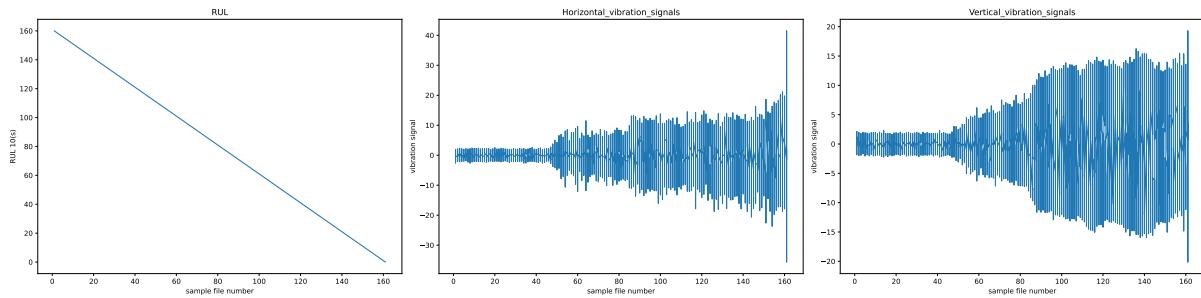
در شکل های «۶-۲» و «۷-۲» نمونه ای از داده های این مجموعه داده آورده شده است.



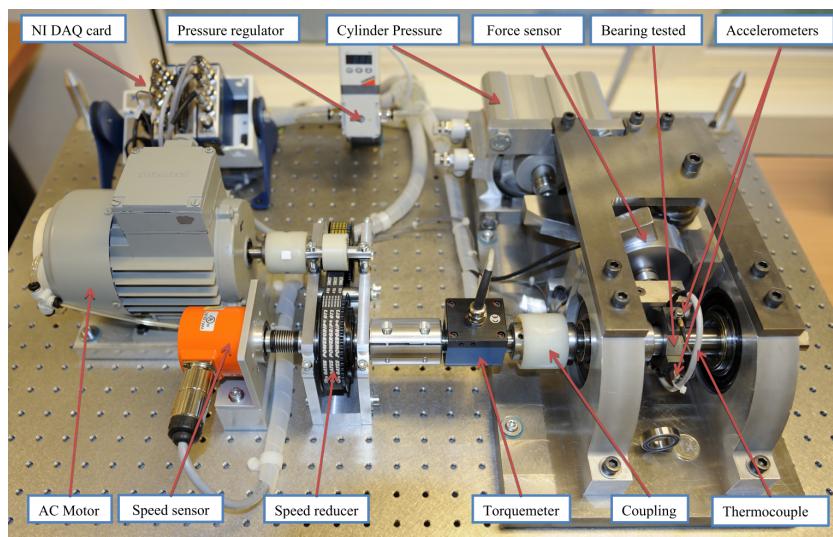
شکل ۲-۶: سیگنال های ارتعاش افقی و عمودی و RUL بلبرینگ ۱

۲-۲-۲ مجموعه داده PRONOSTIA

همانند مجموعه داده XJTY-SY این مجموعه داده [۳] نیز، داده های ثبت شده از بلبرینگ توسط بستر آزمایشی نشان داده شده در شکل «۸-۲» است و امکان انجام آزمایش از ابتدای کار بلبرینگ تا زمان خرابی و شکست کامل را فراهم می آورد.



شکل ۷-۲: سیگنال‌های ارتعاش افقی و عمودی و RUL بلبرینگ ۲-۲



شکل ۸-۲: بستر تهیه مجموعه داده PRONOSTIA [۴]

در این آزمایش، شرایط عملیاتی با اندازه‌گیری‌های لحظه‌ای ۱) نیروی شعاعی^{۱۰} اعمال شده بر بلبرینگ، ۲) سرعت چرخش شفتی که بلبرینگ را جای‌جا می‌کند و ۳) گشتاور اعمال شده به بلبرینگ تعیین می‌شوند. هر یک از این سه اندازه‌گیری که به صورت آنالوگ انجام شده است، با فرکانسی برابر با ۱۰۰ هرتز به دست آمده است.

در این مجموعه داده، از دو نوع حسگر^{۱۱} برای تشخیص و توصیف عملیات تخریب استفاده شده است. سنسورهای:

- ارتعاش

- دما

همانند مجموعه داده XJTU-SY، حسگرهای ارتعاش در این مجموعه آزمایشی نیز شامل دو شتاب‌سنج

¹⁰Radial Force

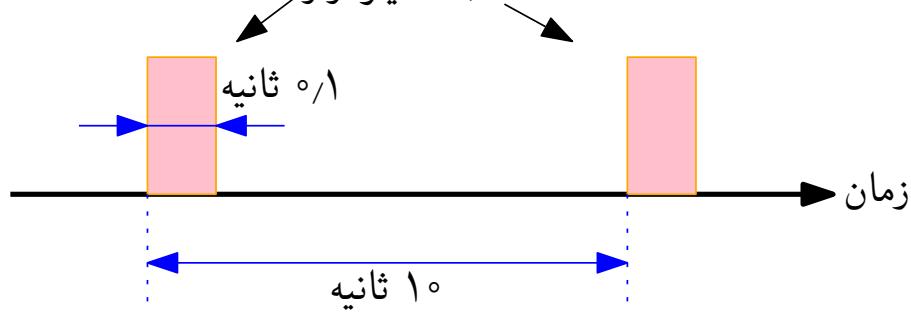
¹¹Sensor

هستند که با زاویه 90° درجه نسبت به هم قرار دارند (اولی به صورتی عمودی و دومی به صورت افقی). این دو شتاب‌سنج به صورت شعاعی روی مسیر خارجی بلبرینگ قرار گرفته‌اند.

در این مجموعه آزمایشی علاوه بر حسگرهای ارتعاش، حسگرهای دما نیز وجود دارد. حسگر دما یک پروب RTD از جنس پلاتینین است که در داخل سوراخی نزدیک به حلقه خارجی بلبرینگ قرار می‌گیرد.

حسگرهای شتاب، نمونه‌ها را با فرکانس $25/6$ کیلوهرتز و حسگر دما با فرکانس 10° هرتز ثبت می‌کنند. یعنی سیگنال‌های ارتعاشی 2560 نمونه ($\frac{1}{10}$ ثانیه) در هر 10° ثانیه ضبط می‌شوند و سیگنال‌های دما، 600 نمونه هر دقیقه ضبط می‌شوند.

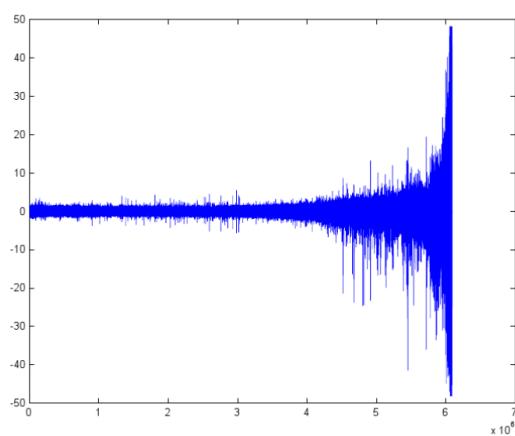
مراحل ضبط سیگنال با فرکانس نمونه‌برداری $25/6$ کیلوهرتز



شکل ۹-۲: تنظیمات نمونه‌برداری سیگنال‌ها در این آزمایش

در تهیه این مجموعه داده، برای جلوگیری از انتشار آسیب به کل بستر آزمایش، آزمایش‌ها زمانی متوقف شدند که دامنه سیگنال ارتعاش از $20g$ فراتر رفت.

شکل «۱۰-۲» مثالی است از آنچه می‌توان بر روی اجزای بلبرینگ قبل و بعد از یک آزمایش مشاهده کرد و همچنین یک سیگنال خام ارتعاش جمع‌آوری شده در طول یک آزمایش کامل را نشان می‌دهد.



شکل ۱۰-۲: سیگنال خام تهیه شده در این آزمایش و بلبرینگ تخریب شده، [۳]

ذکر این نکته الزامی است که بلبرینگ‌ها رفتارهای بسیار متفاوتی را در زمان تخریب نشان می‌دهد که منجر

به تفاوت در زمان آزمایش و تخریب می‌شود.

این آزمایش نیز تحت سه شرط عملیاتی مختلف انجام شده است:

• ۱۸۰° دور در دقیقه و بار دینامیکی ۴ کیلونیوتن

• ۱۶۵° دور در دقیقه و بار دینامیکی ۴/۲ کیلونیوتن

• ۱۵۰° دور در دقیقه و بار ۵ کیلونیوتن

داده‌های جمع‌آوری شده از این آزمایش به صورت جدول «۳-۲» تقسیم‌بندی می‌شود

جدول ۳-۲: تقسیم‌بندی داده‌های جمع‌آوری شده در مجموعه داده PRONOSTIA

مجموعه داده‌ها	شرایط عملیاتی ۱	شرایط عملیاتی ۲	شرایط عملیاتی ۳
مجموعه آموزشی	بلبرینگ ۱_۲	بلبرینگ ۱_۳	بلبرینگ ۱_۳
	بلبرینگ ۲_۲	بلبرینگ ۲_۳	بلبرینگ ۲_۳
مجموعه تست	بلبرینگ ۳_۲	بلبرینگ ۴_۲	بلبرینگ ۴_۲
	بلبرینگ ۵_۲	بلبرینگ ۵_۱	بلبرینگ ۵_۲
	بلبرینگ ۶_۲	بلبرینگ ۶_۱	بلبرینگ ۶_۲
	بلبرینگ ۷_۲	بلبرینگ ۷_۱	بلبرینگ ۷_۲

هر دو مجموعه داده‌های آموزشی و آزمایشی حاوی فایل‌هایی csv ارتعاش به نام acc_xxxxxx.csv و فایل‌های دما به نام temp_xxxxxx.csv هستند.

۳-۲-۲ مجموعه داده C-MAPSS

فصل ۳

کارهای پیشین

در فصل سوم پایان نامه، کارهای پیشین انجام شده روی مسئله به تفصیل توضیح داده می‌شود. نمونه‌ای از فصل کارهای پیشین در زیر آمده است.^۱

۱-۳ مسائل خوشبندی

مسئله‌ی خوشبندی^۲ یکی از مهم‌ترین مسائل در زمینه‌ی داده‌کاوی به حساب می‌آید. در این مسئله، هدف دسته‌بندی تعدادی شیء به‌گونه‌ای است که اشیاء درون یک دسته (خوش)، نسبت به یکدیگر در برابر دسته‌های دیگر شبیه‌تر باشند (معیارهای متفاوتی برای تشابه تعریف می‌گردد). این مسئله در حوزه‌های مختلفی از علوم کامپیوتر از جمله داده‌کاوی، جست‌وجوی الگو^۳، پردازش تصویر^۴، بازیابی اطلاعات^۵ و رایانش زیستی^۶ مورد استفاده قرار می‌گیرد [؟].

تا کنون راه حل‌های زیادی برای این مسئله ارائه شده است که از لحاظ معیار تشخیص خوشها و نحوه انتخاب یک خوش، با یکدیگر تفاوت بسیاری دارند. به همین خاطر مسئله‌ی خوشبندی یک مسئله‌ی بهینه‌سازی چندهدفه^۷ محسوب می‌شود.

همان طور که در مرجع [؟] ذکر شده است، خوش در خوشبندی تعریف واحدی ندارد و یکی از دلایل

^۱ مطالب این فصل نمونه از پایان نامه‌ی آقای بهنام حاتمی گرفته شده است.

Clustering^۸

Pattern recognition^۹

Image analysis^{۱۰}

Information retrieval^{۱۱}

Bioinformatics^{۱۲}

Multi-objective^{۱۳}

وجود الگوریتم‌های متفاوت، همین تفاوت تعریف‌ها از خوش‌های است. بنابراین با توجه به مدلی که برای خوش‌های ارائه می‌شود، الگوریتم متفاوتی نیز ارائه می‌گردد. در ادامه به بررسی تعدادی از معروف‌ترین مدل‌های مطرح می‌پردازیم:

- مدل‌های مرکزگرا: در این مدل‌ها، هر دسته با یک مرکز نشان داده می‌شود. از جمله معروف‌ترین روش‌های خوش‌بندی بر اساس این مدل، خوش‌بندی k -مرکز، خوش‌بندی k -میانگین^۸ و خوش‌بندی k -میانه^۹ است.

- مدل‌های مبتنی بر توزیع نقاط: در این مدل، دسته‌ها با فرض پیروی از یک توزیع احتمالی مشخص می‌شوند. از جمله الگوریتم‌های معروف ارائه شده در این مدل، الگوریتم بیشینه‌سازی امید ریاضی^{۱۰} است.

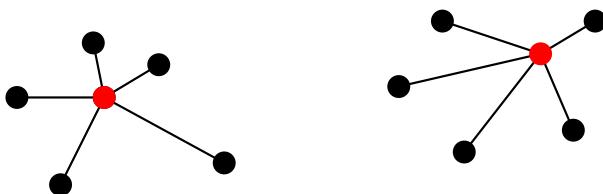
- مدل‌های مبتنی بر تراکم نقاط: در این مدل، خوش‌های متناسب با ناحیه‌های متراکم نقاط در مجموعه داده مورد استفاده قرار می‌گیرد.

- مدل‌های مبتنی بر گراف: در این مدل، هر خوش‌های مجموعه از رئوس گفته می‌شود که تمام رئوس آن با یک دیگر همسایه باشند. از جمله الگوریتم‌های معروف این مدل، الگوریتم خوش‌بندی HCS^{۱۱} است.

الگوریتم‌های ارائه شده تنها از نظر نوع مدل با یک دیگر متفاوت نیستند. بلکه، می‌توان آن‌ها را از لحاظ نحوی تخصیص نقاط بین خوش‌های نیز تقسیم‌بندی کرد:

- تخصیص قطعی داده‌ها: در این نوع خوش‌بندی هر داده دقیقاً به یک خوش‌های اختصاص داده می‌شود.
- تخصیص قطعی داده‌ها با داده‌ی پرت: در این نوع خوش‌بندی ممکن است بعضی از داده‌ها به هیچ خوش‌های اختصاص نیابد، اما بقیه داده‌ها هر کدام دقیقاً به یک خوش‌های اختصاص می‌یابد.
- تخصیص قطعی داده: در این نوع خوش‌بندی هر داده دقیقاً به یک خوش‌های اختصاص داده می‌شود.
- خوش‌بندی همپوشان: در این نوع خوش‌بندی هر داده می‌تواند به چند خوش‌های اختصاص داده شود. در گونه‌ای از این مدل، می‌توان هر نقطه را با احتمالی به هر خوش‌های اختصاص می‌یابد. به این گونه از خوش‌بندی، خوش‌بندی نرم^{۱۲} گفته می‌شود.

k-Means ^۸	
k-Median ^۹	
Expectation-maximization ^{۱۰}	
Highly Connected Subgraphs ^{۱۱}	
Soft clustering ^{۱۲}	



شکل ۱-۳: نمونه‌ای از مسئله‌ی ۲-مرکز

- خوشبندی سلسه‌مراتبی: در این نوع خوشها، داده‌ها به گونه‌ای به خوشها تخصیص داده می‌شود که دو خوشه یا اشتراک ندارند یا یکی به طور کامل دیگری را می‌پوشاند. در واقع در بین خوشها، رابطه‌ی پدر فرزندی برقرار است.

در بین دسته‌بندی‌های ذکر شده، تمرکز اصلی این پایان‌نامه بر روی مدل مرکزگرا و خوشبندی قطعی با داده‌های پرت با مدل k -مرکز است. همان‌طور که ذکر شد علاوه بر مسئله‌ی k -مرکز که به تفصیل مورد بررسی قرار می‌گیرد، k -میانگین از جمله معروف‌ترین خوشبندی‌های مدل مرکزگرا هستند. در خوشبندی k -میانه، هدف افزایش نقاط به k خوش است به گونه‌ای که مجموع مربع فاصله‌ی هر نقطه از میانه نقاط آن خوش، کمینه گردد. در خوشبندی k -میانگین، هدف افزایش نقاط به k خوش است به گونه‌ای که مجموع فاصله‌ی هر نقطه از میانگین نقاط داخل خوش (یا مرکز آن خوش) کمینه گردد.

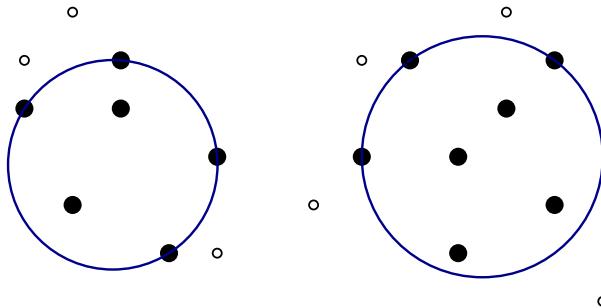
۲-۳ خوشبندی k -مرکز

یکی از رویکردهای شناخته‌شده برای مسئله‌ی خوشبندی، مسئله‌ی k -مرکز است. در این مسئله هدف، پیدا کردن k نقطه به عنوان مرکز دسته‌ها است به‌طوری‌که شعاع دسته‌ها تا حد ممکن کمینه شود. مثالی از مسئله‌ی ۲-مرکز در شکل ۱-۳ نشان داده شده است. در این پژوهش، مسئله‌ی k -مرکز با متريک‌های خاص و برای k -های کوچک مورد بررسی قرار گرفته است و هر کدام از تعریف‌ رسمي مسئله‌ی k -مرکز در زیر آمده است:

مسئله‌ی ۱-۳ (k -مرکز) گراف کامل بدون جهت $G = (V, E)$ با تابع فاصله‌ی d ، که از نامساوی مثلثی پیروی می‌کند داده شده است. زیرمجموعه‌ی $V \subseteq S$ با اندازه‌ی k را به‌گونه‌ای انتخاب کنید که عبارت زیر را کمینه کند:

$$\max_{v \in V} \left\{ \min_{s \in S} d(v, s) \right\} \quad (1-3)$$

گونه‌های مختلفی از مسئله‌ی k -مرکز با محدودیت‌های متفاوت توسط پژوهشگران مورد مطالعه قرار گرفته است. از جمله‌ی این گونه‌ها، می‌توان به حالتی که در بین داده‌های ورودی، داده‌های پرت وجود دارد، اشاره کرد.



شکل ۲-۳: نمونه‌ای از مسئله‌ی ۲-مرکز با داده‌های پرت

در واقع در این مسئله، قبل از خوشبندی می‌توانیم تعدادی از نقاط ورودی را حذف نموده و سپس به خوشبندی نقاط بپردازیم. سختی این مسئله از آن جاست که نه تنها باید مسئله‌ی خوشبندی را حل نمود، بلکه در ابتدا باید تصمیم گرفت که کدام یک از داده‌ها را به عنوان داده‌ی پرت در نظر گرفت که بهترین جواب در زمان خوشبندی به دست آید. در واقع اگر تعداد نقاط پرتی که مجاز به حذف است، برابر صفر باشد، مسئله به مسئله‌ی k -مرکز تبدیل می‌شود. نمونه‌ای از مسئله‌ی ۲-مرکز با ۷ داده‌ی پرت را در شکل ۲-۳ می‌توانید ببینید. تعریف دقیق‌تر این مسئله در زیر آمده است:

مسئله‌ی ۲-۳ (۲-۳ مرکز با داده‌های پرت) یک گراف کامل بدون جهت $G = (V, E)$ با تابع فاصله‌ی d ، که از نامساوی مثلثی پیروی می‌کند داده‌شده است. زیرمجموعه‌ی $Z \subseteq V$ با اندازه‌ی z و مجموعه‌ی $S \subseteq V - Z$ با اندازه‌ی k را انتخاب کنید به طوری که عبارت زیر را کمینه کند:

$$\max_{v \in V - Z} \left\{ \min_{s \in S} d(v, s) \right\} \quad (2-3)$$

گونه‌ی دیگری از مسئله‌ی k -مرکز که در سال‌های اخیر مورد توجه قرار گرفته است، حالت جویبار داده‌ی آن است. در این گونه از مسئله‌ی k -مرکز، در ابتدا تمام نقاط در دسترس نیستند، بلکه به مرور زمان نقاط در دسترس قرار می‌گیرند. محدودیت دومی که وجود دارد، محدودیت حافظه است، به طوری که نمی‌توان تمام نقاط را در حافظه نگه داشت و بعضی حتی امکان نگه‌داری در حافظه‌ی جانبی نیز وجود ندارد و به طور معمول باید مرتبه‌ی حافظه‌ای کمتر از مرتبه‌ی حافظه‌ی خطی^{۱۳} متناسب با تعداد نقاط استفاده نمود. از این به بعد به چنین مرتبه‌ای، مرتبه‌ی زیرخطی^{۱۴} می‌گوییم. مدلی که ما در این پژوهش بر روی آن مرکز داریم مدل جویبار داده تک‌گذره^{۱۵} [?] است. یعنی تنها یک بار می‌توان از ابتدا تا انتهای داده‌ها را بررسی کرد و پس از عبور از یک داده، اگر آن داده در حافظه ذخیره نشده باشد، دیگر به آن دسترسی وجود ندارد. علاوه بر این، در هر لحظه باید بتوان به پرسمان (برای تمام نقاطی از جویبار داده که تاکنون به آن دسترسی داشته‌ایم) پاسخ داد.

Linear^{۱۳}

sublinear^{۱۴}

Single pass^{۱۵}

مسئله‌ی ۳-۳ (k-مرکز در حالت جویبار داده) مجموعه‌ای از نقاط در فضای d -بعدی به مرور زمان داده می‌شود. در هر لحظه از زمان، به ازای مجموعه‌ی U از نقاطی که تا کنون وارد شده‌اند، زیرمجموعه‌ی $S \subseteq U$ با اندازه‌ی k را انتخاب کنید به طوری که عبارت زیر کمینه شود:

$$\max_{u \in U} \left\{ \min_{s \in S} d(u, s) \right\} \quad (3-3)$$

از آنجایی که گونه‌ی جویبار داده و داده پرت مسئله‌ی k -مرکز به علت بهروز بودن مبحث داده‌های حجیم^{۱۵}، به تازگی مورد توجه قرار گرفته است. در این تحقیق سعی شده است که تمرکز بر روی این گونه‌ی خاص از مسئله باشد. همچنین در این پژوهش سعی می‌شود گونه‌های مسئله را برای انواع متريک‌ها و برای n ‌هاي کوچک نيز مورد بررسی قرار داد.

۳-۳ مدل جویبار داده

همان‌طور که ذکر شد مسئله‌ی k -مرکز در حالت داده‌های پرت و جویبار داده، گونه‌های تعمیم‌یافته از مسئله‌ی k -مرکز هستند و در حالت‌های خاص به مسئله‌ی k -مرکز کاهش پیدا می‌کنند. مسئله‌ی k -مرکز در حوزه‌ی مسائل ان‌پی-سخت^{۱۶} قرار می‌گیرد و با فرض $P \neq NP$ الگوریتم دقیق با زمان چندجمله‌ای برای آن وجود ندارد [؟]. بنابراین برای حل کارای^{۱۷} این مسائل از الگوریتم‌های تقریبی^{۱۸} استفاده می‌شود.

برای مسئله‌ی k -مرکز، دو الگوریتم تقریبی معروف وجود دارد. در الگوریتم اول، که به روش حریصانه^{۱۹} عمل می‌کند، در هر مرحله بهترین مرکز ممکن را انتخاب می‌کند به طوری تا حد ممکن از مراکز قبلی دور باشد [؟]. این الگوریتم تقریبی با ضریب تقریب ۲ ارائه می‌دهد. در الگوریتم دوم، با استفاده از مسئله‌ی مجموعه‌ی غالب کمینه^{۲۰}، الگوریتمی با ضریب تقریب ۲ ارائه می‌گردد [؟]. همچنین ثابت شده است، که بهتر از این ضریب تقریب، الگوریتمی نمی‌توان ارائه داد مگر آنکه $P = NP$ باشد.

برای مسئله‌ی k -مرکز در حالت جویبار داده برای ابعاد بالا، بهترین الگوریتم موجود ضریب تقریب $\epsilon + 2$ دارد [؟، ？، ？] و ثابت می‌شود الگوریتمی با ضریب تقریب بهتر از ۲ نمی‌توان ارائه داد. برای مسئله‌ی k -مرکز با داده‌ی پرت در حالت جویبار داده نیز، بهترین الگوریتم ارائه شده، الگوریتمی با ضریب تقریب $\epsilon + 4$ است که با کران پایین ۳ هنوز اختلاف قابل توجهی دارد [？].

Big data^{۱۶}

NP-hard^{۱۷}

Efficient^{۱۸}

Approximation algorithm^{۱۹}

Greedy^{۲۰}

Dominating set^{۲۱}

جدول ۱-۳: نمونهایی از کران پایین تقریب‌پذیری مسائل خوشبندی

مسئله	کران پایین تقریب‌پذیری
k - مرکز	$[?] 2$
k - مرکز در فضای اقلیدسی	$[?] 1/822$
۱- مرکز در حالت جویبار داده	$[?] \frac{1+\sqrt{2}}{2}$
k - مرکز با نقاط پرت و نقاط اجباری	$[?] 3$

برای k ‌های کوچک به خصوص، $1, 2 = k$ ، الگوریتم‌های بهتری ارائه شده است. بهترین الگوریتم ارائه شده برای مسئله ۱- مرکز در حالت جویبار داده برای ابعاد بالا، دارای ضریب تقریب $1/22$ است و کران پایین $\frac{1+\sqrt{2}}{2}$ نیز برای این مسئله اثبات شده است [؟، ?]. برای مسئله ۲- مرکز در حالت جویبار داده برای ابعاد بالا، اخیرا راه حلی با ضریب تقریب $\epsilon + 1/8$ ارائه شده است [؟]. برای مسئله ۱- مرکز با داده‌ی پرت، تنها الگوریتم موجود، الگوریتمی با ضریب تقریب $1/73$ است [？].

۴-۳ تقریب‌پذیری

یکی از راهکارهایی که برای کارآمد کردن راه حل ارائه شده برای یک مسئله وجود دارد، استفاده از الگوریتم‌های تقریبی برای حل آن مسئله است. یکی از عمدترين دغدغه‌های مطرح در الگوریتم‌های تقریبی کاهش ضریب تقریب است. در بعضی از موارد حتی امکان ارائه الگوریتم تقریبی با ضریبی ثابت نیز وجود ندارد. به طور مثال، الگوریتم تقریبی با ضریب تقریب کمتر از ۲، برای مسئله k - مرکز وجود ندارد مگر این‌که $P = NP$ باشد. برای مسائل مختلف، معمولاً می‌توان کران پایینی برای میزان تقریب‌پذیری آن‌ها ارائه داد. در واقع برای برخی مسائل انپی-سخت، علاوه بر این که الگوریتم کارآمدی وجود ندارد، بعضًا الگوریتم تقریبی با ضریبی تقریب کم و نزدیک به یک نیز وجود ندارد. در جدول ۱-۳ میزان تقریب‌پذیری مسائل مختلفی که در این پایان‌نامه مورد استفاده قرار می‌گیرد را می‌بینید.

فصل ۴

چالش‌ها و نوآوری‌ها

در این فصل نتایج جدید به دست آمده در پایان‌نامه توضیح داده می‌شود. در صورت نیاز می‌توان نتایج جدید را در قالب چند فصل ارائه نمود. همچنین در صورت وجود پیاده‌سازی، بهتر است نتایج پیاده‌سازی را در فصل مستقلی پس از این فصل قرار داد.

فصل ۵

نتیجه‌گیری

در این فصل، ضمن جمع‌بندی نتایج جدید ارائه شده در پایان‌نامه یا رساله، مسائل باز باقی‌مانده و همچنین پیشنهادهایی برای ادامه‌ی کار ارائه می‌شوند.

Bibliography

- [1] Y. Wei and D. Wu. Conditional variational transformer for bearing remaining useful life prediction. *Advanced Engineering Informatics*, 59:102247, 2024.
- [2] B. Wang, Y. Lei, N. Li, and N. Li. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability*, 69(1):401–412, 2018.
- [3] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, and C. Varnier. Pronostia: An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM’12.*, pages 1–8. IEEE Catalog Number: CPF12PHM-CDR, 2012.

واژه‌نامه

ت

experimental	تجربی	الف
density	تراکم	ابعاد بالا
approximation	تقريب	اریب
partition	تقسیم‌بندی	آستانه
mesh	توری	اصل لانه‌ی کبوتری
distributed	توزيع شده	ان‌پی-سخت
		NP-Hard
		انتقال

ج

separable	جداپذیر	ب
black box	جعبه سیاه	برخط
data stream	جویبار داده	برنامه‌ریزی خطی
		بهینه
		maximum

ح

extreme	حدی	پ
greedy	حریصانه	پرت
		پرسمان
		پوشش
cluster	خوشه	پیچیدگی
linear	خطی	complexity

ف

distance	فاصله
space	فضا

ق

deterministic	قطعی
---------------------	------

د

داده	داده
داده‌کاوی	داده‌کاوی
داده‌پرت	داده‌پرت
دوبابرازی	دوبابرازی
دودویی	دودویی

ک

efficient	کارا
candidate	کاندیدا
minimum	کمینه

ر

vertex	رأس
formal	رسمی

ز

sublinear	زیرخطی
-----------------	--------

م

set	مجموعه
coreset	مجموعه هسته
planar	مسطح
parallelization	موازی‌سازی
buffer	میانگیر

س

amortized	سرشکن
hierarchichal	سلسه‌مراتبی

ش

pseudocode	شبه کد
object	شیء
center point	نقطه‌ی مرکزی
half space	نیم‌فضا

ص

satisfiability	صدق‌پذیری
----------------------	-----------

ن

inversion	نابه‌جایی
invariant	ناوردا

ش

object	شبه کد
dominate	شبه کد

ه

price of anarchy (POA)	هزینه‌ی آشوب
------------------------------	--------------

غ

dominate	غلبه
----------------	------

ی

edge	یال
------------	-----

آپیوست

مطالب تکمیلی

پیوست‌های خود را در صورت وجود می‌توانید در این قسمت قرار دهید.

Abstract

In industrial environments and factories, one of the most important concerns is the maintenance and repair of tools and industrial equipment. Each piece of equipment plays a crucial role in the production line of any company, and the smallest damage or malfunction can disrupt the entire production process. Therefore, it is essential to continuously monitor the current health status of the equipment and predict the time of potential failures and the remaining useful life.

The prediction of Remaining Useful Life (RUL) is a critical task in the field of Prognostics and Health Management (PHM) for industrial tools and equipment. Accurate RUL prediction is vital because it enables timely maintenance and repair, reduces downtime, and allows for the optimization of operational efficiency by loading the equipment according to its remaining useful life.

Keywords: Remaining Useful Life, RUL, Transformer, FPGA



Amirkabir University of Technology

(Tehran Polytechnic)

Department of Computer Engineering

M.Sc. Seminar Report

FPGA-Based Hardware Acceleration of Remaining Useful Life Prediction of Rotating Machinery Using Transformer Neural Network

By:

Reza Adinepour

Supervisor:

Prof. Morteza Saheb Zamani

August 2024