نشریه علمی "پدافند الکترونیکی و سایبری"
سال هفتم، شماره ۴، زمستان ۱۳۹۸، ص ۱۴۵-۱۳۱

# زمان‌بندی گردش‌کار در محیط ابر ترکیبی با در نظر گرفتن امنیت کارها و ارتباطات با الگوریتم ازدحام ذرات بهبودیافته

مائده مهرآوران[۱]، محمدرضا پژوهان [۲*]، فضل الله ادیب نیا[۳]
۱- دانشجوی دکتری مهندسی کامپیوتر، دانشگاه یزد، ۲ و ۳- استادیار، دانشکده مهندسی کامپیوتر، دانشگاه یزد

(دریافت: ۹۷/۰۶/۱۳، پذیرش: ۹۷/۱۲/۱۴)

## چکیده

در حالی که منابع ابر خصوصی برای اجرای درخواست‌ها، کاهش هزینه و امنیت بیشتر اطلاعات را به دنبال دارد، استفاده از ابر عمومی علاوه‌بر هزینه، مخاطرات احتمالی در حفاظت از اطلاعات سازمان را نیز به همراه دارد. اما نیاز سازمان‌ها به منابع با کارایی و ظرفیت ذخیره‌سازی بالا، آن‌ها را ناگزیر به استفاده از ابر عمومی می‌کند. بنابراین زمان‌بندی درخواست‌ها به منابع، یکی از مسائل مهم در محاسبات ابری است. در این مقاله روش جدیدی پیشنهاد می‌شود که به زمان‌بندی کارها با در نظر گرفتن ملاحظات امنیتی می‌پردازد. ملاحظات امنیتی شامل حساسیت برای کارها که در پژوهش‌های اخیر در نظر گرفته شده، حساسیت برای داده‌های انتقالی بین کارها و همچنین ایده اصلی در نظر گرفتن قدرت امنیتی برای منابع و مسیرهای ارتباطی بین آن‌ها می‌باشد. سناریوی پیشنهادی توسط الگوریتم PSO بهبودیافته (PSO-WSCS) پیاده‌سازی می‌شود. تابع هدف، حداقل‌کردن فاصله امنیتی کارها و داده‌ها از قدرت امنیتی منابع و ارتباطات است؛ به‌طوری‌که دو محدودیت زمان و هزینه نیز برآورده شود. الگوریتم پیشنهادی PSO-WSCS که تغییراتی روی الگوریتم PSO اصلی می‌دهد، با سه الگوریتم دیگر زمان‌بندی مطرح و مشابه VNPSO، MPSO و MPSO-SA با در نظر گرفتن امنیت در محیط ابر ترکیبی مقایسه می‌گردد. نتایج ارزیابی حاکی از مؤثر بودن الگوریتم پیشنهادی در یافتن منابع با شباهت امنیتی نزدیک به نیازهای امنیتی می‌باشد. به‌طور متوسط، بهبود ۴۰ درصدی در نمونه‌های در نظر گرفته‌شده این مهم را نشان می‌دهد.

**کلیدواژه‌ها**: ابر ترکیبی، زمان‌بندی جریان کارها، نیاز امنیتی کار، نیاز امنیتی داده، قدرت امنیتی منابع، قدرت امنیتی مسیر ارتباطی

## ۱. مقدمه

در محیط محاسبات ابری کاربران می‌توانند از منابع موجود بر اساس تقاضا استفاده کنند. مسئله زمان‌بندی کارها[۱]، راهبرد مناسب در نحوه پاسخ‌دهی به درخواست‌ها و توزیع صحیح کارها میان منابع موجود با توجه به ویژگی‌ها و توانایی منابع و معیارهای کارایی است. مجموعه کارها با یک گراف جهت‌دار مشخص می‌شود که گره‌های آن مجموعه کارها و یال‌ها، نشان‌دهنده وابستگی بین کارها است. معمولاً میزان داده برای انتقال بین دو کار به‌وسیله وزن یال مشخص می‌شود [۲- ۱].

محیط ابری را به سه دسته ابر خصوصی[۲]، عمومی[۳] و ترکیبی[۴] دسته‌بندی می‌کنند که هر کدام از دسته‌ها به شرح زیر بیان می‌شود[۳].

ابر خصوصی: ابر خصوصی که به ابر داخلی هم مشهور است، جهت استفاده اختصاصی سازمان در نظر گرفته شده و سازمان مالک منابع آن است. منابع ابر خصوصی محدود می‌باشد.

ابر عمومی: در محیط ابر عمومی منابع نامحدود موجود می‌باشد اما در آن، بر اساس میزان استفاده کاربر، هزینه محاسبه می‌شود.

ابر ترکیبی: ابر ترکیبی، ترکیب ابر عمومی و خصوصی است. درصورتی‌که منابع موجود در ابر خصوصی کافی نباشد از منابع ابر عمومی استفاده می‌شود. از لحاظ دسترسی به منابع، ابر ترکیبی از همه بهتر است چون علاوه بر مزایای ابر خصوصی (عدم پرداخت هزینه و امنیت در صورت موجود بودن منابع)، ویژگی نامحدود بودن منابع ابر عمومی را نیز دارد.

مسئله زمان‌بندی در ابر ترکیبی با چالش‌هایی روبرو است. درواقع استفاده از منابع خارج از سازمان چالش‌هایی چون ملاحظات امنیتی، حفظ حریم خصوصی و نحوه تخصیص‌دهی منابع به کارها [۴] را به وجود می‌آورد [۵]. یکی از چالش‌های

*رایانامه نویسنده مسئول: pajoohan@yazd.ac.ir

[1] Task Scheduling
[2] Private Cloud
[3] Public Cloud
[4] Hybird Cloud

اصلی که سازمان‌ها در مواجه شدن با محیط ابر دارند، مسئله امنیت است. در IDC[1] بحث امنیت و محرمانگی یکی از مهم‌ترین چالش‌هایی است که محیط ابر را تهدید می‌کند. همچنین یکی از مسائل قابل توجه و پیچیده، مسئله زمان‌بندی کارها در محیط ابر ترکیبی است [۶].

مسئله زمان‌بندی کارها راهبرد تخصیص کارها به منابع است که در این مقاله با در نظر گرفتن ملاحظات امنیتی که در ادامه توضیح داده می‌شود، در محیط ابر ترکیبی انجام خواهد شد.

در پژوهش‌های شریف، ابریشمی و رضاییان [۷-۹] مبحث امنیت در کارها لحاظ شده است و بر این اساس، کارها به دودسته حساس و غیر حساس تقسیم می‌شوند. کارهای حساس ترجیحاً باید در داخل سازمان و ابر خصوصی انجام شود و کارهای غیر حساس می‌تواند در ابر خصوصی یا در ابر عمومی انجام شود. با توجه به اینکه حساسیت کارها وابسته به داده‌هایی است که با آن کار می‌کنند، در اینجا علاوه بر تعریف حساسیت برای کارها، برای داده‌ها نیز حساسیت تعریف می‌شود. درواقع داده‌های ورودی و خروجی هر کار می‌تواند درجه حساسیت مختلفی داشته باشد. همچنین با توجه به حملات امنیتی متفاوتی که در دنیای شبکه وجود دارد، مثل حملات شنود و افشا، تخریب و حملات جعل هویت، می‌توان قدرت امنیتی برای منابع و مسیرهای ارتباطی بین آن‌ها تعریف کرد. قدرت امنیتی، بر اساس خدمات مختلف امنیتی که ارائه می‌کنند (محرمانگی، صحت و احراز هویت) مشخص می‌شود[۱۰]. برای ارائه سرویس‌های امنیتی، الگوریتم‌های گوناگونی به وجود آمده است که این الگوریتم‌ها ازنظر قدرت امنیتی و میزان سرباری که ایجاد می‌کند، متفاوت هستند. به‌طور مسلم هرچه الگوریتم قدرت امنیتی بالاتری داشته باشد، سربار بیشتری برای اجرای آن الگوریتم به وجود می‌آید و درنتیجه زمان افزایش پیدا می‌کند.

با توجه به اینکه جهت مسئله زمان‌بندی راه‌حل چندجمله‌ای وجود ندارد [۱۱]، الگوریتم‌های مختلفی برای مسئله زمان‌بندی پیشنهادشده است[۱۲]. این الگوریتم‌ها به سه دسته تقسیم می‌شوند[۱۳]. الف- الگوریتم‌های مکاشفه‌ای ب- الگوریتم‌های فرا مکاشفه‌ای ج- الگوریتم‌های ترکیبی.

آنچه این مقاله را از کارهای پیشین متمایز می‌کند، در نظر گرفتن نیاز امنیتی برای کارها و داده‌های انتقالی بین آن‌ها و همچنین قدرت امنیتی برای منابع و مسیرهای ارتباطی بین منابع است. همچنین تابع هدف مسئله میزان تشابه امنیتی را مدنظر دارد که از الگوریتم فرا مکاشفه‌ای بهبودیافته PSO برای حل مسئله استفاده‌شده است.

ادامه این مقاله به‌صورت زیر ساخت‌یافته است: در بخش دوم کارهای مرتبط که بحث امنیت را در نظر گرفته‌اند مرور می‌شود. روش پیشنهادی در بخش سوم تشریح می‌گردد. در بخش چهارم روش پیشنهادی و روش‌های مشابه مورد ارزیابی و مقایسه قرار می‌گیرد.

## ۲. کارهای پیشین

لیو و همکاران [۱۴] مدل محدودیت امنیتی[2] تعریف کردند که سه حالت امنیتی برای زمان‌بندی کارها روی منابع در نظر می‌گیرد. بر اساس این مدل برای کارها و منابع سطح امنیتی تعریف می‌شود. اگر کار فقط روی منابع با سطح امنیتی بالاتر زمان‌بندی شود، آنگاه حالت امنیتی، امن گفته می‌شود. اگر کارها برای زمان‌بندی محدودیتی نداشته باشند، حالت امنیتی ریسکی گفته می‌شود و اگر کارها روی منابعی با احتمال ریسک $\gamma$ $(0<\gamma<1)$ زمان‌بندی شود، حالت امنیتی با ریسک $\gamma$ گویند. این مدل با روش فرا مکاشفه‌ای ازدحام ذرات حل می‌شود.

لیو و همکاران [۱۵] در ادامه سرویس‌های امنیتی جهت محافظت از داده‌های حساس در جریان کار تعریف می‌کنند. موارد تهدید برای داده‌ها عبارت از تغییر[3]، شنود[4] و جاسوسی[5] داده‌ها هستند که برای مقابله با این تهدیدها از سرویس احراز هویت[6]، سرویس محرمانگی[7] و سرویس جامعیت[8] استفاده‌شده است. مدلی نیز جهت سرویس‌های امنیتی مراکز داده و سرویس‌های امنیتی موردنیاز کارها ارائه می‌کند که کارها روی مراکز داده‌ای متناسب با نیازهای امنیتی زمان‌بندی می‌شود.

چن و همکاران[۱۰] روشی ارائه کردند که با اجرای تکراری کارها روی منابع از انتقال داده‌های حساس بین منابع جلوگیری می‌شود و همچنین با اجرای کارها روی منبع واحد نیازی به رمزنگاری داده‌های میانی نیست. بدین ترتیب زمان موردنیاز برای رمزنگاری از زمان کل اجرای کارها کسر و زمان کل کاهش می‌یابد.

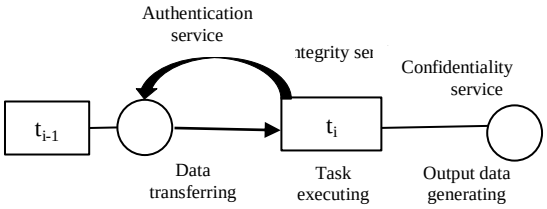
در ادامه لی و همکاران [۱۶] مدلی دیگر جهت سربار امنیتی کارها و مبحث مربوط به تهدیدات امنیتی که ممکن است

[1] International Data Corporation

[2] Security constraint
[3] Alteration
[4] spoofing
[5] snooping
[6] Authentication
[7] Confidentiality
[8] Integrity Service

مجموعه کارها را در محیط ابر مورد تهدید قرار دهد، مطرح می‌کنند. در این سیستم از سه سرویس امنیتی شامل سرویس احراز هویت، محرمانگی و جامعیت استفاده‌شده است.

در شکل (۱) سرویس‌های مختلف ارائه‌شده در فرآیند اجرای کار نشان داده‌شده است.




**شکل (۱)**: سرویس‌های امنیتی[۱۶]

همان‌طور که در شکل (۱) می‌بینیم، هر کار به سه نوع سرویس شامل سرویس احراز هویت، جامعیت و محرمانگی نیاز دارد. این سرویس‌ها با سطوح امنیتی مختلفی که باید توسط کاربر مشخص شود، تعیین می‌شوند. مجموعه $sr_i = [sr_i^1, sr_i^2, \ldots, sr_i^l, \ldots, sr_i^q]$ سطوح امنیتی کار i ام را نشان می‌دهد که q برابر ۳ است. (سه سطح را نشان می‌دهد). با در نظر گرفتن سرویس‌های امنیتی گفته‌شده سیستم دارای سربار می‌شود که برای اعمال این سرویس‌ها زمان و هزینه افزایش پیدا می‌کند. جداول ۱، ۲ و ۳ الگوریتم‌های سرویس امنیتی را نشان می‌دهد. در جدول (۱) نمونه‌ای از الگوریتم‌های رمزنگاری را که برای سرویس محرمانگی استفاده می‌شود، نشان داده است. جدول (۲) توابع درهم‌سازی[1] را برای بررسی سرویس جامعیت نشان می‌دهد و جدول (۳) الگوریتم‌های احراز هویت را برای سرویس احراز هویت مشخص می‌کند. ضرایب بیان‌شده در جداول بین صفر و یک قرار دارد که صفر کمترین و یک بیشترین قدرت الگوریتم را نشان می‌دهد. هرچه قدرت الگوریتم بیشتر باشد، نیاز به محاسبات پیچیده‌تری دارد و هزینه اجرای آن بیشتر و سرعت اجرا کمتر می‌شود.

**جدول (۱)**: الگوریتم رمزنگاری [۱۶]

| الگوریتم رمزنگاری | قدرت امنیتی | سرعت (Mb/s) | سربار (KB/ms) |
|---|---|---|---|
| IDEA | ۱/۰ | ۱۷/۳۴ | ۱۳/۵ |
| DES | ۰/۹ | ۱۸/۲۱ | ۱۵ |
| Rijndael | ۰/۶۴ | ۳۹/۸۸ | ۲۱/۰۹ |
| RC4 | ۰/۳۶ | ۳۲/۹۸ | ۹۶/۴۳ |

[1] Hash

**جدول (۲): الگوریتم احراز هویت** [۱۰]

| الگوریتم احراز هویت | قدرت امنیتی | سرعت(Mb/s) |
|---|---|---|
| CBC-MAC-AES | ۰/۹ | ۱۶۳ |
| HMAC-SHA-1 | ۰/۶ | ۱۴۸ |
| HMAC-MD5 | ۰/۳ | ۹۰ |

**جدول (۳)**: توابع درهم سازی [۱۰]

| نام تابع | قدرت امنیتی | سرعت(Mb/s) |
|---|---|---|
| TIGER | ۱/۰ | ۴۸/۳ |
| RIFDMD-160 | ۰/۷۷ | ۷۱/۲۷ |
| SHA-1 | ۰/۶۳ | ۸۰/۶ |
| RIFDMD-128 | ۰/۳۶ | ۸۶/۹۷ |
| MD5 | ۰/۲۶ | ۱۳۸/۱۲ |

لی در سال ۲۰۱۶ [۱۶] مسئله زمان‌بندی را با محدودیت‌های زمان تعریف‌شده توسط کاربر و مخاطره امنیتی بیان می‌کند. هدف در این مسئله، حداقل کردن هزینه‌ها است. در این مقاله از الگوریتم مورچگان که از دسته الگوریتم‌های فرا مکاشفه‌ای است، استفاده‌شده است.

در مقاله ابریشمی و همکاران [۱۷] زمان‌بندی کارها در محیط ابر ترکیبی شامل منابع ابر عمومی و ابر خصوصی انجام می‌شود. آقای ابریشمی با در نظر گرفتن حساسیت برای کارها، مسئله زمان‌بندی را مطرح می‌کند. کارهای حساس باید در محیط ابر خصوصی انجام شود و کارهای غیر حساس می‌تواند در محیط ابر عمومی یا خصوصی زمان‌بندی شود. مدل پیشنهادی به‌وسیله روش مکاشفه‌ای حل می‌شود.

یکی از روش‌های موفق در زمان‌بندی سامانه‌های ابر ترکیبی که برای پردازشگرهای چندهسته‌ای طراحی‌شده است، الگوریتم زمان‌بندی "بهینه‌شده هزینه برای ابرهای ترکیبی"(HCOC) است که توسط بیتکورت و همکاران [۱۸] ارائه گردیده است. این روش مهلت زمانی دریافت کرده و سعی می‌کند که با حداقل هزینه زمان‌بندی، گردشکاری را قبل از مهلت تعیین‌شده، انجام دهد.

شریف و همکاران [۷] مبحث محافظت از محرمانگی داده‌ها

را در مسئله زمان‌بندی مطرح می‌کند. در این مسئله علاوه‌بر در نظر گرفتن محدودیت زمان و هزینه، سه سطح محرمانگی برای کارها و منابع تعریف می‌شود. سطح اول محرمانگی شامل کارهایی است که قابلیت اجرا در هر دو محیط ابر را دارد. سطح دوم محرمانگی مربوط به کارهایی می‌شود که روی منابع خصوصی یا بعضی منابع عمومی قابل انجام هستند. سطح سوم، کارهایی هستند که فقط قابلیت زمان‌بندی روی منابع خصوصی را دارند. مجموعه سطوح محرمانگی به‌وسیله مجموعه $\Gamma_t = \{\tau_{t1}, \tau_{t2}, \tau_{t3}\}$ نشان داده‌شده است. همچنین سه دسته محرمانگی $\Gamma_s = \{\tau_{s1}, \tau_{s2}, \tau_{s3}\}$ برای منابع در نظر گرفته می‌شود. دسته اول منابع عمومی، دسته دوم و سوم منابع خصوصی و بعضی منابع عمومی را شامل می‌شود.

در شکل (۴) مشاهده می‌کنیم که منابع دسته اول فقط می‌تواند کارهای سطح اول محرمانگی را اجرا کند. منابع دسته دوم می‌تواند کارهای موجود در سطح دو و سه محرمانگی را اجرا کنند. منابع دسته سوم قابلیت اجرای همه کارها را دارد. در ادامه از الگوریتم‌های مسیر بحرانی [۱۹] استفاده می‌شود تا زمان‌بندی انجام شود.

**جدول (۴)**: نحوه تخصیص کارها به منابع [۷]

| | $\tau_{t1}$ | $\tau_{t2}$ | $\tau_{t3}$ |
|---|---|---|---|
| $\tau_{s1}$ | ۱ | ۰ | ۱ |
| $\tau_{s2}$ | ۱ | ۱ | ۰ |
| $\tau_{s3}$ | ۱ | ۱ | ۱ |

ابریشمی و همکاران روش جدیدی پیشنهاد کردند [۸] که با استفاده از معماری ابر ترکیبی، محرمانگی کارهای خصوصی و حساس را برآورده می‌کند. در این روش، کارهای حساس روی ابر خصوصی که تحت کنترل سازمان است انجام می‌شود و کارهایی که حساسیت خاصی ندارند، مختار به انتخاب بین منابع ابر خصوصی و ابر عمومی هستند. در این مسئله هدف مینیمم کردن پارامترهای زمان کل کارها و هزینه است.

در مقاله سوزی و همکاران [۲۰] الگوریتم زمان‌بندی در محیط ابر چندگانه مطرح‌شده است. ابر چندگانه که با ابر ترکیبی متفاوت است، شامل چند ابر مختلف است که در هر ابر منابع مختلفی قرار دارد. منابع موجود در یک ابر خاص، دارای پهنای باند بیشتری برای تبادل داده‌ها هستند. بنابراین سعی می‌شود کارهایی که تبادل داده‌های بیشتری دارند ترجیحاً در یک محیط ابر اجرا شوند. این الگوریتم ابتدا مسیرهایی که تبادل داده‌های زیادی دارند را به‌عنوان مسیر بحرانی انتخاب و سعی می‌کند تا میزان تبادل داده‌ها بین مسیرهای بحرانی زیاد نباشد. درنهایت مسیرهای به‌دست‌آمده به ابر مناسب اختصاص داده می‌شوند. هدف این مسئله کاهش هزینه اجرای کارها در محدودیت زمانی تعریف‌شده توسط کاربر است.

فرناندز و همکاران [۲۱] با استفاده از الگوریتم فرا مکاشفه‌ای ژنتیک سعی در حداقل کردن انرژی مصرفی زمان‌بندی کارها روی منابع و همچنین مقدار زمان اجرای کل کارها را دارد. در این مقاله محدودیت امنیتی برای کارها و منابع در نظر گرفته‌شده است.

ون و همکاران [۲۲] سعی در حداقل کردن دو هدف مهم الگوریتم‌های زمان‌بندی یعنی زمان اجرای کارها و هزینه اجرای کارها روی منابع را دارد. جهت دست‌یابی به این اهداف از الگوریتم ژنتیک چندهدفه استفاده می‌کند که محدودیت تابع احتمال خطر امنیتی برای کار و منبع در نظر گرفته‌شده است.

الگوریتم زمان‌بندی ارائه‌شده در مقاله ابراهام و همکاران [۲۳] که با استفاده از الگوریتم ازدحام ذرات همسایگی متغییر پیاده‌سازی شده است، هدف را حداقل کردن زمان اجرای کل کارها در نظر می‌گیرد. این الگوریتم سال ۲۰۱۶ با الگوریتم بهبودیافته ازدحام ذرات ارائه‌شده توسط نایدو و همکاران [۲۴] مقایسه می‌شود. نتایج نشان می‌دهد که الگوریتم جدید اهداف مسئله را با در نظر گرفتن محدودیت امنیتی بهبود قابل‌توجهی داده است.

در پژوهش‌های قبلی روی مبحث امنیت و محرمانگی گردش کارها تمرکز شده و اغلب مقالات امنیت را به‌عنوان محدودیت برای مسئله تعریف کرده‌اند. با توجه به این‌که مبحث امنیت برای منابع و مسیرهای ارتباطی نیز بسیار بااهمیت است، در این مقاله قصد داریم با در نظر گرفتن حساسیت برای کارها و داده‌هایی که به‌عنوان ورودی و خروجی دارند (حساسیت همان نیاز امنیتی است)، همچنین در نظر گرفتن قدرت امنیتی برای منابع و مسیرهای ارتباطی، منابعی را جهت زمان‌بندی در نظر بگیریم که سطح امنیتی مناسبی داشته باشند. روش پیشنهادی در این مقاله روی این موضوع تمرکز می‌کند و مسئله را با الگوریتم PSO بهبودیافته پیاده‌سازی می‌کند. در ادامه روش پیشنهادی و ملاحظات امنیتی بیان می‌شود.

## ۳. روش پیشنهادی

با توجه به آنچه در کارهای پیشین مطرح‌شده، مبحث امنیت در زمان‌بندی کارها در ابر بسیار بااهمیت است، در برخی تحقیقات پیشین [۱۰, ۱۴, ۱۵] با در نظر گرفتن امنیت برای کارها و مراکز داده، به این موضوع پرداخته شده است. در این مقاله

علاوه‌بر در نظر گرفتن امنیت برای کارها و داده‌ها، برای منابع و مسیر ارتباطی بین آن‌ها، قدرت امنیتی تعریف می‌شود. همچنین سعی می‌شود در محیط ابر ترکیبی، کارها روی منابع با امنیت سازگار زمان‌بندی شود. جهت پیاده‌سازی روش پیشنهادی از الگوریتم فرا مکاشفه‌ای PSO (Particle Swarm Optimization) بهبودیافته، استفاده‌شده که در ادامه به بیان فضای مسئله و ملاحظات امنیتی پرداخته می‌شود.

## ۳-۱. فضای مسئله

برای بیان مسئله زمان‌بندی کارها در محیط ابر ترکیبی نیاز به تعریف یکسری مفاهیم در فضای مسئله داریم.

**محیط ابر**: همان‌طور که قبلاً بیان شد، در محاسبات ابری، محیطی فراهم آمده است تا کاربران بتوانند از منابع موجود، بر اساس تقاضا استفاده کنند. ابر ترکیبی شامل ابر خصوصی و ابر عمومی، علاوه بر مزایای ابر خصوصی (عدم پرداخت هزینه و امنیت بالا)، نامحدود بودن منابع ابر عمومی را نیز دارد. در این مقاله محیط مسئله زمان‌بندی روی ابر ترکیبی بیان‌شده است.

**منابع** (ماشین‌های مجازی (VM)): ماشین‌های مجازی محاسباتی موجود در فضای ابر جزء منابع ابر به‌حساب می‌آید. ویژگی‌های منابع شامل قدرت پردازشی W(ti,vmj)، قیمت واحد ساعت $Cost(vm_j)$، سطح امنیتی $SS_{l\epsilon\{c,i,a\}}(vm_i)$و حافظه- که می‌تواند برای سرویس‌های محرمانگی، جامعیت و احراز هویت فراهم کند- است. قابل‌ذکر است منابع ابر خصوصی، به علت آنکه در اختیار سازمان هستند از بالاترین سطح امنیتی برخوردار هستند و در ضمن به غیر از هزینه‌های پشتیبانی، هزینه دیگری ندارند. اما به علت محدود بودن، ناگزیر بایستی از منابع ابر عمومی نیز استفاده شود.

**مجموعه کارها**[1]: مجموعه درخواست‌های کاربران است که به‌وسیله یک گراف جهت‌دار نشان داده می‌شود. گره‌های گراف که نشان‌دهنده کارها است، را به‌وسیله $T=\{t_1,t_2,..t_n\}$ و یال‌های گراف که نشانه داده‌های انتقالی بین کارها است را به‌وسیله $d_{i,j}$ نمایش داده می‌شود. اگر از کار ۱ به ۲ یالی وجود داشته باشد، کار ۱ پیش‌نیاز کار ۲ است و کار ۲ تا تمام شدن کار ۱ باید منتظر بماند. در گراف کارها، برای یال‌ها علاوه بر حجم داده، نیاز امنیتی داده‌های انتقالی نیز باید مشخص شود.

**مسئله زمان‌بندی**: در مسئله زمان‌بندی، نگاشت کارها روی منابع صورت می‌گیرد، به‌گونه‌ای که اهداف و محدودیت‌های تعریف‌شده توسط کاربر، برآورده شود. در این مقاله مسئله زمان‌بندی، با در نظر گرفتن ملاحظات امنیتی بررسی می‌شود.

[1] Workflow

با توجه به تعاریف گفته‌شده و همچنین در ادامه این مقاله از نشانه‌ها و نمادهایی که در جدول (۵) ارائه‌شده استفاده می‌شود. در بخش بعدی ملاحظات امنیتی در مسئله زمان‌بندی موردنظر در این مقاله توضیح داده می‌شود.

**جدول (۵)**: نشانه‌گذاری مسئله

| تعریف | نشانه |
|---|---|
| کار i در جریان کارها | $t_i$ |
| کار ورودی در جریان کارها | $t_{entry}$ |
| کار نهایی در جریان کارها | $t_{exit}$ |
| حجم داده انتقالی بین دو کار i و j | $d_{i,j}$ |
| پهنای باند بین دو منبع i و j | $bw_{i,j}$ |
| گره‌های پدر گره i | $Parrent(t_i)$ |
| گره‌های فرزند گره i | $Child(t_i)$ |
| منبع مجازی i | $vm_j$ |
| هزینه محاسباتی به ساعت روی منبع i | $Cost(vm_j)$ |
| زمان محاسباتی کار i روی منبع j | $W(t_i,vm_j)$ |
| زمان اجرای کار i ام روی سریع‌ترین منبع | $W(t_i)$ |
| سریع‌ترین زمان تمام شدن کار i | $EFT(t_i)$ |
| سریع‌ترین زمان شروع شدن کار i | $EST(t_i)$ |
| هزینه اجرای کار i روی منبع j | $TaskCost(t_i,vm_j)$ |
| قدرت امنیتی منبع i ام بر اساس سرویس‌های امنیتی | $SS_{l\epsilon\{c,i,a\}}(vm_i)$ |
| قدرت امنیتی مسیر ارتباطی بین دو منبع i,j | $SS_{l\epsilon\{c,i,a\}}(vm_i,vm_j)$ |
| سربار قدرت امنیتی روی منبع i | $OS_{l\epsilon\{c,i,a\}}(vm_i)$ |
| نیاز امنیتی داده‌های انتقالی بین دو کار i و j | $SR_{l\epsilon\{c,i,a\}}(t_i,t_j)$ |
| نیاز امنیتی کار i ام بر اساس سرویس‌های امنیتی | $SR_{l\epsilon\{c,i,a\}}(t_i)$ |
| فاصله امنیتی کار i از منبع j | $\partial_{i,vm_j}$ |
| سرویس احراز هویت | a |
| سرویس جامعیت | I |
| سرویس محرمانگی | C |
| حداکثر فاصله امنیتی | α |
| ضریب اتلاف منبع | W |

## ۳-۲. ملاحظات امنیتی

برای مسئله زمان‌بندی کارها در محیط ابرترکیبی نیاز به تعریف مفاهیمی است که ملاحظات امنیتی را به‌عنوان اهداف مسئله مشخص کند.

**تعریف ۱**: قدرت امنیتی منابع

قدرت امنیتی برای منابع با سه مؤلفه امنیتی برای هر منبع تعریف می‌شود. این مؤلفه‌ها حداکثر قدرت امنیتی منبع در برآورده شدن سه سرویس محرمانگی، جامعیت و احراز هویت را مشخص می‌کند. در این مقاله برای منابع سرویس‌های مختلف امنیتی تعریف می‌شود. برای ایجاد سطوح مختلف برای سرویس‌های مشخص امنیتی محرمانگی، جامعیت و احراز هویت از الگوریتم‌های جداول (۳- ۱) استفاده می‌شود. جدول (۶) نمونه‌ای از منابع را که درروش پیشنهادی استفاده می‌شود، نشان می‌دهد. درواقع برای منابع علاوه بر ظرفیت پردازش و قیمت، قدرت سرویس‌های امنیتی نیز در نظر گرفته می‌شود. مثلاً منبع ۲ و ۳ دارای سرعت پردازشی یکسانی هستند اما قدرت امنیتی متفاوتی ارائه می‌کنند، بنابراین قیمت متفاوتی دارند. منبع ۱ نیز به علت سرعت پردازشی کمتر، دارای قیمت کمتری است. قیمت‌ها بر اساس ساعت محاسبه می‌شود، به‌طوری‌که اگر میزان استفاده از منبع یک ساعت و نیم باشد، به‌اندازه ۲ ساعت، محاسبه می‌شود.

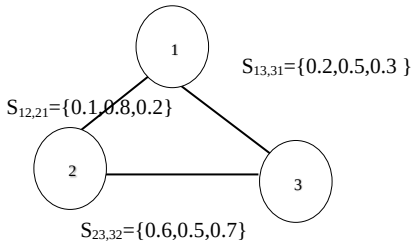
**جدول (۶)**: نمونه منابع با ویژگی قدرت سرویس‌های امنیتی

| منبع | ظرفیت پردازش (MIPS) | قیمت واحد ساعت $ | قدرت سرویس محرمانگی | قدرت سرویس تشخیص هویت | قدرت سرویس جامعیت |
|---|---|---|---|---|---|
| 1. m4.2xlarge | ۸۸۰۰ | ۰/۳۵ | ۰/۲ | ۰/۲ | ۰/۵ |
| 2. m4.xlarge | ۱۷۶۰۰ | ۰/۱ | ۰/۲ | ۰/۳ | ۰/۳ |
| 3. m4.large | ۱۷۶۰۰ | ۰/۲ | ۰/۵ | ۰/۳ | ۰/۵ |

**تعریف ۲**: قدرت امنیتی برای مسیر ارتباطی بین منابع

برای نشان دادن قدرت امنیتی مسیر ارتباطی بین منابع از سه مؤلفه استفاده می‌شود که هرکدام قدرت امنیتی مسیر را در برآورده کردن سرویس‌های امنیتی محرمانگی، جامعیت و احراز هویت را نشان می‌دهد.

در کارهای انجام‌شده پیشین برای مسیر ارتباطی بین منابع، امنیت در نظر گرفته نشده است. این در حالی است که در مسیر ارتباطی بین منابع امکان حملات وجود دارد. بنابراین، در این مقاله برای مسیر ارتباطی بین منابع، قدرت امنیتی تعریف می‌شود. عوامل تأثیرگذار روی این پارامتر امنیتی می‌تواند شامل مکان قرار گرفتن منابع ازنظر جغرافیایی، فاصله منابع از یکدیگر و تعداد سوییچ‌های مابین منابع باشد. قابل‌ذکر است مسیر ارتباطی لزوماً کانال ارتباطی مستقیم نخواهد بود و ممکن است از چندین سوییچ عبور کند. بنابراین قدرت امنیتی را به‌طورکلی برای مسیر ارتباطی تعریف می‌کنیم. با توجه به الگوریتم‌های مختلفی که در مسیر ارتباطی استفاده می‌شود، پارامترهای قدرت امنیتی تعریف می‌شود. برای مثال ممکن است الگوریتم رمزنگاری RC4 و RC5 در یک مسیر استفاده شود. شکل (۲) قدرت امنیتی مسیر ارتباطی بین سه منبع را نشان می‌دهد. واضح است که امنیت مسیر ارتباطی بین منبع ۱ و ۲ و بالعکس می‌تواند ضرایب متفاوتی داشته باشد. این، به این دلیل است که مسیر ارتباطی رفت‌وبرگشت بین دو منبع می‌تواند، متفاوت باشد. در این مقاله ضرایب مسیر ارتباطی بین منابع، دررفت و برگشت یکسان فرض می‌شود.




**شکل (۲)**: قدرت امنیتی مسیر ارتباطی منابع

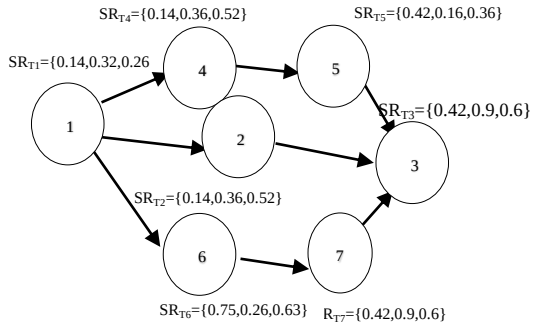
در کارهای قبلی انجام‌شده [۱۰، ۱۴و ۱۵]، برای کارها حساسیت تعریف می‌شود. درحالی‌که علاوه بر حساسیت کار، حساسیت داده‌های انتقالی بین کارها نیز باید در نظر گرفته شود. درواقع منظور از حساسیت کار، در نظر گرفتن نفوذ حمله‌کننده به اجرای کار است و منظور از حساسیت داده، حساسیت داده‌هایی است که کار به‌عنوان ورودی دریافت یا به‌عنوان خروجی تولید می‌کند. بنابراین در این مقاله حساسیت برای کار و داده آن ورودی-خروجی در نظر گرفته می‌شود.

**تعریف ۳**: نیاز امنیتی کارها

نیاز امنیتی برای کارها، به‌صورت نیاز امنیتی کار به سه سرویس امنیتی محرمانگی، جامعیت و احراز هویت بیان می‌شود.

نیاز امنیتی به‌صورت ضرایبی مشخص می‌شده که هر ضریب میزان نیاز امنیتی کار به سرویس امنیتی خاص است. شکل (۳)

نمونه گرافی را نشان می‌دهد که نیازهای امنیتی کارها به سه سرویس امنیتی محرمانگی، جامعیت و احراز هویت، با ضریب‌هایی مشخص‌شده است. برای مثال کار شماره ۴ نیاز امنیتی $s_{T4}$={0.14,0.36,0.52} به سرویس‌های مختلف دارد.

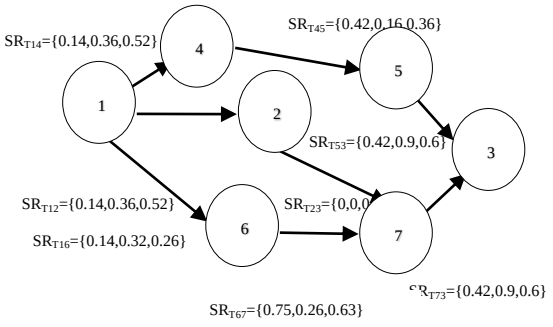



**شکل (۳)**: نیازهای امنیتی کارها [۱۵]

**تعریف ۴**: نیازهای امنیتی برای داده

نیاز امنیتی برای داده، به‌صورت نیاز داده به سه سرویس امنیتی محرمانگی، جامعیت و احراز هویت بیان می‌شود. داده‌ها علاوه بر بررسی صحت و محرمانگی باید مشخص باشد که از چه کاربری آمده است.

نیاز امنیتی، به‌صورت ضرایبی مشخص می‌شود که هر ضریب میزان نیاز امنیتی داده به سرویس امنیتی خاص است. شکل (۴) نمونه گرافی را نشان می‌دهد که نیازهای امنیتی داده به سه سرویس امنیتی محرمانگی، جامعیت و احراز هویت، با ضریب‌هایی مشخص‌شده است.




**شکل (۴)**: نیاز داده انتقالی بین کارها به سرویس‌های امنیتی

با توجه به تعاریفی که بیان شد، برای کار و منبع سطح امنیتی تعریف می‌شود و هدف زمان‌بندی کارها روی منابع به‌گونه‌ای است که سطح امنیتی داده انتقالی، منطبق بر قدرت امنیتی منابع باشد و مخاطره امنیتی حداقل شود. آقای Li و همکاران [۱۶] برای زمان‌بندی کارها از پارامتر محاسبه مخاطره امنیتی[1] استفاده کرده‌اند. مخاطره امنیتی را از تابع پوآسن، رابطه (۱)، محاسبه می‌کنند که یک تابع تصادفی است. اشکال استفاده از تابع پواسون علاوه بر محاسبات پیچیده، این است که برای بعضی موارد درست جواب نمی‌دهد. این ایراد را در قالب مثال ۱ بیان می‌شود. در رابطه (۱)، p مخاطره امنیتی، sr نیاز امنیتی کار و sl قدرت امنیتی منبع را نشان می‌دهد. مجموعه {a, q, c} به ترتیب سرویس‌های محرمانگی، جامعیت و احراز هویت را مشخص می‌کند.

$$p(t_i, sl_i^l) = 1 - \exp\left(-\gamma\left(sr_i^l - sl_i^l\right)\right), l \in \{a, g, c\}$$

$$p(t_i) = 1 - \prod_{l\in\{a,g,c\}} p(t_i, sl_i^l) \qquad (۱)$$

**مثال ۱:** ایراد تابع پواسون در بیان مخاطره امنیتی فرض کنید دو منبع باقدرت امنیتی $SS_1$={0.14, 0.36, 0.52} و $SS_2$={0.3, 0.36, 0.36} باشند. اگر داده‌های با نیاز امنیتی SR={0.14, 0.36, 0.52} وجود داشته باشد، بر اساس رابطه (۱) تابع پواسون مخاطره امنیتی را برای هر دو منبع مقدار یک، محاسبه می‌کند. از آنجا که تفاضل دو مقدار یکسان در توان رابطه (۱) مقدار صفر را مشخص می‌کند، بنابراین تابع احتمال خطر امنیتی یک به‌دست می‌آید. درصورتی‌که کاملاً واضح است که منبع شماره یک، به‌طور کامل نیاز امنیتی داده را برآورده می‌کند، درحالی‌که منبع شماره ۲ با نیاز امنیتی داده فاصله دارد.

با توجه به ایراد مطرح‌شده، درروش پیشنهادی برای محاسبه فاصله امنیتی از فاصله منهتن[2] استفاده می‌کنیم که در ادامه توضیح داده می‌شود.

**تعریف ۵**: محاسبه فاصله امنیتی

فاصله امنیتی عبارت از میزان انطباق نیاز امنیتی باقدرت امنیتی است. برای محاسبه فاصله امنیتی در این مقاله از فاصله منهتن استفاده می‌شود. فاصله منهتن، تفاضل دوبه‌دو ابعاد را نشان می‌دهد. رابطه (۲)، فاصله امنیتی کار i ام را از منبع j محاسبه می‌کند. هدف از به دست آوردن فاصله امنیتی، یافتن منبعی است که بیشترین انطباق را ازلحاظ امنیتی با کار داشته باشد. پس هر چه پارامتر $\partial$ کمتر باشد از لحاظ امنیتی بیشتر منطبق هستند.

$$\partial = \sum \beta * |SS_l - \mathrm{SR}_l \mid l \in \{\mathrm{a, c, i}\} \qquad (۲)$$

$$= \beta * |(\mathrm{SS_c} - \mathrm{SR_c}| + \beta * |(\mathrm{SS_a} - \mathrm{SR_a})| + \beta * |(\mathrm{SS_i} - \mathrm{SR}_i|$$

همان‌طور که در رابطه (۲) مشخص می‌شود، برای محاسبه فاصله امنیتی کار از منبع، باید فاصله امنیتی داده‌های ورودی (خروجی) کار را از منبع محاسبه کنیم. دلیل در نظر گرفتن

[1] ریسک امنیتی عبارت است از میزان خطر امنیتی در زمانبندی کار روی منبع
[2] Manhattan Distance

فاصله امنیتی داده‌های ورودی (خروجی) این است که منبع برای انجام کار، نیاز به داده‌های ورودی دارد که نیاز امنیتی داده‌های ورودی باید در سطح امنیتی منبع باشد. همچنین بعد از اجرای کار روی منبع، داده‌های خروجی تولید می‌شود که در اختیار منبع است بنابراین باید نیاز امنیتی داده‌های خروجی نیز در سطح امنیتی منبع باشد. در رابطه (۲) از پارامتر β استفاده‌شده است. برای توضیح این پارامتر به دو تعریف نیاز داریم که در ادامه می‌آید.

**تعریف ۶**: حداکثر فاکتور امنیتی α

درصورتی‌که قدرت امنیتی منبع از نیاز امنیتی داده کمتر باشد، حداکثر فاکتور امنیتی به‌صورت حداکثر میزان مجاز تفاضل این دو مقدار تعریف می‌شود. این میزان با پارامتر α در رابطه (۳) مشخص‌شده است.

به‌عبارت‌دیگر اگر بخواهیم زمان‌بندی امنی داشته باشیم، باید داده‌ها را روی منابع با قدرت امنیتی بیشتر زمان‌بندی کنیم اما این همواره امکان‌پذیر نیست. بنابراین جهت ساده‌تر شدن مسئله، امکان انتخاب منابع با قدرت امنیتی کمتر داده می‌شود به‌شرط آنکه تفاضل قدرت امنیتی منبع با نیاز امنیتی داده از حداکثر ضریب امنیتی تعریف‌شده، تجاوز نکند. برای مثال اگر نیاز امنیتی داده روی سرویس محرمانگی ۰/۵ باشد و حداکثر ضریب امنیتی ۰/۱ باشد، منابعی می‌تواند انتخاب شود که قدرت امنیتی سرویس محرمانگی آن از ۰/۴ کمتر نباشد.

**تعریف ۷**: ضریب اتلاف منبع (w)

ضریب اتلاف منبع عبارت از میزان اتلاف در انتخاب منبع با قدرت امنیتی بیشتر از نیاز امنیتی داده در مقابل منبعی با قدرت امنیتی کمتر از نیاز امنیتی داده، است.

در رابطه (۲)، β در محاسبه فاصله امنیتی استفاده می‌شود. زمانی که قدرت امنیتی منبع از نیاز امنیتی داده بیشتر باشد ضریب β یک در نظر گرفته می‌شود و زمانی که قدرت امنیتی منبع از نیاز امنیتی داده کمتر باشد، مقدار β برابر با ضریب اتلاف منبع w در نظر گرفته می‌شود. تفاضل بین نیاز امنیتی داده و قدرت امنیتی منبع نباید از α (حداکثر ضریب امنیتی) کمتر باشد.

فرض کنید داده‌ای با نیاز امنیتی SR={0.4, 0.3, 0.3} و دو منبع با قدرت امنیتی $SS_1$ ={0.6, 0.3, 0.3} و $SS_2$={0.3, 0.3, 0.3} داشته باشیم. همان‌طور که مشاهده می‌شود نیاز امنیتی سرویس جامعیت و تشخیص هویت با قدرت امنیتی منابع در همین سرویس‌ها برابر است. با در نظر گرفتن ۲/۵ برای ضریب اتلاف، بر اساس رابطه (۲) منبع با قدرت امنیتی سرویس محرمانگی ۰/۶ انتخاب می‌شود و ۰/۲ افزایش نسبت به ۰/۱ کاهش ترجیح داده می‌شود.

$$\beta = \begin{cases} w \; if \; (SS_c(vm_j) - SR_c(t_p, t_i) < \alpha \\ 1 \; else \end{cases} \qquad (۳)$$

بر اساس تعاریف بیان‌شده الگوریتم پیشنهادی در ادامه می‌آید.

## ۳-۳. الگوریتم پیشنهادی

با توجه به اینکه راه‌حل چندجمله‌ای برای مسئله زمان‌بندی وجود ندارد [۱۱]، یکی از الگوریتم‌های مناسب برای حل مسئله، الگوریتم فرا مکاشفه‌ای PSO (ازدحام ذرات) است [۲۵]. دلیل انتخاب این الگوریتم، پارامترهای کم و سرعت بالای اجرای آن است.

الگوریتم PSO یکسری ذرات برای جمعیت اولیه در فضای جستجوی جواب در نظر می‌گیرد و برای هر ذره سه پارامتر موقعیت فعلی ذره، سرعت ذره و بهترین موقعیت ذره لحاظ می‌شود. این الگوریتم مشابه الگوریتم‌های فرا مکاشفه‌ای با چندین تکرار همراه است که در هر تکرار طبق رابطه (۴) موقعیت جدید هر ذره به دست می‌آید. سرعت هر ذره بر اساس رابطه (۵)، به‌وسیله بهترین موقعیت هر ذره و بهترین موقعیت کل ذرات محاسبه می‌شود.

population[i]=velocity[i]+population[i] (۴)

velocity[i]=a*velocity[i]+c1*r1*(pbest[i]-population[i])+c2*r2*(gbest-population[i]) (۵)

که در این رابطه‌ها:

velocity[i]: سرعت ذره i ام

pbest[i]: بهترین موقعیت ذره i ام

gbest: بهترین موقعیت بین تمام ذرات

population[i]: ذره i ام

c1 وc2 : اعداد ثابتی هستند که شتاب حرکت را کنترل می‌کنند و

a: پارامتری که مقادیر ۰/۹ تا ۰ را دارد و

r1 و r2: دو عدد تصادفی بین صفر و یک است.

الگوریتم پیشنهادی و بهبودیافته PSO در ادامه شرح داده می‌شود. این الگوریتم در ابتدا و در مرحله ایجاد جمعیت اولیه با بررسی محدودیت‌های مسئله سعی در تولید جواب‌های قابل‌قبول می‌کند. درواقع با بررسی قابلیت امکان‌پذیری ذره‌ها تا رسیدن به

یک مقدار ممکن با مقدار تصادفی، ذره‌ها را جایگزین می‌کند. همچنین تفاوت الگوریتم پیشنهادی در تعیین مقدار متغیر a وزن اینرسی (باقی ماندن در سرعت قبلی) است که به‌وسیله رابطه (۱۲) محاسبه می‌شود. در مرحله ایجاد جمعیت جدید اگر بهبود نسبت به بهترین حالت ذره در تعدادی از ذره‌ها (درصدی از تعداد کل ذره‌ها) صورت نگرفته باشد، به‌اندازه تعداد حالات بهبود نیافته، ذره‌های تصادفی با مقدار تصادفی جایگزین می‌شوند.

برای مدل‌سازی مسئله نیاز به ایجاد یک جمعیت اولیه است. ازآنجاکه قصد داریم زمان‌بندی کارها را روی منابع انجام دهیم، جهت ایجاد جمعیت اولیه، برای هر ذره، فهرستی به‌اندازه تعداد کارها در نظر می‌گیریم. برای هر عنصر لیست، عدد تصادفی که شماره منبع را مشخص می‌کند، قرار می‌دهیم. سپس کارها را بر اساس رابطه (۶)، اولویت‌بندی و به ترتیب به اعضای لیست اختصاص می‌دهیم. بنابراین فهرستی داریم که شماره کار را به شماره منبع نگاشت می‌کند. مثال ۲ نمونه‌ای از ایجاد یک جمعیت اولیه را نشان می‌دهد.

$$rank(t_i) = \begin{cases} w_{i,r} & t_i = t_{exit} \\ w_{i,r} + max_{t_{p\in childs\,of(t_i)}}(rank(t_i)+c_{i,p}) & otherwise \end{cases} \quad (۶)$$

**مثال ۲**: ایجاد جمعیت اولیه

فرض کنید مجموعه‌ای از کارها شامل ۹ کار و ۴ منبع داشته باشیم. برای ایجاد جمعیت اولیه هر پاسخ به‌صورت یک مجموعه ۹ تایی مشخص می‌شود به‌طوری‌که هر عضو مجموعه یکی از اعداد ۱ الی ۴ -نشان‌دهنده شماره منبع- است. در شکل (۵) لیست کارها که اولویت‌بندی شده‌اند بر روی لیست پاسخ - شامل شماره منابع- نشان داده‌شده است. Vm نشان‌دهنده ماشین مجازی (منبع) است.

| 1 | 3 | 4 | 2 | 8 | 7 | 5 | 6 | 9 |
|---|---|---|---|---|---|---|---|---|
| Vm1 | Vm3 | Vm2 | Vm1 | Vm4 | Vm2 | Vm1 | Vm3 | Vm4 |

**شکل (۵)**: نمونه ذره برای ایجاد جمعیت اولیه

همان‌طور که در رابطه (۴) مشاهده شد، جهت به دست آوردن سرعت هر ذره باید بهترین موقعیت هر ذره و بهترین موقعیت همه ذرات محاسبه شود. برای محاسبه بهترین ذره تابع هدف را بیان می‌کنیم.

### ۳-۳-۱. تابع هدف

برای تخصیص کارها روی منابع با در نظر گرفتن ملاحظات امنیتی، هدف را حداقل کردن فاصله امنیتی کارها و منابع در نظر گرفته می‌شود. فاصله امنیتی کار و منبع شامل فاصله نیاز امنیتی کار و قدرت امنیتی منبع، فاصله نیاز امنیتی داده‌های ورودی و خروجی کار و قدرت امنیتی منبع و فاصله نیاز امنیتی داده‌ها و مسیر ارتباطی منابع است. هر چه میزان این تابع کمتر باشد به معنای شباهت بیشتر کارها به منابع ازلحاظ امنیتی است. رابطه (۷) تابع هدف را نشان می‌دهد.

$\partial_i$: فاصله امنیتی کار i ام را از منبع زمان‌بندی‌شده $vm_j$، نشان می‌دهد.

$$\text{Min } \partial = Min\,(\sum_{t_i\in all\,Tasks} \partial_i) \quad (۷)$$

$$\partial_i = \sum_{t_k\in all\,childs\,of\,t_i} \beta * |SS(vm_j) - SR(t_i,t_k)| + \sum_{t_p\in all\,parrents\,of\,t_i} \beta * |SS(vm_j) - SR(t_p,t_i)| + \beta * |SS(vm_j) - SR(t_i)| + Relations\_Security$$

$$Relation\_Security = \sum_{t_k\in all\,childs\,of\,t_i} \beta * |SS(vm_j,vm_i) - SR(t_i,t_k)| + \sum_{t_p\in all\,parrents\,of\,t_i} \beta * |SS(vm_i,vm_j) - SR(t_p,t_i)|$$

Relation_Security: میزان فاصله نیاز امنیتی داده‌های ورودی (خروجی) کار را با قدرت امنیت مسیر ارتباطی بین منابع زمان‌بندی‌شده را محاسبه می‌کند. فرض شده کار i روی منبع $vm_j$ و کارهای پدر و فرزند روی منبع $vm_i$زمان‌بندی‌شده است.

### ۳-۳-۲. محدودیت‌های مسئله

در این الگوریتم برای زمان و هزینه محدودیت در نظر گرفته‌شده است. بدین معنی که زمان اجرای کل کارها نباید از حداکثر زمان مشخص‌شده توسط کاربر بیشتر شود و همچنین هزینه اجرای کل کارها روی منابع نیز از حداکثر هزینه مشخص‌شده توسط کاربر نبایستی بیشتر باشد. محدودیت‌های مسئله به‌صورت زیر مشخص می‌شود.

Subject to: Makespan< Deadline

TotalCost<MaxCost

Makespan زمان انجام کل کارها را نشان می‌دهد که باید از Deadline حداکثر زمان تعریف‌شده توسط کاربر کمتر باشد. همچنین TotalCost هزینه انجام کارها توسط منابع است که باید از MaxCost (حداکثر هزینه تعریف‌شده توسط کاربر) کمتر باشد. محاسبه هزینه‌بر اساس رابطه (۸) برابر با زمان اجرای کار ضرب در میزان هزینه منبع به ساعت است.

$$\text{TaskCost}(t_i, vm_j) = W(t_i,vm_j) * \text{Cost}(vm_j)$$

$$\text{TotalCost} = \sum_{t_i\in workflow} \text{TaskCost}(t_i, vm_j) \quad (۸)$$

برای محاسبه زمان اجرای کل کارها (Makespan) باید زمان اجرای هر کار را محاسبه نمود و بیشترین زمان اتمام، زمان کل اجرا است.

زمان اتمام اجرای هر کار برابر است با زمان شروع اجرا (EST) به‌اضافه‌ی زمان اجرای کار روی منبع $W(t_i,vm_j)$ و همچنین زمان مورد نیاز برای محاسبات امنیتی[1] (SC). رابطه (۹) زمان اتمام کار i ام را روی منبع j ام نشان می‌دهد.

$$EFT(t_i,vm_j)= EST(t_i)+ W(t_i,vm_j)+ +SC(t_i,vm_j) \quad (۹)$$

زمان شروع هر کار برابر است با ماکزیمم زمان اتمام کارهای پیش‌نیاز (پدر)، به‌اضافه زمان انتقال داده‌های مورد نیاز (DT) که طبق رابطه (۱۰) محاسبه می‌شود. در این عبارت زمان انتقال داده‌ها، برابر است با نسبت حجم داده به پهنای باند منابعی که کارها را زمان‌بندی می‌کنند. درصورتی‌که کارها روی منابع یکسان زمان‌بندی شوند، زمان انتقال صفر در نظر گرفته می‌شود.

$$EST(t_i)= \max_{t_p \in parrents\ of(t_i)} \left(EFT(t_p) + DT_{p,i}\right) \quad (۱۰)$$

$$DT_{p,i} = \begin{cases} \frac{d_{p,i}}{bw_{n,m}} & \text{if } t_p \text{ on } vm_n \text{ and } t_i \text{ on } vm_m \\ 0 & \text{if the resource is the same} \end{cases}$$

زمان موردنیاز برای محاسبات امنیتی (SC) از رابطه (۱۱) محاسبه می‌شود. در این عبارت برای هر سرویس (محرمانگی، جامعیت و احراز هویت) میزان سربار زمانی که الگوریتم ایجاد می‌کند، برای داده‌های ورودی (خروجی) به دست می‌آید.

$d_{p,i}$: حجم داده‌هایی است که باید بین دو کار p و i منتقل شود.

$OS_l(vm_i)$: سربار امنیتی که الگوریتم‌های سرویس امنیتی قابل‌اجرا روی منبع i، برای داده‌ها ایجاد می‌کند. l می‌تواند هرکدام از سرویس‌های امنیتی باشد.

$$SC(t_i,vm_j) = \sum_{l\in\{c,i,a\}} SC_l(t_i,vm_j) \quad (۱۱)$$

$$SC_l(t_i,vm_j) = \sum_{t_p \in parrents\ of(t_i)} d_{p,i} * OS_l(vm_i) + \sum_{t_p \in childs\ of(t_i)} d_{i,p} * OS_l(vm_i)$$

### ۳-۳-۳. الگوریتم پیشنهادی PSO-WSCS

الگوریتم ۱، الگوریتم ازدحام ذرات بهبودیافته را برای n کار و m ماشین مجازی نشان می‌دهد. روش پیشنهادی PSO-WSCS [2] که در چندین مرحله با الگوریتم PSO معمولی متفاوت است در ادامه توضیح داده می‌شود.

```
PSO-WSCS Algorithm for Security Task Scheduling
BEGIN
1. Set the number of particles to p and dimension of the particle to n;// p=20
2. Set the gbest and pbest to zero.
3. for each particle i=1 to p
4. Randomly create particle population[i] and velocity[i]
5 pbest[i]= population[i];
6. gbest= the best solution on current population
7. end for
10. t=0;
11. while (t<number of iteration)
12. for each particle i=1 to p
13. Cycle=0;
Update the velocity and position of particle on Eq. (4,5)
14. Check Bound of position of particle
while Bound of position is violated,
the position must be updated by random values
15. Calculate fitness function for each particle on Eq.6;
16. If (fitness function(current particle i) is better than fitness function(pbest[i]))
17. { pbest[i]= current particle;
}
Else
Cycle++;
18. end for
19. If Cycle > t/10
20. for(int q=1;q<=cycle;q++)
        do{
        int z=create random number (0,p)
  population[z]= CreatePopulation();
  }while(CheckBound of position [z])!=true);

  21. calculate the gbest;// the best between all pbest;
  22. t=t+1;
  23. end while
      END
```

**شکل (۶)**: الگوریتم زمان‌بندی گردشکارهای پیشنهادی PSO-WSCS

الگوریتم زمان‌بندی گردشکارها بر اساس الگوریتم ازدحام ذرات در شکل (۶) نوشته‌شده است. در خط ۱ الگوریتم، تعداد ذرات ۲۰ فرض شده و طبق شکل (۵)، اندازه هر ذره برابر تعداد کارها در نظر گرفته‌شده است. خط ۳ تا ۷ مقداردهی جمعیت اولیه الگوریتم را نشان می‌دهد. این جمعیت بر اساس یکسری مقادیر تصادفی و همچنین استفاده جواب الگوریتم‌های HEFT[3] و DHEFT[4] ایجاد می‌شود. استفاده از الگوریتم‌های ذکرشده در ایجاد جمعیت اولیه، باعث بهبود عملکرد الگوریتم و همچنین سریع‌تر همگرا شدن آن می‌شود. خطوط ۱۱ تا ۲۲ تعداد تکرار الگوریتم را به‌وسیله حلقه کنترل می‌کند. در این مقاله تعداد تکرار برابر ۱۰۰۰ در نظر گرفته‌شده است. در هر مرحله از تکرار حلقه

[1] Security Computing
[2] PSO Workflow Scheduling Considering Security
[3] Heterogeneous Earliest Finish Time
[4] Deadline Earliest Finish Time

ابتدا سرعت و موقعیت جدید هر ذره بـر اسـاس رابطـه (۴) و (۵) محاسبه می‌شود. همان‌طور که بیان شد موقعیت جدید هر ذره بر اساس بهترین موقعیت همه ذرات و بهتـرین موقعیـت همـان ذره محاسبه می‌شود. جهت وزن دهی به‌سرعت قبلی ذره از پـارامتر a استفاده‌شده که مقادیر ۰/۹ تا صـفر را در تکرارهـای مختلـف بـه خود اختصاص می‌دهد. این مقدار بر اساس تعداد تکرار الگـوریتم تعیین می‌شود. رابطه (۱۲) تعداد کل تکـرار را iteration در نظـر گرفته و در اجرای تکرار k ام، مقدار a را تعیین می‌کند.

a=(a*(iteration-k))/iteration (۱۲)

خط ۱۴ تابع بررسی محدودیت‌ها را مشـخص مـی‌کنـد تـا موقعیت جدید ذرات، شرایط مسئله را نقـض نکننـد و در صـورت نقض، دوباره ذره ایجاد می‌شود. این عمل باعـث مـی‌شـود در هـر مرحله جواب‌های حاصله قابل‌قبول باشند. در خط ۱۵ تابع هدف، برای همه ذرات محاسبه می‌شود و بهتـرین موقعیـت هـر ذره در صورت بهتر بودن مقدار تابع هدف، بـا مقـدار ایـن ذره جـایگزین مـی‌شـود. در ایـن الگـوریتم درصـورتی‌کـه درصـدی از ذرات موقعیتشان بهبود پیدا نکند موقعیـت ذره را بـا بهتـرین موقعیـت همان ذره جـایگزین مـی‌کنـد. درنهایـت بهتـرین موقعیـت ذرات gbest از بین ذراتی که تابع هدف بهتری دارند، انتخاب می‌شـود. در انتهای اجرای حلقه تکرار الگـوریتم، مقـدار بهتـرین موقعیـت ذرات، بهترین پاسخ مسئله خواهد بود.

همان‌طور که بیـان شـد در تـابع هـدف شـکل (۷)، انطبـاق امنیتی بین کارها و منابع در نظر گرفتـه شـده اسـت. اسـتفاده از منابعی کـه ازلحـاظ امنیتـی بسـیار نزدیـک کارهـا باشـند بـرای زمان‌بندی کارها، باعث می‌شود تا خطر حمله‌های امنیتـی کمتـر شود و در ضمن سطح امنیتی کار و منبع نزدیک به هم باشد. این باعث می‌شود تا تابع هدف، زمانی که قدرت امنیتی منبـع از نیـاز امنیتی کار بیشتر باشد ضریب یـک و زمـانی کـه قـدرت امنیتـی منبع از نیاز امنیتی کار کمتـر باشـد در ضـریب ۲ ضـرب شـود و درنتیجه تابع هدف مقدار بیشتری می‌گیرد. بنابراین حداقل کردن تابع هدف، سطح امنیتی منطبقی را ایجاد می‌کند.

```
Calculate Security function (Task t, Resource r)
1-  Begin
2-Calculate sum of the S1, S2, S3
      S1= Security distance between task t and resource r
      S2= Security distance between input (output) data
        of task t and resource r.
      S3= Security distance between data and resource
        channel
3-  End
```

**شکل (۷):** پیاده‌سازی تابع هدف

## ۴. ارزیابی روش پیشنهادی

الگوریتم پیشنهادی زمان‌بندی کارهـا در محـیط ابـر ترکیبـی بـا ملاحظات امنیتی، در محـیط شـبیه‌سـازی WorkflowSim [۲۶] پیاده‌سازی شده است. ازنظر مرتبه زمانی اگر اندازه جمعیت را M و تعداد کارها را N در نظـر بگیـریم، آن‌گـاه بـرای محاسـبه تـابع هدف، پیچیدگی O(N^2) است. این مقدار بر اساس بعـد هـر ذره که در اینجا به‌انـدازه تعـداد کارهـا اسـت، محاسـبه شـده اسـت. همچنین به‌اندازه تعداد جمعیت، این زمـان تکـرار مـی‌شـود کـه مرتبه زمانی کل برابر با O(M*N^2) است.

با توجه به این‌که فرضـیات مسـئله دقیقـاً در کارهـای قبلـی استفاده نشده است، سعی می‌شود تا در مقایسه از الگـوریتم‌هـایی استفاده شود که در کارهای تقریباً مشابه روش پیشنهادی استفاده شده‌اند.

الگوریتم VNPSO [۲۳] الگوریتم بر پایه الگوریتم PSO اصلی است با این تفاوت که اگر در تعدادی حالات بهبود تابع هـدف رخ ندهد، سرعت را بر اساس رابطه (۱۳) تغییر می‌دهد.

$$V=\begin{cases} v_{ij} if\ |v_{ij}|>v_c \\ \frac{u(-1,1)v_{max}}{\rho} \quad if\ |v_{ij}|<v_c \end{cases} \qquad (۱۳)$$

الگوریتم MPSO-SA [۲۴] مجموعه ذراتی در نظر می‌گیرد و در هر مرحلـه کـه موقعیـت ذرات بهبـود پیـدا نکنـد، بـر اسـاس پارامتری افزایشی، موقعیت تعدادی از ذرات را به‌صـورت تصـادفی تغییر می‌دهد. بـدین ترتیـب سـرعت همگرایـی الگـوریتم را بـالا می‌برد. البته در مقاله Li تابع هدف حـداقل کـردن زمـان اجـرای کارها است و احتمال خطر امنیتی بـه‌عنـوان محـدودیت در نظـر گرفته‌شده است.

ابتـدا روش پیشـنهادی و الگـوریتم‌هـای PSO، VNPSO و MPSO-SA در محـیط شـبیه‌سـازی پیـاده‌سـازی شـد. سـپس پارامترهای الگوریتم به شرح بعدی تعیین گردید.

برای انتخاب یک حداکثر زمان برای اجرای کارهـا، مـی‌تـوان الگوریتم سریع‌ترین زمان اجرا (HEFT) را محاسبه کـرد و مقـدار آن را برای Deadline در نظر گرفت. همان‌طور که می‌دانیم هرچه این مقدار بیشتر باشد، زمان انجام کارها نیز می‌تواند بیشتر باشـد و این، امکان انتخاب منابع با ویژگی‌های بهتر را فـراهم مـی‌کنـد. به‌عبارت‌دیگر اگر Deadline نقض نشود، می‌توان منابع با سـرعت پایین‌تر، امنیت منطبق‌تر و درنتیجه هزینه کمتر، انتخاب نمود.

بقیه پارامترهای موردنیاز در الگوریتم PSO بـه شـرح بعـدی است.

مقادیر c1 و c2 را بـه ترتیـب ۱/۴۹ و ۱/۴۹ قـرار داده‌شـده و

تعداد ذرات و تعداد تکرار الگوریتم را ۲۰ و ۱۰۰۰ گرفته شد (مشابه مقاله Liu [۱۴]).

مقدار حداکثر فاصله امنیتی ($\alpha$) توسط کاربر مشخص می‌شود. لازم به ذکر است که این مقدار نباید از حداقل تفاضل نیاز امنیتی و قدرت امنیتی در هر سرویس کمتر باشد. در صورت بیشتر بودن، مسئله قابل‌حل نیست.

مقادیر مجاز برای ضریب اتلاف منبع (w) عبارت از مقادیر بین ۱ تا ۳ است. با توجه به توضیحاتی که در قسمت تعریف ضریب اتلاف آمده است، در این مقاله مقدار w، ۲ فرض شده است. در پاراگراف بعدی ارزیابی و شبیه‌سازی الگوریتم پیشنهادی آمده است.

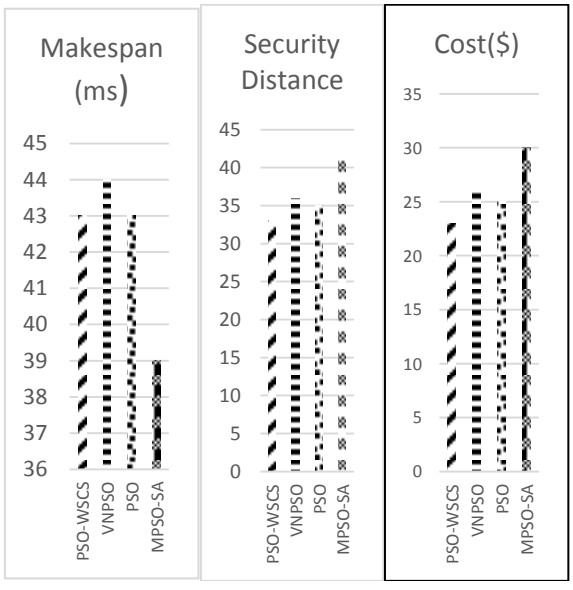
در محیط شبیه‌سازی برای ارزیابی، الگوریتم پیشنهادی PSO-WSCS[1] و الگوریتم‌های PSO، VNPSO و MPSO-SA را روی گردشکارهای دنیای واقعی شامل Inspiral,HEFT و Montage[2] اجرا می‌شود [۲۷]. هرکدام از این گردش‌ها ساختار خاصی دارد و اطلاعات کامل در رابطه با این گردش‌های کاری در مقاله جوو [۲۷] بیان شده است. برای ارزیابی بهتر الگوریتم، نیازهای امنیتی به گردش‌های کاری که به فرمت XML است، اضافه‌شده است.

همان‌طور که ذکر شد برای ارزیابی الگوریتم‌ها از سه نمونه گراف استفاده شده است. نمونه اول گراف HEFT با ده کار، روی سه منبع در نظر گرفته شد که منبع اول در ابر خصوصی بدون هزینه اجرایی و در بالاترین سطح امنیتی قرار دارد و دو منبع دیگر با سرعت پردازشی یکسان و قدرت امنیتی متفاوت که منبع با قدرت امنیتی بیشتر دارای هزینه اجرایی بالاتری بود. نمونه دوم گراف Montage با ۲۵ کار است که روی ۵ منبع که در جدول (۶) نشان داده شده است زمان‌بندی شدند. نمونه سوم Inspiral با ۵۰ کار که از همان منابع جدول (۷) برای زمان‌بندی استفاده شد. قابل‌ذکر است که منابع به‌صورت کاملاً متصل هستند و قابلیت ارتباط بین هر دو منبع وجود دارد.

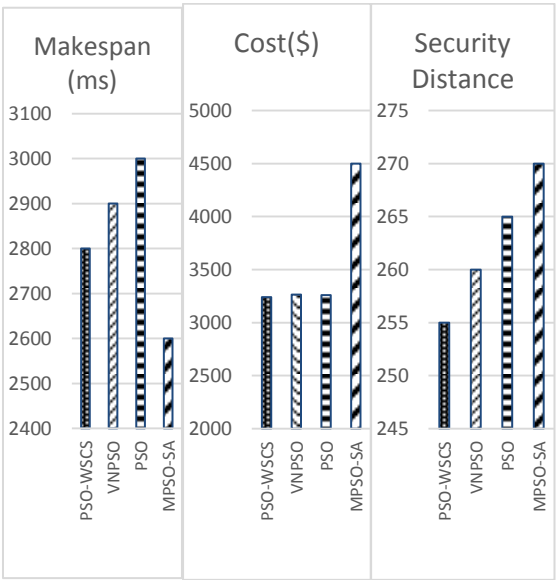
نتایج الگوریتم پیشنهادی PSO-WSCS و بقیه الگوریتم‌ها در شکل (۱۰-۸) آمده است.

**جدول (۷)**: ویژگی‌های منابع در ارزیابی روش پیشنهادی

| | Processing capacity (MIPS) | Cost per hour $ | Security strength for confidentially | Security strength for integrity | Security strength for authentication |
|---|---|---|---|---|---|
| VM1 | ۱۰۰۰ | ۰/۳ | ۰/۲ | ۰/۳ | ۰/۵ |
| VM2 | ۲۰۰۰ | ۰/۶ | ۰/۱ | ۰/۲ | ۰/۳ |
| VM3 | ۱۰۰۰ | ۰/۴ | ۰/۳ | ۰/۴ | ۰/۶ |
| VM4 | ۱۰۰۰ | ۰ | ۱ | ۱ | ۱ |
| VM5 | ۱۵۰۰ | ۰ | ۱ | ۱ | ۱ |

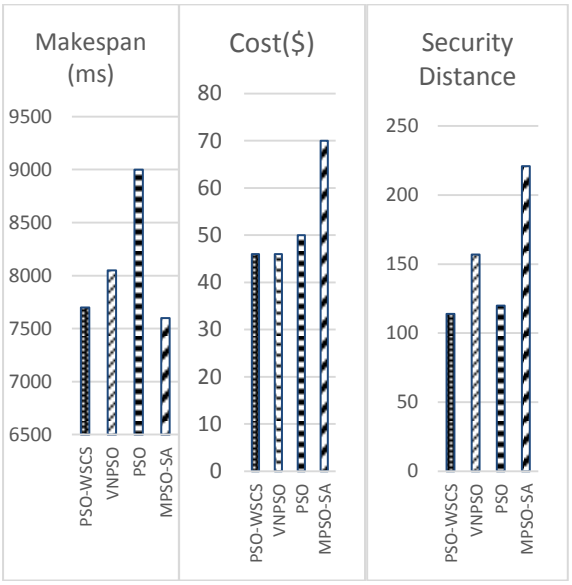



**شکل (۸)**: تابع هدف، زمان اجرا و هزینه روی گراف HEFT




**شکل (۹)**: تابع هدف، زمان اجرا و هزینه روی گراف Inspiral

[1] PSO Workflow Scheduling considering Security
[2] http://pegasus.isi.edu/schema/DAX




**شکل (۱۰)**: تابع هدف، زمان اجرا و هزینه روی گراف Montage

معیارهای موردبررسی در الگوریتم‌ها شامل زمان و هزینه اجرای کل کارها - محدودیت‌های مسئله در مقاله - و همچنین تابع هدف فاصله امنیتی است. معیارها روی سه نمونه توسط الگوریتم‌های مختلف بررسی‌شده است.

با توجه به نتایج به‌دست‌آمده در شکل (۸) روی گراف HEFT، تابع هدف در الگوریتم پیشنهادی PSO-WSCS کاهش ۳۰ درصدی نسبت به الگوریتم PSO داشته است. همین‌طور در شکل (۹) روی گراف Inspiral، تابع هدف در الگوریتم پیشنهادی PSO-WSCS کاهش ۵۳ درصدی نسبت به الگوریتم‌های دیگر داشته است. شکل (۱۰) روی گراف Montage، تابع هدف را نشان می‌دهد که در الگوریتم پیشنهادی PSO-WSCS کاهش ۴۵ درصدی نسبت به الگوریتم‌های دیگر داشته است. کاهش تابع هدف (فاصله امنیتی) به معنای این است که برای کارها، منابعی با قدرت امنیتی مشابه انتخاب‌شده است. درواقع الگوریتم PSO با قابلیت کنترل روی حداکثر زمان اجرا و حداکثر هزینه، کمترین فاصله امنیتی را پیدا می‌نماید. یکی دیگر از محاسن الگوریتم پیشنهادی، در نظر گرفتن حساسیت کارها است. به‌طوری‌که کارهایی که درجه حساسیت بالایی دارند یا با داده‌های حساس کار می‌کنند، روی منابع خصوصی و تحت کنترل سازمان زمان‌بندی می‌شوند.

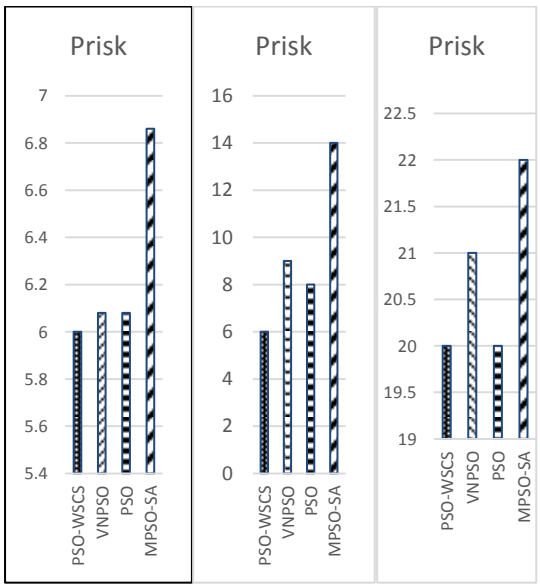
معمولاً با افزایش سرعت اجرای منبع، هزینه اجرا نیز افزایش پیدا می‌کند. اما در رابطه با قدرت امنیتی منبع نمی‌توان به‌طورقطع اظهارنظر نمود. بنابراین الگوریتم پیشنهادی تا زمانی که حداکثر زمان مشخص‌شده توسط کاربر اجازه بدهد، منبعی را انتخاب می‌کند که علاوه‌بر کاهش هزینه‌ها ازلحاظ امنیتی منطبق به کار باشد.

برای مقایسه بیشتر الگوریتم پیشنهادی با الگوریتم‌های بیان‌شده از معیار احتمال خطر امنیتی[۲۴] نیز استفاده می‌کنیم. این معیار از رابطه (۱۴) محاسبه می‌شود.

$$Prob_{risk}\begin{cases} 0 \; if \; S_T - S_C \le 0 \\ 1 - e^{-0.5(S_T - S_C)} \; if \;\; 0 < S_T - S_C \le 1 \\ 1 - e^{-0.5(S_T - S_C)} if \;\; 1 < S_T - S_C \le 2 \\ 1 \qquad\qquad\qquad if \;\; 2 < S_T - S_C \le 5. \end{cases} \quad (۱۴)$$

در رابطه (۱۴)، $S_T$ میزان سطح امنیتی کار و $S_C$ میزان سطح امنیتی منبع را نشان می‌دهد.

درروش پیشنهادی و الگوریتم‌های ارائه‌شده روی نمونه گراف‌های بیان‌شده، نتایج شکل (۱۱) حاصل شد که نشان می‌دهد روش پیشنهادی احتمال خطر امنیتی را بیشتر کاهش می‌دهد و این به معنای دست‌یابی به امنیت بالاتر در زمان‌بندی است. برای مثال در شکل (۱۱) احتمال خطر امنیتی روی نمونه اول، گراف HEFT با ده کار، در الگوریتم پیشنهادی نسبت به الگوریتم‌های دیگر کاهش چشم‌گیری داشته است.




**شکل (۱۱)**: مقایسه احتمال خطر امنیتی در روش‌ها

## ۵. نتیجه‌گیری

با توجه به بالا بودن حجم کارها در محیط‌های سازمانی و استفاده از ابر عمومی علاوه بر ابر خصوصی، مبحث زمان‌بندی با در نظر گرفتن پارامترهای امنیتی موردتوجه است. درواقع خروج کارها از محیط‌های سازمانی، مستلزم رعایت یکسری نکات امنیتی است تا داده‌هایی که از حساسیت خاصی برخوردار هستند ترجیحاً در داخل سازمان استفاده شوند و یا اگر به محیط ابر عمومی می‌روند، حفظ محرمانگی و صحت را داشته باشند. بنابراین، در

این مقاله سعی شد که زمان‌بندی کارها روی منابع به‌گونه‌ای انجام شود که با در نظر گرفتن نیاز امنیتی برای کارها و داده‌های انتقالی، قدرت امنیتی منابع و مسیرهای ارتباطی، حداکثر تشابه امنیتی بین کارها و داده‌ها با منابع و مسیرهای ارتباطی وجود داشته باشد. همان‌طور که در قسمت ارزیابی مشاهده کردیم، با لحاظ کردن محدودیت زمان و هزینه، نیاز امنیتی کارها نزدیک به قدرت امنیتی منابع است. درواقع منابعی برای زمان‌بندی کارها انتخاب می‌شوند که ازلحاظ امنیتی حداکثر تشابه را داشته باشند و درعین‌حال زمان و هزینه را مدنظر بگیرند. مزیت دیگر این الگوریتم علاوه بر موارد ذکرشده، در نظر گرفتن منابع ابر خصوصی است که با تعریف هزینه منبع برابر با صفر و قدرت امنیتی برابر با (1, 1, 1)- نشان‌دهنده حداکثر قدرت امنیتی- در نظر گرفته می‌شود.

برای ادامه کار می‌توان از زمان‌بندی امن بین کارها [۲۸] روی انواع دیگر ابر، ابرهای چندگانه [۲۰] استفاده کرد. همچنین می‌توان برای چند گردشکاری (چند مجموعه گراف کارها)، مثل گردشکاری که امکان تبادل داده داشته باشد [۲۹]، نیز زمان‌بندی امن انجام داد. بهبود بر روی روش PSO و همچنین استفاده از روش‌های تابع چندهدفه [۲۳] نیز ازجمله کارهای آتی خواهد بود.

## ۶. مراجع


[1] M. Naghibzadeh, "Modeling Workflow of Tasks and Task Interaction Graphs to Schedule on the Cloud," Cloud Computing 2016, p. 81, 2016.

[2] C. Jianfang, C. Junjie, and Z. Qingshan, "An optimized scheduling algorithm on a cloud workflow using a discrete particle swarm," Cybernetics and Information Technologies, vol. 14, pp. 25-39, 2014.

[3] S. Singh and I. Chana, "A survey on resource scheduling in cloud computing: Issues and challenges," Journal of Grid Computing, vol. 14, pp. 217-264, 2016.

[4] Sh. Jamali and S. Hourali, "Decentalized load balancer in cloud enviroment by usig multi attribute decision making policy," Tabriz Journal of Electrical Engineering, pp. 95-106, 2016. (In Persian)

[5] R. Gupta, "Above the Clouds: A View of Cloud Computing," Asian Journal of Research in Social Sciences and Humanities, vol. 2, pp. 84-110, 2012.

[6] H. Abrishami, A. Rezaeian, and M. Naghibzadeh, "Scheduling in hybrid cloud to maintin data privacy," 20th National CSI Computer Conference, 2015. (In Persian)

[7] S. Sharif, J. Taheri, A. Y. Zomaya, and S. Nepal, "Mphc: Preserving privacy for workflow execution in hybrid clouds," in 2013 International Conference on Parallel and Distributed Computing, Applications and Technologies, pp. 272-280, 2013.

[8] H. Abrishami, A. Rezaeian, and M. Naghibzadeh, "A novel deadline-constrained scheduling to preserve data privacy in hybrid Cloud," in Computer and Knowledge Engineering (ICCKE), 2015 5th International Conference on, pp. 234-239, 2015.

[9] A. Rezaeian, H. Abrishami, S. Abrishami, and M. Naghibzadeh, "A Budget Constrained Scheduling Algorithm for Hybrid Cloud Computing Systems Under Data Privacy," in Cloud Engineering (IC2E), 2016 IEEE International Conference on, pp. 230-231, 2016.

[10] H. Chen, X. Zhu, D. Qiu, L. Liu, and Z. Du, "Scheduling for workflows with security-sensitive intermediate data by selective tasks duplication in clouds," IEEE Transactions on Parallel and Distributed Systems, 2017.

[11] M. L. Pinedo, "Scheduling: theory, algorithms, and systems," Springer, 2016.

[12] F. Wu, Q. Wu, and Y. Tan, "Workflow scheduling in cloud: a survey," The Journal of Supercomputing, vol. 71, pp. 3373-3418, 2015.

[13] M. Masdari, S. ValiKardan, Z. Shahi, and S. I. Azar, "Towards workflow scheduling in cloud computing: a comprehensive analysis," Journal of Network and Computer Applications, vol. 66, pp. 64-82, 2016.

[14] H. Liu, A. Abraham, V. Snášel, and S. McLoone, "Swarm scheduling approaches for work-flow applications with security constraints in distributed data-intensive computing environments," Information Sciences, vol. 192, pp. 228-243, 2012.

[15] W. Liu, S. Peng, W. Du, W. Wang, and G. S. Zeng, "Security-aware intermediate data placement strategy in scientific cloud workflows," Knowledge and information systems, vol. 41, pp. 423-447, 2014.

[16] Z. Li, J. Ge, H. Yang, L. Huang, H. Hu, H. Hu, et al., "A security and cost aware scheduling algorithm for heterogeneous tasks of scientific workflow in clouds," Future Generation Computer Systems, 2016.

[17] H. Abrishami, A. Rezaeian, and M. Naghibzadeh, "Workflow Scheduling on the Hybrid Cloud to Maintain Data Privacy under Deadline Constraint," Journal of Intelligent Computing Volume, vol. 6, p. 93, 2015.

[18] L. F. Bittencourt and E. R. M. Madeira, "HCOC: a cost optimization algorithm for workflow scheduling in hybrid clouds," Journal of Internet Services and Applications, vol. 2, pp. 207-227, 2011.

[19] S. Abrishami, M. Naghibzadeh, and D. H. Epema, "Deadline-constrained workflow scheduling algorithms for Infrastructure as a Service Clouds," Future Generation Computer Systems, vol. 29, pp. 158-169, 2013.

[20] N. Sooezi, S. Abrishami, and M. Lotfian, "Scheduling Data-Driven Workflows in Multi-cloud Environment," in Cloud Computing Technology and Science (CloudCom), 2015 IEEE 7th International Conference on, 2015, pp. 163-167.

[21] D. Fernández-Cerero, A. Jakóbik, D. Grzonka, J. Kołodziej, and A. Fernández-Montes, "Security supportive energy-aware scheduling and energy policies for cloud environments," Journal of Parallel and Distributed Computing, vol. 119, pp. 191-202, 2018.

[22] Y. Wen, J. Liu, W. Dou, X. Xu, B. Cao, and J. Chen, "Scheduling workflows with privacy protection constraints for big data applications on cloud," Future Generation Computer Systems, 2018.

[23] A. Abraham, H. Liu, and T.-G. Chang, "Variable neighborhood particle swarm optimization algorithm," in Genetic and Evolutionary Computation Conference (GECCO-2006), Seattle, USA, 2006.

[24] P. S. Naidu and B. Bhagat, "Secure workflow scheduling in cloud environment using modified particle swarm optimization with scout adaptation," International Journal of Modeling, Simulation, and Scientific Computing, vol. 9, p. 1750064, 2018.

[25] K. Pradeep and T. P. Jacob, "CGSA scheduler: A multi-objective-based hybrid approach for task scheduling in cloud environment," Information Security Journal: A Global Perspective, vol. 27, pp. 77-91, 2018.

[26] "Work flow Simulator code," https://github.com/WorkflowSim.

[27] G. Juve, A. Chervenak, E. Deelman, S. Bharathi, G. Mehta, and K. Vahi, "Characterizing and profiling scientific workflows," Future Generation Computer Systems, vol. 29, pp. 682-692, 2013.

[28] A. Mohsenzadeh, H. Motameni, J. Vahidi, "A fuzzy trust evaluation mode to enhance security of cloud system entities with petri net," Journal of Electonic and Cyber Defence, vol. 4, 2016. (In Persian)

[29] M. Naghibzadeh, "Modeling and scheduling hybrid workflows of tasks and task interaction graphs on the cloud," Future Generation Computer Systems, vol. 65, pp. 33-45, 2016.

Journal of Electronical & Cyber Defence
Vol. 7, No. 4, 2020, Serial No. 28

# Secure and Confidential Workflow scheduling in hybrid cloud with improved Particle Swarm optimization algorithm

**M. Mehravaran, M. R. Pajoohan*, F. Adibnia**

*School of Computer Engineering, Yazd University

(Received: 12/09/2018, Accepted: 05/03/2019)

## ABSTRACT

*While private clouds provide high security and low cost for scheduling workflow, public clouds in addition to higher costs, are potentially exposed to the risk of data and computation breach. However, in real world, needs for high performance resources and high capacity storage devices encourage organizations to use resources in public clouds. Task scheduling, therefore, is one of the most important challenges in cloud computing. In this paper, whilst considering the security issue, a new scheduling method is proposed for workflow applications in hybrid cloud. Specifically, sensitivity of tasks, which has been considered in recent works, as well as security requirement for data and security strength for both resources and channels are taken into account. The proposed scheduling method is implemented in improved Particle Swarm Optimization (PSO-WSCS) algorithm. The goal function, is minimizing the security distance of data and workflow from security strengths of resources and channels such that time and budget constraints are observed. The proposed PSO-WSCS algorithm, which is based on original PSO with some modifications, is compared with three similar scheduling algorithms, namely VNPSO, MPSO and MPSO-SA, in hybrid cloud. Evaluations show the effectiveness of our algorithm in finding resources whose security aspects have high resemblance to the security requirements. This is displayed by an average improvement of 40% in the studied samples.*

**Keywords:** Hybrid cloud, Task scheduling, Security requirements of task and data, Security strength of resource and communication paths

* Corresponding Author Email: pajoohan@yazd.ac.ir