

# ضرورت کاربرد آمار در تحقیقات و آزمایشات

شماره ۱

## مقدمه

در سالهای اخیر متخصصان با واژه تحقیق، در حوزه کاری خود، رادیو، تلویزیون، روزنامه، اینترنت و حتی در کوچه و خیابان و بازار زیاد برخورد کرده اند و شنیده اند که موفقترین تولید و/یا تجارت در دنیای امروزه، که حرفه های موجود به سرعت در حال فرارویدن، فروپاشیدن و ادغام میباشند، آن است که همواره به نهاد پژوهش اهمیت لازم را بدهد.

در دنیای امروز برای سپردن فرزندان خود به یک سیستم آموزشی که در منطقه شما وجود دارد احتیاج به تحقیق علمی دارید و الا چندی میگذرد و متوجه خواهید شد که با تمام زیرکی، توجه و سرمایه گذاری تان رودست خورده اید و نتیجه مطلوب را نگرفته اید، اگرچه حتی بر اساس استانداردهای موجود بین المللی هزینه خوبی هم پرداخت کرده اید. میدانیم که وجود چنین فضایی، هم احتیاج به نهادهای مورد نیاز، بستر تاریخی و فرصت تمرین، توسعه، آزمایش و خطا دارد، و هم احتیاج به فرهنگ، علم، علاقه، شکیبایی و دلسوزی لازم برای این تمرین.

امروزه هر شرکت ولو کوچک، یک دایره یا بخش با نام **تحقیق و توسعه** (اصطلاحاً **R&D**) دارد. اگر این بخش به هیچ کار نیاید و به هیچ دردی نخورد، برای بازاریابی که خوب است، برای جلب مشتری که خوب است، برای شناسایی نهادهای تولیدی مورد نیاز، قیمت ها و جهت بازار که خوب است! اکنون شاید نگرشی دیگر لازم است (صرف نظر از اینکه آیا در گذشته نیاز بوده است یا خیر)!

در دهکده جهانی امروز برای هر فعالیت تولیدی و خدماتی باید تحقیقات لازمه در حوزه های فنی و بازرگانی (و دیگر حوزه های لازم) آن حرفه انجام گیرد، زیرا رقابت در این دهکده کوچک، بسیار پیچیده و تحلیل اطلاعات در آن به مراتب سخت تر از آن دنیای قدیم است که سقف آن بالای پشت بام خانه ها بود.

به هر حال پژوهش یا تحقیق وسیله خوبی است ولی احتیاج به ابزارهای مناسبی نیز دارد و الا ان قلت های ناشتایی میشود که گره کور تحقیقات را کورتر میکند و آن سقف یاد شده را پایین تر می آورد. در این مقاله اشاره مختصری به انواع تحقیق میشود و سپس به زبان ساده نقش آمار در تحقیقات علوم کامپیوتر مورد بحث قرار میگیرد.

شبکه های کامپیوتری از حیث مفاهیم اولیه برای عموم قابل درک هستند (هرچند به آن صورت تجربه کار با محیط های پیچیده و تخصصی موجود نباشد)، لذا در این مقاله نیز برای تسهیل درک، شبکه های کامپیوتری را بحث می کنیم و موضوع عملکرد مسیریابی را به صورت میکروسکوپیکی ولی قابل فهم مورد نظر قرار می دهیم، بنابراین از پرداختن به مفاهیم یا سناریوهای پیچیده اجتناب می کنیم، لذا درک این مقاله احتیاج به دانش فنی چندانی نیز نخواهد داشت.

## تعریف تمقیق و انواع آن

تحقیق طبق تعریف "فرایند جست و جوی منظم و هدفدار برای مشخص نمودن یک موقعیت نامعین" است. در این تعریف چهار وجه متمایز وجود دارد:

۱. تحقیق فرایند است،

۲. فرآیند از نوع جست و جو است،

۳. این فرآیند منظم است،

۴. این فرآیند هدفدار است،

۵. این فرآیند برای روشن شدن یک موقعیت نامعین (وضعیت یا چیز ناشناخته) بکار می رود.

بدیهی است که تحقیق احتیاج به "مهارت پژوهشی" دارد. مهارت پژوهشی به معنی فراهم کردن "اطلاعات پژوهشی" (وظیفه CIO) و "ملزومات، لوازم و فناوری پژوهشی" (وظیفه CTO) و "دانش پژوهشی" (وظیفه CKO) در پیشبرد تحلیل

است. اطلاعات پژوهشی در نتیجه "انجام یا اجرای طرح تحقیقی" و "گردآوری طرح ها و داده ها" بدست می آید. تحقیق بر چهار نوع است:

۱. **تحقیق تاریخی:** در این راهبرد، داده‌ها و اطلاعات تحقیق از رویدادها و وقایع گذشته بدست می آید. اکنون در علوم کامپیوتر تنها میتوان به این اطلاعات مراجعه کرد و با در نظر گرفتن رشد سریع علم در حوزه کامپیوتر نمی توان به این اطلاعات بسنده کرد، بلکه شاید تنها به عنوان یک نقطه شروع یا یافتن ذهنیت مناسب مطرح باشند. محقق توان تغییر رویدادهای گذشته را ندارد. برای "بررسی وضعیت اقتصادی سکنه شهرها در دوره ساسانیان یا نوع تغذیه آنها در این دوره" هیچ ابزاری جز مراجعه به متون و سنگ نبشته‌ها یا روشهای احتمالی دیگر وجود ندارد، در نتیجه این تحقیق تاریخی نام می گیرد. در این نوع تحقیق، صحت و سقم اطلاعات تاریخی و پیرایش آنها از اهمیت خاصی برخوردار است. لازم به ذکر است که متاسفانه در ایران، بیشتر به دلیل وجود ماهیت "انتقال علم" (به جای تولید علم)، اکثرا تحقیقات نوعا بازنگری، بازسازی، تلخیص یا تلقیح است، لذا شاید بتوان آنها را تماما در دسته تحقیقات تاریخی جای داد.

۲. **تحقیق توصیفی:** در این شیوه عمدتا به بررسی وضع موجود پرداخته می شود. در این نوع تحقیق، متغیرهای زیادی در حوزه موضوع مورد تحقیق دخالت دارند که بسیاری از آنها متغیرهای ناخواسته هستند که احتمالا قابل کنترل نیز نمی باشند، مثلا "ارتباط سن ورود به مدرسه دانش آموزان ابتدایی و گسترش مهارت‌های یادگیری آنها" را در نظر بگیرید. یادگیری بچه‌ها تحت تاثیر عوامل متعدد موجود در محیط غیر مدرسه، همچون خانه است (مثل ارتباط با والدین، رفتار والدین و نزدیکان با کودک، نوع تغذیه و غیره) که به عنوان یک متغیر ناخواسته هیچ کنترلی بر روی آن نمی توان داشت.

وجود چنین شرایطی بر روی نوع پژوهش و روش تحقیق مورد استفاده، موثر است. در تحقیقات توصیفی روشهای تحقیق پیمایشی، همبستگی، مقایسه‌ای، مطالعه موردی و اقدام پژوهی وجود دارد که هر یک به فراخور مسئله و وضعیت ناشناخته موجود بکار می رود. بعلاوه در حوزه علوم کامپیوتر، در مدرن ترین و پویاترین دسته بندی تحقیقی از گزارش نویسی، مستندسازی، گزارش روند، گزارش چرخه و ... به عنوان یک تحقیق تلقی می شود که در زمره تحقیقات توصیفی جای میگیرند. در حوزه هایی که اساس فعالیت بر مبنای رفتار و سپس تفسیر آن رفتار است، تحقیقات توصیفی بسیار پر کاربرد و بلاعوض می باشند، مثلا حوزه هوش مصنوعی و زیر شاخه های آن از این دسته اند. در این راستا بکارگیری علمی چون ریاضی رفتار (تئوری بازی)، منطق فازی، رویداد ناقص، مقادیر حجیم و ... مفید فایده واقع می شوند و حتی می توانند به عنوان یک کارگروه نیز مورد نظر قرار گیرند.

۳. **تحقیق آزمایشی:** در تحقیقات آزمایشی، متغیرهای ناخواسته طبقه بندی و کنترل می شوند و بسته به نوع طبقه بندی و نوع مسئله ناروشن، روش تحقیق مناسب استعمال می شود. تحقیقات آزمایشی بیشتر در حوزه مسائل ملموس و قابل تعمق بکار می رود (مثلا علوم تجربی، یا در علوم کامپیوتر بررسی لایه ای یک موضوع خاص). لازم است متغیری را که میخواهیم اثر آن را بررسی کنیم به نوعی منطقی طبقه بندی کنیم و از طرح آماری مناسبی استفاده کنیم و داده‌های بدست آمده را تحلیل کنیم. مستندات، گزارشات و طرح های (proposal) گذشته و فعلی نمونه ای از منابع این تحقیق هستند.

توجه شود که این نوع تحقیق رفته رفته با تحقیق تاریخی (که رویکردی همسان دارد) هم پوشانی نکند و نهایتا با آن جایگزین نشود!

۴. **تحقیق مکاشفه ای:** اصول و فرضیات این نوع تحقیق اصولا نرم و قابل تغییر هستند. در این نگرش تحقیقی، هدف به عنوان یک آرمان انگار می شود. ابزار، فرضیات و اصول، تماما تحت الشعاع آرمان قرار می گیرند. لذا شاید مفروضات در زمان  $x$ ، در زمان  $x+1$  بی اعتبار باشند. لازم به ذکر است که در دیگر راهبردهای تحقیقی، هدف به عنوان یک مسئله قابل دسترس و ملموس مطرح است، اما در این نوع راهبرد تصویر کاملا واضحی از هدف در

دسترس نیست و تحت الشعاع قرار گرفتن تمام ملزومات تحقیق نیز دقیقا به همین دلیل انجام می شود که تصویر روشنی از هدف بدست آید (با هر میزان بها و هر میزان انرژی). لذا دو مورد یاده شده ی آخر را می توان به عنوان فاکتورهای اساسی در تفاوت این نوع تحقیق با دیگر راهبردهای تحقیقی دانست. لازم به ذکر است که فعالیت در قالب این راهبرد اندکی مشکل و غیر قابل تحمل و درک می نماید و این موضوع دقیقا به دلیل عدم آشنایی محقق با فضا و ضمائ این راهبرد است. اما نتایج بدست آمده همواره با ارزش بوده و حتی می توانند به عنوان یک ایده، پایان نامه، طرح یا سبک نیز شناخته شوند. ذکر جمله ای از آقای ویلیام بلیک خالی از لطف نیست:

**I must create my own System, or be enslaved by another Man's. I will not Reason and Compare, my business is to Create! (William Blake)**

در هر حال، هر گونه تحقیق در دسته های اخیر مستلزم به کار بردن آمار است. نمیتوان ادعای علمی داشت ولی آن ادعا بدون کاربرد زبان آمار باشد.

## آمار در تمقیقات

آمار روش تحلیل مبتنی بر علم ریاضی است که بسته به نوع مسئله ی نامعین، امکانات مالی، توان تحقیقی و غیره میتوان با استفاده از آن و با دقت ریاضی مناسب، تفاوت بین پدیده ها را بیان کرد، چه پدیده های تحقیقی، چه پدیده های روزمره. در آمار دو نکته محوری باید مورد نظر واقع گردد:

- تفاوت و ارتباط جمعیت و نمونه
- شاخصهای آماری مناسب در آمار

### جمعیت و نمونه

در آمار به جای جمعیت از نمونه استفاده میشود. معنی این حرف چیست؟ فرض کنید بخواهیم اثرات و نتایج یک عمل مشخص را در مسیر یابهای مرزی (border) شبکه های سوئیچی (Switched) با یک محدوده بار ترافیکی مشخص بررسی کنیم (که هر کدام توپولوژی و پیاده سازی خاص خود را دارند). بدیهی است که نمی توانیم تمام شبکه های سوئیچی را بررسی کنیم (که واقعا دارای بار حقیقی هستند، نه بار آزمایشی) و نهایتا به پاسخ خود برسیم. نمی توانیم، چون نه از نظر اقتصادی ممکن است و نه از لحاظ بعد مکان، و اگر لابراتوار یا تجهیزات مورد نیاز در اختیار نداشته باشیم، نه از نظر مدیران شبکه ای که مایل به همکاری باشند! تمام شبکه هایی که در حوزه تعریف ما صدق می کنند، "جمعیت" نامیده میشود. با استفاده از مفاهیم معتبر و علمی آمار، به جای استفاده از کل جمعیت شبکه ها، "نمونه" مناسبی از آن جمعیت را مورد بررسی قرار می دهیم و به بررسی اثر آن می پردازیم و نتایج بدست آمده را در مورد این جمعیت خاص بکار می بریم و به عنوان گزارش منتشر می کنیم. بدیهی است که بر اساس معیارهای ما در حوزه تعریف شبکه های سوئیچی و همچنین میزبان ترافیک و نوع ترافیک، تعداد شبکه های موجود در نمونه نباید از حدی کمتر و از حدی بیشتر باشند.

اگر شبکه های نمونه کم باشند، نمی توان نتایج تحصیل شده را در مورد جمعیت، عمومی دانست. اگر شبکه های نمونه زیاد باشد و کار آزمایشی و روند فراوری پیچیده باشد، نمی توان یکنواختی آزمایش را تضمین کرد. به طور کلی هر چه تعداد نمونه های آزمایش بیشتر باشد بهتر است، به شرط اینکه یکنواختی طرح دستخوش تغییر نشود. این مطلب بدان معنی نیست که استفاده از تعداد نمونه های کم غلط و نا به جا است، بلکه انتخاب شبکه های نمونه نیز خود دارای روش های آماری خاص است. آمار به ما فرمان می دهد که اگر خواستیم آزمایشی انجام دهیم حداقل اندازه نمونه چقدر است که کمتر از آن آزمایش را دچار خطا می کند و بیشتر از آن از نظر علمی خاصیت و شاخصی در بر ندارد.

## شاخص‌های آماری میانگین و پراکنش

در آمار از تعاریف و شاخصهای آماری مناسب استفاده میکنیم. بهترین مشخصات آماری برای پی بردن به ماهیت یک صفت یا اندازه‌گیری در یک جمعیت "میانگین" و "پراکنش" (انحراف از میانگین) است. با میانگین به اندازه کافی آشنا هستیم و میدانیم که برای بدست آوردن آن کافی است ارقام را با هم جمع و بر تعداد آنها تقسیم کنیم. پراکنش میزان انحراف از میانگین را نشان میدهد. فرض کنیم سه شبکه سوئیچی در اختیار داریم که در سه مسیر یاب مرزی خود ترافیک باری معادل زیر دارند (بدیهی است که به صورت انتزاعی اندازه تمام بسته‌ها ثابت و یکسان و واحد بار نیز مثلا هزار فرض شده):

۱. شبکه A با  $2/3$ ، ۲ و  $1/7$  بار بر بسته

۲. شبکه B با  $2/1$ ، ۲ و  $1/9$  بار بر بسته

۳. شبکه C با ۲، ۲ و ۲ بار بر بسته

برای محاسبه میانگین در هر دسته سه عدد را با هم جمع و تقسیم بر سه میکنیم. نتیجه این محاسبه برای هر سه شبکه عدد ۲ است. به بیان دیگر میانگین بار مرزی هر سه شبکه، یکسان و برابر ۲ است.

اما از حیث موازنه بار (Load Balancing) که شاید یکی از علل بررسی ترافیک بار در این شبکه نیز باشد، این سه شبکه تفاوت مشهودی با هم دارند. در یک شبکه بار هر سه مسیریاب برابر ۲ است، در شبکه دیگر با وجود اینکه میانگین بار مرزی ۲ و برابر شبکه قبلی است، اما بار تک تک مسیریاب‌ها با هم تفاوت دارد و در شبکه سوم تفاوت‌ها باز هم بیشتر می‌شود. در زبان آمار بیان می‌شود که گرچه سه شبکه میانگین بار مرزی یکسانی دارند، اما تفاوت آنها در پراکنش بار مسیریابها است. سه مسیریاب موجود در شبکه A دارای پراکنش باری بیشتری هستند، در حالیکه شبکه C، پراکنش بار ندارد و هر سه مسیریاب دارای بار یکسانی می‌باشند. در آمار هرگاه صحبت از میانگین یک جامعه یا گروه یا نمونه است، حتما پراکنش نیز بیان می‌شود تا تعریف دقیق تری از آن جامعه بدست داده باشیم.

حال ببینیم یک توده ی شبکه چه وضعیتی دارد. در یک توده شبکه نیز ممکن است میانگین بار مرزی ۲ باشد، ولی پراکنش توده متفاوت باشد. این توده متفاوت از توده ای است که تک تک مسیریاب‌های مرزی، ترافیک بار واحد ۲ داشتند. پدیده میانگین و پراکنش در مقایسه دو جمعیت نیز از اهمیت بسزایی برخوردار است. فرض کنیم یک توده شبکه، ترافیک بار مرزی واحد ۲ داشته باشد و تک تک مسیریابهای مرزی در این توده نیز بار واحد ۲ دارند. توده دیگری وجود دارد که میانگین ترافیک بار مرزی  $1/9$  دارد و تک تک مسیریابهای مرزی نیز بار واحد  $1/9$  دارند. چون در هر یک از توده‌ها پراکنش یا انحراف از میانگین نداریم، لذا بلافاصله می‌توانیم قضاوت کنیم که توده شبکه ای دوم بهتر از توده شبکه ای اول است. ولی در واقعیت هیچ گاه با چنین مواردی برخورد نخواهیم داشت. هیچ گاه توده ای نخواهیم دید که مسیریابهای موجود در آن دقیقا از نظر عملکرد یکسان باشند، هرچند صفات یکسانی داشته باشند، از جمله این صفات میتوان به پروتکل مورد استفاده، واسط‌های مورد استفاده، سیستم عامل مسیریاب و ... اشاره کرد. حتی لازم به ذکر است که اگر تمام مسیریابها از تمامی وجوه یکسان و معادل باشند، باز هم انتظار معادل بودن بار اشتباه است، من جمله به دلیل الگوریتم (های) مسیریابی مورد استفاده، وضعیت و حجم جداول‌های مسیریابی، تعداد مسیریابهای همسایه، میزان و لایه مورد نظر در شناخت مسیریابهای همسایه، تعدد مسیرها، حجم خطوط، وضعیت خطوط و بسیاری از فاکتورهای ریز و درشت دیگر.

در آزمایشات دقیق به خصوص در حوزه سنجش و نمره گذاری الگوریتم‌ها و/یا مسیریابها، معمولا از مسیریابهای یکسان در یک محیط ایزوله استفاده می‌شود. به همین دلیل هیچ گاه نمی‌توان دو توده را در چنین وضعیتی مقایسه کرد. بهتر است جمله بالا را مجدداً تکرار کنیم که اگر تمام مسیریابها یکسان باشند، باز هم انتظار معادل بودن بار اشتباه است. برای مقایسه دو جمعیت حتما باید معیارهای متناسب و مربوط در مورد هر جمعیت استعمال شود.

گفتنی است که راندمان بار مرزی در توده شبکه یک شاخص بررسی عملکرد است و می‌توان این شاخص را بررسی

کرد و به دلایل محیطی، ماهیتی و پیاده سازی در توده پی برد، ولی مقایسه دو توده با استفاده از راندمان بار مرزی معمولا مقایسه درستی نیست. برای این منظور لازم است دلایل ماهیتی در اسکلت شبکه را بررسی کرد و آنرا در فرمولی گذاشت و "شاخص بار" را بدست آورد. شاخص بار دو توده قابل مقایسه است، چون علاوه بر راندمان بار مرزی، شاخص های مهم دیگری نیز در آن منظور شده است و امکان مقایسه توده ها را فراهم می آورد. در واقع استفاده از این شاخص باعث میشود که دو جامعه (توده) مورد نظر استاندارد شود. وقتی این استاندارد سازی انجام شد، به طبع امکان مقایسه نیز فراهم است.

### مقایسات توده در یک آزمایش

در آزمایشهای تحقیقی مرسوم است که مسیر یابها را در سطح کوچکی ایزوله می کنند که ظرفیت تراکنش بار مشخصی داشته و دارای کمترین مسیر یاب همسایه باشد (بسته به نوع و عمق آزمایش حتی می توان این تعداد را به صفر رساند). برای هر معیار قابل مقایسه (که آزمایش برای یافتن اختلاف آنها طراحی شده است)، نمودهایی در نظر گرفته می شود و در هر نمود چند مسیر یاب قرار می گیرد. این بسیار مهم است که تعداد مسیر یابهای موجود در نمود و تعداد نمودهای اختصاص یافته به هر معیار، از نظر آماری تعریف شود. این تعداد نباید کم و نیز نباید خیلی زیاد باشد. لازم به ذکر است که معیار قابل مقایسه، در اصطلاح کامپیوتر "سایه" نام دارد. می گویند دو یا چند سایه در فلان آزمایش مقایسه شده است. برای مثال، دو پروتکل یکی بهمراه خاصیت  $X$  (در الگوریتم های) مورد استفاده در پروتکل و دیگری بدون خاصیت ذکر شده متوازن می شوند و هر یک در اختیار چند نمود قرار می گیرد. می توان پس از یک دوره زمانی مشخص، میزان تراکم بسته ها، ضریب فقدان بسته، ضریب گذردهی بسته، ضریب عملکرد قوانین (rule)، میزان اتصال، زمان توقف یا عکس العمل (latency) و بسیاری از موضوعات دیگر را مقایسه آماری کرد و دلیل آورد که پروتکل مورد نظر بهمراه خاصیت  $X$  بهتر عمل کرده، یا بدون خاصیت  $X$ ، یا اینکه گاه دو پیاده سازی موجود از پروتکل از منظر ترافیک شبکه کار آیی یکسانی داشته اند (اگرچه مورد اخیر بعید به نظر می رسد، اما سناریوهای این چنینی هم می توانند وجود داشته باشند).

در این آزمایش وجود یا عدم وجود خاصیت  $X$  در پروتکل، سایه ی آزمایش قلمداد می شود و دو سایه که گویی **دو** پروتکل هستند، با هم مقایسه می شوند. معمولا تحقیقات به همین منوال اجرا می شوند، منتها علاوه بر دو سایه ی یاد شده، از یک سایه ی شاهد نیز استفاده می گردد. سایه شاهد همان سایه ای است که معمولا تا کنون از آن استفاده می کرده اند. در مثال بالا اگر تا کنون از خاصیت  $X$  در پروتکل مورد نظر استفاده می کرده اند، سایه شاهد یا سایه کنترل آزمایش، سایه ای است که از خاصیت  $X$  استفاده می کند و سایه آزمایشی، همان سایه ای است که از خاصیت  $X$  استفاده نمی کند.

اکنون سوال می کنیم که در این نمایه ها چه چیز مقایسه می شود؟ جواب: میانگین ها. لذا برای ارائه دلایل آماری لازم است که میانگین سایه های مورد آزمایش در شرایط / استاندارد مقایسه شوند. شرایط استاندارد آزمایش، در واقع مجموعه خطاهای آزمایش است. خطاهای آزمایش یک عبارت آماری است و به این معنی نیست که سعی کنیم به طریقی آزمایش را با خطا انجام دهیم (این روند به طور جداگانه در مراحل جداگانه انجام می شود)! اگرچه خطای عمدی نیز جزو خطای آزمایش تلقی می شود. خطای آزمایش در واقع مجموعه عوامل ناشناخته ای است که در آزمایش وجود دارند و محقق قادر به کنترل آنها نیست (مثلا وجود noise در کابل شبکه و منتهی به کارت های واسط). هر یک از این عوامل به نوعی ضریب عملکرد را تحت تاثیر قرار می دهد و موجب تغییر میانگین نموده می شود. این عوامل در واقع خطای آزمایش می باشند. همانگونه که قبلا نیز ذکر شد اگر خطای آزمایش وجود نداشته باشد، کمترین تفاوت میانگین بین دو توده یا دو سایه حائز اهمیت می گردد. ولی همان طور که گفتیم، همیشه بین سایه ها و توده ها تفاوت وجود دارد، همان توده ها یا نمودهایی که از یک پروتکل مشخص استفاده می کنند!

نمودهایی که از یک پروتکل استفاده می کنند، از نظر آماری **تکرار (iteration)** های آن آزمایش را شکل می دهند. نکته قابل توجه این است که هر چه خطای آزمایش کمتر باشد، احتمال اینکه تفاوت های کمتر بین دو سایه، حائز اهمیت گردد، بیشتر می شود.

استفاده از تیم های تحقیقی که دارای متخصصین موضوع، متخصص تحقیق، تکنسینهای آشنا و سه نهاد اولیه مذکور در پیشبرد تحقیق هستند، این خطاها را کاهش می دهد. تصور کنید اگر تحقیقی انجام دهیم که تکنسین های مورد استفاده، تجربه کار کردن با مسیریابها (یا حداقل brand مسیریاب مورد نظر) را نداشته باشند، حساسیت و توانایی لازم را نداشته باشند یا اینکه عمدا خود را به اشتباه بیاندازند (شاید به دلیل حس راحت طلبی)! این موارد (به همراه بسیاری از عوامل دیگر) شرایطی را تشکیل می دهند که خطای آزمایش را تغییر می دهند و تصمیم گیری در مورد تصادفی یا واقعی بودن میانگین ها و نتایج بدست آمده در آزمایش را تحت تاثیر قرار می دهند.

فرض کنیم در همان آزمایش وجود خاصیت  $X$  در پروتکل، تعداد ۱۰ نمود در اختیار داریم. به هر یک از پروتکلها (پروتکل بدون خاصیت  $X$  و پروتکل همراه با خاصیت  $X$ )، ۵ نمود اختصاص یافته است و بسته به اندازه هر نمود، در هر نمود، ۵ الی ۱۰ مسیریاب قرار داده ایم. از ابتدا پروتکل هر نمود مشخص می شود و از ابتدا تا انتهای دوره آزمایش بدون اینکه اشتباهی صورت گیرد از آن پروتکل استفاده شده و در انتها عملکرد نمودهها را سنجیده و میانگین ضریب عملکرد مسیریابها و پروتکل مورد استفاده را تعیین می کنیم. نکته بسیار مهم اینکه در انتهای دوره آزمایش برای هر یک از پروتکل های دارای خاصیت  $X$  و بدون خاصیت  $X$ ، ۵ میانگین بدست می آوریم (که طبق توضیحات ارائه شده، به احتمال قریب به یقین با یکدیگر تفاوت دارند)، که هر میانگین، خود از ۵ الی ۱۰ مسیریاب بدست آمده است. حال با استفاده از طرح آماری مناسب می توان این دو پروتکل را مقایسه کرد و تفاوت آنها را از نظر آماری اعلام داشت.

در آزمایشات تحقیقی معمولا کلیه شرایط و عواملی را که به نوعی می توانند بر روی نتایج آزمایش تاثیر داشته باشند تحت کنترل در می آوریم. اگر این شرایط قابل کنترل نباشند، سعی می کنیم طرح آزمایشی مناسبی را برای آن بکار ببریم. در مثال فوق لازم است که مسیریابها فقط از نظر نوع سایه (در مثال بالا پروتکل مورد استفاده با/بدون خاصیت مورد نظر) با یکدیگر تفاوت داشته باشند و شرایط دیگر برای آنها و نمودهها و غیره کاملا یکسان باشد (هر چند ذکر کردیم که لاجرم اندکی ضریب خطا وجود خواهد داشت). اگر این کار صورت گیرد می توان داده های حاصله را آنالیز کرد و نتایج بدست آمده را از نظر آماری به عنوان ضریب عملکرد اعلام داشت.

نویسنده: سعید بیکی (cephexin@secumania.net)

© Secumania Security & Vulnerability Research Lab  
www.secumania.net

دی ماه ۱۳۸۶