مجموعه مقالات پانزدهمین کنفرانس بین‌المللی انجمن رمز ایران
دانشکده مهندسی برق، دانشگاه تربیت دبیر شهید رجایی، تهران، ۶-۷ شهریور ۱۳۹۷

ارائه معیاری جهت ارزیابی میزان تفکیک پذیری برای شناسایی تروجان سخت‌افزاری با اثر نشت کلید الگوریتم رمزنگاری

مسعود زیوری اخلاص[۱]، محمدعلی دوستاری[۲]، حامد یوسفی[۳و۴]

[۱] کارشناسی ارشد مهندسی کامپیوتر، دانشگاه شاهد، تهران، m.zivari@shahed.ac.ir

[۲] استادیار گروه مهندسی کامپیوتر، دانشگاه شاهد، تهران، doostari@shahed.ac.ir

[۳] دانشجوی دکتری مهندسی الکترونیک، دانشگاه شاهد، تهران، h.yusefi@shahed.ac.ir

[۴] پژوهشگر گروه رمز و امنیت، پژوهشگاه خواجه‌نصیرالدین طوسی، تهران، h.yusefi@rcdat.ir

چکیده

تروجان سخت‌افزاری (دستکاری خرابکارانه در مدارهای مجتمع) یکی از مهمترین مشکلات امنیتی در مدارهای مجتمع است. اهمیت این موضوع در سامانه‌های با کاربرد حیاتی، مثل سامانه‌های کنترل صنعتی یا زیرساخت‌های سایبری و پزشکی بیشتر است. تروجان سخت‌افزاری ممکن است در هرکدام از مراحل روند طراحی تا ساخت، در مدار مجتمع درج شود. برای ساخت مدار مجتمع، طرح چینش مداری به کارخانه‌های ساخت تراشه فرستاده می‌شود. این کارخانه‌ها تحت نظارت مستقیم قرار ندارند و ممکن است خرابکاری‌های عامدانه‌ای در مدار صورت گیرد. تاثیر تروجان‌های سخت‌افزاری از نشت کلید الگوریتم رمزنگاری تا تخریب کامل دستگاه را شامل می‌شود. در این مقاله، معیاری برای بیان میزان تفکیک روش تحلیل در سیگنال‌های توان مصرفی برای تشخیص نمونه‌ی تروجان‌دار از نمونه‌ی طلایی ارائه شده است. با استفاده از این معیار، می‌توان روش‌های کانال جانبی برای شناسایی تروجان سخت‌افزاری را با یکدیگر مقایسه کرد و همچنین برای رسیدن به شناسایی بهتر، پارامترهای موثر از جمله توزیع بردار ورودی را بهبود داد. برای ارزیابی عملی معیار ارائه شده، تروجان سخت‌افزاری AEST100 با اثر نشت کلید الگوریتم رمزنگاری AES بر روی مدار SAKURA پیاده‌سازی شده است. توان مصرفی این تراشه در مقایسه با نمونه‌ی بدون تروجان مورد تحلیل و ارزیابی قرار گرفته است. به منظور تحلیل سیگنال‌های توان از دو روش تشخیص الگوی SVD و PCA استفاده شده است.

کلمات کلیدی

شناسایی تروجان سخت‌افزاری، تحلیل کانال جانبی، محک تروجان AEST100

۱- مقدمه

پیچیدگی در مراحل طراحی و ساخت مدارهای مجتمع باعث شده است تا بستری جدید برای حمله در سامانه‌های الکترونیکی ایجاد شود. در این شکل جدید حمله، موجودیت‌های نامطمئن به صورت مستقیم یا غیرمستقیم در قطعه‌ی الکترونیکی یا مدار مجتمع جاسازی می‌شوند [۱]. همین نگرانی‌ها باعث شده که مساله‌ی تروجان‌های سخت‌افزاری در دهه‌ی اخیر مورد توجه بسیاری از پژوهشگران، صنعت‌گران و دولتمردان قرار گیرد. این موضوع در کشور ما با توجه به تهدیدات بالقوه امنیتی بسیار مهم است.

تروجان‌های سخت‌افزاری ممکن است درون تراشه‌های ASIC، ریزپردازنده‌ها، میکروکنترلرها، پردازنده‌های شبکه و یا حتی به عنوان ثابت‌افزار درج شوند. هدف و تاثیر این تروجان‌ها بسیار گسترده است. تروجان سخت-افزاری دارای تاثیرهایی از جمله افت کارایی سامانه، عدم خدمات‌دهی و یا نشت اطلاعات حیاتی است [۲]. روش‌های متفاوتی برای شناسایی تروجان سخت‌افزاری ارائه شده است. از بین این روش‌ها، روش‌های مبتنی بر کانال جانبی بیشتر مورد توجه قرار گرفته‌اند. این روش‌ها غیر مخرب هستند و

قابلیت پیاده‌سازی بیشتری نسبت به سایر روش‌ها دارند. در سال‌های اخیر معیارهای تحلیلی و تجربی برای شناسایی تروجان ارائه شده است. این معیارها بیشتر روی کوچک‌ترین تروجان قابل کشف و یا احتمال کشف تروجان تمرکز داشته اند [۳, ۴]. در [۵] میزان فعالیت و تاثیر تروجان روی پارامترهای کانال جانبی مورد بررسی قرار گرفته است و با استفاده از افزایش فعالیت تروجان، میزان شناسایی بهبود داده شده است.

حفره‌ی تحقیقاتی مهم در این حوزه، عدم وجود معیار شایستگی[۲]مناسب برای مقایسه نتایج روش‌های مبتنی بر کانال جانبی است. نبود امکان مقایسه روش‌های مختلف باهم، باعث می‌شود تا نتوان عوامل موثر در شناسایی تروجان را بهبود داد و میزان شناسایی را به صورت دقیق گزارش کرد. در واقع معیاری مشخص برای بیان میزان تفکیک‌پذیری نمونه تروجان‌دار از نمونه طلایی وجود ندارد. این پژوهش، تلاش دارد معیاری کاربردی برای بیان میزان تفکیک ایجاد شده در روش تحلیل شناسایی تروجان، ارائه دهد. با استفاده از این معیار، می‌توان با بهبود پارامترهای مختلف که در شناسایی موثر هستند، میزان تفکیک را افزایش داد و شناسایی بهتر و با درصد خطای کمتری داشت.

در ادامه‌ی این مقاله، در بخش ۲ به مبانی اولیه در مورد تروجان سخت-افزاری و اقدامات متقابل آن اشاره می‌شود. در بخش ۳ کارهای صورت گرفته در زمینه‌ی شناسایی تروجان سخت‌افزاری با روش‌های کانال جانبی بیان می-شود. در بخش ۴ راه‌حل پیشنهادی این مقاله و نتایج بدست آمده آورده می-شود. در بخش ۵ نیز جمع‌بندی اقدامات انجام شده در این پژوهش، آورده شده است.

۲- مبانی تروجان سخت‌افزاری

هرگونه دستکاری و تغییر در مدار اصلی توسط مهاجم برای نفوذ به سخت‌افزار یا استفاده از سازوکارهای سخت‌افزاری برای بدست آوردن داده‌ها یا کنترل عملکرد نرم‌افزار در حال اجرا روی تراشه، تروجان سخت‌افزاری گفته می‌شود. تروجان‌های سخت‌افزاری را می‌توان به طور عمده بر اساس ۶ دسته طبقه-بندی کرد. این دسته‌بندی، تروجان سخت‌افزاری را بر اساس شکل، فاز طراحی، سطح انتزاعی توصیف، نحوه‌ی فعال شدن، تاثیر، موقعیت و ویژگی-های فیزیکی طبقه‌بندی می‌کند [۲]. توسعه و استفاده دقیق از مدل‌های حمله برای پیشرفت در تحقیقات امنیتی حیاتی است. تحقیق در مورد تروجان‌های سخت‌افزاری در این مورد استثنا نیست. بنابراین هنگام توسعه‌ی یک تروجان سخت‌افزاری یا به طور کلی اقدام علیه تروجان‌ها باید ابتدا مدل حمله مشخص شود. مدل حمله به عنوان یک راهنما برای تحقیقات سایر افراد هم می‌تواند موثر باشد. به طور کلی مدل‌های حمله‌ای در تروجان سخت‌افزاری در [۱] ارائه شده‌است. این دسته‌بندی بر اساس اینکه کدام یک از مراحل زنجیره طراحی تا ساخت مدار مجتمع، امن و کدام ناامن است، انجام شده است.

بیشتر تحقیقات در حال حاضر روی اقدام متقابل برای کاهش تهدیدات بالقوه درج تروجان‌های سخت‌افزاری در زنجیره‌ی طراحی سامانه متمرکز شده است. به طور کلی برای مقابله با تهدید امنیتی تروجان سخت‌افزاری سه دسته اقدام صورت می‌گیرد. این سه دسته اقدام عبارتند از: ۱. شناسایی تروجان ۲. طراحی مدار مطمئن مدار و جلوگیری از درج تروجان۳. تقسیم فرآیند ساخت برای ساخت مطمئن مدار [۱].

یک روش قدرتمند برای شناسایی دستکاری‌های نامطلوب در مدار مجتمع استفاده از سیگنال‌های کانال جانبی است. در این روش، ویژگی فیزیکی، مثل مصرف توان (توان نشتی یا گذرا)، امواج الکترومغناطیس یا زمان تاخیر برای دستگاه تحت آزمون با نمونه‌ی طلایی مقایسه می‌شود. در روش کانال جانبی تقریبا همیشه به نمونه‌ی طلایی نیاز داریم. این روش از این مفهوم استفاده می‌کند که به طور کلی هرگونه تغییر در مدار سبب ایجاد تفاوت در پارامترهای آن می‌شود. مهمترین محدودیت روش‌های مبتنی بر کانال جانبی دو مورد است: ۱. ویژگی‌های فیزیکی اندازه‌گیری شده فقط با درج تروجان سخت-افزاری تغییر نمی‌کنند، بلکه ممکن است اختلاف ایجاد شده در پارامترهای مدار، به دلیل تفاوت در فرآیند ساخت باشد. ۲. اندازه‌گیری دقیق ویژگی فیزیکی انتخاب شده ممکن است دشوار باشد، مثلا اندازه‌گیری دقیق تاخیر برای یک مسیر خاص در مدارکار دشواری است [۱]. در بخش بعدی کارهای صورت گرفته در زمینه شناسایی تروجان با تمرکز بر روش‌های کانال جانبی مرور شده است.

۳- کارهای انجام شده در این زمینه

پژوهش‌های مختلفی در زمینه شناسایی تروجان سخت‌افزاری صورت گرفته است. مقاله [۶] نقطه آغاز مطالعات در حوزه شناسایی تروجان سخت‌افزاری با استفاده از پارامترهای کانال جانبی است. در [۶] از توان مصرفی به عنوان شناسه مدار استفاده شده است و با استفاده از روش تحلیل سیگنال PCA[۳]، تروجان‌هایی با ابعاد دو تا سه برابر کوچکتر از مدار اصلی به صورت عملی شناسایی شده‌اند. درپژوهش‌های بعدی سایر پارامترهای کانال جانبی مورد استفاده قرار گرفتند از جمله در [۷] که از جریان نشتی و [۸] از جریان دینامیک و [۹] از تاخیر و در [۱۰] از امواج الکترومغناطیسی استفاده شده است. در [۱۱] با استفاده از الگوریتم‌های یادگیری ماشین شناسایی صورت گرفته است. همچنین کارهای مختلفی در زمینه‌ی بهبود الگوی ورودی انجام شده است [۱۲–۱۴]. در ادامه به پژوهش‌هایی که از روش‌های تحلیل توان مصرفی و روش تشخیص الگو استفاده کرده‌اند، اشاره می‌شود.

در [۱۵] دو نوع تروجان که هر دو به صورت بمب‌زمانی هستند(۳۲ و ۱۹۲ بیتی)، در DES-64 bit درج شده‌اند. این دو تروجان به ترتیب ۱,۳ و ۷,۸ درصد ار فضای تراشه را اشغال کرده‌اند. از روش SVD[۴] برای پردازش سیگنال‌های توان استفاده شده است. در [۱۶] تروجان به صورت شمارنده و یک فرستنده‌ی AM برای ارسال سیگنال توان است. تروجان درج شده در این مقاله، ۱,۱ درصد از سطح تراشه را اشغال کرده است. مدار اصلی به صورت AES-128bit است و از الگوریتم PCA برای ایجاد تفکیک بین نمونه‌های توان استفاده شده است. در [۱۶] از مدارهای QFP استفاده شده است. در پژوهش [۱۷] از AEST100 [۱۸] به عنوان تروجان استفاده شده است. این تروجان ۲,۴ درصد از کل تراشه را اشغال کرده است. با استفاده از روش PCA و یک آزمون آماری، فرآیندی برای احراز هویت تراشه ارائه شده است. همچنین از فاصله‌ی ماهانالوبیس به عنوان معیاری برای فاصله استفاده شده است. برای مجموعه داده $(\vec{x})$ مقدار فاصله ماهانالوبیس در رابطه (۱) آورده شده است. در این رابطه ماتریس S ماتریس کوواریانس است. در این مقاله نیز از مدارهای QFP[۵] برای درنظر گرفتن تفاوت در فرآیند ساخت استفاده شده است.

$$D_M(\vec{x}) = \sqrt{(\vec{x}-\vec{\mu})^T S^{-1}(\vec{x}-\vec{\mu})}$$
$$(\vec{x}) = (x_1, x_2, \dots, x_N)^T \qquad (۱)$$

۴- معیار ارائه شده برای بیان میزان تفکیک‌پذیری روش شناسایی تروجان سخت‌افزاری

ایده‌ی اساسی در روش‌های شناسایی تروجان سخت‌افزاری با استفاده از پارامتر کانال جانبی به این صورت است که ابتدا در فاز آموزش از تعدادی تراشه که آلودگی یا سلامت آن‌ها احراز شده است، نمونه‌گیری می‌شود. سپس نمونه‌های دریافت شده با استفاده از روش‌های تحلیل سیگنال مورد پردازش قرار می‌گیرند و یک مقدار آستانه برای تفکیک بدست می‌آید. این مقدار آستانه در فاز آزمایش برای احراز هویت یک تراشه تحت آزمون بکار می‌رود. در فاز آزمون مساله اساسی قرار دادن تراشه تحت بررسی در یکی از دو دسته تروجان‌دار یا بدون تروجان است. برای دسته‌بندی نمونه‌های تروجان‌دار و بدون تروجان در انتهای فاز آموزش لازم است فاصله میان همه‌ی نمونه‌ها بدست آید. این مقادیر فاصله در فاز آزمون برای احراز هویت تراشه تحت بررسی بکار گرفته می‌شود. برای اندازه‌گیری فاصله میان نمونه‌ها، تمام نمونه‌ها را بر اساس اینکه تروجان‌دار یا بدون تروجان هستند، برچسب‌گذاری کرده و میزان فاصله آن‌ها از هم را در یک ماتریس n*n قرار می‌دهیم. در این ماتریس n تعداد نمونه‌ها است و $x(i,j)$ نشان‌دهنده فاصله نمونه‌ی iام از نمونه‌ی jام است. رابطه (۲) این ماتریس را نشان می‌دهد. معیار فاصله‌ای برای اندازه‌گیری اختلاف بین نمونه‌ها مناسب است که منجر به بیشترین تفکیک و کمترین درصد خطای تشخیص شود. معیار فاصله همبستگی می‌تواند بیشترین تفاوت بین دو نمونه‌ی تروجان‌دار و بدون تروجان را مشخص کند. رابطه (۳) نشان دهنده این معیار فاصله است. در این رابطه σ بیانگر انحراف از معیار و μ مقدار میانگین را نشان می‌دهد. بنابراین هر درایه ماتریس Distance بیانگر میزان فاصله همبستگی بین دو نمونه است (همه‌ی نمونه‌های تروجان‌دار و بدون تروجان).

$$Distance = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}_{n*n} \qquad (۲)$$

$$Corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \qquad (۳)$$

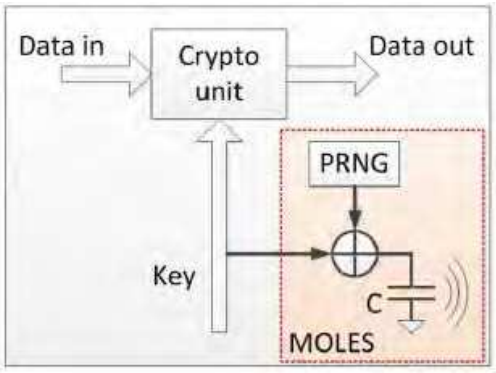
به هر میزان که فاصله بین دسته تروجان‌دار با دسته بدون تروجان بیشتر باشد، عملیات قرار دادن تراشه تحت بررسی در یکی از این دسته‌ها با کارایی بیشتر و درصد خطای کمتری انجام خواهد شد. پس از بدست آمدن ماتریس Distance می‌توان حدآستانه‌ای برای فاز آزمون تعیین کرد. این حد آستانه مشخص می‌کند که تراشه تحت بررسی در گروه تروجان‌دار یا بدون تروجان قرار می‌گیرد. هدف این مقاله ارائه معیاری برای بیان میزان فاصله بین دسته بدون تروجان و تروجان‌دار است. هرمقدار فاصله دسته تروجان‌دار با دسته بدون تروجان بیشتر باشد، حد آستانه‌ای که با استفاده از ماتریس Distance بدست می‌آید دقیق‌تر می‌شود و فاز آزمون خطای کمتری خواهد داشت. با فرض اینکه نمونه‌های کانال جانبی استخراج شده دارای توزیع نرمال هستند، میانگین μ و انحراف معیار σ مربوط به نمونه‌ها بدست می‌آید. مقدار $\mu+4\sigma$ به ازای هر دو دسته تروجان‌دار و بدون تروجان در نظر گرفته می‌شود. از آنجایی که می‌توان داده‌ها را در محدوده $\mu-4\sigma$ تا $\mu+4\sigma$ قرار داد، فاصله دو دسته تروجان‌دار و بدون تروجان از رابطه (۴) بدست می‌آید. بنابراین می‌توان معیار شایستگی برای میزان تفکیک‌پذیری شناسایی تروجان را به صورت رابطه (۴) بیان کرد.

$$FOM = |(\mu_{troj} + 4\sigma_{troj}) - (\mu_{free} - 4\sigma_{free})| \qquad (۴)$$

در این رابطه μ_{troj} و σ_{troj} میانگین و انحراف معیار مربوط به نمونه‌های تروجان‌دار، μ_{free} و σ_{free} میانگین و انحراف معیار مربوط به نمونه‌های بدون تروجان را نشان می‌دهد. لازم به ذکر است که اندازه‌گیری فاصله برای دو دسته بعد از اجرای روش تحلیل سیگنال روی پارامتر کانال جانبی صورت می‌گیرد.

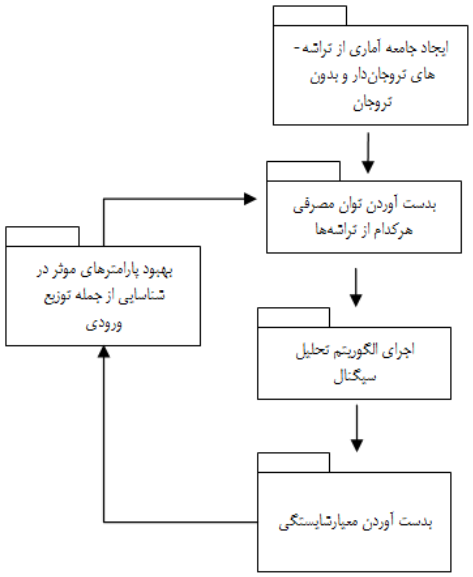
۵- ارزیابی عملی معیار ارائه شده

در این پژوهش از محک استاندارد AEST100 [۱۸] استفاده شده است. این تروجان در الگوریتم رمزنگاری AES در سطح RTL درج شده است و به صورت همیشه فعال است. اثر آن نشت کلید الگوریتم رمزنگاری است و ۲,۴ درصد از فضای کل تراشه را اشغال می‌کند. این تروجان از انباره‌ی Trust-hub دریافت شده است. شکل ۱ نشان دهنده این نوع تروجان موسوم به MOLES است. نکته‌ی اساسی در طراحی MOLES این است که، فقط و فقط حمله‌کننده باید بتواند کلید را استخراج کند. پارامترهای کانال جانبی(توان) نشتی توسط آزمون‌کننده مدار قابل دریافت است اما این پارامترها به صورتی هستند که تنها حمله‌کننده قادر به استخراج کلید از آنها است. در واقع، یک مولد عدد شبه‌تصادفی در مدار قرار داده شده است. چون فقط حمله کننده از مقدار اولیه‌ی این تولیدکننده‌ی مقدار شبه‌تصادفی آگاه است، موجودیت دیگری نمی‌تواند از این طیف بهره‌برداری موثر کند [۱۹].



شکل (۱) عملکرد تروجان نشت دهنده کلید الگوریتم رمزنگاری از طریق توان مصرفی [۱۹]

شکل ۲ روش ارائه شده برای ارزیابی عملی این پژوهش را نشان می‌دهد. گام اول در این فلوچارت تهیه یک جامعه آماری از تراشه‌های بدون تروجان و تروجان‌دار است. در مرحله بعد الگوی توان مصرفی هرکدام از این تراشه‌ها، تحت یک بردار ورودی بدست می‌آید و سپس برای ایجاد تفکیک بین نمونه‌ها از روش‌های تحلیل سیگنال استفاده می‌شود. روش‌های پیشرفته تحلیل سیگنال مانند PCA، SVD و SVM داده‌های تکراری را حذف می‌کنند و داده‌های اصلی را در فضایی جدید نمایش می‌دهند تا عملیات تحلیل و مقایسه راحت‌تر صورت گیرد. پژوهش‌های انجام شده، خروجی تحلیل سیگنال‌ها را به عنوان شاخص مقایسه نمونه مشکوک با طلایی درنظر می‌گیرند. اما روی میزان تفکیک‌پذیری نظری ارائه نمی‌دهند. پس از بدست آمدن نتایج مربوط به تحلیل سیگنال با استفاده از معیار شایستگی بدست آمده، میزان شناسایی صورت گرفته، گزارش می‌شود.



شکل (۲) : فلوچارت ارائه شده برای ارزیابی عملی

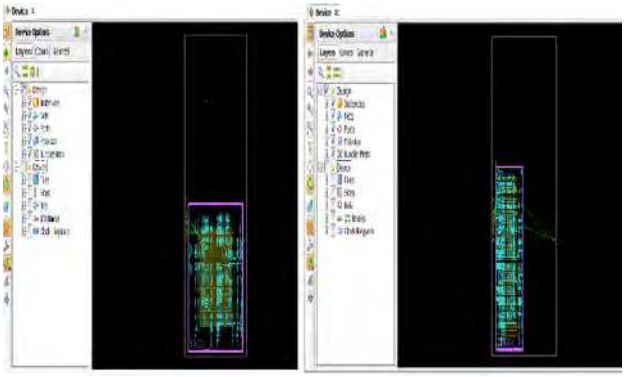
به دلیل محدودیت‌های موجود در ساخت تراشه و استفاده از مدارهای QFP، در این پژوهش روشی متفاوت برای شبیه‌سازی تفاوت در فرآیند ساخت و ایجاد جامعه آماری از تراشه‌ها، بکار گرفته شده است. مبنای این روش، قراردادن طرح در مکان‌های مختلف FPGA است. استفاده از این روش به دلیل تفاوت در مسیریابی و تفاوت طول سیم‌بندی دارای درصدی از خطا می‌باشد؛ اما از این درصد خطا صرف‌نظر شده است تا ایده ارائه شده ارزیابی شود. در پژوهش‌های بعدی می‌توان با استفاده از مدارهای QFP میزان این خطار را به حداقل رساند. هرچند که استفاده از این روش برای در نظر گرفتن تفاوت در ساخت، فرآیند شناسایی را دشوارتر می‌کند.

روند ساخت جامعه آماری از یک تراشه به صورت زیر است: ابتدا کد HDL مدار سنتز می‌شود. از ابزار ISE 14.7 برای سنتز مدار استفاده شده است. مرحله‌ی بعد از سنتز، ایجاد یک بلوک در فایل UCF برای قرار گرفتن تمام مدار در داخل آن است. به این منظور می‌توان از ابزار PlanAhead استفاده کرد. در گام بعدی تمام مدار سنتز شده در بلوک مورد نظر قرار داده می‌شود و مدار در مکانی از سطح FPGA جایگذاری می‌شود و عملیات مسیریابی اجرا می‌شود. در انتها رشته‌ی بیتی مدار تولید و روی FPGA پیاده‌سازی می‌شود.

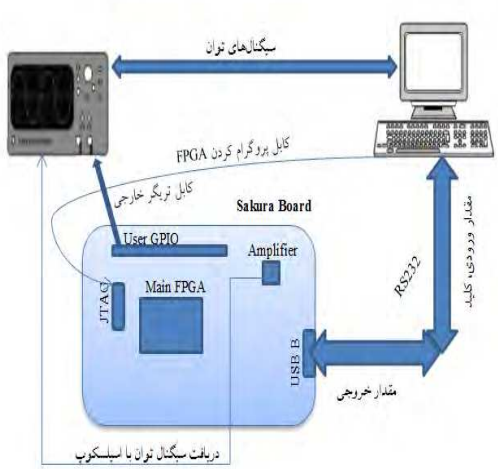
در FPGA های کمپانی Xilinx مدار پیاده‌سازی شده را می‌توان در قالب فایلی با پسوند XDL[۸] درآورد که در این فایل قابلیت جابجایی نقاط اتصال سیم‌ها (PIP) وجود دارد. می‌توان با استفاده از برنامه‌ریزی PIP، جابجایی بلوک‌ها را انجام داد که در این صورت مسیریابی ثابت می‌ماند.

شکل ۳ تغییر در جانمایی طرح در سطح FPGA را نشان می‌دهد. در پنج حالت مختلف مکان‌های متفاوتی برای مدار انتخاب شده است. با توجه به اینکه FPGA دارای ساختاری به صورت سطرها و ستون‌های با منابع یکسان است، این جابجایی در مکان‌های مختلف، تفاوت‌هایی را در ویژگی‌های اصلی مدار از جمله حداکثر فرکانس و حداقل تاخیر، توان مصرفی استاتیک و ... ایجاد می‌کند. این تفاوت‌ها را می‌توان به عنوان تفاوت در فرآیند ساخت در نظر گرفت. به عبارت دیگر چون منابع در مکان‌های مختلف FPGA یکسان هستند، در صورتی که در هر بار پیاده‌سازی از گروهی از آنها استفاده کنیم، پارامترهای مربوط به آن قسمت استفاده می‌شود. این نوع از شبیه‌سازی تفاوت در فرآیند ساخت، تفاوت در ویژگی‌های روی یک die (intra- die) را نشان می‌دهد.

برای دریافت سیگنال کانال جانبی در این پژوهش از مدار SAKURA [۲۰] استفاده شده است. این مدار دارای درگاه خروجی جریان مصرفی با اثر تقویت‌کنندگی است. شکل ۴ نحوه‌ی نصب تجهیزات را در این پژوهش نشان می‌دهد.



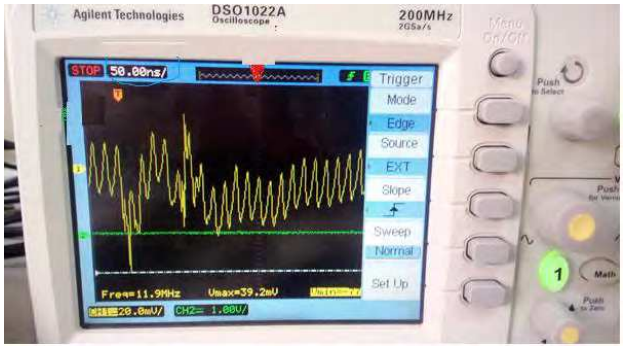
شکل (۳) : تغییر جانمایی طرح با استفاده از PlanAhead



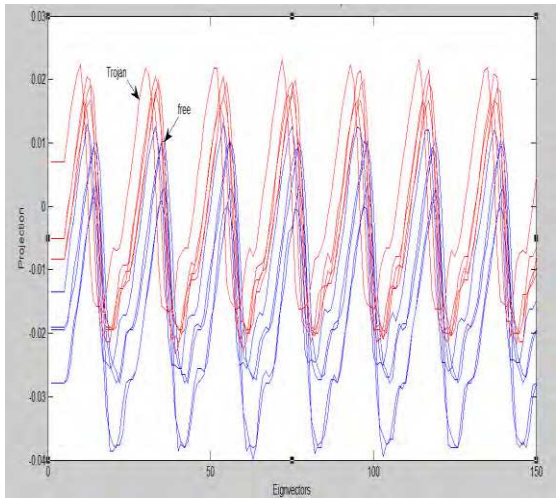
شکل (۴) : نحوه نصب تجهیزات به منظور دریافت توان مصرفی تراشه

در گام برنامه‌ای نرم‌افزاری توسعه داده شد که عملیات ارسال plain و کلید به FPGA، دریافت نتیجه الگوریتم رمزنگاری و ذخیره توان مصرفی مدار به ازای همان ورودی را به صورت دقیق کنترل می‌کند. شکل ۵ سیگنال توان مصرفی دریافت شده توسط اسیلسکوپ را نشان می‌دهد. تمام ۱۰ مرحله الگوریتم رمزنگاری AES در یک پالس کلاک انجام شده است. همچنین فرکانس کاری سخت‌افزار آزمون در این آزمایش ۴۸ مگاهرتز بوده است. نرخ نمونه‌برداری 1Gsa/s است.

پس از بدست آوردن نمونه‌های توان در هر دو جامعه آماری، تحلیل SVD با استفاده از ۱۵۰ بردار ویژه صورت گرفته است که تفکیک مناسبی را صورت داده است. شکل ۶ مربوط به نمایش سیگنال توان در جامعه آماری بدون تروجان و تروجان‌دار است. خطوط آبی رنگ نشان‌دهنده‌ی جامعه آماری بدون تروجان و خطوط قرمز مربوط به نمونه تروجان‌دار است. این نمونه توان در حالت plain و کلید تصادفی و ثابت بدست آمده است. منظور از ثابت بودن این است که در طول فرآیند نمونه‌برداری مقدار کلید و plain تغییر داده نمی‌شوند.



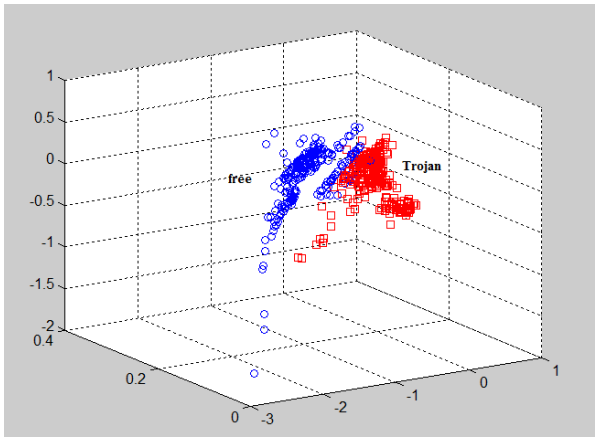
شکل (۵) : دریافت نمونه توان مصرفی



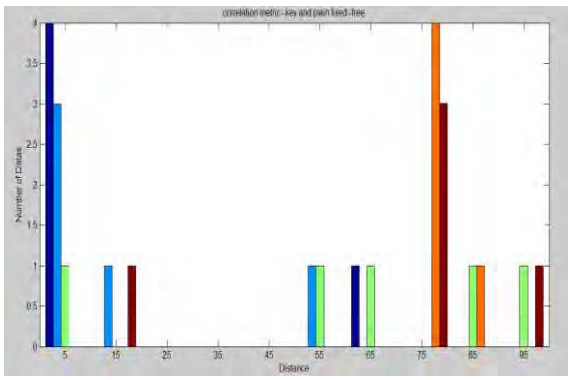
شکل (۶) : تحلیل سیگنال توان مصرفی با استفاده از روش SVD

در شکل ۷ اعمال روش تحلیل سیگنال PCA در سه بعد در حالت مشابه با حالت قبل دیده می‌شود. نقاط آبی رنگ مربوط به نمونه‌های بدون تروجان و نقاط قرمز رنگ نمونه‌های تروجان‌دار هستند. پس از اعمال روش‌های تحلیل سیگنال در انتهای فاز آموزش، لازم است حد آستانه برای شناسایی مشخص شود. برای بدست آوردن ماتریس فاصله از تابع pdist2 با معیار فاصله همبستگی در نرم‌افزار MATLAB استفاده شده است.

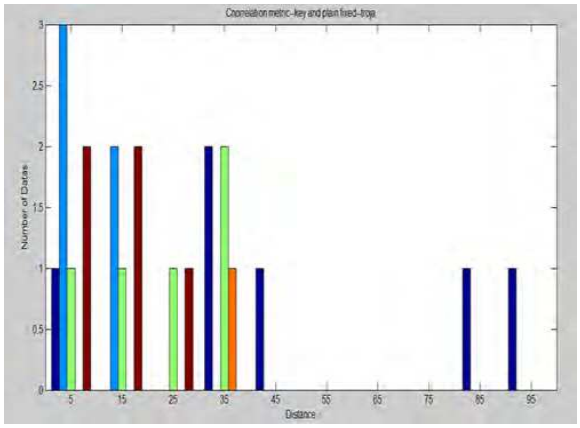
شکل ۸ نمایش هیستوگرامی ماتریس Distance در حالت کلید و plain ثابت بعد از تحلیل SVD است. همانطور که مشخص است توزیع میزان فاصله در نمونه‌ی بدون تروجان (۸– الف) بعد از ۵۵ می‌باشد و در نمونه‌ی تروجان‌دار (۸– ب) قبل از ۵۵ می‌باشد. از این مقدار در فاز آزمون استفاده می‌شود.



شکل (۷) : تحلیل سیگنال توان مصرفی با استفاده از روش PCA



الف



ب

شکل (۸) : نمایش توزیع میزان فاصله الف) بدون تروجان ب) تروجان‌دار

می‌توان فاصله نمونه‌ها در شکل ۷ را به طریق مشابه بدست آورد. معیار شایستگی در این حالت با استفاده از رابطه ۴ برای نتیجه شناسایی شکل ۶ و ۷ در جدول ۱ آورده شده است.

بنابر نتایج جدول ۱ می‌توان نتیجه گرفت که استفاده از روش PCA و الگوی ورودی کلید و plain ثابت می‌تواند منجر به شناسایی بهتری در تروجان سخت‌افزاری AEST100 شود. با استفاده از معیار شایستگی بدست آمده می‌توان روش‌های مختلف شناسایی تروجان را با هم مقایسه کرد، حد آستانه دقیق‌تری برای فاز آزمون بدست آورد و پارامترهای موثر در شناسایی از جمله توزیع ورودی را بهبود داد.

.

جدول (۱) : معیار شایستگی بعد از تحلیل سیگنال توان مصرفی بر اساس روش تشخیص الگو و توزیع ورودی

روش تشخیص الگو / توزیع ورودی	PCA	SVD
کلید ثابت، plain متغیر	0.0241	0.0103
کلید متغیر، plain ثابت	0.0200	0.0136
کلید متغیر، plain متغیر	0.0141	0.0075
کلید ثابت، plain ثابت	0.0286	0.0175

۶- نتیجه

در این مقاله ابتدا مبانی اولیه و کارهای صورت گرفته در حوزه‌ی شناسایی تروجان سخت‌افزاری با تمرکز بیشتر بر روش‌های کانال جانبی و تشخیص الگو بیان شد. در ادامه حفره تحقیقاتی موجود در این زمینه مطرح شد که امکان مقایسه بین روش‌های کانال جانبی و بهبود شرایط آزمایش برای تفکیک بیشتر وجود نداشت. به این منظور با استفاده از معیار فاصله همبستگی، میزان فاصله در نمایش سیگنال‌های توان مصرفی مربوط به نمونه‌های تروجان‌دار و نمونه طلایی بدست آمد. با استفاده از این فاصله می‌توان در فاز آزمون، نمونه تحت بررسی را در گروه تروجان‌دار یا بدون تروجان دسته‌بندی کرد. سپس معیار شایستگی ارائه و ارزیابی شد که توسط آن می‌توان روش‌های مختلف شناسایی تروجان با استفاده از پارامتر کانال جانبی را مقایسه کرد و حد آستانه برای تشخیص را به صورت دقیق‌تری تعیین نمود. این معیار با استفاده از روش تحلیل سیگنال PCA و SVD در محک AEST100 ارزیابی شد. در کارهای آینده لازم است تروجان‌های بیشتر و روش‌های تحلیل سیگنال بیشتری در توزیع‌های مختلف ورودی بررسی شوند.

مراجع

[1] Xiao, K., et al., *Hardware Trojans: Lessons Learned after One Decade of Research.* ACM Transactions on Design Automation of Electronic Systems (TODAES), 2016. 22(1): p. 6.

[2] Tehranipoor, M. and F. Koushanfar, *A survey of hardware Trojan taxonomy and detection.* IEEE Design and Test of Computers, 2010. 27(1): p. 10-25.

[3] Popat, J. and U. Mehta. *Transition probabilistic approach for detection and diagnosis of Hardware Trojan in combinational circuits*. in *India Conference (INDICON), 2016 IEEE Annual*. 2016 .IEEE

[4] Zhou, B., et al., *Cost-efficient acceleration of hardware Trojan detection through fan-out cone analysis and weighted random pattern technique.* IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2016. 35(5): p. 792-793

[5] Salmani, H., M. Tehranipoor, and J. Plusquellic, *A novel technique for improving hardware trojan detection and reducing trojan activation time.* IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2012. 20(1): p. 112-125.

[6] Agrawal, D ,.et al. *Trojan detection using IC fingerprinting*. in *Security and Privacy, 2007. SP'07. IEEE Symposium on*. 2007. IEEE

[7] Potkonjak, M., et al. *Hardware Trojan horse detection using gate-level characterization*. in *Proceedings of the 46th Annual Design Automation Conference*. 2009. ACM

[8] Salmani, H. and M. Tehranipoor, *Layout-aware switching activity localization to enhance hardware Trojan detection.* IEEE Transactions on Information Forensics and Security, 2012. 7(1): p. 76-87.

[9] Jin, Y. and Y. Makris .*Hardware Trojan detection using path delay fingerprint*. in *Hardware-Oriented Security and Trust, 2008. HOST 2008. IEEE International Workshop on*. 2008. IEEE

[10] Soll, O., et al. *EM-based detection of hardware trojans on FPGAs*. in *Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium on*. 2014. IEEE

[11] Shende, R. and D.D. Ambawade. *A side channel based power analysis technique for hardware trojan detection using statistical learning approach*. in *Wireless and Optical Communications Networks (WOCN), 2016 Thirteenth International Conference on*. 2016. IEEE

[12] Chakraborty, R.S., et al., *MERO: A statistical approach for hardware Trojan detection*, in *Cryptographic Hardware and Embedded Systems-CHES 2009*2009, Springer. p. 396-410

[13] Huang, Y., S. Bhunia, and P. Mishra, *MERS: Statistical Test Generation for Side-Channel Analysis based Trojan Detection*

[14] Banga, M. and M.S. Hsiao. *A novel sustained vector technique for the detection of hardware Trojans*. in *2009 22nd International Conference on VLSI Design*. 2009. IEEE.

[15] Wang, L.-W. and H.-W. Luo. *A power analysis based approach to detect Trojan circuits*. in *Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2011 International Conference on*. 2011. IEEE.

[16] Wang, L., H. Xie, and H. Luo. *Malicious circuitry detection using transient power analysis for IC security*. in *Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE), 2013 International Conference on*. 2013. IEEE

[17] Wang, L., et al. *Detection of information-leak hardware Trojan in AES cryptographic circuits*. in *Reliability, Maintainability and Safety (ICRMS), 2014 International Conference on*. 2014. IEEE

[18] Salmani, H., M. Tehranipoor, and R. Karri. *On design vulnerability analysis and trust benchmarks development*. in *Computer Design (ICCD), 2013 IEEE 31st*

[19] Lin, L., W. Burleson, and C. Paar. *MOLES: malicious off-chip leakage enabled by side-channels*. in *Proceedings of the 2009*

[20] *SAKURA-G (Side-channel AttacK User Reference Architecture)*, in *Version 1.0*, M.T. CO..LTD, Editor August 1, 2013.

زیرنویس‌ها

[1] Application Specific Integrated Circuit
[2] Figure Of Merit (FOM)
[3] Principal Component Analysis
[4] Singular Value Decomposition
[5] Quad Flat Package
[6] Malicious off-chip leakage enabled by side-channel
[7] User Constraint File
[8] Xilinx Design Language
[9] Programmable Interconnect Point