

پالایش ترکیبی خروجی ابزارهای خودکارشناسایی الگوهای طراحی در تشخیص الگوهای با ساختار مشابه

زینب اسمعیل پور^۱ و اشکان سامی^۲

^۱ دانشجوی کارشناسی ارشد، دانشگاه شیراز، دانشکده مهندسی برق و کامپیوتر، شیراز
esmailpoor1991@gmail.com

^۲ استادیار، دانشگاه شیراز، دانشکده مهندسی برق و کامپیوتر، شیراز
asami@ieee.org

چکیده

الگوهای طراحی، ساختارهای اثبات شده و قابل اطمینانی هستند که برای حل برخی از مشکلاتی که به طور مکرر در طراحی نرم افزار شی گرا رخ می دهد، بکار برده می شوند. هر الگو هدف و ساختار مشخصی دارد. بنابراین شناسایی خودکار الگوهای طراحی می تواند هدف و طرح کد نرم افزاری را بازیابی کند و هزینه و زمان نگهداشت پذیری را کاهش دهد.

تاکنون ابزارها و روش های زیادی برای شناسایی خودکار الگوهای طراحی ارائه شده است. خروجی تولید شده توسط این ابزارها هنوز هم دارای مثبت کاذب زیادی است (خصوصاً در شناسایی الگوهای با ساختار مشابه). از اینرو روش های پالایش خودکار نتایج ابزارها نیز پیشنهاد شدند. روش های پالایش سعی بر شناسایی مثبت کاذب نتایج، برای ارائه ی یک خروجی مطمئن تر دارند. در این مقاله یک پالایش ترکیبی ارائه شده است. ابداع این پالایش نسبت به روش های قبلی، استخراج معیارهای جدید با ثباتی است که به خوبی می توانند، یک الگو را در نظر گرفتن الگوهای با ساختار مشابه پالایش کند. خودکار سازی این پالایش با استفاده از روش های داده کاوی می باشد. این معیارها اولین دسته ای هستند که خصوصیت ثابت الگوریتمیک الگوی استراتژی را جستجو کرده اند. نتایج حاصل از روش پالایش ارائه شده، با دقت یادگیری ۹۸.۲٪ تا ۱۰۰٪ الگوها را بصورت صحیح شناسایی می کند.

کلمات کلیدی

الگوهای طراحی، نگهداشت پذیری، شناسایی خودکار، مثبت کاذب، ساختارهای مشابه، پالایش ترکیبی.

شده توسط [6] پت نام دارد. پت تنها الگوهای ساختاری، مثل الگوی Adapter^۴ را (برای فهم الگو های نامبرده به زیر نویس مراجعه شود) شناسایی می کند. پت یک سری قوانین پرولوگ از اطلاعات هر الگوی طراحی تولید و اطلاعات کد منبع را نیز در قالب پرولوگ ارائه می کند. سپس با پرس و جو نویسی الگوهای ساختاری را شناسایی می کند. در این ابزار، کار پالایش روی نتایج برای تشخیص مثبت های کاذب به صورت دستی انجام می گیرد. ابزار ارائه شده دارای ۴۰ درصد دقت می باشد.

در سال ۱۹۹۸ [10] معیارهای شی گرا عام نظیر تعداد صفت ها، رابطه ها و ... را برای نقش های هر الگوی طراحی و هم چنین برای کلاس های کد منبع محاسبه می کند. سپس براساس جستجوی معیارهای استخراج شده برای نقش ها در کلاس ها، کاندیدهای هر نقش را شناسایی و از مساله ی کوتاه ترین مسیر برای پیوند کاندیدها و تشکیل الگوی مورد جستجو استفاده می کند. به دلیل استفاده از

۱- مقدمه

طراحی یک نرم افزار شی گرا دشواری های خاص خود را دارد. دشواری از آن، طراحی یک نرم افزار شی گرا با قابلیت استفاده ی مجدد است. الگوهای طراحی دارای اهداف منحصر بفرد می باشند [1].

مهم ترین مسئله ی نگهداشت پذیری سیستم های نرم افزاری خصوصاً سیستم های قدیمی، نبود سند کامل از طرح سیستم و اهداف آن می باشد. بنابراین شناسایی الگوهای طراحی به صورت خودکار یا نیمه خودکار، باعث تسهیل در نگهداشت پذیری و قابلیت استفاده ی مجدد می گردد. الگوهای طراحی بر حسب کاربرد استفاده، در سه دسته ی ساختاری^۱، رفتاری^۲، و ایجادکننده^۳ قرار می گیرند.

ابزارها و روش های زیادی جهت خودکار سازی شناسایی الگوهای طراحی ارائه شده است. یکی از اولین کارها برای خودکار سازی شناسایی الگوهای طراحی در سال ۱۹۹۶ [6] ارائه شده است. ابزار تولید