

УДК 330.322; 004.42

*В. В. Колдовський,
к. е. н., доцент кафедри економічної кібернетики,
Українська академія банківської справи НБУ (м. Суми)
О. С. Поліщук,
студент магістратури відділу комп'ютерної інженерії, Східноєвропейський
Середземноморський Університет (Північний Кіпр, м. Фамагуста)*

МЕТОДИКА ОЦІНКИ ІНВЕСТИЦІЙНОЇ ПРИВАБЛИВОСТІ ВІДКРИТОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ МЕТРИК ПРОГРАМНОГО КОДУ

У статті запропоновано оригінальну методику оцінки привабливості інвестицій у відкрите програмне забезпечення на основі метрик програмного коду. Розглянуті метрики коду, придатні для вирішення задачі. Здійснено приклад використання методики для порівняння декількох відкритих програмних проектів ERP-класу.

The paper presents an original method of estimation of the attractiveness of investment in open source software metric software. Code metrics are considered suitable for solving the problem. An example of using methods for comparing several open software projects of ERP-class.

*Ключові слова: відкрите програмне забезпечення, інвестиції, метрики коду.
Key words: open source software, investment, code metrics.*

ВСТУП

У сучасних умовах економічний розвиток значною мірою залежить від використання інформаційних технологій, які, в свою чергу, засновані на програмному забезпеченні (ПЗ), що фактично являє собою головну складову будь-якої інформаційної системи. Ще десять-п'ятнадцять років назад практично все ПЗ, що використовувалося для автоматизації роботи підприємств купувалося у вигляді готового рішення у постачальників програмних продуктів чи розроблялося під замовлення. У такому разі замовник отримував, як правило, як сам товар (ПЗ), так і послуги з його підтримки, що дозволяло достатньо надійно захиститися від ймовірних проблем, які неминуче виникають внаслідок помилок і недоліків у програмному коді. Водночас протягом останніх декількох років для компаній відкриваються перспективи використання у все більших обсягах так званого "відкритого" ПЗ, яке являє собою безкоштовний програмний

продукт, що розробляється силами ентузіастів чи некомерційними, а іноді й комерційними організаціями за участі певних зацікавлених інвесторів, які обирають нехарактерні для традиційної, заснованої на товарно-грошових відносинах ринкової економіки, методи розповсюдження. Таким чином, на відміну від традиційного товару чи послуги, у яких є своя ціна, а взаємовідносини між клієнтом і постачальником формуються на основі класичних товарно-грошових відносин, у випадку використання відкритого ПЗ такі взаємовідносини часто відсутні, а традиційні методи оцінки інвестиційної привабливості на основі розрахунку грошових потоків, термінів окупності та ін. є погано пристосованими.

ПОСТАНОВКА ЗАВДАННЯ

Не вдаючись у розгляд питань, що стосуються цілей і методів моделі роботи розробників відкритого ПЗ, що може слугувати темою для окремого дослідження,

автори даної роботи вважають за доцільне запропонувати альтернативний підхід до оцінки інвестиційної привабливості подібних проектів, що не повинен повністю замінити існуючі підходи, однак доповнити їх і використовуватися паралельно. У якості об'єкта дослідження були обрані відкриті програмні продукти, які використовуються для вирішення задач автоматизації роботи підприємства.

Відкрите ПЗ поставляється у такому вигляді, що надає можливість безперешкодно ознайомитися з вихідним кодом. На нашу думку, найбільш прийнятним способом оцінки інвестиційної привабливості відкритого ПЗ, має бути розрахунок метрик коду.

Метрики коду — кількісні характеристики, які відображають певні властивості програмного коду [2; 3]. За їх допомогою існує можливість достатньо успішно отримувати оцінки стосовно якості коду і відповідно передбачати можливі проблеми, які він в собі приховує.

Існує достатньо велика кількість метрик, використання яких може бути доцільним на різних стадіях роботи над проектом. На стадії проектування кількісними методами оцінки якості можуть бути оцінка складності та відповідності шаблонам проектування UML-діаграм в цілому та окремих класів, об'єктно-орієнтовані метрики (NORM, RFC, WMPC1, WMPC2, LOCOM1, LOCOM2, LOCOM3). На стадії кодування для кількісної оцінки якості можуть бути використані метрики Холстеда, різні варіації метрики цикломатичної складності МакКейба (модифікована, строга, спрощена, актуальна цикломатична складність, метрика складності глобальних даних), інтервальна метрика Майерса, метрика Хансена, метрика Пивоварського та ін. Важливе значення мають абсолютні показники кількості рядків коду та інших структурних блоків, із яких складається комп'ютерна програма, а також відносні показники, які дозволяють виявити характерні співвідношення між структурними складовими типами для якісного ("здорового") коду та неякісного ("хворого"), який потребує оперативного втручання. Крім того, необхідні виявляти такі складові вихідного коду ПЗ, як дублювання коду, "мертвий код", не документований код, особливо складні для розуміння ділянки коду, ділянки коду, які є осередками підвищеного ризику. Для розрахунку метрик може бути використаний відповідний інструментарій, що дозволяє виконати

Таблиця 1. Метрики складності

№ п/п	Назва метрики	Формула розрахунку	Пояснення
1	Максимальна глибина блоків у методах	$BD = \frac{\sum MAX(BD)}{NMeth}$	BD – глибина блоків (Block Depth); NMeth – кількість методів (Number of Methods); MAX(BD) – максимальна глибина блоків методу (Maximum Block Depth of Method).
2	Цикломатична складність методів	$CC_{NMeth} = 1 + \sum Path$	CC _{NMeth} – цикломатична складність методу (Cyclomatic Complexity); Path – окремі шляхи виконання методу, викликаний виразами вибору, умов чи переходу.
3	Зважені методи	$WM = \sum CC_{NMeth}$	WM – зважені методи (Weighted Methods); CC _{NMeth} – цикломатична складність методу (Cyclomatic Complexity); NMeth – кількість методів.

обчислювання у автоматичному режимі. Таким чином, найбільш очевидним способом оцінки якості і відповідно ризиків, які може нести неякісний програмний код, у відкритому ПЗ, є використання метрик коду, обрахувати значення яких доцільно для деяких популярних відкритих проектів.

РЕЗУЛЬТАТИ

Для дослідження були обрані наступні популярні продукти ERP-класу (Enterprise Resource Planning): Compiere (<http://www.compiere.com>), ADempiere ERP (<http://www.adempiere.com>), Openbravo (<http://www.openbravo.com>). З метою визначення можливих відмінностей у значеннях метрик для різних версій програмних продуктів було проаналізовано також декілька різних випусків вихідного коду для кожного із продуктів.

Розрахунок метрик здійснювався за чотирма групами, які, на наш погляд, достатньо добре дозволяють охарактеризувати стан ПЗ:

- складності: максимальна глибина блоків у методах, цикломатична складність, зважені методи (табл. 1);
- залежності: рівень абстрактності, аферентні і еферентні зв'язки (табл. 2);
- метрики Холстеда: складність, зусилля, кількість операндів, кількість операторів, кількість унікальних

Таблиця 2. Метрики залежності

№ п/п	Назва метрики	Формула розрахунку	Пояснення
1	Абстрактність	$A_s = \frac{NAT}{NT} \cdot 100\%$	A _s – абстрактність (Abstractness); NAT – кількість абстрактних типів [класів та інтерфейсів] (Number of Abstract Types); NT – кількість типів (Number of Types).
2	Аферентні зв'язки	$AC = \sum NT_{out}$	AC – аферентні зв'язки (Afferent Couplings); NT _{out} – зовнішній тип, залежний від внутрішніх типів (Type Outside Target Elements Depending on Types Inside Target Elements).
3	Еферентні зв'язки	$EC = \sum NT_{in}$	EC – еферентні зв'язки (Efferent Couplings); NT _{in} – внутрішній тип, залежний від зовнішніх типів (Type Inside Target Elements Depending on Types Outside Target Elements).
4	Нестабільність	$I = \frac{EC}{AC + EC}$	I – нестабільність (Instability); EC – еферентні зв'язки (Efferent Couplings); AC – аферентні зв'язки (Afferent Couplings).
5	Відстань	$D = \frac{A_s}{100\%} + I - 1$	D – відстань (Distance); A _s – абстрактність (Abstractness); I – нестабільність (Instability).

Таблиця 3. Метрики Холстеда

№ п/п	Назва метрики	Формула розрахунку	Пояснення
1	Складність	$Diff = \frac{NUOpts \cdot NOpds}{2 \cdot NUOpds}$	Diff – складність (Difficulty); Eff – зусилля (Effort); NOpds – кількість операндів (Number of Operands); NOpts – кількість операторів (Number of Operators); NUOpds – кількість унікальних операндів (Number of Unique Operands); NUOpts – кількість унікальних операторів (Number of Unique Operators); Len – довжина програми (Program Length); Voc – словник програми (Program Vocabulary); Vol – обсяг програми (Program Volume); Opd – операнд (Operand); Opt – оператор (Operator); UOpd – унікальний операнд (Unique Operand); UOpt – унікальний оператор (Unique Operator).
2	Зусилля	$Eff = Diff \cdot Vol$	
3	Кількість операндів	$NOpds = \sum Opd$	
4	Кількість операторів	$NOpts = \sum Opt$	
5	Кількість унікальних операндів	$NUOpds = \sum UOpd$	
6	Кількість унікальних операторів	$NUOpts = \sum UOpt$	
7	Довжина програми	$Len = NOpts + NOpds$	
8	Словник програми	$Voc = NUOpts + NUOpds$	
9	Обсяг програми	$Vol = Len \cdot \log_2(Voc)$	

Таблиця 4. Метрики наслідуваності

№ п/п	Назва метрики	Формула розрахунку	Пояснення
1	Глибина наслідування	$DI = 1 + DI_{sc}$	DI – глибина наслідування (Depth of Inheritance); DI _{sc} – глибина наслідування суперкласу (Depth of Inheritance of Superclass).
2	Кількість підтипів	$NT_{Sub} = \sum T_{Sub}$	NT _{Sub} – кількість підтипів (Number of Subtypes); T _{Sub} – підтип (Subtype).

Таблиця 5. Розраховані значення метрик складності

Назва ПЗ	Максимальна глибина блоків у методах	Цикломатична складність	Зважені методи
Compiere 2.4.4	0,91	2,47	22 101
Compiere 2.6.0	1,09	2,47	78 248
Compiere 3.0.0	1,10	2,49	86 348
Compiere 3.2.0	1,04	2,16	81 743
Adempiere 2.5.3	1,08	2,46	70 516
Adempiere 3.1.0	1,08	2,47	70 934
Adempiere 3.4.0	0,84	2,04	117 552
Adempiere 3.5.4	0,72	1,80	138 369
Openbravo 2.5 alpha	1,34	4,19	24 970
Openbravo 2.5 beta	1,35	4,12	26 217
Openbravo 2.5 final	1,37	3,84	28 382

Таблиця 6. Розраховані значення метрик залежності

Назва ПЗ	Абстрактність, %	Аферентні зв'язки	Еферентні зв'язки	Нестабільність	Відстань
Compiere 2.4.4	6,8	23,23	48,61	0,98	0,05
Compiere 2.6.0	3,0	22,41	34,40	0,99	0,02
Compiere 3.0.0	3,6	24,44	35,09	1,00	0,03
Compiere 3.2.0	3,0	39,16	18,17	1,00	0,03
Adempiere 2.5.3	3,4	26,47	41,59	1,00	0,03
Adempiere 3.1.0	3,3	26,58	41,88	1,00	0,03
Adempiere 3.4.0	17,2	26,13	17,09	1,00	0,17
Adempiere 3.5.4	18,1	21,17	15,82	1,00	0,18
Openbravo 2.5 alpha	4,2	32,85	36,37	1,00	0,04
Openbravo 2.5 beta	4,5	33,48	37,46	1,00	0,04
Openbravo 2.5 final	5,9	36,17	12,91	1,00	0,06

операндів, кількість унікальних операторів, довжина програми, словник програми, обсяг програми (табл. 3);

— наслідуваності: глибина наслідуваності, кількість підтипів (табл. 4).

Автоматизований розрахунок метрик коду здійснювався за допомогою інструмента CodePro AnalytiX. Отримані результати для кожної з груп метрик (складності, залежності, Холстеда та наслідуваності) наведені відповідно у табл. 5—8.

Виходячи з даних табл. 5, можна зробити висновок, що в цілому показники метрик складності по продуктам, які розглядаються, мають близькі значення і не виходять за граничні межі (наприклад, для цикломатичної складності граничною величиною є значення, що наближається до 7). Водночас нескладно помітити, що в цілому Adempiere має нижчі значення максимальної глибини блоків у методах і цикломатичної складності ніж інші продукти, а Openbravo — навпаки, вищі. Це означає, що за інших рівних обставин Adempiere є більш привабливим продуктом з погляду складності його підтримки.

Відповідно до табл. 6 мож-

Таблиця 7. Розраховані значення метрик Холстеда*

Назва ПЗ	Складність	Зусилля	Кількість операндів	Кількість операторів	NUOpds	NUOpts	Довжина програми	Словник програми	Обсяг програми
Compiere 2.4.4	272	1 600 903 926	282 464	119 959	24 890	48	402 423	24 938	5 877 814
Compiere 2.6.0	359	6 650 054 998	799 580	376 653	54 524	49	1 176 233	54 573	18 509 085
Compiere 3.0.0	363	7 433 283 286	876 525	414 989	59 152	49	1 291 514	59 201	20 474 803
Compiere 3.2.0	341	6 765 214 231	864 186	383 147	62 166	49	1 247 333	62 215	19 863 747
Adempiere 2.5.3	345	5 647 209 792	710 708	339 640	49 478	48	1 050 348	49 526	16 381 121
Adempiere 3.1.0	340	5 656 470 046	719 565	343 509	50 736	48	1 063 074	50 784	16 618 065
Adempiere 3.4.0	430	13 000 000 000	1 300 000	580 000	75 000	50	1 860 000	75 000	30 000 000
Adempiere 3.5.4	390	14 000 000 000	1 500 000	670 000	100 000	52	2 200 000	100 000	36 000 000
Openbravo 2.5 alpha	305	2 481 117 011	390 587	166 844	25 004	39	557 431	25 043	8 145 249
Openbravo 2.5 beta	303	2 591 378 898	407 788	174 052	26 210	39	581 840	26 249	8 541 397
Openbravo 2.5 final	294	2 770 533 306	445 103	189 438	29 523	39	634 541	29 562	9 423 858

Примітка: значення для Adempiere 3.4.0 і Adempiere 3.5.4 отримані наближено.

Таблиця 8. Розраховані значення метрик наслідуваності

Назва ПЗ	Глибина наслідуваності	Кількість підтипів
Compiere 2.4.4	4,01	0,72
Compiere 2.6.0	3,64	0,81
Compiere 3.0.0	3,59	0,81
Compiere 3.2.0	3,49	0,75
Adempiere 2.5.3	3,70	0,85
Adempiere 3.1.0	3,70	0,85
Adempiere 3.4.0	3,31	0,95
Adempiere 3.5.4	3,24	0,96
Openbravo 2.5 alpha	3,95	0,70
Openbravo 2.5 beta	3,95	0,69
Openbravo 2.5 final	3,87	0,69

на побачити, що значення метрики "Нестабільність" у продуктах, що розглядаються, є рівним або наближається до 1. Подібне значення свідчить про слабку стійкість наведених продуктів до змін, достатньо високу вартість супроводу під час усунення недоліків чи реалізації нових функцій. Для продукту Adempiere версій 3.4.0 та 3.5.4 помітно суттєве зростання показника абстрактності (із 3,3% у версії 3.1.0 до 17,2% та 18,1% відповідно). Саме по собі значення даного показника не дозволяє говорити про переваги чи недоліки самого продукту, однак динаміка зростання свідчить про серйозні архітектурні зміни у його внутрішній будові у порівнянні із попередніми версіями. Подібні зміни, як правило, направлені на спрощення розуміння і підтримки продукту, що й підтверджується попередніми висновками на основі метрик складності табл. 5.

За значеннями метрик Холстеда (табл. 7) продукти Adempiere версій 3.4.0 та 3.5.4 є найбільшими за обсягами. За інших рівних умов більші за обсягами рішення є складнішими з точки зору підтримки, водночас, виходячи з даних попередніх метрик, вдала архітектура і внутрішня будова рішення може нівелювати ці недоліки.

За значеннями метрик наслідуваності (табл. 8) найменші і відповідно найпривабливіші показники глибини наслідуваності є у продукту Adempiere версій 3.4.0 та 3.5.4, хоча середній показник кількості підтипів має найбільше значення. Ці показники свідчать про достатньо вдалу архітектуру об'єктно-орієнтованого рішення, покладеного в основу проекту, оскільки зі зростан-

ням глибини наслідуваності ускладнюється підтримка рішення.

Таким чином, за сукупністю метрик достатньо очевидно є перевага продукту Adempiere, зокрема його останніх версій. Незважаючи на достатньо високі показники обсягу програмного коду (за метриками Холстеда), продукт має найкращі показники складності і наслідуваності.

ВИСНОВКИ

Відкрите ПЗ від традиційних комерційних рішень відрізняється, насамперед, високою доступністю, яка у багатьох випадках не передбачає початкові інвестиції у придбання продукту. Водночас вартість його підтримки може виявитися вирішальною у процесі подальшої експлуатації. На нашу думку, керуватися в процесі вибору з-поміж декількох схожих рішень слід не лише базуючись на їх функціональності, а й виходячи й з оцінки внутрішньої будови на основі метрик програмного коду, що може надати необхідні дані для більш точної оцінки інвестиційної привабливості. Напрямо подальших досліджень з даної тематики доцільно присвятити розробці інтегральних показників, що включають і економічні характеристики програмного проекту [1], а також поєднання запропонованого підходу із класичними методиками оцінки інвестиційної привабливості.

Література:

1. Визначення економічних параметрів інноваційних процесів на етапах життєвого циклу програмного забезпечення / В.В. Колдовський // Вісник Української академії банківської справи. — 2005. — № 1. — С. 105—113.
 2. Науково-методичні аспекти оцінки якості коду програмних проектів банківсько-фінансової сфери / Колдовський В.В. // Збірник тез доповідей XIII Всеукраїнської науково-практичної конференції "Проблеми і перспективи розвитку банківської системи України". — Суми: ДВНЗ "УАБС НБУ", 2010. — Т. 2. — С. 92—93.
 3. Разработка ПО: метрики программных проектов / В.В. Колдовский // Компьютерное обозрение, 2007. — № 13. — С. 69—72.
- Стаття надійшла до редакції 12.05.2011 р.