

## ТЕМА 15. МЕТРИКА ЯК ОСНОВА ВИМІРЮВАННЯ ЯКОСТІ

Атрибути програмної системи, що характеризують її якість, вимірюються з використанням метрик якості.

**Метрика** – це комбінація конкретного методу вимірювання (способу отримання значень) атрибута сутності та шкали вимірювання (засобу, використовуваного для структурування одержуваних значень).

Метрика визначає **міру атрибуту** – змінну, якій присвоюється значення в результаті вимірювання. Наприклад, сутність – «звіт про виявлені дефекти », атрибут – «список дефектів», метод вимірювання – «підрахунок кількості дефектів у списку», шкала – «цілі числа більше нуля», міра атрибуту – «загальне число дефектів», ім'я метрики (зазвичай однойменне мірою) – «загальне число дефектів».

Міра атрибута може бути безпосередньою, якщо вона не залежить від заходів інших атрибутів, або непрямую, отриманої по заходам інших атрибутів.

За визначенням стандарту ISO/IEC 9126-2 метрика якості програмної системи являє собою «модель вимірювання атрибута, який пов'язує з характеристикою якості ПЗ. Служить індикатором одного або багатьох атрибутів.

Метрика називається **базовою**, якщо в її основі лежить елементарний метод (примітив) вимірювання атрибута.

Для того щоб правильно користуватися результатами вимірювань, для кожної міри потрібно ідентифікувати шкалу вимірювання.

Стандарт ISO/IEC 9126-2 рекомендує застосовувати 5 видів шкали вимірювання значень (впорядкованих від менш суворих до більш суворим):

– номінальна шкала. Це класифікаційна шкала, що виконує категоризацію властивостей об'єкта, що оцінюється. Категорії не впорядковані. Наприклад, дефекти можуть класифікуватися на дефекти інтерфейсу, логіки, оголошення даних та ін.;

– порядкова шкала. Дозволяє упорядковувати характеристики за зростанням або убуттям шляхом порівняння їх з базовими значеннями. Наприклад, для рівня серйозності наслідків події шкала може містити значення «низький», «середній», «високий»,

– «критичний». Для рівнів СММ – {1, 2, 3, 4, 5}. Відстань між значеннями за шкалою не грає ролі. Характеристики, що мають номінальну або порядкову шкалу вимірювання, називаються якісними.

– інтервальна шкала. Відзначає істотні відмінності властивостей об'єкта, «дистанцію» між ними (наприклад, календарні дати або значення щільності дефектів. Використовується в арифметичних операціях і операціях порівняння. Нульове значення не припустимо;

– відносна шкала. Значення за цією шкалою розрізняються по відношенню до вибраної одиниці (наприклад, часу, змінюваному від 0 до безкінечності, або вартості). Застосовуючи цю шкалу можна розрахувати, наприклад, час між відмовами, розмір програмного компонента в SLOC та ін. Вважається найбільш кращою шкалою вимірів. Дозволяє застосовувати широкий спектр інструментів вимірювання (гістограм, діаграм Парето та ін.);

– абсолютна шкала. Це спеціальний випадок відносної шкали. У цій шкалі вказується абсолютне значення величини. Наприклад: «розмір програми дорівнює 2К», «число виявлених помилок одно 20». Виміряне значення метрики саме по собі не несе інформації про рівень задоволення вимог до якості. Для цих цілей шкала повинна бути розділена на області (ранги), які відповідають різним ступеням задоволення вимог. Приклади розподілу шкали:

- розподіл значень за двома категоріями – задовільні і незадовільні значення;
- поділ шкали за чотирма категоріями, обмеженим трьома рівнями значень – поточним, гіршим і плановим.

По мірі накопичення практики вимірювань і знань про вивчаємі атрибути шкал їх вимірювання можуть еволюціонувати від менш інформативних (номінальної і порядкової) до більш інформативним (відносної чи абсолютної).

## **2. Класифікація мір якості**

Для розробки процедур збирання даних, інтерпретації заходів та їх нормалізації з метою порівняння, потрібно розрізнити такі типи міри, визначених метриками.

**Міри розміру.** Показують розмір ПЗ в різних одиницях виміру, наприклад:

- функціональний розмір – облік функціональних можливостей ПЗ;

– розмір програми – облік кількості рядків вихідного коду, кількості модулів, кількості операторів на мові програмування;

– обсяг ресурсів, що використовуються працюючою програмою – облік об'єму оперативної пам'яті, дискової пам'яті або мережевого трафіку, завантаженості процесора – кількості оброблюваних інструкцій у секунду і ін.;

**Міри часу.** Наводять періоди реального часу (в секундах, хвилинах або годинах), процесорного часу (в секундах, хвилинах або годинах роботи процесора) або календарного часу (в робочих годинах, календарних днях, місяцях, роках), наприклад:

– час функціонування системи з безперервною або дискретною роботою компонентів програмного забезпечення (облік минулого часу (при роботі програмного забезпечення в системі протягом встановленого періоду часу), облік часу з моменту включення (при безперервному використанні вбудованого програмного забезпечення системи реального часу), облік нормалізованого часу (при спільній роботі кількох комп'ютерів, що обмінюються даними);

– час виконання (час, необхідний працюючому програмному компоненту системи для завершення розв'язання певної задачі в певних умовах);

– час використання (час, що витрачається певним користувачем на вирішення завдань з допомогою ПЗ), наприклад: час сеансу роботи (від початку до завершення), час вирішення задачі і ін.;

**Міри зусиль.** Наводять корисний (продуктивний) час, пов'язаний з певним завданням проекту, наприклад:

– продуктивність праці (при індивідуальній роботі);

– трудомісткість (при колективній роботі);

**Міри інтервалів між подіями.** Це інтервали часу між настанням подій, що відбуваються в певний період спостереження, наприклад, час між послідовними відмовами. Замість цієї міри може використовуватися частота настання подій;

**Лічильні міри.** Являють собою статичні лічильники (для обліку певних елементів у робочих продуктах ПЗ (документах)) або кінетичні (динамічні) лічильники (для обліку подій або дій людини), наприклад:

– кількість виявлених помилок (облік помилок в ході інспекції, тестування, функціонування або супроводу ПЗ);

- структурна складність програми (кількість шляхів у програмі, цикломатична складність та ін.);
- кількість несумісних елементів (облік помилок узгодження (вимог, стандартів, форматів та ін.), облік відповідей типу «так» / «ні» у опитувальних);
- кількість змін – облік елементів конфігурації, в яких відбулися зміни;
- кількість виявлених відмов – облік відмов при тестуванні, функціонуванні чи супроводі ПЗ;
- кількість спроб – облік спроб корегування дефектів або помилок;
- ергономічні лічильники – число натиснень клавіш, клацань на кнопках та ін.
- лічильники-оцінки – облік результатів, зведених в опитувальних листах, контрольних листах та ін.)

При виконанні вимірювань базові міри розміру, часу і розрахунки можуть використовуватися в різних комбінаціях. Вони служать основою для нормалізації та забезпечують можливість зіставлення метрик.

### **3. Класифікація метрик якості**

Для зручності застосування загальних прийомів вимірювань метрики зазвичай класифікують як:

- об'єктивні/суб'єктивні. Об'єктивні метрики містять підрахунки елементів, які можуть бути незалежно перевірені (кількість рядків коду, кількість помилок, складність та ін.) Вони знижують вплив особистої думки на обчислення й аналіз метрик. Суб'єктивні метрики ґрунтуються на індивідуальному чи колективному розумінні або перевазі певних характеристик або умов (рівень складності проблем, вартісні коефіцієнти та ін.);
- примітивні/обчислювані. Примітивні (базові) метрики можна безпосередньо спостерігати (розмір програми в KSLOC, кількість дефектів, знайдених при тестуванні та ін.) Обчислювані метрики не можуть безпосередньо спостерігатися, але можуть обчислюватися по примітивним метрик (число дефектів, що припадають на SLOC, трудомісткість та ін.);
- динамічні/статичні. Динамічним метрикам властивий компонент часу. Значення змінюються з плином часу, починаючи з моменту збирання даних (наприклад, кількість

помилки в місяць). Статичні метрики інваріантні до часу (кількість виявлених дефектів, загальна трудомісткість робіт тощо);

- прогнозуючі/пояснюючі. Значення прогнозуючих метрик можуть бути отримані заздалегідь, наприклад, оцінювана інтенсивність відмов. Значення пояснюючих метрик з'являються пост-фактум, наприклад, реальна інтенсивність відмов.

По відношенню до виду об'єкта вимірювання (працююча програма або сукупність документів) міри та відповідні метрики підрозділяються на зовнішні, внутрішні і метрики використання ПЗ.

Зовнішні метрики використовують міри працюючого на комп'ютері програмного продукту, отримані в результаті вимірювання його поведінки в ході тестування і функціонування.

Зовнішні метрики розробляються з метою:

- демонстрації якості програмного продукту, за допомогою характеристик і підхарактеристик якості, на стадії тестування і експлуатації;
- використання для підтвердження (валідації) того, що програмний продукт задовольняє зовнішнім вимогам до якості;
- передбачення реальної експлуатаційної якості;
- визначення ступеня, в якій програмний продукт буде задовольняти встановленим і передбачуваним вимогам користувачів у ході реальної експлуатації.

Можна сказати, що сукупність зовнішніх метрик призначена для оцінювання зовнішньої якості – міри, в якій продукт задовольняє встановленим (заявленим) і що мають на увазі потребам при використанні в певних умовах.

Розробка зовнішніх метрик ґрунтується на виконанні наступних вимірювань:

- поведінки програмного продукту при тестуванні та функціонуванні в поєднанні з іншими програмними продуктами, апаратним забезпеченням або системою обробки інформації в цілому;
- поведінки користувача (сценаріїв використання ПЗ).

Під виміром (оцінкою) поведінки розуміється оцінка масштабів можливих наслідків неадекватної поведінки, які загрожують життю або здоров'ю людей, природним ресурсам, можуть призвести до руйнування даних, неузгодженості або недостовірності інформації, втрати безпеки, деградації сервісу (послуг), економічних втрат та ін.

Прикладами зовнішніх метрик для такої характеристики як надійність, можуть бути середній час між відмовами, кількість усунених дефектів при тестуванні, інтенсивність відмов та ін.

Внутрішні метрики забезпечують можливість користувачам, розробникам, тестувальникам і менеджерам оцінювати якість проміжних і кінцевих продуктів ПЗ безпосередньо за їх властивостями, без виконання на комп'ютері.

Внутрішні метрики розробляються таким чином, щоб вони могли:

- відображати якість проміжних і кінцевих програмних продуктів по тим характеристикам і підхарактеристикам якості, які визначені в моделі якості ПЗ;
- служити посібником до дії при плануванні та поліпшенні процесів, які впливають на проміжні та кінцеві продукти;
- використовуватися при верифікації того, що проміжні і кінцеві продукти задовольняють вимогам до внутрішньої якості ПЗ, передбаченим планами вдосконалення процесів;
- передбачати зовнішні метрики якості.

Можна сказати, що сукупність внутрішніх метрик призначена для оцінювання внутрішнього якості – безлічі атрибутів продукту, що визначає його здатність задовольняти встановленим або реальним потребам при використанні в певних умовах.

Розробка внутрішніх метрик ґрунтується на виконанні вимірювань статичних атрибутів, які визначені і можуть бути оцінені по тексту вихідного коду, графічному або табличному поданням потоків управління і даних, структур переходу станів або за документами ПЗ.

Прикладами внутрішніх метрик для надійності можуть бути кількість помилок, знайдених при інспекції коду, кількість усунених дефектів в результаті інспекції коду, прогнозована кількість помилок, що залишилися, та ін.

Метрики якості у використанні (метрики експлуатаційної якості) вимірюють ступінь, в якій програмний продукт, встановлений і експлуатований в певному середовищі, задовольняє потребам користувачів в ефективному, продуктивному і безпечному рішенні задач.

Метрики якості у використанні допомагають оцінити не властивості самої ПЗ, а видимі результати її експлуатації – експлуатаційну якість.

Очевидно, що для правильного вимірювання експлуатаційної якості важливо враховувати контекст застосування ПЗ – особливості категорій її користувачів, специфіку розв'язуваних ними завдань, а також фізичні та соціальні фактори середовища їх роботи.

Прикладами експлуатаційних метрик якості можуть бути точність і повнота досягнення цілей користувачів, продуктивність праці, ресурси, витрачені у зв'язку з досягненням ефективного вирішення завдань, думку користувачів щодо властивостей ПЗ та ін.

Внутрішні, зовнішні та експлуатаційні метрики якості взаємопов'язані. Досягнення експлуатаційної якості залежить від задоволення критеріїв зовнішньої якості, заснованих на зовнішніх мірах і метриках якості, які, в свою чергу, залежать від задоволення відповідних критеріїв внутрішньої якості, заснованих на внутрішніх мірах та метриках, пов'язаних із зовнішніми.

Зазвичай вимоги користувачів до якості специфікуються за допомогою зовнішніх метрик і експлуатаційних метрик якості, а внутрішні метрики вибираються таким чином, щоб вони могли використовуватися для передбачення значень зовнішніх метрик.

Побудувати строгу теоретичну модель, що встановлює взаємозв'язок зовнішніх і внутрішніх метрик, складно, тому, як правило, будується гіпотетична модель, взаємозв'язок метрик в якій моделюється статистично в ході використання метрик.

#### **4. Проектування метрик якості**

Для створення моделі якості ПЗ необхідно:

1. Визначити цілі вимірювання і виконувати всі наступні дії з урахуванням факторів, що впливають на якість.
2. Ідентифікувати компоненти ПЗ, якість яких повинна моделюватися окремо.
3. Виділити і класифікувати найбільш істотні компоненти, які несуть відповідальність за якість властивості (атрибути) для кожного компонента.
4. Сформулювати аксіоми для зв'язування властивостей ПЗ з характеристиками якості.
5. Визначити систему внутрішніх і зовнішніх метрик якості ПЗ, адекватну переліком характеристик якості і атрибутів ПЗ.

6. Перевірити модель якості.

Процес, що виконується при розробці власних або виборі існуючих метрик якості ПЗ, що відповідають цільовим вимогам до якості ПЗ (або її компоненту), містить наступні кроки:

**Крок 1.** Визначення понять. Визначення всіх використовуваних понять повинні відповідати зазначеним у застосовуваних стандартах або приводитися в документі, що описує прийняту модель якості ПЗ. Не можна допускати неоднозначності тлумачення термінів різними категоріями осіб, які використовують метрики (користувачами, замовниками, розробниками).

**Крок 2.** Визначення внутрішньої структури (моделі) кожної метрики. Значення базових метрик вимірюються безпосередньо і модель таких метрик – це найменування відповідної змінної. Значення складних метрик являють собою математичні моделі, що використовують базові або інші складні метрики. Переважний прагматичний підхід до моделювання метрик – метрики повинні відображати найбільш важливі аспекти вимірюваного атрибуту і не бути занадто складними. Моделі метрик можуть бути запозичені з стандартів, наукової літератури, з досвіду інших організацій і т.д. і повинні накопичуватися і при необхідності адаптуватися до потреб конкретних вимірювань. Для розроблення нових (оригінальних) метрик корисно залучати експертів в проблемній області проекту та інженерії якості ПЗ.

**Крок 3.** Формулювання методу обчислення метрики (критерію оцінювання). Модель кожної метрики декомпозується до рівня метрик-примітивів (базових метрик) і далі для цих примітивів визначається механізм отримання значення (критерій оцінювання). Приклад метрики-примітиву зі складним механізмом отримання значення – SLOC.

Метрики-примітиви та критерії їх оцінювання утворюють перший рівень необхідних даних, що збираються.

Методи визначення значень метрик, рекомендовані стандартом ДСТУ 2850, такі:

– вимірювальний. Метод ґрунтується на отриманні інформації про властивості та характеристики ПЗ з використанням вимірювальних технічних і програмних засобів (розмір завантажувального модуля, час виконання гілки програми та ін.);



– реєстраційний. Метод ґрунтується на отриманні інформації під час випробувань або при безпосередньому використанні ПЗ за призначенням, коли реєструються і підраховуються певні події (моменти часу збоїв, кількість відмов і ін.);

– розрахунковий. Метод заснований на використанні теоретичних і емпіричних залежностей на ранніх стадіях розробки, а також статистичних даних, що накопичуються при випробуваннях, експлуатації та супроводі. Цим методом прогнозуються характеристики і підхарактеристики на ранніх стадіях розробки (точність, стійкість, надійність та ін.) Розрахунковий метод використовується і для визначення фактичних значень характеристик за результатами тестування;

– експертний. Метод полягає у визначенні значень характеристик групою фахівців-експертів. Застосовується в тому випадку, якщо завдання визначення значень не може бути вирішена іншими методами. Для проведення експертизи встановлюються контрольовані ознаки, що містяться у вигляді запитань в аналітичній опитувальній анкеті експертів.

**Крок 4.** Визначення критерію «хорошого» значення метрики. Коли відомо, що вимірювати і як вимірювати, потрібно визначити, як інтерпретувати результати вимірювань, і якими мають бути межі їх змін.

Критерієм істини у багатьох випадках служить думка замовника і користувача. Можуть також використовуватися опубліковані відомості про прийнятні значення тих чи інших метрик, отримані в галузевій виробничій практиці значення та ін. Так, наприклад, відомо, що Цикломатична складність модулів повинна бути ≤ 10. Найбільш достовірне джерело інформації – власні історичні дані про вимір схожих проектів.

**Крок 5.** Документування метрик. На цьому кроці потрібно прийняти рішення про формат звіту, про цикл витягу та реєстрації даних (частоті вимірювань), про механізми обліку даних (тверда копія, електронна копія), про порядок їх розподілу (напрямки) різним учасникам процесу вимірювання, про обмеження (визначають готовність метрики до використання) та ін.

**Крок 6.** Визначення додаткових кваліфікаторів метрик. Метрика вважається добре спроектованою, якщо вона є узагальненням різних поглядів і рівнів вимірювання і якщо її можна застосовувати, наприклад, на рівні класу продуктів, одного продукту або

компонентів цього продукту ПЗ. В якості додаткових кваліфікаторів може вказуватися контекст застосування й інші обмеження.

Існує безліч різних метрик, які можуть мати цінність для керування сучасним процесом.

Якісна метрика володіє наступними основними характеристиками:

1. Метрика має сенс для замовника, менеджера та виконавця. Якщо хоча б одна із зацікавлених сторін не розглядає метрику як осмислену, вона не буде використовуватися.

2. Метрика показує кореляцію між змінами в процесі розроблення та ходом бізнесу. Єдиними реальними цілями та завданнями організації є фінансові: зниження витрат, підвищення доходів і збільшення прибутку.

3. Метрика повинна бути об'єктивною та її визначення недвозначно. Об'єктивність слід трансформувати в певну форму кількісної уяви (наприклад, певні значення, відсотки, частки). Невизначеність мінімізується за допомогою використання зрозумілих одиниць виміру (таких, як людино-місяці, SLOC, зміни, функціональні точки, класи, сценарії, вимоги), які не так просто визначити точно.

4. Метрика має показувати тенденції. Це дуже важлива характеристика. Розуміння змін значення метрики в залежності від часу, від проекту до проекту, від версії до версії і т.д. виявляється надзвичайно важливим, особливо для сучасних моделей ітераційного розроблення. Рідко трапляється, щоб деяка метрика безпосередньо приводила до виявлення яких-небудь конкретних дій. Швидше, метрика демонструє погляд з певної точки зору, та інтерпретувати цю метрику та визначати необхідні дії повинні менеджери проектів.

5. Метрика є природним побічним продуктом процесу. Це означає, що заради метрики не додаються нові робочі продукти або додаткові види діяльності; вона виводиться безпосередньо з самих робочих процесів.

6. Метрика може бути автоматизована. Досвід показує, що найбільш вдалими метриками є ті, інформація про які збирається та виявляється автоматично. Частково це так, тому що програмний інструментарій вимагає строгих визначень для даних, що оброблюються.