Хамза Гаркус, Рамез Рахман, Боян Карлас і Карл Аберер

**Цікавий випадок PDF- конвертера** **, якому подобається Моцарт: аналіз і зниження рівня конфіденційного ризику персональних хмарних додатків**

Анотація: деякі програми, що працюють на вершині окремих хмарних сервісів, таких як Google Drive та Drop-box, вимагають доступу до даних Користувача, для забезпечення певної функціональності. Завдяки детальному аналізу ста популярних додатків Google Drive із Google Chrome store ми з’ясували, що існуюча модель дозволів досить часто використовується неправильно: близько двох третин додатків analyzed є надмірно привілейованими, тобто вони отримують доступ до більшої кількості даних, ніж це необхідно для їх функціонування.

В цій роботі проаналізовано три різні моделі дозволів, які спрямовані на те, щоб перешкоджати користувачам встановлювати надмірно привілейовані програми. В експериментах з 210 реальними користувачами ми виявили, що найбільш успішною моделлю дозволів є наш новий метод, який ми називаємо далекосяжним розумінням. Далекосяжні попередження інформують користувачів про те, як такі додатки можуть вплинути на певні сторони їх конфіденційності (наприклад, їх топіки інтересів, колаборацій і структури активності тощо). Таким чином, вони прагнуть подолати розрив між тим, що треті сторони можуть фактично знати про користувачів, і сприйняттям користувачами їх витоку конфіденційності. Ефективність далекоглядних ідей у подоланні цього розриву демонструється нашими результатами, оскільки далекосяжні ідеї виявляються в середньому в два рази ефективніше, ніж поточна модель, в тому, що стосується відмови від встановлення надмірно привілейованих додатків. Прагнучи підвищити загальну обізнаність про конфіденційність, ми створили PrivySeal, загальнодоступний магазин додатків, орієнтований на конфіденційність, який використовує далекосяжні ідеї ґрунтуючись на знаннях, взятих з даних користувачів магазину (понад 115 гігабайт даних Google Drive від 1440 користувачів з 662 встановленими додатками), ми також обкреслюємо екосистему для сторонніх хмарних додатків з точки зору розробників і хмарних провайдерів. Нарешті, ми представляємо кілька загальних рекомендацій, які можуть направляти інші майбутні роботи в області конфіденційності хмари. Наскільки нам відомо, це перша робота, в якій так докладно розглядається ризик, пов'язаний з сторонніми додатками на хмарних платформах.

Ключові слова: хмарні обчислення, корисна конфіденційність, додатки

1. Вступ

Хмарні сервіси, такі як Google Drive, Dropbox, OneDrive та ін стають все більш популярними в останні роки. У той же час, такі послуги підвищили питання конфіденційності даних користувачів. Але небезпека серйозніша ніж здається. Поки такі хмарні сервіси мають чітко певну конфіденційність політики, вони також служать в якості платформ, які дозволяють працювати з безліччю інших додатків для роботи даними користувача. Ці сторонні додатки надають певні функціональні можливості для цього вони вимагають доступ до даних користувачів. Простіше кажучи, користувачі жертвують частиною своєї конфіденційності, щоб отримати функціональні можливості. Однак, схоже, що часто такі додатки отримують більше даних, ніж необхідно щоб вони функціонували. Як приклад ми зробили аналіз сотні додатків на одному з найпопулярніших персональних хмар сервісу, а саме Google Drive (240 млн активних користувачів у 2014 році [1]), і виявили, що майже дві третини (64%) з цих додатків вимагають більше дозволів, ніж їм фактично потрібно для роботи. Таким чином, користувачі часто в кінцевому підсумку надають більше даних, ніж необхідно. Наприклад, улюблений PDF конвертер користувача дуже ймовірно, може отримати доступ до його музичної бібліотеки і знайти його смак до музики Моцарта або отримати гео-прикріпленні фото і дізнатися що він відвідував на вихідних. Протягом цієї статті ми називаємо такі програми як надпривілейовані програми. Як спостерігається в інших екосистемах сторонніх додатків, надання таких надмірно привілейованим додаткам надлишкового доступу потенційно може призвести до зловживання даними користувачів. Нещодавно був випадок на ринку додатків для здоров'я, де топ-20 найбільш відвідуваних додатків були викриті у обміні даних користувачів з 70 аналітичними та рекламними компаніями.

Тим не менш, екосистема хмарних додатків має унікальний особливості, які вимагають особливого вивчення цієї екосистеми.

На відміну від досліджень з екосистем мобільних додатків, де дозволи стосуються списку контактів користувача, поточного розташування або фотографії, хмарні дозволяють іншим додаткам отримати доступ до будь-якого файлу у хмарці збереженого користувачем. Таким чином, замість профілювання контексту поточного користувача, такі додатки можуть отримати далекосяжні висновки з його документів. Наприклад що стосуються його фінансових, юридичних або, перспективи, пов'язані зі здоров'ям. Простіше кажучи, масштаб і якість даних, які можуть бути зібрані являє собою кошмар конфіденційності для необізнаних користувачів і золота жила для рекламодавців. По-друге, збір дозволів хмарних додатків є складним завданням. На відміну від інших екосистем, де дозволи додатку із тисячі додатків можуть бути легко добуті через традиційний пошук в інтернеті, кожний інший хмарний додаток має унікальний інтерфейс, який пов'язує з постачальниками послуг.

Отже, це обмежує корпус додатків, які можна вивчати. У цій статті ми прагнемо проаналізувати і пом'якшити загрози порушення конфіденційності, що породжуються такими більш привілейованими хмарними додатками. Ми вивчаємо екосистему додатків Google Drive з точки зору всіх відповідних сторін, а саме користувачі, розробники додатків і хмарні провайдери. Взагалі, ми прагнемо до характеристики впливових факторів, які можуть стримувати користувачів до установки більш привілейованих додатків, умови визначення неправильної поведінки розробників, кроки, які хмарні провайдери можуть зробити для зниження ризику конфіденційності користувачів.

Наскільки нам відомо, це наша перша робота яка вивчає ризик конфіденційності сторонніх додатків в особистих хмарних екосистемах. З цією метою ми представляємо три різних дозволи моделі, а саме: (а) Дельта дозволу, (б) негайно Прозріння і (с) далекосяжні прозріння. Перша модель, тобто різницеві дозволи, інформує користувачів про непотрібні дозволи, які використовують привілейовані додатки.

Друга модель, безпосередні висновки представлені у випадковому порядку вибрані приклади з даних Користувача, такі як частини тексту або файлів зображення, фото місця і т. д., для яких привілейовані додатки можуть отримати доступ. Далекосяжне розуміння, було мотивоване концепцією новели про Зворотної конфіденційності Зворотна конфіденційність відноситься до ситуації, коли користувач не в курсі того що зовнішній об'єкт має інформацію про Користувача. Ґрунтуючись за цим визначенням, далекосяжні ідеї спілкуються з користувачам висновками, які можуть бути зроблені додатками із зайвими дозволами, використовуючи розширений текст і методи аналізу зображень. Вони включають, але не обмежується спільної роботою користувачів і моделями діяльності; верхні грані, розташування і поняття, які відображаються в фотографії і т. д.

В цілому, ми вносимо наступні конкретні вклади у цій роботі:

Я. далекосяжні ідеї підвищують чутливість користувачів інтимні деталі, і сприяти конфіденційності обізнаної поведінки: завдяки великим користувальницьким експериментам ми виявляємо що наша перша проста модель, Дельта дозволу провалює утримання користувачів від установки надмірно привілейованих додатків. Прямо, кажучи користувачам, що їх конфіденційність порушується не допомагає. Друга модель, негайної проникливості, роблять двічі також у бентеженні користувачів від установка привілейованих додатків. Однак явний переможець наша романна модель, далекосяжні проникливості, яка може будьте в два рази ефективніше в утриманні користувачів від установки за правами додатків в негайні висновки. Наприклад, ми виявляємо що в рамках далекосяжних ідей, реляційних ідей (які показують відносини користувачів з іншими людьми) зменшилась установка привілейованих додатків у 2 рази, в порівнянні до особистих прозрінь (які розкривають інформацію про самого користувача).

PrivySeal допомагає нам профілювати поведінку розробників і допомагає користувачам захистити свої дані: зусилля задля підвищення обізнаності громадськості про конфіденційність, і щоб допомогти користувачам захистити свою конфіденційність, ми представляємо PrivySeal, орієнтований на конфіденційність магазин додатків, який використовує Далекі досягнення, що попереджає користувачів про надмірні привілеї застосування. Цей магазин доступний для громадського користування і був використанний більш ніж 1440 зареєстрованими користувачами. Значна частина з цих користувачів має попередній досвід використання сторонніх додатків Google. Автоматично отримуючи метадані раніше встановлених додатків ми розкриваємо поточні поведінки розробників, які вказують на потенційні можливості нещастя, і представляємо пропозиції щодо стримування поганої поведінки. Загальна мудрість: нарешті, на основі нашого аналізу ми представляємо кілька простих у реалізації практичних пропозицій це може бути прийнято хмарними провайдерами і ті, хто працює в області конфіденційності для захисту конфіденційність користувачів у хмарі. частина паперу організована наступним чином.

Розділ 2 описує екосистему сторонніх хмарних додатків, його модель загроз і конкретний випадків Google Drive.

У розділі 3 ми детально описуємо наші дозволи додатків конфіденційності огляд процесу і результатів. У розділі 4, ми уявляємо наші три моделі дозволів, перш ніж оцінювати їх в розділі 5. На основі магазину, орієнтованого на конфіденційність ми розробили (Розділ 6), ми аналізуємо App developer поведінку в Розділі 7. Нарешті, ми даємо наші рекомендації для спільноти в розділі 8.

2 сторонніх хмарних додатки Екосистема

Є три сутності, які взаємодіють у системах сторонніх хмарних додатків:

(1) розробник, який програмує і управляє стороннім додатком, (2) користувач, який використовує цей додаток для досягнення певного сервісу і (3) постачальник хмарного сховища (CSP), що розміщує дані користувача. Використовуючи криптопровайдери.

API, додаток отримує доступ до підмножини даних користувача після авторизації користувача, яка заснована на прийнятті користувачем список дозволів, що визначає ця підмножина.

2.1 Модель загроз

При використанні сторонніх хмарних додатків, які отримують доступ до своїх даних, користувачі жертвують частиною своєї конфіденційності для отримання деяких послуг цей компроміс між конфіденційністю та послугами була названа дилемою конфіденційності проти послуг в літературі[5].

Однак, як ми побачимо в наступному розділі, є багато додатків, які вимагають більше дозволів, ніж необхідні для їх функціонування. Ми називаємо такі додатки заможними додатки (на відміну від найменш привілейованих додатків, які запитувати тільки дозволи, необхідні для їх функціональності). Ці програми представляють ризик, який може бути потенційно використовуватися, наприклад, шляхом продажу даних до сторонніх рекламодавців. Багато з широко використовуваних сторонніх додатків не мають політику конфіденційності або обґрунтованого запитування на дозвіл. Крім того, користувачі зазвичай не знають деталей API або функціональності програми, особливо до установки додатків. Таким чином, вибір установки не добре поінформований з точки зору конфіденційності. У цій роботі, ми розглядаємо сторонні додатки в якості противника (а не DCP). Ми прагнемо боротися з ризиком поставленими надмірно привілейованими додатками з а рахунок поліпшення показників ризику(зокрема, моделі дозволів), що користувачі представлено під час процесу авторизації.

2.2 Справа Google диску

З цією метою ми взяли Google диск в якості тематичного дослідження, і ми анатомували цю екосистему в деталь. Проте, висновки, отримані в результаті аналізу застосовані і до інших хмарних платформ.

Насамперед, будь-який розробник може зареєструвати додаток, який звертається до Google диску API в консолі розробників Google безкоштовно. Потім вона отримує ідентифікатор клієнта і секрет клієнта який повинен бути включений в код програми для доступу API Google. Потім розробник може вказати в своєму коді набір дозволів Google (або scopes) які вона хоче отримувати.

Сам додаток може бути розміщено на будь-якому веб-сайті за вибором розробника; тобто воно не розміщується самим Google. Розробник також може подати заявку на участь у додатку Google Chrome Web Store, в якому є розділ для додатків, що працюють з Google Диском. В магазині, додатки представлені разом зі скріншотами і описом їх функціоналу (надається розробником). Магазин також дозволяє користувачам оцінювати і переглядати додатки. Програми також можуть бути відправлені в інші розміщені веб-магазини від Google, такі як магазин аддонів для Google документи, Google таблиці або Google презентації та додатки Google

Ринок для підприємств. Тим не менш, є багато додатків, які існують поза цими магазинами. Додаток може запросити дозвіл на доступ до даних Google Диску в будь-який час його роботи, і не обов'язково на початку. Наприклад, користувач може бути представлений з кнопкою в бічному меню, яке читає " імпорт файл з Google Диска", і натиснувши на цю кнопку перенаправляє на сторінку Google, на якій представлений набір дозволів, запитаних додатком, як показано на малюнку 1. Користувач повинен прийняти всі ці дозволи задля підключення додатку до його Google диску. Він не може вибрати їх підмножину під час установки або пізніше.

Однак, він може повністю скасувати авторизацію програми від його облікового запису Google. Як ми побачимо пізніше, відсутність стандартного розташування і інтерфейсу для розміщення додатків і запуск запиту дозволів є однією з причин, яка робить автоматизований, широкомасштабний аналіз конфіденційності додатків нездійсненним. Основні дозволи, що відносяться для Google Диска представлені в таблиці 1, поряд з описом, наданим Google для кожного. Цей короткий опис також представлений користувачеві і доступне більш довге пояснення , натиснувши на кнопку “Інформація” i далі до кожного дозволу.

Що стосується даних файлів, додаток може запросити доступ до всіх файлів (дозволи drive і drive\_readonly) або на основі кожного файлу (drive\_file). Коли розробники запросять тільки файл drive\_file, явне твердження для кожного нового файлу(ів) опосередковується наданим інтерфейсом на Google. Наприклад, розробник представляє користувачу спливаюче вікно вибору яке розміщено Google. 2) щоб він зміг обрати (і таким чином затвердити доступ) до файлу. Крім того, файл може бути відкритим з інтерфейсу Google Диска через опцію «Відкрити за допомогою» у контекстному меню файлу. У випадку повного доступ, додаток може отримати доступ до будь-якого файлу безпосередньо через Google Drive API без необхідності втручання користувача.

Наприклад, цей тип доступу включає додаток щоб отримати всі файли користувача і завантажити їх фон. Розробник може альтернативно запросити доступ до метаданих файлів тільки через drive\_metadata або drive\_metadata\_readonly (дозволяє отримати доступ до імен файлів, редагування дат, EXIF-інформації фотографій і т. п.).

Крім того, розробник може запросити доступ до списку додатків користувач авторизований черезdrive\_apps\_readonly. Варто відзначити, що список дозволів не обмежується Google Drive API і що це, як правило, включає дозволи від інших сервісів Google API, такі як доступ до інформації профіля користувача, електронна пошта адреса, список контактів і т. дп.

3 Ризик конфіденційності третьої сторони Google Drive Apps

Наступне питання, яке приходить у голову : «який ступінь ризику, якому піддаються фактичні користувачі?» Щоб відповісти на це питання, ми розглянули зразок сторонніх дискових додатків, щоб визначити відсоток додатків, які вимагають додаткові дозволи. Ми перейшли до Google Веб-магазину Chrome, в якому є розділ для додатків, які працюють з Google Диском.

У магазині є додатки на його головній сторінці, які змінюють час. В цілому, там близько 420 додатків в магазині, які позначені як «Працює з Google Диском». Ми вибрали 100 зв’язаних додатків з випадкової головної сторінки (протягом травня 2015), і ми вручну переглядали їх один за іншим. Отже, наш зразок представляє навколо 1/4 з усього набору додатків в магазині, який є одним з основних напрямів пошуку додатків Google Диска. Як ми обговоримо далі в розділі 7.1, ми виявили, що в нашому прикладі реального світу з 1440 користувачів, близько у однієї чверті встановлені додатки з магазину Google Chrome. Більш детальну інформацію про

набір даних додатків представлений в додатку A.

3.1Процес перевірки дозволів

Тепер ми пояснюємо огляд дозволів програми (APR)методології ми дотримувалися, і відсилаємо читача до малюнка 15 в додатку A для відповідної блок-схеми. Наша методологія натхненна керівництвом Google Drive для вибору областей автентифікації. Кожен АТР прагне отримати: (a) набір P запитуваних дозволів, (b) набір s достатніх дозволи для програми. Для кожної програми ми спочатку знаходимо крок, до якого додаток може бути підключений Google Диск (якщо це не при первісні реєстрація). Потім ми записуємо набір P запитаних дозволів і авторизації програми для доступу до тестового облікового запису Google Drive створений для цієї мети, і ми записуємо запрошені дозволи.

Якщо drive\_file (тобто мінімальний доступ на файл) це єдиний запрошений дозвіл Google Drive то, перевірка додатку буде завершена (S = {drive\_file}). В іншому випадку, ми продовжуйте перевіряти інтерфейс програми для всіх засобів вибору файлів що дозволяє імпортувати файли з Google Диска (майже у всіх випадках є не більше одного засобу вибору файлів).

У першому випадку, коли додаток використовує виключно офіційний вибір Google файлів як на малюнку 2 (наприклад, додаток, який дозволяє користувачам конвертувати певні файли у формат PDF), ми встановлюємоS = {drive\_file}.

У другому випадку, коли ми знаходимо що в інтерфейсі немає засобів вибору файлів і що функціональність додатку не вимагає доступу до файлів, додаток позначено як не вимагає будь-яких дозволів на файли (S = {}). У разі якщо додаток включає в себе користувальницкий вибір файлів, ми вирішили, що (a) S = {drive}, якщо оголошена функціональність додатку вимагає файловий зміст (наприклад, додаток для колажу фотографій з користувацьким браузером фотографій) або (b) S = {drive\_metadata}, якщо функціональність не потрібна (наприклад, програма, яка візуалізує, хто має доступ до вибраної папки). Аналогічно, якщо додаток не має засобу вибору файлів, ми вирішили, що (a) S = {drive}, оголошена функціональність додатку вимагає файловий зміст (наприклад, шкідливі програми сканування додатків для Google Drive, яким не потрібно засіб вибору файлів) або (b) S = {drive\_metadata} иякщо функціональність вимагає метаданих файлу (наприклад, додаток, який візуалізує всіх співробітників з доступом до файлів користувача). Нарешті, ми позначаємо додаток як надмірно привілейований a) s порожнього напрям S є порожнім а P ні, або б) якщо напрям P містить хоча б один дозвіл , що є більш вимогливим ніж всі дозволи в S. Ми також визначаємо набір непотрібних дозволів U, що складається з кожного дозволу в P, який є більш вимогливим, ніж всі дозволи в S. Набір необхідних дозволів дається по N = P \ U.

3.2 Результати аналізу

Аналізуючи APRs, ми з’ясували, що 64 із 100 додатків запитують непотрібні дозволи. Іншими словами, розробники могли б запросити менш інвазійні дозволи з поточним API, наданим Google. У підсумку, 76 з 100 додатків запросили повний доступ до усіх файлів в Google диску користувача. Крім того, 64 надпривілейованих додатків, насправді запросили повний доступ. Таким чином, в нашій вибірці, близько 84% (64/76) додатків, що запитують повний доступ, є надмірно привілейованими. Верхній дозвіл, який запитується без необхідності представляє повний доступ на читання і запис на Google Диск (в 55 додатках), а потім повний дозвіл на читання (в 17 додатках). Це далі буде збільшувати обсяг даних, які можна використовувати з додатковими дозволами. З іншого боку, доступ по одному файлу є найкращим доступом який насправді потрібен при запиті. Це відбувається у 41 програмах. Однак у 16 з них 41 додатки, ми виявили, що розробник також просить повний доступ до даних користувача. Відповідно, розробники іноді змішуючи повний доступ з частковим доступом (який є підмножиною першого). Відзначимо, що таке змішання дозволів може бути результатом некомпетентності розробника, або це може бути спрямоване на обман користувача. Незважаючи на це, такі програми становлять ризик, який може бути потенційно експлуатований. Інший результат APR був drive\_file, який був кращим альтернативним дозволом (в 48 додатках), які можуть замінити непотрібні запити дозволу. drive\_metadata\_readonly був альтернативою тільки для однієї програми. Це вказує на те, що просто правильне використання поточного Google Drive API (який забезпечує доступ в-файлу), може виключити більшу частину ризику конфіденційності. Тим не менш, очевидно, що розробники, як правило, винні в тому, що не роблять цього.

3.3 Автоматизація процесу АТР

Будучи третьою стороною, ми не маємо доступу до повного список додатків Google Drive з їх дозволами. Отож, перше завдання, треба було знайти місце в кожен додаток, де запитуються дозволу Google Диска. Це не завжди на головній сторінці програми, і іноді для його пошуку потрібна навігація декількох меню і / або сторінок.

Автоматизація цього завдання включає побудову розширеного веб-шукача, який може отримати дозволи з великої кількості таких додатків шляхом розумного пошуку для кнопки входу в систему. Другим завданням була перевірка функціональності програми, щоб побачити, чи відповідає вона поданим дозволам. Автоматизація процесу занадто привілейованого пошуку або пошуку витоку особистих даних в реальному часі було вирішено в сценарії мобільних додатків (наприклад, в [10] і [9]). Проте в мобільному сценарії (або будь-якої подібної архітектури), пристрій користувача хостує дані, сторонні програми та виявлення / моніторинг рішення. Хмарні додатки представляють зовсім інший сценарій оскільки дані розміщуються у CSP, додаток третьої сторони подається в зазначеному розробником місці, і будь-які виявлення / моніторинг додатка буде працювати ззовні.

Єдина частина коду, яка надається сторонньому додатку це код на стороні клієнта. Отже, всі методи, які перевіряють код програми (наприклад, за допомогою статичного / динамічного аналізу) або свої вхідні сигнали / виходи не можна трансплантувати до хмарних програм, оскільки вони, очевидно, недооцінюють дозволи які можуть знадобитися API додатку.

Єдиний автоматичний спосіб виявлення надмірних привілеїв полягає в кластеризації додатків аналогічної функціональності та ідентифікації ті, які запитують більше дозволів, ніж інші той же кластер. Навіть тоді дані, які збиралися вручну будуть використовуватися в якості наземного оцінювання автоматизованого методу. Виявлення фактичного витоку даних є набагато складнішим завданням з хмарним додатком. Додаток може відправляти дані користувачів іншим сторонам на стороні сервера (який неможливо контролювати через зовнішне рішення.) Зіткнувшись з цими обмеженнями, відгуки експерта це найближче що ми можемо отримати для оцінки необхідних дозволів додатку. Тим не менш, ми не стверджуємо, що цей метод є абсолютно точним, як розробник міг би працювати над неоголошеною функцією, яка вимагає новий дозвіл. Однак ми припускаємо, що APRs є точною з переважною більшістю розглянутих apps1. Нарешті, нашою основною метою в даній роботі є характеристика екосистеми і запропонування альтернативних дозволів моделей, що автоматизують колекцію дозволів додатків і завдання виявлення понад-привілеї падають з

обсяг цієї роботи. Ми відзначаємо, однак, що ми одночасно працюємо над конкретною дослідницькою проблемою проектування автоматизованих APRs.

4 Нові моделі дозволу

У світлі ризику, який представляють надмірно привілейовані додатки, ми пропонуємо в цьому розділі три альтернативи існуючим моделям дозволу в Google Диск, перш ніж давати оцінку їх ефективності у зниженні ризику в наступному розділі.

4.1 Дельта дозволу

Наша перша модель заснована на наступній гіпотезі: «Коли користувачі отримують інформацію про непотрібні дозволи будучи запитаними додатками, вони з меншою ймовірністю авторизують такі додатки. «Отже, ця модель замінює поточний дозвільний інтерфейс, показаний на малюнку 1 на новий інтерфейс, представлений на малюнку 3. Ми називаємо це моделлю Дельта дозволу( DP), і вона показує для користувача відмінність між дозволами, які необхідні для функціональності програми і тих, інших ( Дельта), які надмірно запитуються.

4.2 Безпосередні висновки

Друга модель заснована на наступній гіпотезі: «Коли користувачам показані зразки даних, які можуть витягуватися з непотрібних дозволів, наданих додатками, вони з меншою ймовірністю авторизують ці додатки.» Відповідно, ми показуємо користувачам випадково вибрані приклади даних, взяті безпосередньо з Google Диска, наприклад уривки з текстових або графічних файлів, місцеположень фотографій або людей з якими вона співпрацювала.

Примірник цієї моделі, який ми називаємо миттєвим прозрінням (IM), наведеним на малюнку 4. Зліва у нас той самий попередній інтерфейс DP. Справа у нас є область, де ми показуємо питання, яке говорить: «що непотрібні дозволи кажуть про тебе?», за яким слідує відповідь-форма - наочна з коротким пояснювальним текстом.

У цьому малюнку, область проникливості візуалізує місце, де була зроблена випадково обрана фотографія користувача. В наступному, ми описуємо IM дизайн:

Зображення: ми показуємо зображення, вибране випадковим чином набір файлів зображень користувача.

Розташування: ми випадковим чином вибираємо фотографію з набору файлів зображень користувача, такі, де він включає в себе місце розташування GPS в його Exif даних. Потім ми покажемо це фото на карті по центру в цьому місці (як на малюнку 4).

Текст: показуємо користувачеві уривок з початку випадково вибраного текстового файлу.

Співробітник: ми показуємо зображення профілю і ім'я випадково обраного співробітника.

4.3 Далекосяжні висновки

Третя модель заснована на наступній гіпотезі:

«Коли користувачам показують далекосяжну інформацію з якої можна зробити висновок про непотрібні дозволи надані додаткам, вони з меншою ймовірністю авторизуються в цих додатках. Це ідеї, які виходять за рамки прикладів і включають в себе те, що можна зробити, запустивши більш ускладнені алгоритми, такі як настрої по відношенню до сутностей, об'єктів виявлені у фотографіях, виявлених осіб і т. п. Отже, позначимо цю модель, як далекосяжну ідею (або коротко FR Інтуїтивна.) Інтерфейс такий же, як на Малюнку 4, але з областю висновку, що містить FR висновок замість безпосереднього висновку. У цій роботі, ми розробили 6 типів FR, які можуть бути взяті з даних користувачів. Нижче ми коротко опишемо кожен з них, і ми відсилаємо читачів до додатку B для використовування алгоритму і для їх генерації: Сутності, концепції та теми (ECT): перший тип ідей, який ми формуємо, заснований на застосуванні різних Методів НЛП для вилучення іменованих сутностей (E) , понять (C) і теми (T) з текстових файлів користувачів. Сутність, витягнута за допомогою розпізнавання іменованих сутностей, можуть включати людей, з якими працює користувач, компанії, про які він говорить, місця які він планує відвідати і т. п. Концепції являють собою витягнуті теги із документів користувачів Наприклад, речення «Мої улюблені бренди-BMW, Ferrari і Porsche», буде позначено як поняття «автомобільна промисловість». Теми абстракції більш високого рівня (наприклад, технології, мистецтво, бізнес і т. п.), які можна використовувати для класифікації документів користувачів. Ми об’єднуємо їх разом за подібною поведінкою цих прозрінь. Коли ми використовуємо ECT, одне з E, T або C випадково показано споживачеві всередині області висновків. Для цілей аналізу тексту в цій роботі, ми частково використовували службу AlchemyAPI з випадковими фрагменти тексту, витягнутого з документів. Користувачі в наших експериментах були проінформовані про це на головній сторінці веб-програми, в яку вони входять. Сентименти: ми використовували аналіз настроїв для того, щоб виявити осіб з найбільш позитивним або негативним почуття і показати їх користувачеві у картинці 6. Кращі співробітники: наступне розуміння ми додали на дисплей кращих співробітників, які є у користувача, на основі проаналізованого файлу. До них зазвичай відносяться близькі колеги по роботі, близькі друзі або люди, з якими користувач відпочиває і потім ділиться фотографіями. Загальні інтереси: в цьому сенсі ми представляємо загальні теми інтересів з групою людей. Як показано на рис. 7, панель висновків буде містити список користувачів і різні розділи документа, з якими ділилися ці люди. Особи з контекстом: це користувальницькі образи.

Цей перший висновок такого типу показує групу граней, які представляють постійні люди, які з’являються в звичайних зображеннях, поряд з поняттями, які з’являються в тому ж зображенні (див. 8). Ви можете собі уявити, що така інформація це може бути корисно, наприклад, для рекламодавців які спрямовані на отримання інтересів користувача в певних людях, продуктах або послугах. Особи на карті: крім змісту зображення, метадані зображення також можуть бути чутливими, особливо географічне місце зйомки зображення. Так, це розуміння, показане на малюнку 9, передає користувачеві місця, де його фотографують, крім особи та об'єктів на цих фотографіях.